TOME CXII 2024

Multilingualism and Language Variation in English across Genres and Registers

Mémoires de la Société Néophilologique de Helsinki

Mémoires de la Société
Néophilologique de Helsinki
Tome CXII

éditées par Päivi Pahta,
Meri Larjavaara et
Christopher M. Schmidt

Mémoires de la Société
Néophilologique de Helsinki
Tome CXII

Multilingualism and Language Variation in English across Genres and Registers

A Festschrift in Honour of Päivi Pahta

edited by
Jenni Räikkönen, Carla Suhr, Minna Palander-Collin, Arja Nurmi,
Minna Nevala & Turo Hiltunen

Helsinki 2024
Société Néophilologique

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Société Néophilologique de Helsinki
https://edition.fi/ufy
www.ufy.fi

Layout Design: Camilla Pentti

ISBN
978-951-9040-71-4
ISSN: 2984-0961 (online
ISSN: 0355-0192 (print)
Société Néophilologique de Helsinki
Helsinki 2024

htpps://doi.org/10.51814/ufy.1041

Content

Authors and editors

Anna Havinga
Senior Lecturer in Sociolinguistics, University of Bristol

Turo Hiltunen
Senior University Lecturer of English language, University of Helsinki

Alpo Honkapohja
Lecturer of English, Tallinn University

Andreas H. Jucker
Professor Emeritus of English Linguistics, University of Zurich

Hanna Limatius
Postdoctoral Researcher, University of Helsinki

María José López-Couso
Professor of English Philology, University of Santiago de Compostela

Belén Méndez-Naya
Professor of English Philology, University of Santiago de Compostela

Martti Mäkinen
Associate Professor of English Language, Hanken School of Economics

Minna Nevala
Professor of English Philology, University of Helsinki

Terttu Nevalainen
Professor Emerita of English Philology, University of Helsinki

Juhani Norri
University Lecturer of English Language, Tampere University

Arja Nurmi
University Lecturer of English Translation, Tampere University

Minna Palander-Collin
Professor of English Philology, University of Helsinki

Jenni Räikkönen
Postdoctoral Researcher, University of Helsinki

Delia Schipor
Associate Professor, University of South-Eastern Norway

Janne Skaffari
University Lecturer of English Language, University of Turku

Carla Suhr
University Lecturer of English Philology, University of Helsinki

Irma Taavitsainen
Professor Emerita of English Philology, University of Helsinki

Laura Wright
Professor of English, University of Cambridge

Tabula gratulatoria

Karin Aijmer, University of Gothenburg
Marc Alexander, University of Glasgow
Lois Xacobe Atanes Barciela, Tampere University
Anita Auer, University of Lausanne
Joan C Beal, University of Sheffield
Kristin Bech, University of Oslo
Maija Belliard, Tampere University
Ylva Biri, University of Helsinki
Laurel J. Brinton, University of British Columbia
Nicholas Brownlees, University of Florence
Marcus Callies, University of Bremen
Nuria Calvo Cortés, Complutense University Madrid
Claudia Claridge, University of Augsburg
Begoña Crespo, University of A Coruña
Isabel de la Cruz-Cabanillas, University of Alcalá
David Denison, University of Manchester
Marina Dossena, University of Bergamo
Hannele Dufva, University of Jyväskylä
María José Esteve-Ramos, Universitat Jaume I
Teresa Fanego, University of Santiago de Compostela
Anita Fetzer, University of Augsburg
María F. García-Bermejo Giner, University of Salamanca
Mehdi Ghasemi, University of Turku
Pentti Haddington, University of Oulu
Soili Hakulinen, Tampere University
Tuomo Hiippala, University of Helsinki
Risto Hiltunen, University of Turku
Turo Hiltunen, University of Helsinki
Maija Hirvonen, Tampere University
Sebastian Hoffmann, Trier University
Sari Hokkanen, Tampere University
Niina Hynninen, University of Helsinki
Juhani Härmä, University of Helsinki
Giovanni Iamartino, University of Milan
Andreas H. Jucker, University of Zurich
Leena Kahlas-Tarkka, University of Helsinki

Samuli Kaislaniemi, University of Eastern Finland
Mark Kaunisto, Tampere University
Nely Keinänen, University of Helsinki
Anne Ketola, Tampere University
Zea Kingelin-Orrenmaa, Tampere University
Tuija Kinnunen, Tampere University
Monika Kirner-Ludwig, University of Innsbruck
Marja Kivilehto, Tampere University
Juhani Klemola, Tampere University
Sini Knuutila, Tampere University
Johanna Koivisto, Tampere University
Joanna Kopaczyk, University of Glasgow
Kaisa Koskinen, Tampere University
Dinah Krenzler-Behm, Tampere University
Igor Kudashev, Tampere University
Roman Kushnir
Merja Kytö, Uppsala University
Samu Kytölä, University of Jyväskylä
Leila Kääntä, University of Jyväskylä
Mikko Laitinen, University of Eastern Finland
Daniela Landert, Heidelberg University
Meri Larjavaara, Åbo Akademi University
Unni Leino, Tampere University
Sirpa Leppänen, University of Jyväskylä
Annikki Liimatainen, Tampere University
Aatu Liimatta, University of Helsinki
Niina Lilja, Tampere University
Hanna Limatius, University of Helsinki
Pekka Lintunen, University of Turku
Elisabetta Lonati, University of Eastern Piedmont
María José López-Couso, University of Santiago de Compostela
Lucía Loureiro-Porto, University of the Balearic Islands
Kathryn Lowe, University of Glasgow
Ursula Lutzky, Vienna University of Economics and Business
Sirje Lällä, Tampere University
Kevin J. McGinley, Tampere University
Belén Méndez-Naya, University of Santiago de Compostela
Mikhail Mikhailov, Tampere University
Colette Moore, University of Washington

Isabel Moskowich, University of A Coruña
Liisa Mustanoja, Tampere University
Katja Mäntylä, University of Jyväskylä
Minna Nevala, University of Helsinki
Terttu Nevalainen, University of Helsinki
John G. Newman, University of Texas Rio Grande Valley
Tarja Nikula-Jäntti, University of Jyväskylä
Sara Norja, University of Turku
Juhani Norri, Tampere University
Paloma Núñez-Pertejo, University of Santiago de Compostela
Arja Nurmi, Tampere University
Jopi Nyman, University of Eastern Finland
Henrik Oksanen, Tampere University
Mari Pakkala-Weckström, University of Helsinki
Minna Palander-Collin, University of Helsinki
Outi Paloposki, University of Turku
Matti Peikola, University of Turku
Eppu Penttilä, University of Eastern Finland
Carol Percy, University of Toronto
Mika Perkiömäki, University of Helsinki
Elizabeth Peterson, University of Helsinki
Bo Pettersson, University of Helsinki
Laura Pihkala-Posti, Tampere University
Arja Piirainen-Marsh, University of Jyväskylä
Anne Pitkänen-Huhta, University of Jyväskylä
Merja Polvinen, University of Helsinki
Elina Ranta, Tampere University
Saara Ratilainen, Tampere University
Paula Rautionaho, University of Eastern Finland
Ewald Reuter, Tampere University
Paul Rickman, Tampere University
Javier Ruano-García, University of Salamanca
Juhani Rudanko, Tampere University
Jenni Räikkönen, University of Helsinki
Tiina Räisänen, University of Oulu
Veera Saarimäki
Maria Salenius, University of Helsinki
Markku Salmela, Tampere University
Christina Samson, University of Florence
Mónica Sánchez Torres, Tampere University

Christopher Schmidt, Åbo Akademi University
Elena Seoane, University of Vigo
Elina Siltanen, University of Eastern Finland
Janne Skaffari, University of Turku
Jeremy J. Smith, University of Glasgow
Erik Smitterberg, Uppsala University
Hilkka and John Stotesbury, University of Eastern Finland
Cristina Suárez-Gómez, University of the Balearic Islands
Carla Suhr, University of Helsinki
Tanja Säily, University of Helsinki
Irma Taavitsainen, University of Helsinki
Sanna-Kaisa Tanskanen, University of Helsinki
Elina Tergujeff, University of Jyväskylä
The Modern Language Society of Helsinki
Ingrid Tieken-Boon van Ostade, University of Leiden
Olga Timofeeva, University of Zurich
Jarkko Toikkanen, University of Oulu
Johanna Vaattovaara, Tampere University
Elina Valovirta, University of Turku
Mari-Liisa Varila, University of Turku
Turo Vartiainen, University of Helsinki
Anu Viljanmaa, Tampere University
Tuija Virtanen-Ulfhielm, Åbo Akademi University
Matylda Włodarczyk, Adam Mickiewicz University in Poznań
Laura Wright, University of Cambridge
Brita Wårvik, Åbo Akademi University
Nuria Yáñez-Bouza, University of Vigo
Aleksandr Zelenin, Tampere University

Introduction

1. Multilingualism and language variation in English across genres and registers

A Festschrift by its nature reflects the career, accomplishments and legacy of the recipient. In the case of such a versatile researcher as Päivi Pahta, this creates a plethora of options. Her work on the history of English alone has provided many landmark studies, starting from her doctoral dissertation *Medieval Embryology in the Vernacular: The Case of* De spermate (Pahta 1998). The range from manuscript studies to corpus linguistics shows the kind of adaptability that few possess, and through participation in many projects, the focus of Pahta's research has moved from the history of English to more contemporary linguistic patterns of use and their social context. The chapters in the Festschrift are written to honour Pahta's work in the areas of multilingualism and multilingual practices, scientific and medical writing, and corpus approaches to genres and registers.

Pahta's key areas of interest are related to how language works in meaning-making practices in various social and communicative contexts, with a specific focus on multilingualism; specialised languages, especially scientific writing; and digitally enhanced methodologies, especially corpus linguistics. Characteristic of her research is a historical or long-term diachronic view on language practices, which is evident in her most important publications. Recent international co-publications, showing Pahta's broad area of competence, include *The Cambridge Handbook of English Historical Linguistics* (Kytö & Pahta 2016), *Multilingual Practices in Language History: English and Beyond* (Pahta, Skaffari, & Wright 2018), *Challenging the Myth of Monolingual Corpora* (Nurmi, Rütten, & Pahta 2018) and *Corpus Approaches into World Englishes and Language Contrasts* (Parviainen, Kaunisto, & Pahta 2019).

In particular, Pahta has advanced the study of multilingualism as one of the pioneers in the field of code-switching in Middle English (Pahta 2004) and in that of Early Modern English medical writing (Pahta 2011). For Middle English, her work has also highlighted the role of translation in a multilingual society and in the way it impacts linguistic usage (Pahta 1998; Pahta & Carillo

https://doi.org/10.51814/ufy.1041.c1453

Linares 2006). Beyond that, Pahta has expanded on the study of multilingualism over the course of the history of English from Old English to Late Modern English, covering a variety of genres and social contexts, such as early correspondence and literature (Nurmi & Pahta 2004; Pahta & Nurmi 2006; Nurmi, Tyrkkö, Petäjäniemi, & Pahta 2018).

Pahta's work on multilingualism on a societal level has also reached contemporary usage through her work towards implementing a major survey of English in Finland (Leppänen, Pitkänen-Huhta, Nikula, Kytölä, Törmäkangas, Nissinen, Kääntä, Räisänen, Laitinen, Pahta, Koskela, Lähdesmäki, & Jousmäki 2011). Her other scholarly contributions in this area include an exploration into language ideological stances (Leppänen & Pahta 2012) and a pioneering appraisal of the global and local meanings of English in the Finnish context (Taavitsainen & Pahta 2008).

A big bulk of Pahta's research has focused on the study of scientific and medical language, starting from editing a Middle English manuscript on embryology for her doctoral dissertation, accompanied with a comprehensive scholarly commentary and elucidation of its transmission of history (Pahta 1998). This area of her studies has also paved the way for an ambitious corpus compilation project, culminating in the publication of three corpora in collaboration with the Scientific Thought-styles team: *Middle English Medical Texts (MEMT)* (Taavitsainen, Pahta, & Mäkinen 2005), *Early Modern English Medical Texts (EMEMT)* (Taavitsainen, Pahta, Hiltunen, Mäkinen, Marttila, Ratia, Suhr, & Tyrkkö 2010) and *Late Modern English Medical Texts (LMEMT)* (Taavitsainen, Hiltunen, Lehto, Marttila, Pahta, Ratia, Suhr, & Tyrkkö 2019). Along with having an instrumental role in the overall design of the corpora and the thoroughgoing documentation of the compilation principles, Pahta's main contribution in the project was to draw up the text categories representing midwifery (Ratia & Pahta 2010; Pahta 2019).

Building on her expertise in medical language, genres and registers have provided one of the throughlines in Pahta's work. This includes exploring the emerging conventions of particular medical genres (Taavitsainen & Pahta 2000; Pahta 2001) and specific features in medical language, such as intensifiers in medical language (Mendéz-Naya & Pahta 2010) and circumstantial adverbials in a variety of genres (Meurman-Solin & Pahta 2006). In much of her work on multilingualism, Pahta has shown a keen awareness of the social context, but there are also specific studies where this aspect has been brought to the fore. These focus, for example, on the construction of the social roles different individuals take on in the course of interaction (Pahta, Palander-Collin, Nevala, & Nurmi 2010) and on the ways in which the process of vernacularisation in the scientific and medical fields proceeds (Pahta & Taavitsainen 2004). While much of her research is typically carried out using electronic corpora, Pahta's understanding of manuscripts and their many characteristics has always provided a depth of understanding in historical research (Pahta & Jucker 2011).

Much of Pahta's research has taken place under the auspices of projects she has led. The scope and quality of her research is indicated by the fact that several of these projects were successful in gaining highly competed funding from the Research Council of Finland. Her two most recent projects are Multilingual Practices in the History of Written English (2012–2016) and the research consortium Democratization, Mediatization and Language Practices in Britain, 1700–1950 (2016–2020). Pahta's projects have provided funding for many of her postgraduate students, including a few of the editors of this volume. Sociocultural Reality and Language Practices (2005–2007), in particular, was a significant early project led by Pahta that provided a springboard for new project ideas and funding opportunities for younger researchers and set us an example of cooperative leadership.

Most recently, Pahta has taken up leadership roles in university administration and research administration. She was the Dean of the Faculty of Education and Culture at Tampere University (2019–2023). Prior to that, she was Dean of the Faculty of Communication Sciences, and Dean of the School of Language, Translation and Literary Studies at the former University of Tampere. Pahta also shares her expertise as a member of the editorial board or advisory editorial board of several scholarly publications, such as *European Journal of English Studies, English Language & Linguistics*, and *ESP Across Cultures*, and publication series such as the *Mémoires de la Société Néophilologique*. She is currently Editor-in-Chief of the *Neuphilologische Mitteilungen* and was instrumental in turning the journal into an Open Access electronic journal. Pahta's dedication to furthering research and publishing research has been recognised in 2015 in the invitation to join the Finnish Academy of Science and Letters, where she is currently a Board Member and Secretary of the Section of Humanities. She was also recently elected a board member of the European Federation of Academies of Sciences and Humanities (ALLEA) to represent the Council of Finnish Academies. Finally, Pahta has been the President of the Modern Language Society of Helsinki (Uusfilologinen yhdistys) since 2015.

Last, but not the least, we want to mention Pahta's role in mentoring and guiding young researchers. Pahta has supervised a number of doctoral dissertations and provided her supervisees not only help in completing their dissertations but also in navigating the academic world. Pahta's experience in research and administration, combined with her active role in various academic associations, has provided her with wide-ranging academic expertise, making her an excellent mentor for young as well as more experienced scholars.

2. Studies in the volume

The studies in this volume cover various aspects of the study of multilingualism and language variation in English across genres and registers – as well as across centuries. The studies relate to Pahta's work and build upon it in a variety of ways.

In the first article of the volume, Anna Havinga, Delia Schipor and Laura Wright provide an overview of research on medieval multilingualism carried out over the past decades. They show how the field has grown into maturity as new insights have developed from an increasingly broad perspective. Both multilingual practices and the process of vernacularisation have been extensively studied over the past decades, bringing to light varying practices across genres and topic domains, involving a variety of languages and purposes, and covering not only multilingual texts but also multilingual individuals and communities.

Continuing the theme of multilingualism, Hanna Limatius studies multilingual practices among social media influencers, particularly related to the choice between Finnish and English. Her work is in the field of discourse-centred online ethnography, focusing on attitudes and multilingual practices. Beyond multilingualism, another relevant concept in Limatius's work is genre, as influencers communicate across various genres; this interaction includes, for example, stylistic and community-specific practices as well as terminology-related cases. Limatius pays particular attention to the motives of her informants, as language choice can be based on financial gain rather than just personal identity. Beyond that, there is also tailoring of language to make it accessible to a specific audience and the ever-present choice of multilingual practices as creative language play.

Looking at a wide range of manuscripts from the long twelfth century, Janne Skaffari explores the various ways code-switching has been flagged, that is, made visible or notable through means beyond merely switching languages. He outlines the variety of both visual and metalinguistic means through which the scribes of medieval manuscripts highlighted the changes between Latin and English. Skaffari produces a typology of the instances of flagging observed in the data, and ties that in with both the potential perceived differences between the languages used in the manuscripts, and the textual functions of the switched passages in, for example, organising the different levels of discourse.

Irma Taavitsainen and Andreas H. Jucker explore medical discourse from the Late Middle Ages using the MEMT corpus. Their focus is on instructional texts and particularly on the speech act of giving advice. They use two meta-illocutionary expressions, *rede* and *counsel*, to identify locations of particularly rich advice giving and track these through three different traditions of writing: surgical texts, specialised texts and remedy books. Their analysis shows that giving advice was tied to general interpersonal aspects of the texts studied, and the speech act even extends to advice on how to interact with patients and colleagues in terms of courteous behaviour.

Juhani Norri presents findings on the lexis of a specific aspect of medicine, dentistry, in the late medieval period in England, which was an important period for shaping English medical vocabulary in general. Norri's treatment of the topic comprises a variety of perspectives, starting with how practitioners of dentistry are described in medieval medical writings and going on to discuss the different treatments available, before focusing on the specific vocabulary related to teeth, the related ailments and treating them. Norri notes on the role of influential medical writers, such as Guy de Chauliac, on the formation of the vocabulary and points out the central role of translations in this field.

Another contribution exploring medical terminology is by Alpo Honkapohja and Carla Suhr. Their focus is on human reproduction in Late Medieval and Early Modern English, and their study is based on selections from the MEMT and EMEMT corpora. Here, too, vernacularisation emerges as a key determinant of linguistic usage, as many synonymous terms were introduced in an effort to discuss medicine in English. Honkapohja and Suhr specifically focus on references to the womb, genitalia and afterbirth, paying special attention to the varying contexts of use of different terms. They identify a trend of the terms having a broader meaning in Middle English and becoming more specific by the early modern texts. There is also an influx of new terms throughout the period under study.

Martti Mäkinen and Turo Hiltunen discuss modal auxiliaries in Late Modern English medical writing, relying on a tagged version of the LMEMT corpus. They apply phraseological register analysis to the data, looking specifically at the combinations of modal auxiliary and main verb. Their data is formed by recipes, surgical texts and medical periodicals, allowing them to observe generic variation in the different branches of medical writing, and combining both text-linguistic and variationist perspectives. They identify central textual functions for each modal auxiliary in connection with specific verbs in particular categories of writing, teasing out typical usages of individual modals in each of them, using the metrics of attraction and reliance.

Introducing yet another genre of texts to this volume, María José López-Couso and Belén Méndez-Naya look at contemporary American English using the COCA corpus. Their focus is on parentheticals featuring a noun phrase subject with the noun *study*, and they identify two different patterns for these constructions, as well as change over the past thirty decades. A key finding of the study is closely related to the notion of genre: these constructions are particularly typical of journalistic texts, and especially popular scientific writing. López-Couso and Méndez-Naya make explicit the evidential function of the parentheticals, which they take to be indicative of a potential path of grammaticalisation in evidence.

The volume is concluded by Terttu Nevalainen’s chapter, which explores linguistic variation in terms of social practices. Her main focus is on register variation, whether related to a particular context of communication, the topic domain or the writer–recipient relationship. Nevalainen explores specific linguistic features, the varied uses and implications of chosen terms of address in the correspondence of Elizabeth Stuart, and involvement features, such as the choice of pronouns or nouns, in the letters and public writings of

Thomas Browne, showing how linguistic differences between registers are often subtle and not immediately obvious.

Helsinki & Tampere, June 2024
Arja Nurmi, Minna Nevala, Turo Hiltunen, Carla Suhr, Jenni Räikkönen, & Minna Palander-Collin

Bibliography

Kytö, Merja & Päivi Pahta (eds.) 2016. *The Cambridge Handbook of English Historical Linguistics*. Cambridge: Cambridge University Press.

Leppänen, Sirpa & Päivi Pahta 2012. Finnish culture and language endangered – Language ideological debates on English in the Finnish press from 1995 to 2007. *Dangerous Multilingualism: Northern Perspectives on Order, Purity and Normality*, ed. by Jan Blommaert, Sirpa Leppänen, Päivi Pahta, & Tiina Räisänen. 142–175. Basingstoke and New York: Palgrave Macmillan.

Leppänen, Sirpa, Anne- Pitkänen-Huhta, Tarja Nikula, Samu Kytölä, Timo Törmäkangas, Kari Nissinen, Leila Kääntä, Tiina Räisänen, Mikko Laitinen, Päivi Pahta, Heidi Koskela, Salla Lähdesmäki, & Henna Jousmäki 2011. *National Survey on the English Language in Finland: Uses, Meanings and Attitudes* (Studies in Variation, Contacts and Change in English 5). Helsinki: VARIENG. https://varieng.helsinki.fi/series/volumes/05/

Mendéz-Naya, Belen & Päivi Pahta 2010. Intensifiers in competition: The picture from early English medical writing. *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen & Päivi Pahta. 191–213. Amsterdam: John Benjamins.

Meurman-Solin, Anneli & Päivi Pahta 2006. Circumstantial adverbials in discourse: a synchronic and diachronic perspective. *The Changing Face of Corpus Linguistics*, ed. by Antoinette Renouf & Andrew Kehoe. 117–141. Leiden: Rodopi.

Nurmi, Arja & Päivi Pahta 2004. Social stratification and patterns of code-switching in early English letters. *Multilingua* 23 (4): 417–456. https://doi.org/10.1515/mult.2004.23.4.417

Nurmi, Arja, Tanja Rütten, & Päivi Pahta (eds.) 2018. *Challenging the Myth of Monolingual Corpora*. Leiden: Brill.

Nurmi, Arja, Jukka Tyrkkö, Anna Petäjäniemi, & Päivi Pahta 2018. The social and textual embedding of multilingual practices in Late Modern English: A corpus-based analysis. *Multilingual Practices in Language History: English and Beyond*, ed. by Päivi Pahta, Janne Skaffari, & Laura Wright. 171–198. Berlin: De Gruyter Mouton.

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 2001. Creating a new genre: Contextual dimensions in the production and transmission of early scientific writing. *European Journal of English Studies* 5 (2): 205–220.

Pahta, Päivi 2004. Code-switching in medieval medical writing. *Medical and Scientific Writing in Late Medieval English*, ed. by Irma Taavitsainen & Päivi Pahta. 73–99. Cambridge: Cambridge University Press.
Pahta, Päivi 2011. Code-switching in Early Modern English medical writing. *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen & Päivi Pahta. 115–134. Cambridge: Cambridge University Press.
Pahta, Päivi. 2019. Midwifery. *Late Modern English Medical Texts. Writing Medicine in the Eighteenth Century*, ed. by Irma Taavitsainen & Turo Hiltunen. 271–218. Amsterdam: John Benjamins.
Pahta, Päivi & María José Carrillo Linares 2006. Translation Strategies: De spermate and De humana natura. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS R.14.52, Its Texts, Language, and Scribe* (Medieval and Renaissance Texts and Studies 292, 2 vols.), ed. by M. Teresa Tavormina. 95–117. Tempe: Arizona Center for Medieval and Renaissance Studies.
Pahta, Päivi & Andreas H. Jucker (eds.) 2011. *Communicating Early English Manuscripts*. Cambridge: Cambridge University Press.
Pahta, Päivi & Arja Nurmi 2006. Code-switching in the Helsinki Corpus: A thousand years of multilingual practices. *Medieval English and its Heritage: Structure, Meaning and Mechanisms of Change*, ed. by Nikolaus Ritt, Herbert Schendl, Christiane Dalton-Puffer, & Dieter Kastovsky. 203–220. Frankfurt: Peter Lang.
Pahta, Päivi, Minna Palander-Collin, Minna Nevala, & Arja Nurmi 2010. Language practices in the construction of social roles in Late Modern English. *Social Roles and Language Practices in Late Modern English* (Pragmatics & Beyond 195), ed. by Päivi Pahta, Minna Nevala, Arja Nurmi, & Minna Palander-Collin. 1–27. Amsterdam & Philadelphia: John Benjamins.
Pahta, Päivi, Janne Skaffari, & Laura Wright (eds.) 2018. *Multilingual Practices in Language History: English and Beyond*. Boston: De Gruyter Mouton.
Pahta, Päivi & Irma Taavitsainen 2004. Vernacularisation of scientific and medical writing in its sociohistorical context. *Medical and Scientific Writing in Late Medieval English*, ed. by Irma Taavitsainen & Päivi Pahta. 1–22. Cambridge: Cambridge University Press.
Parviainen, Hanna, Mark Kaunisto, & Päivi Pahta (eds.) 2019. *Corpus Approaches into World Englishes and Language Contrasts*. Helsinki: VARIENG. https://varieng.helsinki.fi/series/volumes/20/
Ratia, Maura & Päivi Pahta 2010. Treatises on specific topics. *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen & Päivi Pahta. 73–102. Amsterdam: John Benjamins.
Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr, & Jukka Tyrkkö (eds.) 2019. *Late Modern English Medical Texts (LMEMT)*. CD-ROM. Amsterdam: John Benjamins.
Taavitsainen, Irma & Päivi Pahta 2000. Conventions of professional writing: The medical case report in a historical perspective. *Journal of English Linguistics* 28 (1): 60–76.
Taavitsainen, Irma & Päivi Pahta 2008. From global language use to local meanings: English in Finnish public discourse. *English Today* 24 (3): 25–38. https://doi.org/10.1017/S0266078408000266

Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, & Jukka Tyrkkö (eds.) 2010. *Early Modern English Medical Texts (EMEMT)*. CD-ROM. Amsterdam: John Benjamins.
Taavitsainen, Irma, Päivi Pahta, & Martti Mäkinen (eds.) 2005. *Middle English Medical Texts (MEMT)*. CD-ROM. Amsterdam: John Benjamins.

Multilingual practices in medieval Britain
Reflections on the scholarship of the last twenty years

Anna D. Havinga
ORCID: 0000-0002-6470-423X
Delia Schipor
Laura Wright
ORCID: 0000-0002-9953-6483

This chapter considers research published over the last two decades on the linguistics of medieval multilingualism as evidenced in all sorts of text types, including work published by our honorand and her collaborators. As historical multilingualism has now become such a burgeoning field, Pahta's contribution to the use of corpus-building and searching for evidence of historical multilingualism can be seen as pioneering. The research considered in this chapter reveals the ubiquity of multilingual practices, the discourse-organizing functions of code-switching, the functional properties of switches, the relevance of abbreviations and their frequency, and the major contribution of Romance borrowings to nuancing the English lexicon. Code-switching, in particular, played an important role in vernacularisation processes, accompanying the shift from French to English in the fifteenth century, and as an intermediate stage in the shift from Latin to English in civic records. Further, work has been carried out on multilingual practices involving Middle Dutch, Old Norse,

https://doi.org/10.51814/ufy.1041.1452

and Celtic languages. The field of historical multilingualism in Britain has advanced considerably in recent years, not least as a result of Pahta's inspirational work on the topic.

Keywords: code-switching, historical multilingualism, multilingual practices, vernacularisation

1. Introduction

In 2001 Schendl provided a resumé of scholarship on code-switching in medieval British texts, observing that "the study of code-switching in the history of English [is] a still largely neglected field of historical linguistics" (Schendl 2001: 305, 308–310). Although medieval writing in Britain was usually multilingual, it was the late nineteenth and early twentieth-century norm to deprecate code-switching as due to a faulty grasp of Medieval Latin or Anglo-Norman (Wright 1997). Our chapter aims to continue Schendl's initiative from 2001 to the present, as today historical code-switching in the history of English, and historical multilingualism more widely, can no longer be said to be neglected – due not least to Päivi Pahta's work on medieval multilingual practices.

Pahta's initial training was in Middle English manuscript studies, resulting in her PhD thesis, an edition of the Middle English text *De spermate*, ff 28r–36v, Trinity College, Cambridge, MS R.14.52 (Pahta 1998). *De spermate* is a translation from Medieval Latin, and it had previously been suggested that the translator was not a native speaker of English. Pahta proposed the alternative view that *De spermate*'s awkward English could have been the work of "an extremely literal translator" using "translation practices which are heavily conditioned by the Latin source copy" (1998: 122), a suggestion which Barron and Wright (2021: 36), analysing another work copied by the same scribe, were later to agree with.[1] Pahta was to develop her investigation into the

1 "We interpret the Hammond scribe's wording not as a bad translation of Anglo-Norman but as a legal, administrative one. The originator of the text invented an English-language register of governance whereby the original Anglo-Norman wording can be reverse-engineered. It was a matter of proceeding, as lawyers do, conservatively and with caution: the first step in establishing English as a medium for written legal record was to translate the morphology (the smallest units of

relationship between Latin and English in *De spermate* into a career-long body of work on the integration of languages in medieval, early modern and late modern texts by analysing multilingual practices in the *Corpus of Middle English Medical Texts* (Pahta 2004), the *Helsinki Corpus* (Pahta & Nurmi 2006), and other text types (Pahta 2011; 2012; Pahta & Nurmi 2009; 2011; Nurmi & Pahta 2010; 2013). In 2012, Pahta won funding from the Research Council of Finland for her project Multilingual Practices in the History of Written English (resulting in Pahta, Skaffari, & Wright 2017) and she authored the entry for "Code-switching in English of the Middle Ages" in *The Oxford Handbook of the History of English* (Pahta 2012). All this, plus participation from the outset in corpus-building and corpus-analysis, enabled her, and subsequently other scholars also, to see what other builders of English corpora had largely overlooked: namely that all corpora of English from all periods, and of any text type or purpose contain multilingual discourse (Nurmi, Rütten, & Pahta 2018).

Pahta's work has had a significant impact on historical linguistic scholarship in the last twenty years, as this chapter illustrates. Our aim is not to provide a complete list of publications on multilingual practices in medieval Britain but to give scholars interested in this field an overview of more recent relevant publications. On a terminological note, studies of historical multilingualism have traditionally employed the term 'code-switching' to refer to various types of multiple-language use in a single communication act, such as a text (see Kopaczyk 2013; Pahta 2004; Schendl 2000; Schendl & Wright 2011). More recently, Pahta *et al.* (2017: 11–12) have suggested that 'multilingual practices' is a more appropriate umbrella term for a wide palette of linguistic phenomena where several languages are involved. This includes, for example, 'code selection', understood as the initial choice of a certain language for a certain text (Schipor 2018: 43), whereby individual monolingual texts may form a multilingual text collection. However, in many text types, for instance late medieval business accounts and inventories, the selected code itself is created by mixing linguistic input from Latin, English and Anglo-Norman (Wright 2000; 2013). While the language mixing in such texts is systematic, with identifiable patterns, it is not always possible to identify switching points between the different languages involved (see Wright 2000; 2011). While our

grammar, the word-endings, prepositions, negators) but to keep the rest of the text – the wordstock – as faithful as possible to the Anglo-Norman exemplar." (Barron & Wright 2021: 36) chapter is divided into sections that concentrate on different phenomena of multilingual practices, based on terms used in the publications discussed, it is important to note that these practices are interconnected and not strictly separable. We start with a focus on historical code-switching (Section 2) before discussing research on vernacularisation (Section 3). Section 4 features research on languages beyond what Putter (2016: 126) calls "the Big Three (English, French [or Anglo-Norman], and Latin)". Section 5 offers concluding remarks.

2. Research on historical code-switching

In the medieval period, writing was characterised by multilingualism (Schendl 2000: 77). Latin, French, and English were written discretely in separate sections (be it paragraphs, sentences, or clauses) in the same text, but also fused together (that is, switching from word to word). Consequently, code-switching was a common phenomenon in a variety of text types.

Research on historical code-switching initially mirrored code-switching theory based on synchronic speech and focussed on switchpoints. During this initial stage, linguistic characteristics and pragmatic functions of code-switching were typically analysed in association with text type. As the field expanded, scholars continued to investigate multilingual phenomena by including syntactic and visual aspects of code-switching, and the complex relationship between code-switching and borrowing.

The discourse functions of code-switching in scientific and medical texts were analysed by Pahta (2004) by surveying texts dated from c. 1375 to 1500, ranging from recipes to specialised treatises. In her investigation, Pahta focussed mainly on "where and why vernacular writers switched to Latin" (2004: 74). Her findings indicated that Latin was used for three main categories of switches: 1) specialised medical terminology representing professional discourse in texts from all medical traditions, 2) religious discourse, such as prayers, quotations, blessings, and charms, especially in remedy books, and 3) textual-organising devices, such as titles, headings, and efficacy phrases in the form of metatextual comments at the end of recipes. Further, Pahta placed her findings in a larger context of multilingual practices in the scientific and medical genre. First, she noted that many switches had multiple functions at the same time, so that, for example, the Latin name of a prayer would also have a text-organising function, namely signalling the end of a text. Second, she explained that, although the discourse community of vernacular medicine seemed to be multilingual, the functions of the three languages involved were not mutually exclusive. The vernaculars (English or French) were also used in some of the contexts where Latin had been attested, and certain texts, such as vernacular translations of learned special treatises, contained no switches to Latin. Notably, Pahta's work allowed her to infer that the medical discourse community operating in the vernacular possessed multilingual literacy skills, which enabled them to interact with specialised texts in a variety of complex ways (2004: 97).

Analyses of the functions and types of code-switching in Middle English letters, poetry, and macaronic sermons were carried out by Schendl (2001; 2002; 2012; 2013a). Schendl's work (2004; 2017) also confirmed the presence of multilingual practices in the Old English period, when at least two text types (charters and mixed poems) contained switching between Old English and Latin. Unsurprisingly, multilingual practices in the Middle English period were more diversified in terms of languages, functions, and text types (Schendl 2004). One of the main conclusions of his work is that code-switching plays an important role both at local and global textual levels, where the local dimension may be represented by textual and pragmatic functions of individual switches, such as indicating changes of footing, while the global level may indicate the cumulative significance of multiple switches in a certain text type, social context, or community of practice. In addition, Schendl (2013b) explored the relevance of code-switching in the development of the English lexicon by analysing a range of late medieval mixed-language text samples (a poem, a sermon, a letter, wills, and accounts). He concludes that these texts "provide important insights into the way in which code-switching must have acted as a mechanism of contact-induced lexical change in the history of English" (Schendl 2013b: 56).

The frequency and functions of code-switching in women's English correspondence from 1400–1800 were investigated by Nurmi and Pahta (2012). Their study combined quantitative and qualitative methods and was based on data from the *Corpus of Early English Correspondence*. Nurmi and Pahta's findings indicate that 49 out of 292 female writers used mainly French, Latin, and Italian alongside English, for various textual functions. Unsurprisingly, switching to French increased progressively throughout the period, and

French was also used in a more creative manner than Latin and Italian (Nurmi & Pahta 2012: 52). By contrast, switching to Latin was less consistent chronologically, and switching to Italian was limited to the eighteenth century. Multilingual practices in women's letters were motivated by various factors: the topic of correspondence, women's identities and social status, their access to foreign language education, and the linguistic skills and preferences of their addressees (Nurmi & Pahta 2012). This study showed that women used their multilingual repertoires in complex ways, to achieve a wide range of communication goals and pragmatic functions, in accordance with genre conventions. Schendl and Wright (2011) also analysed functions of switchpoints and explained that the use of code-mixing served the multilingual identity of both authors and users of multilingual texts (Schendl & Wright 2011: 20). Schendl and Wright (2011) took the part-of-speech of words functioning as switchpoints into account. More specifically, they explained that code-mixing found in inventories and accounts was rule-governed, meaning that closed-class words were typically provided in the matrix language, namely Latin, while open-class words were generally in the vernacular, English or French. The use of English gradually increased in certain positions, such as noun and verb-root positions, and later as modifiers in noun phrases (Schendl & Wright 2011: 34). Further, Wright showed how prepositions in the matrix language, Latin, could be followed by gerunds in either Latin or the vernacular (Wright 2011: 205). Similarly, the positions of adjectives in noun phrases followed both Germanic and Romance rules in the code-mixing system (Wright 2011: 204). In other words, business writing was characterised by mixing three linguistic systems rather than by switching between them (Wright 2011: 204).

Ingham (2009; 2011; 2013; 2018) analysed syntactic properties, and ter Horst and Stam (2017), working with Medieval Latin and Irish, found visual diamorphs (correlating with the concept of homophonous diamorphs in speech (Stam 2017; Wright 2011)) to be of significance at switchpoints. Honkapohja and Liira (2020: 310) note that "[a]bbreviations could function as visual diamorphs, that is, language independent elements which can potentially be expanded in several languages". Their quantitative analysis traced the disappearance of the abbreviation and suspension system in the *Polychronicon* from the late fourteenth to the early sixteenth centuries. It revealed that the reduction of this system was gradual but not linear and that it happened earlier in English than in Latin copies of the text, although parts of the system (such as the ampersand) continued to be used. Honkapohja and Liira thus showed how frequencies of abbreviation and suspension symbols differed according to whether a text was realised in Medieval Latin or in English translation; that is, the visual look of how medieval scripts were presented shifted along with switches in language.

Taking the visual look and visual cues (e.g. line breaks or changes in scribal hands) into account has been found to be important when investigating multilingual practices "as a wholesale multimodal communicative event" (Kopaczyk 2017: 291). Kopaczyk's (2017) analytical framework, which considers the interaction between visual cues and code-switching on different linguistic levels on the page, offers a systematic way to conceptualise historical code-switching.[2] As Kopaczyk (2017: 292–293) says, "[i]f a scribe chooses to underline, capitalize, abbreviate or otherwise indicate the difference in codes within a single lexical unit, we encounter a different communicative situation than when the same information is not conveyed by any visual cues".

The question of contact-induced lexical change in the form of borrowings and its relation to code-switching has been discussed in several publications. Durkin (2020: 348) reported a "huge spike" of Romance vocabulary entering English from the late fourteenth century. He analysed the thousand most frequent words in the *British National Corpus*, finding that just under 50% are borrowings from French or Latin or both, most being first attested in the Middle English period. Anglo-Norman was not a mother-tongue by that point, so these high-frequency words could not have entered through common parlance, but code-switching was just becoming the dominant professional business mode. In a series of papers Trotter (2003; 2006; 2011a; 2011b) made the point that 'code-switching' was not the apposite term for these international lexemes, namely words for commodities which entered numerous European languages and which simply changed morphological shape and orthography to match the language of written record in which they were embedded. Sylvester (2016; 2018a; 2018b; 2020) developed the concept of the lexical technolect, where calques from different languages for the same thing developed technical meanings. This is an important concept because unlike the forces of standardisation where variation reduced, in terms of lexis, the addition of Romance vocabulary in such high numbers (more than 50% of

2 See Kopaczyk (2023) for a discussion on challenges in applying this model.

the English lexicon, according to the Anglo-Norman Hub) enabled nuance of meaning to develop.[3] Her insight was to be further developed by Sylvester, Tiddeman, and Ingham (2020; 2022; 2023), Tiddeman (2018; 2020), and Sylvester and Tiddeman (2023).

3. Research on vernacularisation

Code-switching in medieval Britain formed part of the larger process of shift from Medieval Latin to vernaculars in the languages of written record. Pahta and Taavitsainen (2004) use the term 'vernacularisation' to refer to this process, which Voigts (1996: 813) had defined as "the movement of vernacular languages into domains of written language that were formerly the exclusive preserve of Latin". In the introduction to their edited volume on medical and scientific writing in Late Medieval English, Pahta and Taavitsainen (2004: 2) explain that the increased use of vernaculars was due to various factors, including practical ones (accessibility to information) and attitudinal or ideological ones (ideas about vernaculars and nationalism). In particular, they note that the expansion of English "gathered momentum in the fifteenth century with the nationalistic strivings of the Lancastrian monarchs" (Pahta & Taavitsainen 2004: 10). However, English had already gained traction in professional writing, such as legal texts and instructions, from the late fourteenth century (Pahta & Taavitsainen 2004: 10). Whether the use of vernaculars was instigated by conscious language policies (i.e. 'from above'), e.g. from the Lancastrian monarchs, or whether it was initiated by scribes' language practices (i.e. 'from below') has become a question of recent debate. While some scholars have brought forward arguments for conscious policies (see, in particular, Fisher 1992), more recent work, for example by Dodd (2011) and Schipor (2018), has questioned the idea that the shift to English was driven by language policies from above. Schipor (2018; 2022) found that English was increasingly more commonly used in documentary texts written in the first person, such as testamentary texts, petitions, abjurations, and statements. This indicates that selecting the vernacular rather than Latin was a development

3 Anglo-Norman Hub. 2017. Anglo-Norman Dictionary secures £800,00. https://www.aber.ac.uk/en/news/archive/2017/02/title-197356-en.html#:~:text=A%20form%20of%20French%20brought,are%20derived%20from%20Anglo%2DNorman [accessed 15.11.2023].

'from below', motivated primarily by pragmatic purposes. Beyond the English context, too, it seems that the replacement of Latin happened gradually and was initiated 'from below' rather than 'from above'. Havinga (2023) traced vernacularisation processes in documentary legal texts from Aberdeen (1398–1511) and Lübeck (1430–1451). Based on her comparative analysis, she concludes that language practices of the scribes rather than top-down language policies drove the vernacularisation processes in these sources.

Tracing vernacularisation processes has been a valuable endeavour for understanding the dynamic nature of historical multilingualism. As part of the Middle English Local Documents (MELD) project, Schipor (2018) and Stenroos (2020) explored administrative texts produced outside central government offices, such as "records of cities, churches, manors, local courts and private transactions" (Stenroos 2020: 39), to reveal when and to what extent English replaced French and Latin.[4] Unlike other text types, medieval administrative and legal documents are largely datable and locatable, and unlike other corpora, text was entered directly from manuscript transcription rather than via scholarly editions. The Stavanger team also paid attention to visual and discourse pragmatics around what Schipor termed 'multilingual events' (Schipor 2013; 2018; Stenroos and Schipor 2020), because situation on the folio, size of script, type of script, rubrication, and other visual delineators conditioned language choice (see also Machan 2011). The MELD project found that stretches of English first appeared in bureaucratic documents written in the North (i.e. innovation occurred furthest away from the centres of administration in London, which acted as a strong-tie force for conservatism) and in components that the client needed to understand rather than lawyer-to-lawyer communication. The team also found that Latin remained dominant from 1399 to 1525 and that English replaced the Anglo-Norman component rather than the Latin one, so that the fifteenth century can be characterised as the century when Anglo-Norman became abandoned, rather than when Latin was lost (Schipor 2018; Stenroos 2020). Similarly, Dodd (2019), who analysed administrative texts from central government, notes that Latin was the

---

4 *A corpus of Middle English Local Documents (MELD)*. https://www.uis.no/en/corpus-middle-english-local-documents-meld, compiled by Merja Stenroos, Kjetil Vikhamar Thengs and Geir Bergstrøm with the assistance of Anastasia Khanukaeva, Delia Schipor and Kenneth Solberg-Harestad. University of Stavanger.

"foremost language of the legal written record [...] throughout the Middle Ages" (Dodd 2019: 17) and that English began to replace French (rather than Latin) in the fourteenth century.

Other scholars have focussed on different text types. Carroll (2004), for example, found that recipes from England were vernacularised in the fourteenth century. Alcolado Carnicero (2013) analysed the accounts of the London Mercers' Livery Company (1390–1464) to trace vernacularisation processes. Taking a social networks approach, he observed changes in language use from generations who used Medieval Latin, to generations who used code-switching only, to generations who used monolingual English (that is, an English which had absorbed a great deal of Anglo-Norman vocabulary), via master-apprentice relationships. He was to build on this perception over subsequent publications (Alcolado Carnicero 2015; 2017; 2018; 2020; 2021a; 2021b).

Putting all this work together, it became apparent that rather than a wholesale shift from French to English, with Medieval Latin gradually shifting to Neo Latin and becoming restricted in register before dropping out of use altogether, the Anglo-Norman component reduced from [lexis + grammatical matrix] to content words only, which became absorbed into an English grammatical component, with commensurate code-switching over the century of shift. As Anglo-Norman became absorbed into English over the late-fourteenth and fifteenth centuries, code-switching became default in text types which required itemisation of goods such as customs declarations, business accounts, the testimonial inventories of wills, and personal daybooks such as those kept by Gilbert Maghfeld and William of Worcester (Schendl 2020; Wright 2015; 2018). Anglo-Norman vocabulary entered not only legal registers but also the domain of land-management and stewardship, where it remains in dialectal use (Wright in press).

The concept of supralocalisation was important for the perception of the integration of Anglo-Norman into English. Supralocalisation refers to the process of a feature spreading from its initial centre of usage (both in region and text type) which nevertheless did not then go on to standardise. For example, Thengs (2013) showed that, as Anglo-Norman became abandoned as a monolingual language of written record in the first decades of the fifteenth century, the Anglo-Norman custom of using the letter-graph <x> as a plural marker on words ending in /l/ entered English writing and became widespread in the later fifteenth-century writing, even if the word in question were of English rather than French etymology, such as *shamelx* ‘shambles’ from Old English *sceamel* ‘stall’ (Thengs 2013: 308 (Cheshire, Staffordshire, Shropshire); Stenroos 2020: 63–65 (also Yorkshire); Wright 2020: 526 (also London)). Similarly, the Anglo-Norman -*aun*- trigraph entered words of English origin, such as the place name *Caunterbury* (Trotter 2013: 167; Wright 2020: 525). This is particularly significant as Trotter identified -*aun*- as likely to “be the only exclusively Anglo-Norman graphy that exists”, meaning that -*aun*- sequences, so prevalent in fifteenth-century writing, could only have been ported over from Anglo-Norman. Conde-Silvestre (2020) reported that spelling focussed (reduced variation) in words of Anglo-Norman origin first in an Oxfordshire community of practice of cofeoffees and civil servants writing in English in the mid-fifteenth century. These administrators wrote about matters of law and estate-stewardship and thus used the Latinate vocabulary pertaining to those domains and ported their habit of spelling this Latinate vocabulary (i.e. without much spelling variation) along with the lexemes. This is supralocalisation in action, because although the net effect was to reduce spelling variation, spelling standardisation was still to come. Supralocalisation was thus an important stage on the road to eventual standardisation.

4. Research beyond the “Big Three”

Publications such as *Multilingualism in Medieval Britain* edited by Judith Jefferson and Ad Putter (2013) and *Medieval English in a Multilingual Context* edited by Sara Pons-Sanz and Louise Sylvester (2023) illustrate how fruitful the engagement with multilingual practices has been. Both of these volumes go beyond the focus on English vis-à-vis Latin and French by including discussions on Dutch, Old Norse, and Welsh, amongst other languages. Research on these languages within the context of multilingual Britain has made headway in the last twenty years, as the following examples of recent publications show.

Putter (2021) took inspiration from Christopher Joby’s (2015) book on *The Dutch language in Britain (1550–1702)* to present evidence of the use of Dutch in medieval Britain through three case studies: the writings of Gerald of Wales, the copy of a Dutch letter in the Aberdeen Council Registers, and the Book of Privileges. He highlighted that medieval sources offer a rich basis for studies on the Dutch language in Britain, listing further texts that are worth examining (see also Alban 2014). Alexander Fleming and Roger Mason's (2019) edited volume focussed on the Flemish people in Scotland and included a chapter on the impact of Flemish on the Scots language (Fleming *et al*. 2019). Building on the work by Murison (1971), who provided a list of Dutch elements in the vocabulary of Scots, they found that the "Flemish influence on the Scots language is felt particularly in fields where there were significant areas of contact" (Fleming *et al*. 2019: 134), such as trade and agriculture. Murison's list was also used by Havinga (2020), who assessed the intensity of language contact between Middle Dutch and Scots in late-medieval Aberdeen. Based on an evaluation of trade connections between Aberdeen and the Low Countries, migration from the Low Countries to Scotland, and the use of Middle Dutch in the Aberdeen Council Registers (1398–1511), Havinga (2020: 275) concluded that "the contact between Middle Dutch and Scots in Aberdeen can be classified somewhere between 'casual contact' and 'slightly more intense contact' based on Thomason and Kaufman's (1988) borrowing scale" (see Fox, Grant, and Wright 2023 for a discussion of contact theory and the history of English). Havinga's (2021) more general analysis of multilingual practices and vernacularisation processes in the Aberdeen Council Registers showed the gradual replacement of Latin by Scots in the first eight volumes of the registers, which did not, however, result in a complete loss of Latin. Formulaic entries, in particular, continued to be written in Latin in the sixteenth century. Havinga also investigated the function of codeswitches in one of the volumes and found that such switches were more common at the beginning of the registers' vernacularisation process. Generally, the registers constitute what Kopaczyk (2021) called 'a complex multilingual code' – a term that can be applied to many medieval sources.[5]

That Old Norse also formed part of this multilingual code has been shown by the Gersum Project, led by Richard Dance, Sara Pons-Sanz, and Brittany Schorn (2019). This project examined the Scandinavian influence on English vocabulary, creating an open-access database of 967 lexemes from a corpus of Middle English poems that categorises the likelihood of an Old Norse origin for

5 For an overview of languages used in medieval Scotland, see chapter 3 in Millar (2020). For multilingualism in Scottish poetry, see Machan (2010), who analysed three long poems by Robert Henryson to illustrate the "ordinariness of multilingualism" (2010: 54) in the Middle Ages.

each lexeme.[6] For further scholarship on the Scandinavian influence on English see Barnes (1993), Townend (2002), and Pons-Sanz (2013).

In addition to Scandinavian and Dutch contact impacting the development of English, Celtic languages have also been considered. Klemola (2013) argued that the influence of Celtic language contact can be seen in terms of syntax and phonology rather than lexis. The fact that there are only a few Celtic loanwords in English should not, therefore, be taken as proof that the influence of Celtic languages on English was negligible (Klemola 2013). The multilingual situation in medieval Wales has perhaps attracted the most scholarly attention, even if evidence for multilingualism in medieval Wales is “patchy and scattered” (Russell 2019: 7). Russell (2017) explored the co-existence of and interactions between Welsh and Latin from the Roman period onwards, illustrating “how in some contexts Medieval Welsh became distinctly Latinate and how in certain respects the Latin of medieval Wales arguably became Cambricised” (Russell 2017: 215). While Russell focussed on the relationship between Welsh and Latin, Fulton (2011) discussed uses of the vernacular in Medieval Wales, noting that “vernacular writing in Welsh had a continuous history from the age of the Saxon settlements” (Fulton 2011: 222), even if the “earliest manuscript books containing Welsh texts survive only from the middle of the thirteenth century” (Fulton 2011: 200). Fulton also explained that, alongside Latin, writing in Anglo-Norman was preferred over English, which was partly due to “a certain amount of ‘anti-Saxon’ attitude” (Fulton 2011: 222) at that time. Fulton (2012) elaborated on the cultural and political power connected to languages used in medieval Wales (including mentions of Flemish and Irish). Despite advances in studying the multilingual situation in medieval Wales, discussions on Welsh in volumes that deal with multilingualism in Britain are still underrepresented (Russell 2019: 10); see, however, Smith (2000), Putter (2011), and Fulton (2023).

---

6 Richard Dance, Sara Pons-Sanz, and Brittany Schorn, The Gersum Project: The Scandinavian Influence on English Vocabulary (Cambridge, Cardiff and Sheffield, 2019). https://www.gersum.org/ [accessed 09.11.2023]. The database can be accessed here: https://www.gersum.org/database.

5. Concluding remarks

In conclusion, since Schendl branded historical code-switching a neglected topic, researchers have applied methodologies from dialectology and sociolinguistics (such as social networks, weak and strong ties, text types and communities of practice, as well as the traditional concept of geographical region) to study all sorts of medieval multilingual practices. However, there is still plenty of work to do. Substantial corpora that allow electronic searching for medieval code-switching or vernacularisation processes are still lacking. Ideally, such corpora would include texts from all parts of medieval Britain, which would also facilitate research beyond the "Big Three" languages (Anglo-Norman, English, Latin). Moreover, Honkapohja and Wright (2023) have recently noted that abbreviations and suspensions still do not receive the attention they deserve. That this is a field to be investigated further needs to be kept in mind when building corpora for the study of multilingual practices. Silent expansions are, unfortunately, far too common in transcriptions of historical manuscripts. Much remains to be done, but the 2023 conference of the *Historical Sociolinguistics Network* received more abstracts for papers in the category of historical multilingualism than in any other, so we conclude with the observation that Pahta's work on this topic has been influential and continues to inspire.

Bibliography

Alban, J.R. 2014. Sources in the Norfolk Record Office which relate to the History of Norfolk and the Low Countries. *Dutch Crossing* 38 (2): 101–115. https://doi.org/10.1179/0309656414Z.00000000053

Alcolado Carnicero, José Miguel 2013. *Social Networks and Mixed-language Business Writing: Latin/French/English in the Wardens' Accounts of the Mercers' Company of London, 1390–1464*. University of Castilla-La Mancha, PhD thesis.

Alcolado Carnicero, José Miguel 2015. Dating the shift to English in the financial accounts of some London livery companies: a reappraisal. *Multilingua* 34 (3): 373–404. https://doi.org/10.1515/multi-2014-0068

Alcolado Carnicero, José Miguel 2017. Bridges of innovation and change: the English language around the networks of the Mercery of London. *Merchants of Innovation: the Languages of Traders*, ed. by Esther-Miriam Wagner, Bettina Beinhoff, & Ben Outhwaite. 40–67. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781501503542-003

Alcolado Carnicero, José Miguel 2018. Merchant Taylors of London's shift to business English: New insights on the languages of record in their Master and Wardens'

accounts. *Journal of Historical Sociolinguistics* 4 (1): 1–32. https://doi.org/10.1515/jhsl-017-0037
Alcolado Carnicero, José Miguel 2020. Diachrony of code switching stages in medieval business accounts: The Mercers' livery company of London. *Journal of Historical Linguistics* 9 (3): 378–416. https://doi.org/10.1075/jhl.18001.car
Alcolado Carnicero, José Miguel 2021a. The diffusion of English in late medieval social networks: Henry V, Robert Chicheley, London Grocers, and London Brewers. *Language and History* 64 (2): 1–25. https://doi.org/10.1080/17597536.2021.1904183
Alcolado Carnicero, José Miguel 2021b. Item, pur escrivyng et enrollynge in Englyshe: from Latin and French to English in the medieval business records of the Grocers of London. *Studia Neophilologica* 95 (3): 1–28. https://doi.org/10.1080/00393274.2021.1948352
Armstrong, Jackson W. & Edda Frankot (eds.) 2021. *Cultures of Law in Urban Northern Europe: Scotland and its Neighbours c. 1350–c. 1650*. London and New York: Routledge.
Barnes, Michael P. 1993. Norse in the British Isles. *Viking Revaluations*, ed. by Anthony Faulkes & Richard Perkins. 65–84. London: Viking Society for Northern Research.
Barron, Caroline M. & Laura Wright (eds.) 2021. *The London Jubilee Book, 1376–1387: an edition of Trinity College Cambridge, MS O.3.11, folios 133–157.* London: London Record Society for 2020.
Carroll, Ruth 2004. Middle English recipes: Vernacularisation of a text-type. In Taavitsainen & Pahta (eds.), 174–191.
Conde-Silvestre, Juan Camilo 2020. Communities of practice and proto-standardisation in the Stonor Letters: traces of late Middle English multilingualism in spelling focusing. In Wright (ed.), 443–466. https://doi.org/10.1515/9783110687545-016
Dance, Richard, Sara Pons-Sanz, & Brittany Schorn 2019. *The Gersum Project: The Scandinavian Influence on English Vocabulary.* Cambridge, Cardiff & Sheffield. https://www.gersum.org/ [accessed 09.11.2023].
Dodd, Gwilym 2011. The rise of English, the decline of French: supplications to the English Crown, c. 1420–1450. *Speculum* 86: 117–146. https://doi.org/10.1017/S0038713410003507
Dodd, Gwilym 2019. Languages and law in Late Medieval England: English, French and Latin. *The Cambridge Companion to Medieval English Law and Literature*, ed. by Candace Barrington & Sebastian Sobecki. 17–29. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316848296.002
Durkin, Philip 2020. The relationship of borrowing from French and Latin in the Middle English period with the development of the lexicon of Standard English: Some observations and a lot of questions. In Wright (ed.), 343–364. https://doi.org/10.1515/9783110687545-012
Fisher, John 1992. A language policy for Lancastrian England. *PMLA* 107: 1168–1180. https://doi.org/10.2307/462872
Fleming, Alexander & Roger Mason (eds.) 2019. *Scotland and the Flemish People*. Edinburgh: Birlinn.
Fleming, Alexander, Chris Robinson, & Thomas Clark 2019. Language. In Fleming & Mason (eds.), 133–138.
Fox, Susan, Anthony Grant, & Laura Wright 2023. Contact theory and the history of English. In Pons-Sanz & Sylvester (eds.), 17–47.

Fulton, Helen 2011. Literature of the Welsh Gentry: Uses of the vernacular in medieval Wales. *Vernacularity in England and Wales, c. 1300–1550*, ed. by Helen Fulton, E. Salter, & H. Wicker. 199–223. Turnhout: Brepols. https://doi.org/10.1484/M.USML-EB.3.4948
Fulton, Helen 2012. Negotiating Welshness: Multilingualism in Wales before and after 1066. *Conceptualizing Multilingualism in England, c. 800–c. 1250*, ed. by Elizabeth M. Tyler. 145–170. Turnhout: Brepols. https://doi.org/10.1484/M.SEM-EB.4.8008
Fulton, Helen 2023. Cheapside in Wales: Multilingualism and textiles in medieval Welsh poetry. In Pons-Sanz & Sylvester (eds.), 187–212.
Havinga, Anna D. 2020. Assessing the intensity of language contact between Middle Dutch and Scots in late medieval Aberdeen. *Language on the Move across Domains and Communities. Selected Papers from the 12th Triennial Forum for Research on the Languages of Scotland and Ulster, Glasgow 2018*, ed. by Joanna Kopaczyk & Robert McColl Millar. 251–278. Aberdeen: FRLSU.
Havinga, Anna D. 2021. The vernacularisation of the Aberdeen Council Registers (1398–1511). In Armstrong & Frankot (eds.), 80–102. https://doi.org/10.4324/9780429262869-7
Havinga, Anna D. 2023. Vernacularisation of documentary legal texts in Northern Europe: A comparison of the development of Scots and Low German in the fifteenth century. *Linguistica* 63 (1-2): 19-39. https://doi.org/10.4312/linguistica.63.1-2.19-39
Honkapohja, Alpo & Aino Liira 2020. Abbreviations and standardisation in the Polychronicon: Latin to English and manuscript to print. In Wright (ed.), 269–316. https://doi.org/10.1515/9783110687545-010
Honkapohja, Alpo & Laura Wright 2023. Abbreviation and suspension symbols: a methodological issue for the study of medieval multilingualism. *Mondi plurilingui (Multilingual Worlds)*, ed. by Carmela Perta. 141–159. Alessandria: Edizioni dell'Orso.
Horst, Tom ter & Nike Stam 2017. Visual diamorphs: the importance of language neutrality in code-switching from medieval Ireland. In Pahta, Skaffari, & Wright (eds.), 223–242. https://doi.org/10.1515/9781501504945-011
Hundt, Marianne & Daniel Schreier (eds.) 2013. *English as a Contact Language*. Cambridge: Cambridge University Press.
Ingham, Richard 2009. Mixing languages on the manor. *Medium Aevum* 78 (1): 80–97. https://doi.org/10.2307/43632800
Ingham, Richard 2011. Code-switching in the later English Lay Subsidy Rolls. In Schendl & Wright (eds.), 95–114. https://doi.org/10.1515/9783110253368.95
Ingham, Richard 2013. Language-mixing in medieval Latin documents: Vernacular articles and nouns. In Jefferson & Putter (eds.), 105–121. https://doi.org/10.1484/M.TCNE-EB.1.100796
Ingham, Richard 2018. Medieval bilingualism in England: On the rarity of vernacular code-switching. In Pahta, Skaffari, & Wright (eds.), 319–337. https://doi.org/10.1515/9781501504945-015
Jefferson, Judith & Ad Putter (eds.) 2013. *Multilingualism in Medieval Britain (c. 1066–1520): Sources and Analysis*. Turnhout: Brepols.
Joby, Christopher 2015. *The Dutch Language in Britain (1550–1702). A Social History of the Use of Dutch in Early Modern Britain*. Leiden: Brill. https://doi.org/10.1163/9789004285217

Kleinhenz, Christopher & Keith Busby (eds.) 2011. *Medieval Multilingualism. The Francophone World and its Neighbours*. Turnhout: Brepols.
Klemola, Juhani 2013. English as a contact language in the British Isles. In Hundt & Schreier (eds.), 75–87. https://doi.org/10.1017/CBO9780511740060.005
Kopaczyk, Joanna 2013. Code-switching in the records of a Scottish brotherhood in early modern Poland-Lithuania. *Poznań Studies in Contemporary Linguistics*, 49 (3): 281–319. https://doi.org/10.1515/psicl-2013-0011
Kopaczyk, Joanna 2017. Administrative multilingualism on the page in early modern Poland: In search of a framework for written code-switching. In Pahta, Skaffari, & Wright (eds.), 275–298. https://doi.org/10.1515/9781501504945-013
Kopaczyk, Joanna. 2021. The language of medieval legal record as a complex multilingual code. In Armstrong & Frankot (eds.), 58–79. https://doi.org/10.4324/9780429262869-6
Kopaczyk, Joanna 2023. The challenges of bringing together multilingualism and multimodality: Unpacking the structural model of multilingual practice. *Multilingualism from Manuscript to 3D: Intersections of Modalities from Medieval to Modern Times*, ed. by Matylda Włodarczyk, Jukka Tyrkkö, & Elżbieta Adamczyk. 119–138. London: Routledge. https://doi.org/10.4324/9781003166634-7
Machan, Tim William 2010. Robert Henryson and the matter of multilingualism. *The Journal of English and Germanic Philology* 109 (1): 52–70. https://doi.org/10.1353/egp.0.0105
Machan, Tim William 2011. The visual pragmatics of code-switching in late Middle English literature. In Schendl & Wright (eds.), 303–336. https://doi.org/10.1515/9783110253368.303
Millar, Robert McColl 2020. *A Sociolinguistic History of Scotland*. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9781474448567
Murison, David 1971. The Dutch element in the vocabulary of Scots. *Edinburgh Studies in English and Scots*, ed. by A.J. Aitken, Angus McIntosh, & Hermann Pálsson. 159–176. London: Longman.
Nurmi, Arja & Päivi Pahta 2010. Negotiating interpersonal identities in writing. Code-switching practices in Charles Burney's correspondence. *The Language of Daily Life in England (1400–1800)*, ed. by Arja Nurmi, Minna Nevala, & Minna Palander-Collin. 27–52. Amsterdam: Benjamins. https://doi.org/10.1075/pbns.183.04pah
Nurmi, Arja & Päivi Pahta 2012. Multilingual practices in women's English correspondence 1400–1800. In Sebba, Mahootian, & Jonsson (eds.), 44–67.
Nurmi, Arja & Päivi Pahta 2013. Multilingual practices in the language of the law: Evidence from the Lampeter corpus. *Ex Philologia Lux: Essays in Honour of Leena Kahlas-Tarkka* (Mémoires de la Société Néophilologique de Helsinki XC), ed. by Jukka Tyrkkö, Olga Timofeeva, & Maria Salenius. 187–205. Helsinki: Société Néophilologique.
Nurmi, Arja, Tanja Rütten, & Päivi Pahta (eds.) 2018. *Challenging the Myth of Monolingual Corpora*. Leiden: Brill/Rodopi. https://doi.org/10.1163/9789004276697
Pahta, Päivi 1998. *Medieval Embryology in the Vernacular. The Case of* De spermate. (Mémoires de la Société Néophilologique de Helsinki LIII.) Helsinki: Société Néophilologique.
Pahta, Päivi 2004. Code-switching in medieval medical writing. In Taavitsainen & Pahta (eds.), 73–99.

Pahta, Päivi 2011. Code-switching in Early Modern English medical writing. *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen & Päivi Pahta. 115–134. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.008

Pahta, Päivi 2012. Code-switching in English of the Middle Ages. *The Oxford Handbook of the History of English,* ed. by Terttu Nevalainen & Elizabeth Closs Traugott. 528–537. New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0045

Pahta, Päivi & Arja Nurmi 2006. Code-switching in the Helsinki Corpus: A thousand years of multilingual practices. *Medieval English and its Heritage*, ed. by Nikolaus Ritt, Herbert Schendl, Christiane Dalton-Puffer, & Dieter Kastovsky. 203–220. Frankfurt: Peter Lang.

Pahta, Päivi & Arja Nurmi 2009. Preacher, scholar, brother, friend. Social roles and code-switching in the writings of Thomas Twining. *Social Roles and Language Practices in Late Modern English*, ed. by Päivi Pahta, Minna Nevala, Arja Nurmi, & Minna Palander-Collin. 135–162. Amsterdam: Benjamins. https://doi.org/10.1075/pbns.195.07nur

Pahta, Päivi & Arja Nurmi 2011. Multilingual discourse in the domain of religion in medieval and early modern England: A corpus approach to research on historical code-switching. In Schendl & Wright (eds.), 219–251. https://doi.org/10.1515/9783110253368.219

Pahta, Päivi & Irma Taavitsainen 2004. Vernacularisation of scientific and medical writing in its sociohistorical context. In Taavitsainen & Pahta (eds.), 1–22.

Pahta, Päivi, Janne Skaffari, & Laura Wright 2017. From historical code-switching to multilingual practices in the past. In Pahta, Skaffari, & Wright (eds.), 3–18. https://doi.org/10.1515/9781501504945-001

Pahta, Päivi, Janne Skaffari, & Laura Wright (eds.) 2017. *Multilingual Practices in Language History: English and Beyond*. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9781501504945

Pons-Sanz, Sara 2013. *The Lexical Effects of Anglo-Scandinavian Linguistic Contact on Old English*. Turnhout: Brepols. https://doi.org/10.1484/M.SEM-EB.5.106260

Pons-Sanz, Sara M. & Louise Sylvester (eds.) 2023. *Medieval English in a Multilingual Context: Current Methodologies and Approaches*. London: Springer.

Putter, Ad 2011. Multilingualism in England and Wales, c. 1200: the Testimony of Gerald of Wales. In Kleinhenz & Busby (eds.), 83–105. https://doi.org/10.1484/M.TCNE-EB.3.4604

Putter, Ad 2016. The linguistic repertoire of medieval England, 1100–1500. *Imagining Medieval English: Language Structures and Theories 500–1500*, ed. by Tim William Machan. 126–144. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107415836.007

Putter, Ad 2021. Materials for a social history of the Dutch language in medieval Britain: Three case studies from Wales, Scotland, and England. *Dutch Crossing* 45 (2), 97–111. https://doi.org/10.1080/03096564.2021.1943622

Russell, Paul 2017. "Go and look in the Latin books": Latin and the vernacular in medieval Wales. *Latin in Medieval Britain*, ed. by Richard Ashdowne & Carolinne White. 213–246. London: British Academy. https://doi.org/10.5871/bacad/9780197266083.003.0010

Russell, Paul 2019. Bilingualisms and multilingualisms in medieval Wales: Evidence and inference. *Transactions of the Honourable Society of Cymmrodorion* 25: 7–22.
Schendl, Herbert 2000. Linguistic aspects of code-switching in medieval English texts. In Trotter (ed.), 77–92.
Schendl, Herbert 2001. Code-switching in medieval English poetry. *Language Contact in the History of English*, ed. by Dieter Kastovsky & Arthur Mettinger. 305–335. Frankfurt: Peter Lang.
Schendl, Herbert 2002. Code-choice and code-switching in some early fifteenth-century letters. *Middle English from Tongue to Text*, ed. by Peter J. Lucas & Angela M. Lucas. 247–262. Frankfurt: Peter Lang.
Schendl, Herbert 2004. *Hec sunt prata to wassingwellan*: Aspects of code-switching in Old English charters. *Vienna English Working Papers* 13 (2): 52–68.
Schendl, Herbert 2012. Literacy, multilingualism and code-switching in Early English written texts. In Sebba, Mahootian, & Jonsson (eds.), 27–43.
Schendl, Herbert 2013a. Code-switching in late medieval macaronic sermons. In Jefferson & Putter (eds.), 153–169. https://doi.org/10.1484/M.TCNE-EB.1.100799
Schendl, Herbert 2013b. Multilingualism and code-switching as mechanism of contact-induced lexical change in late Middle English. In Hundt & Schreier (eds.), 41–57. https://doi.org/10.1017/CBO9780511740060.003
Schendl, Herbert 2017. Code-switching in Anglo-Saxon England: A corpus-based approach. In Pahta, Skaffari, & Wright (eds.), 39–60. https://doi.org/10.1515/9781501504945-003
Schendl, Herbert 2020. William Worcester's Itineraria: Mixed-language notes of a medieval traveller. In Wright (ed.), 317–342. https://doi.org/10.1515/9783110687545-011
Schendl, Herbert & Laura Wright 2011. Code-switching in early English: historical background and methodological and theoretical issues. In Schendl & Wright (eds.), 15–45. https://doi.org/10.1515/9783110253368.15
Schendl, Herbert & Laura Wright (eds.) 2011. *Code-Switching in Early English*. Berlin and Boston: De Gruyter Mouton.
Schipor, Delia 2013. *Multilingual practices in Late Medieval English official writing: an edition of documents from the Beverley Town Cartulary*. University of Stavanger, MA thesis.
Schipor, Delia 2018. *A Study of Multilingualism in the Late Medieval Material of the Hampshire Record Office*. University of Stavanger, PhD thesis.
Schipor, Delia 2022. Multilingual events in late medieval personal documentary texts from the Winchester Diocese Collection in 1400–1525. *Language Dynamics in the Early Modern Period*, ed. by Karen Bennett & Angelo Cattaneo. 21–39. New York and London: Routledge. https://doi.org/10.4324/9781003092445-3
Sebba, Mark, Shahrzad Mahootian, & Carla Jonsson (eds.) 2012. *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. New York: Routledge.
Smith, Llinos Beverly 2000. The Welsh and English languages in late-medieval Wales. In Trotter (ed.), 7–21.
Stam, Nike 2017. *A Typology of Code-Switching in the Commentary to the Félire Óengusso*. Utrecht: LOT, PhD thesis.

Stenroos, Merja 2020. The ‘vernacularisation’ and ‘standardisation’ of local administrative writing in late and post-medieval England. In Wright (ed.), 39–85. https://doi.org/10.1515/9783110687545-003

Stenroos, Merja & Delia Schipor 2020. Multilingual practices in Middle English documents. *Records of Real People. Linguistic Variation in Middle English Local Documents*, ed. by Merja Stenroos & Kjetil V. Thengs. 249–277. Amsterdam: John Benjamins. https://doi.org/10.1075/ahs.11.11ste

Sylvester, Louise 2016. Technical vocabulary and medieval text types: A semantic field approach. *Neuphilologische Mitteilungen* 117: 155–176.

Sylvester, Louise 2018a. A semantic field and text-type approach to late-medieval multilingualism. In Pahta, Skaffari, & Wright (eds.), 77–96. https://doi.org/10.1515/9781501504945-005

Sylvester, Louise 2018b. Contact effects on the technical lexis of Middle English: A semantic hierarchic approach. *English Language and Linguistics* 22: 249–264. https://doi.org/10.1017/S1360674318000126

Sylvester, Louise 2020. The role of multilingualism in the emergence of a technical register in the Middle English period. In Wright (ed.), 365–380. https://doi.org/10.1515/9783110687545-013

Sylvester, Louise & Megan Tiddeman 2023. Reframing the interaction between native terms and loanwords: some data from occupational domains in Middle English. In Pons-Sanz & Sylvester (eds.), 159–186.

Sylvester, Louise, Megan Tiddeman, & Richard Ingham 2020. An analysis of French borrowings at the hypernymic and hyponymic levels of Middle English. *Lexis: Journal of English Lexicology* 16: 1–32. https://doi.org/10.4000/lexis.4841

Sylvester, Louise, Megan Tiddeman, & Richard Ingham 2022. Lexical borrowing in the Middle English period: A multi-domain analysis of semantic outcomes. *English Language and Linguistics* 26 (2): 237–261. https://doi.org/10.1017/S1360674321000113

Sylvester, Louise, Megan Tiddeman, & Richard Ingham 2023. Lexical replacement, retention and borrowing in Middle English: a case study. *Language Contact and the History of English: Processes and Effects on Specific Text-Types*, ed. by Gabriella Mazzon. 165–187. Bern: Peter Lang.

Taavitsainen, Irma & Päivi Pahta (eds.) 2004. *Medical and Scientific Writing in Late Medieval English*. Cambridge: Cambridge University Press.

Thengs, Kjetil V. 2013. *English Medieval Documents of the Northwest Midlands*. University of Stavanger, PhD thesis.

Thomason, Sarah Grey & Terrence Kaufman 1988. *Language Contact, Creolization, and Genetic Linguistics.* Berkeley: University of California Press. https://doi.org/10.1525/9780520912793

Tiddeman, Megan 2018. Lexical exchange with Italian in the textile and wool trades in the thirteenth to fifteenth centuries. *Medieval Clothing and Textiles* 14: 113–140. https://doi.org/10.1017/9781787442443.006

Tiddeman, Megan 2020. More sugar and spice: Revisiting medieval Italian influence on the mercantile lexis of England. In Wright (ed.), 381–410. https://doi.org/10.1515/9783110687545-014

Townend, Matthew 2002. *Language and History in Viking Age England: Linguistic Relations between Speakers of Old Norse and Old English*. Turnhout: Brepols. https://doi.org/10.1484/M.SEM-EB.5.106296

Trotter, D.A. (ed.) 2000. *Multilingualism in Later Medieval Britain*. Cambridge: D. S. Brewer.
Trotter, David 2003. Oceano vox: You never know where a ship comes from. On multilingualism and language-mixing in medieval Britain. *Aspects of Multilingualism in European Language History*, ed. by Kurt Braunmüller & Gisella Ferraresi. 15–33. Amsterdam: John Benjamins. https://doi.org/10.1075/hsm.2.02tro
Trotter, David A. 2006. Language contact, multilingualism and the evidence problem. *The Beginnings of Standardization: Language and Culture in Fourteenth-century England*, ed. by Ursula Schaefer. 73–90. Frankfurt: Peter Lang.
Trotter, David A. 2011a. Language labels, language change, and lexis. In Kleinhenz & Busby (eds.), 43–61. Turnhout: Brepols. https://doi.org/10.1484/M.TCNE-EB.3.4602
Trotter, David A. 2011b. Death, taxes and property: some code-switching evidence from Dover, Southampton, and York. In Schendl & Wright (eds.), 155–190. https://doi.org/10.1515/9783110253368.155
Trotter, David 2013. Deinz certains boundes: Where does Anglo-Norman begin and end? *Romance Philology* 67 (1): 139–177. https://doi.org/10.1484/J.RPH.1.103932
Voigts, Linda Ehrsam 1996. What's the word? Bilingualism in late-medieval England. *Speculum* 71: 813–826. https://doi.org/10.2307/2865721
Wright, Laura 1997. The records of Hanseatic merchants: Ignorant, sleepy or degenerate? *Multilingua* 16 (4): 339–350. https://doi.org/10.1515/mult.1997.16.4.339
Wright, Laura 2000. Bills, accounts, inventories: Everyday trilingual activities in the business world of later medieval England. In Trotter (ed.), 149–156.
Wright, Laura 2011. On variation in medieval mixed-language business writing. In Schendl & Wright (eds.), 191–218. https://doi.org/10.1515/9783110253368.191
Wright, Laura 2013. The contact origins of Standard English. In Hundt & Schreier (eds.), 58–74. https://doi.org/10.1017/CBO9780511740060.004
Wright, Laura 2015. On medieval wills and the rise of written monolingual English. *Approaches to Middle English: Variation, Contact and Change*, ed. by Javier Calle-Martin & Juan Camilo Conde-Silvestre. 35–54. Oxford: Peter Lang.
Wright, Laura 2018. A multilingual approach to the history of standard English. In Pahta, Skaffari, & Wright (eds.), 339–358. https://doi.org/10.1515/9781501504945-016
Wright, Laura 2020. Rising living standards, the demise of Anglo-Norman and mixed-language writing, and standard English. In Wright (ed.), 515–532. https://doi.org/10.1515/9783110687545-019
Wright, Laura (ed.) 2020. *The Multilingual Origins of Standard English*. Berlin: Mouton de Gruyter.
Wright, Laura 2023. How multilingualism came to be ignored in the history of Standard English. *Multilingualism and History*, ed. by Aneta Pavlenko. 107–122. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009236287.005
Wright, Laura in press. On supralocalisation and Anglo-Norman lexis in English. *Exploring Supralocalisation Processes: Linguistic Variation in Early Modern Urban and Regional Spaces*, ed. by Anita Auer, Moragh Gordon, & Tino Oudesluijs. Language Science Press.

“I felt like English excluded particular groups” Tracing Finnish Instagram influencers’ language choices and multilingual practices

Hanna Limatius
ORCID: 0000-0002-2842-2358

In order to succeed, social media influencers need to use language in a way that appeals to their audiences and potential business partners. While previous research has investigated influencers’ linguistic practices in terms of self-branding and constructing authenticity, the role of the influencers’ language choice has not been rigorously studied. This chapter utilizes methods of discourse-centered online ethnography to examine language choice and multilingual practices among Finnish beauty influencers. Social media data collected from Instagram illustrate an overall trend of prioritizing Finnish as a content language, while English is used in specific contexts, short-form videos in particular. The commercial and ideological motivations behind such choices are investigated through questionnaire responses from thirteen influencers.

Keywords: computer-mediated communication, Instagram, language ideology, multilingualism, social media influencers

https://doi.org/10.51814/ufy.1041.1454

1. Introduction

This chapter investigates social media influencers' multilingual practices and language choices. More specifically, I examine how Instagram's beauty influencers' multilingual practices evolve over time, adapt to different contexts, and are shaped by their attitudes towards language(s). As previous research has shown, language is an important tool for influencers, as it enables them to construct authenticity and to establish their personal brands (Kováčová 2021; Meer & Staubach 2020). However, the attitudes and motivations behind an influencer choosing particular language(s) for their content creation practices have not yet been rigorously studied. In this chapter, I seek to address this gap in research through a detailed, qualitative study focusing on a group of Finnish beauty and fashion influencers and their language use. The chapter builds on a previous pilot study on multilingualism and multimodality in influencer communication (Limatius 2023).[1]

The data for the study consist of questionnaire responses from thirteen Finnish influencers and multimodal social media data collected from the Instagram accounts of six influencers. The study can be described as a *discourse-centered online ethnography* (Androutsopoulos 2008; Georgalou 2021), as it combines ethnographic observation of social media material and direct contact with influencers. Through this approach, I answer the following research questions: 1) What are Finnish beauty influencers' attitudes towards and motivations for using English and Finnish on social media? 2) How are such motivations and attitudes reflected in influencers' language use on Instagram? 3) How do the influencers' multilingual practices and language choices evolve over time?

The results indicate a shift towards prioritizing predominantly Finnish language social media content among the Finnish online beauty community. English use and multilingual practices, however, are emphasized in specific contexts – most prominently in the recently launched Instagram Reels. These developments reflect the saturation of the online beauty space, and a broader social media trend that prioritizes engagement (i.e., how actively one's followers engage with social media posts) over the number of followers (Greiwe 2018). Such developments can be traced through the Instagram data, while the

1 This research was supported by funding from the Ella and Georg Ehrnrooth Foundation.

questionnaire responses illustrate the motivations behind influencers' language choices.

The chapter is structured as follows. In the literature review, I discuss previous research on multilingualism, language choice and social media. Next, I discuss the concept of the social media influencer and their role in the online beauty space. After these theoretical sections, I present the data and methods of the study, followed by a discussion of the results. The chapter ends with conclusions and recommendations for future research.

2. Literature review

2.1. Multilingualism, language choice and social media

In this chapter, I utilize the related theoretical concepts of *multilingualism* and *language choice* in the context of social media. I follow Blommaert's (2010: 102) definition of multilingualism as "a complex of *specific* semiotic resources, some of which belong to a conversationally defined 'language', while others belong to another 'language'" (emphasis original). These resources include ways of using language in different communicative situations, as well as "the ideas that people have about such ways of using, their languages ideologies" (Blommaert, ibid.). In my analysis of influencer communication, I characterize the ways influencers draw from different languages and linguistic repertoires in their communication as their *multilingual practices* (Pahta, Skaffari, & Wright 2017). In investigating these multilingual practices, my interest is not in studying the influencers' language proficiency – rather, I consider multilingual practices as one "tool" that influencers use to achieve their communicative goals on social media.

By language choice, I refer to the ways in which users choose to utilize different linguistic resources when communicating with others on social media (Lee 2016: 23). According to Lee (2016: 34–35), when choosing which language(s) to use in their computer-mediated interactions, users consider a "number of ecological factors". Such factors include, for example, the users' cultural backgrounds, experiences, attitudes, and perceptions towards different linguistic resources, as well as the intended recipient of the communication, the content of the communication, and the technological affordances of the platform where the communication takes place (Lee, ibid.). Choosing particular linguistic resources for online interactions can also be motivated by purposes of creativity and play (e.g., Dubravac & Skopljak 2020), or identity construction (e.g., Blackwood 2019; Lee & Chau 2018; Pérez-Sabater & Maguelouk Moffo 2019). For example, bilingual speakers may use English in contexts where a "global" identity is emphasized and their local language when they want to highlight an in-group identity (Lee 2016; Pérez-Sabater & Maguelouk Moffo 2019: 35).

While there are numerous studies on multilingual practices and language choice on specific social media platforms such as Facebook (e.g., Androutsopoulos 2014; Pérez-Sabater & Maguelouk Moffo 2019) and Instagram (e.g., Blackwood 2019; Lee & Chau 2018; Mohr 2022), the ways in which professional content creators utilize these practices in their work have not yet received much scholarly interest. In the present chapter, my focus is specifically on the language use of such professionals – namely, influencers who operate in the online beauty space.

2.2. Influencers and language use

In this chapter, I define an influencer as a social media user who posts content regularly and publicly, with the goals of monetizing this content (e.g., through sponsorships), and gaining a large following (see e.g., Abidin 2016; 2018; Hudders & De Jans 2022). The most successful influencers can become celebrities (e.g., Abidin 2018; Jerslev 2016), but many of them are "everyday, ordinary Internet users" (Abidin 2016: 3) interested in creating social media content around a specific topic that is related to their personal expertise (e.g., fashion, parenting, music or technology). Depending on the broader category within which the influencer situates themselves, their content may contain a variety of communicational genres, from product reviews to comedy skits, or even political commentary. However, regardless of genre, maintaining relationships with followers is at the core of influencer work, and this entails projecting a curated image of authenticity, relatability and credibility (e.g., Kováčová 2021; Meer & Staubach 2020; Pöyry *et al.*, 2019; Reinikainen *et al.* 2020). Influencers appeal to their following by blurring the lines between professional and public – for example, they may share parts of their personal lives with their audience, thus creating a sense of intimacy (Meer & Staubach 2020: 251). Indeed, the relationship between a follower and an influencer is often characterized as a *para-social relationship*: an individual's one-sided, emotional attachment to a media personality (Isotalus 2023: 36). To their follower base, an influencer appears as a peer, even a friend, and as a result, followers are more receptive to influencer marketing than traditional advertising (Reinikainen *et al.* 2020).

Beauty and fashion brands utilize social media marketing particularly actively. Instagram especially is considered an important marketing platform due to its ease of access and its attractiveness to "shopping enthusiasts from all socioeconomic backgrounds" (Kim, Hong, & Lee 2021: 179). Since the market for makeup and other beauty products has become extremely saturated, brands are increasingly looking towards influencers to market their products to their followers. According to Statista (Dencheva 2023), in the year 2022, beauty was the third most popular category of Instagram influencers worldwide, lifestyle being number one. Endorsements by popular influencers are valuable to brands, as followers pay close attention to and trust influencers' recommendations (Reinikainen *et al.* 2020: 280). The potential that social media influencers have for creating *Electronic Word-of-Mouth (eWOM)* is enticing to brands, as influencer endorsement can have positive effects on a brand's image and create "buzz" around them (Zhou *et al.* 2021). At the same time, brand deals are equally important to influencers, as they benefit from them both monetarily and in terms of exposure (Reinikainen *et al.* 2020: 281).

While many social media applications, including Instagram, emphasize visuality, verbal communication – spoken or written – is one of influencers' most important tools in maintaining relationships with followers and brands. Thus, the language that influencers use on social media matters. Language can be used to construct the influencer's image as authentic and credible – for example, Kováčová (2021) found that fashion and lifestyle bloggers used informal language to help create a balance between "publicness and privateness" in their Instagram posts. Similarly, according to Meer and Staubach (2020: 254), influencers can use linguistic resources to mark their belonging in a social group – for example, an influencer whose target audience are teenagers can "[testify] her status as a peer" by drawing on linguistic practices typically used by teenagers.

While the role of language use for influencers has already received some scholarly attention, studies that focus specifically on how and why influencers choose to draw from particular languages in their social media content are needed. In a globalized world, an influencer can have followers and brand partners from a variety of countries and language backgrounds. Thus, they also need to decide whether they want to actively "influence" within a smaller, more local community, or on a broader, international level.

In this chapter, I focus on the language use of Finnish influencers who create beauty- and fashion-focused content on Instagram and other social media platforms. Finland is a fairly small country with a population of ca. 5.56 million (Clausnitzer 2023). Out of this population, the majority of ca. 4.9 million people speak Finnish as their first language (Institute for the Languages in Finland). However, English is used extensively across various communicative contexts in Finland, includig social media contexts (e.g., Coats 2019; Kytölä 2023), and it can be considered part of the (socio)linguistic landscape of Finland (Leppänen & Pahta 2012; Mortensen 2023). As influencer practices "[converge] with a framework of social and cultural phenomena" (Hurley 2019: 3), investigating the language choices of Finnish social media influencers can also provide insights into current language ideologies in Finland.

3. Data and methodology

3.1. Overview of data and analytical approach

Discourse-centered online ethnography (Androutsopoulos 2008) combines the systematic, ethnographic observation of online texts with direct contact with the producers of these texts. In this chapter, I combine a study of influencers' multilingual practices on Instagram with a questionnaire that focuses on the influencers' own reflections on their language use. As Androutsopoulos (2008) points out, when investigating language use in digital media, focusing purely on the texts produced is not always enough, as the users' motivations for using specific linguistic practices in their online interactions also need to be considered. In the case of multilingual practices and language choice, contact with the producers of social media texts is necessary to learn about ecological factors like social and cultural backgrounds, attitudes, and experiences (Lee 2016: 34–35). Direct contact with participants can offer new perspectives to the phenomena observed in ethnographic data or reveal completely new phenomena that do not emerge from the textual and (audio)visual data (Georgalou 2021).

Altogether, the data consist of questionnaire responses from thirteen Finnish influencers and multimodal social media data that was collected from six influencers' Instagram accounts. Based on the number of followers (between roughly 1,750–3,800), the six influencers whose Instagram accounts were studied can be described as *nano influencers*[2] (Wiley 2021). These six influencers had been participants in a previous study, as well (Limatius 2023), and I identified more influencers through the lists showing who they are following on Instagram. The questionnaire link was sent to thirty influencers and this group also included the six influencers whose Instagram accounts were observed.[3] Thirteen influencers responded.

The questionnaire was created with the online survey and reporting tool Webropol and it featured one multiple-choice question and four open-ended questions. The language of the questionnaire was Finnish. In the multiple-choice question, the influencers were asked to specify which languages they used on social media. In the open-ended questions, they were asked whether they considered creating content in Finnish and English important, and to explain why, or why not. They were also asked to reflect on any potential changes to their language choices on social media and differences between particular social media platforms.

All six influencers whose Instagram posts were included stated they would prefer to appear under their own names as opposed to pseudonyms. Instagram data was manually collected on two occasions – during a six-month period in 2020, and a three-month period in 2023 – and it contains altogether 526 posts. The data are summarized in Table 1 below.

Table 1: The data.

Data type	Time of collection	Items analysed
Questionnaire responses	2021-2022	13 respondents, five questions
Instagram data	2020 and 2023	Six influencers, 526 posts

For ethical transparency, it should be mentioned that since I am personally active in the online community of Finnish beauty and fashion creators, I was familiar with some of the participants prior to starting the research. This

2 Number of followers in March 2024. *Nano influencer* refers to influencers with less than 10,000 followers. However, the number of the influencers' Instagram followers does not necessarily reflect their total social media following, as some of them had more followers on other platforms (e.g., YouTube, TikTok).

3 As the questionnaire was anonymous, the two datasets were analyzed separately.

issue was mitigated in the research design in two ways. First, I made sure to also include influencers I was unfamiliar with and follow the same coding procedure for all participants. Second, to avoid any bias in the reporting of the questionnaire responses, they were collected anonymously.

3.2. Analysis of questionnaire responses

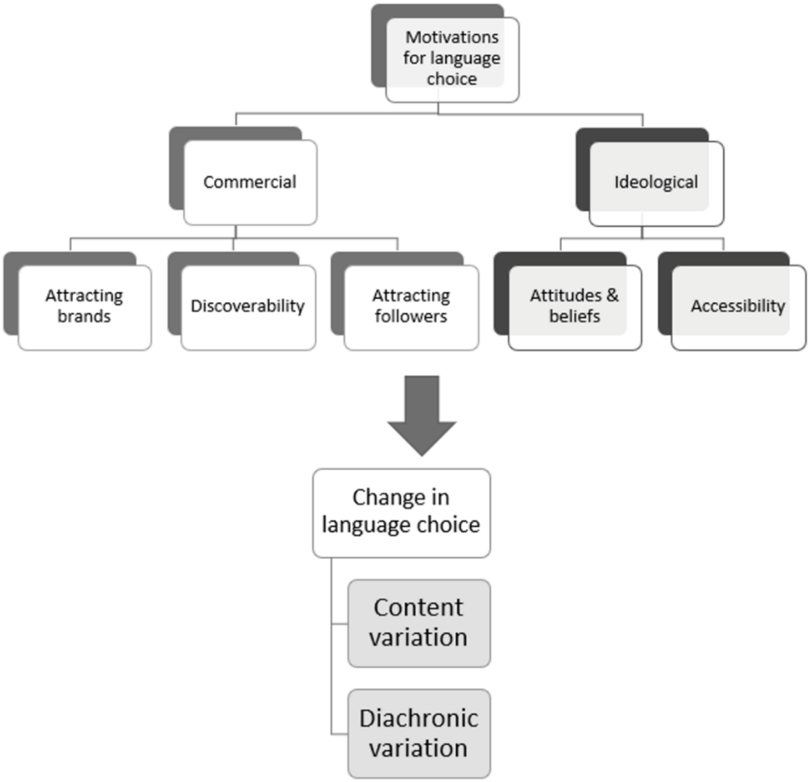


Figure 1: Motivations for language choice in the questionnaire data.

Thematic analysis was used as the analytical method for both datasets (Braun & Clarke 2006). I began the analysis of questionnaire responses by reading each response multiple times and identifying the topics discussed by the influencers. In the next step of coding, I started highlighting specific parts of the response texts that dealt with similar issues, establishing patterns and connections between different respondents. Next, I arranged these patterns into more cohesive themes. These themes can be divided into two broader groups: 1) topics pertaining to the influencers' motivations for using language in specific ways, and 2) topics pertaining to change and/or variation in their language use. Motivation-related themes include the sub-themes of commercial and ideological motivations, and changes in language use include the sub-themes of variation according to social media content and diachronic variation. Figure 1 illustrates the themes established in the analysis of questionnaire responses and the relationships between them.

3.3. Analysis of Instagram data

Initial Instagram data were collected as a part of another research project in 2020. However, as both the questionnaire responses and the previous pilot study (Limatius 2023) had shown diachronic change in the influencers' language choices, I decided to supplement the original dataset with additional data from the first three months of 2023. As a result, a dataset of 526 Instagram posts was formed.

Due to the multimodal nature of the data, I first needed to identify the different types of meaning-making resources that were present. In addition to the written captions of the Instagram posts, I analyzed text that was superimposed on images or video, the influencers' responses to comments from followers, as well as the spoken narration or audio tracks featured in Reels videos. When editing a Reel, an Instagram user can either record their own audio or use the existing audio tracks provided by Instagram. These include both copyrighted material (e.g., pop songs) and tracks recorded by other Instagram users who have chosen to make their audio publicly available. In addition, the user can add captions or other media to their video (see the example in Figure 2).

Again, I began the analysis by carefully reading each post, as well as watching and transcribing the Reels. Next, I documented all instances of code-switching in captions, images, or video (e.g., in audio tracks or superimposed text). Based on this initial analysis, I wrote a short summary of the findings for each influencer and each year. I made separate notes for posts published in 2020 and 2023 to facilitate comparing and contrasting the findings.

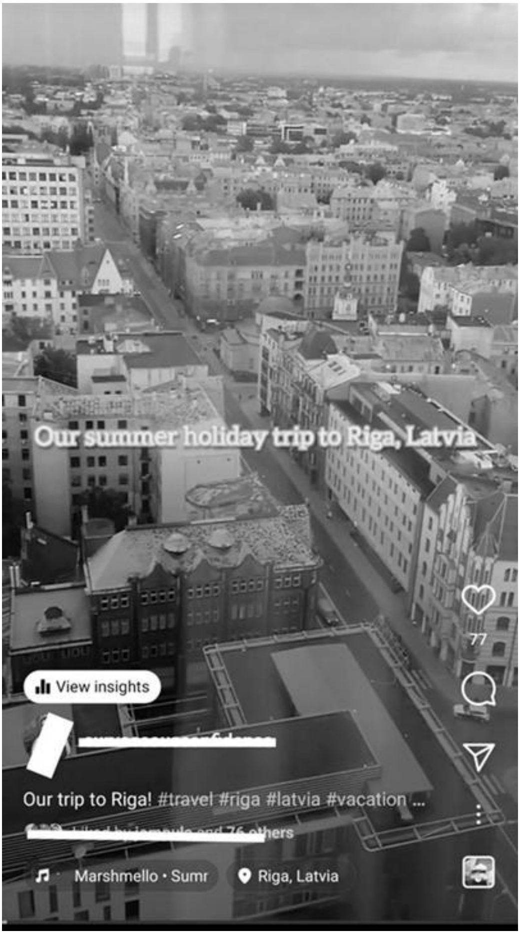


Figure 2: A still image from an Instagram Reel from the author's personal Instagram account.

In the next stage of the analysis, I again used thematic analysis (Braun & Clarke 2006) to identify similarities, differences, and patterns in the influencers' communication and their use of multilingual practices. As a result, four themes or categories of multilingual practices were established: 1) stylistic multilingual practices, 2) instructional multilingual practices, 3) multilingual practices pertaining to beauty terminology, and 4) community- or audience-specific multilingual practices. Table 2 illustrates the presence of these practices in the Instagram posts of each influencer.

Table 2: Multilingual practices in the IG data.

Influencer	Number of posts	Multilingual practices	Differences between 2020 and 2023
Annie @anniemuukka (3,819 followers)	31	Stylistic, instructional, beauty terminology, community-specific	A move towards prioritizing Finnish.
Cece @cece_bombshell (2,754 followers)	130	Community-specific	No apparent differences.
Elliroosa @elliroosa (3,299 followers)	43	Stylistic, beauty terminology, community-specific	English use in Reels, a move towards prioritizing Finnish in post captions.
Enni @ennireme5 (1,748 followers)	140	Stylistic, instructional, community-specific	A move towards prioritizing Finnish.
Hanna @hannamariaw (2,866 followers)	133	Stylistic, instructional, beauty terminology, community-specific	A move towards prioritizing Finnish.
Maria @mariaglow1 (2,413 followers)	49	Stylistic, beauty terminology, community-specific	English use in Reels, otherwise no apparent differences.

4. Discussion of results

In this section, I will present the results of the two analyses. The answers to the questionnaire have been translated from Finnish.

4.1. Questionnaire results

4.1.1 Motivations for language choice

Out of the thirteen influencers who responded to the questionnaire, seven reported they used “mainly Finnish” when creating social media content. Four influencers said they used “only Finnish”, and two “only English”. None of the respondents reported languages other than English or Finnish.

Most respondents considered using Finnish highly significant in terms of their commercial success and growth as social media influencers. It was important for them to respond to the needs of their followers, and most of these followers were Finns. Similarly, the influencers reported that they mostly worked with Finnish brands and businesses, and using Finnish was considered necessary to secure sponsorships. Language was viewed as an important marketing tool, as one of the respondents explained:

> **1** I consider [using Finnish] important because the majority of my followers are Finnish. My commercial partners as well. The language choice is determined by the effectiveness of marketing – in other words, making sure that the message goes through.

The influencer quoted in (1) has clear, commercial motivations for choosing to make content in Finnish. For her, language choice is thus a straightforward, strategic decision. The respondents also explained that they used Finnish on social media because of the saturation of the market. Some influencers considered English beauty content to have more competition, as it is ubiquitous on social media. Finnish use, however, was seen to increase discoverability by the “right” users and brands:

> **2** Yes, I consider [using Finnish] important, although I would prefer to create content in English, as it would feel more natural to me. There is less content available in Finnish in my chosen genre, so it serves a more targeted audience. This is also my full-time job, so if I create content in Finnish, I'm more likely to get Finnish brand partners.

Interestingly, the influencer in (2) stated that even though she would personally prefer to use English, she has chosen not to, as using Finnish is more commercially rewarding. Previous research has also reported that many non-native English speakers, particularly young people, prefer to express themselves in English in certain contexts, including social media (e.g., Dubravac & Skopljak 2020; Mortensen 2023: 118). Thus, the experience described by the respondent in (2) is not exclusive to influencers – however, the need to strategize language choice for financial gain is what is interesting here. Responses like this illustrate that the commercial motivations and the influencers' personal preferences do not always meet.

However, several respondents did state that using Finnish was easier for them personally, as they were able to express themselves in more creative ways in their mother tongue. One influencer even mentioned that she felt it was important to use Finnish to preserve the Finnish language in a time where anglicisms are becoming more and more common on social media. This response echoes the so-called "domain loss" discourse that has been present in language ideological discussions in the Nordic countries since the 1980s (Mortensen 2023: 106). According to Mortensen (ibid.), domain loss refers to "an imagined scenario where English comes to be used in specific societal domains […] *at the expense of* local languages" (emphasis original). The presence of this perspective in the data demonstrates that using Finnish could also be framed as ideologically motivated.

Despite the overall preference for Finnish, some influencers also considered using English important. As with Finnish, if the majority of an influencer's followers were international, English use was prioritized to reach those followers. Some respondents also stated that even though their audience was mostly Finnish, they had friends who did not understand Finnish. Interestingly, most respondents seem to equate a "Finnish" audience with a Finnish-speaking audience – the fact that there are people living in Finland who do not speak Finnish was not explicitly mentioned in the responses, although one influencer described some of her followers as "English-speaking or international".[4] Overall, however, English use appeared to be motivated more by personal preferences or ideological reasons than commercial success:

3 Personally, [creating content in Finnish] does not seem relatable, because in my experience, Finnish social media is not very accepting

4 The position of Swedish as the second official language of Finland was not mentioned by any of the respondents.

of diversity. However, my goal is to move abroad at some point, and then I might consider increasing Finnish content.

While the influencer in (3) also acknowledged that using Finnish on social media could be useful – she responded "Yes and no" to whether she considered Finnish use important – she personally experienced the Finnish-speaking social media space as limiting. For her, the use of the Finnish language was connected to negative experiences of social media use – however, she also felt like her views might change if she moved abroad. While she did not state why that is, one possible reason could be that using Finnish might help her retain a connection to her home country.

Finally, both Finnish and English use were discussed in the context of accessibility. The influencers wanted their content to be easily understood by their target audience – whether that audience was Finnish or international. Some influencers referenced the role of English as a lingua franca, as English was described as "understood by most people around the world", and "an inclusive language choice". However, others made similar arguments for Finnish, as the influencer in (4):

4 I considered [using English] at first, but I felt like English excluded particular groups. After all, my goal is to grow in Finland among Finns, which is why Finnish became established in my content.

In (4), the influencer discusses accessibility in a different context than the respondents who highlighted the role of English as a lingua franca – she focuses on inclusivity "among Finns" in particular. While the respondent does not specify which groups would be excluded by English use, her response acknowledges that all Finns do not have the same level of competence in English. At the same time, however, she constructs a narrow view of "Finns", only including Finnish-speakers. Thus, the influencers also expressed different ideological perspectives in their discussions on language, accessibility, and inclusivity.

4.1.2 Change in language choice

While the questionnaire contained questions specifically about potential changes in the languages the influencers used on social media, the theme of change over time was also evident in other responses throughout the data. Out of the thirteen influencers who responded to the questionnaire, eight reported that they had changed the language(s) they used on social media. Typically, this shift had been towards prioritizing Finnish over English, although one influencer reported that she had changed a Finnish Instagram account to a bilingual one:

> **5** However, on Instagram I have two accounts. On my modelling account, I only create content in English, but on my home renovation account I use both Finnish and English. I made this choice as initially, most of the followers were my Finnish acquaintances. Slowly, I started promoting this other account on my modelling account, and made it bilingual when I got new, international followers.

The response in (5) illustrates that much like language choice in general, change in language choice could be motivated by wanting to respond to one's followers' needs. It also highlights how influencers need to negotiate their language practices in regard to the *context collapse* typical of social media (Androutsopoulos 2014). Context collapse refers to the phenomenon where an audience of a communication consists of people from various social backgrounds and relationships, who may have different communicative expectations (Androutsopoulos 2014; boyd 2011). In the case of the respondent in (5), changes in audience motivated a change in language to facilitate understanding. Thus, depending on how their follower base evolves, influencers strategically adjust their language use. The followers may even be consulted directly – one influencer mentioned that she let her followers vote on which language(s) they preferred and changed her content to Finnish-only as a result. Brand partnerships were also mentioned as a motivation for changing the language of social media content. Four influencers stated that they had switched from English to Finnish because they wanted to secure sponsorships from Finnish brands.

Indeed, most influencers named commercial motivations – increasing follower count, increasing engagement, and maintaining brand relations – as reasons for changing the content language. Ideological reasons or language proficiency did not seem to play much of a role in switching languages.

However, the influencer in (6) discussed the role of content topic, as well as the technological affordances of specific social media platforms:

> **6** When posting about more local topics, I use Finnish, but for more global themes, I often use English. Partly the move towards Finnish has been because it's just easier – using one's mother tongue is more straightforward, although I also use English fluently. On the other hand, these days [Instagram] offers a translation function, which makes it easy for users to translate the content to English if they want to. This has enabled international followers and friends to keep following my account despite the language barrier.

To the respondent in (6), changes in language choice have been made possible through the ease of automatic translation. Social media platforms change rapidly and integrate new features regularly. Some features, like the translation tool mentioned in (6), can help influencers navigate problems caused by context collapse (Androutsopoulos 2014: 64). Influencers may also have different audiences on different platforms. Because of this, I also asked them if the choice of platform affected their language choice.

Six influencers reported differences in their language choice across social media platforms (blogs, Instagram, Snapchat, TikTok, and YouTube). Consistent with other responses, the influencers stated their target audience and potential collaborators on specific platforms as their motivation for choosing the content language. However, the specific affordances of different platforms were also mentioned. One influencer reported that she used English hashtags on TikTok and Instagram for discoverability, and another mentioned that while her original content was in Finnish, she liked to use other people's English audio tracks on TikTok. Like in Instagram Reels, TikTok users can make their audio tracks publicly available, meaning other people can use them to create their own content. Using a "trending" audio track can increase a user's chance of getting views; thus, English tracks can make Finnish influencers' content more widely distributed. Interestingly, however, the influencer in (7) also argued that the shift towards video-based content like Reels had made her use more Finnish:

7 Before, I mostly shared pictures on Instagram, so the language didn't matter as much when the content was focused on images. Now that video material has become more central on Instagram, it is more important to communicate in the language that your followers best understand.

While content type and the technological affordances of social media platforms factor into influencers' language choice, different influencers experience these aspects in different ways and form their communicative strategies accordingly. Next, I will demonstrate how such differences are reflected in the influencers' multilingual practices through the analysis of Instagram data.

4.2 Instagram results

4.2.1 Types of multilingual practices in Instagram posts

The influencers utilized multilingual practices on their Instagram accounts in various ways, which could be categorized under four, broader themes: 1) stylistic use, 2) instructional use, 3) use pertaining to beauty terminology, and 4) community- or audience-specific use. Out of these, the community- and audience-specific multilingual practices were the most common, as all influencers used them. Stylistic multilingual practices were utilized by five influencers, and multilingual beauty terminology by four influencers. Instructional multilingual practices were present in the content of three influencers.

By community- and audience-specific multilingual practices, I refer to multilingual practices that appear to be determined by the recipient of the message. For example, even though Cece's Instagram captions were written in English, she responded to comments from Finnish followers in Finnish. She also used Finnish hashtags in some of her posts:

8 Last weekend was truly amazing. I've missed being on stage so much! This team is 💎Photo: @username #kinkycarnivalburlesque *#kaamoskabaree* #burlesque #dreamteam #burlylife *#burleskiasuomessa* (Cece, October 2020)

Even though the main message in (8) is in English, it is directed at specific Finnish-speaking audiences using the two hashtags (emphasized by the author) #kaamoskabaree ("polar night burlesque"; a name of a burlesque festival held in Northern Finland) and #burleskiasuomessa ("burlesque in Finland"). Although Cece consistently uses English in her Instagram content, and appears to have a large international following, she also wants her posts to be discoverable by the Finnish burlesque community through the relevant hashtags. Thus, the use of multilingual hashtags is motivated by community-specific communication. It also functions as an identity-constructing practice (e.g., Lee & Chau 2018; Mohr 2022) – by adding the Finnish hashtags, Cece highlights her identity as a Finnish burlesque artist.

While the use of hashtags is a common way of directing a message to specific communities on social media (e.g., Zappavigna 2018), the influencers also included community-specific multilingual practices within actual text captions:

9 When I was younger, I never thought I would do some modelling in the future, even if just as a hobby. But here I am, in my 30's, plus sized and posing for a very talented Rovaniemi based young photographer [photographer name], and feeling great about myself. [...] Would you like to see the whole dress soon? I can't wait to show you guys all the beautiful photos [photographer] took! 😍 Thank you [name] for amazing photography! Check out more of her photos on Instagram. ♡ FI *Jos tarvitsette todella taitavaa ammattilaisvalokuvaajaa erityisesti Pohjois-Suomessa, ottakaa [name] yhteyttä!* 😊 *[name] tekee sekä ulko- että studiokuvauksia.* FI Hair pin: @kappahl Earrings: @hm MUAH: me @ennislooks (Enni, July 2020)

The excerpt from Enni's caption in (9) is, overall, typical of her Instagram content at the time. In 2020, Enni posted mostly in English, but in this caption, she has included some Finnish. The translation of the Finnish interjection (italicized by the author) reads: "If you need an extremely talented professional photographer, particularly in Northern Finland, contact [name]! 😊 [name] does both outside and studio photography." The part of the text written in Finnish is thus not a direct translation of the preceding English text, but it gives more information about the photographer for a specific Finnish audience who may be interested in hiring this photographer. Notably, after this Finnish interjection, Enni switches back to English to provide details of her makeup and accessories. The post is multimodally structured so that the Finnish excerpt is separated from the surrounding text by two Finnish flag emoji that are placed before and after.

While the switch into Finnish in Enni's post in (9) is carefully planned and has the specific function of informing a particular audience, there were also multilingual practices that appeared more spontaneous and less specific to particular communicative purposes in the data. These were considered stylistic multilingual practices in my categorization, and they included, for example, using another language in common expressions or for emphasis, as in (10) and (11):

10 *My Twiggy moment*! Tässä yhdistyi monta mitä oon halunnut testata, mutta ei oo _alkuunkaan_ mun tyyliä. (Author's translation, AT from now on: "My Twiggy moment! In this look, I combined many things I've been wanting to test, but that aren't my style _at all_.") (Elliroosa, May 2020).

11 SELKÄONGELMA-*GALORE* (AT: "BACK PROBLEMS GALORE"). (Hanna, July 2020).

In (10), Elliroosa begins her caption of a makeup selfie with an English exclamation ("My Twiggy movement!"), and then goes on to specify some details about her makeup in Finnish. Here, the English phrase introduces the topic (a Twiggy-inspired makeup look) but does not direct the message to a particular community, nor provide any specific instructions for non-Finnish speaking followers. Thus, the use of English appears more of a stylistic choice, perhaps motivated by the topic (Twiggy being a British model). In (11), Hanna similarly introduces the topic of the post with the multilingual noun phrase "back problems galore". She then goes on to discuss a health issue first in Finnish, followed by an English translation of the main points. While providing an English summary of a Finnish text can be considered an instructional or a community-specific multilingual practice, the addition of the English "galore" in the otherwise Finnish title appears more of a stylistic choice, as the reader needs to be proficient in both English and Finnish to understand the multilingual expression.

Interestingly, Hanna was also the only influencer whose communication featured interjections in languages other than Finnish or English; namely, Spanish ("Sí! / Yes!") and Swedish ("Samma här! / Same here!").[5] Such interjections can also be considered stylistic, as they are common expressions, and did not function as translations of Finnish or English content. They could also be considered as instances of multilingual language play; these multilingual practices are creative and their motivations more social than pragmatic (Dubravac & Skopljak 2020: 64). Another way of conceptualizing such integration of words from different languages is the concept of *translanguaging* (e.g., García & Wei 2014; Pennycook 2017). Instead of viewing these practices as belonging to separate, labeled languages, they can be understood as dynamic, semiotic resources that are part of an influencer's linguistic repertoire, and can be applied to different (local or global) contexts of social media communication (García & Wei 2014: 18, 22–23).

The influencers also used multilingual practices in a way that was closely tied to their chosen genre of content creation by incorporating English beauty-specific vocabulary or terminology in otherwise Finnish posts. This sometimes resulted in hybrid, "Finglish" formations:

12 *Shiftaa* eri värisenä eri valaistuksissa. (AT: "[It] shifts into different colours in different lighting.") (Annie, January 2023).

In (12), Annie integrates the English verb "shift" into a Finnish product description. Such multilingual practices may be related to the prevalence of English in the online beauty space. It is worth noting that the influencers themselves may not consider these genre-specific uses as "English" – rather, they may be seen as "new" formations that are part of the shared linguistic repertoire of the international beauty community (García & Wei 2014: 21; Eckert & McConnell-Ginet 1992). In some cases, the use of English beauty terminology may also be motivated by the fact that Finnish translations are cumbersome or simply do not exist (see Limatius 2023).

Finally, I also observed instructional multilingual practices in the influencers' Instagram posts. Similarly to community-specific multilingual practices, instructional practices could be observed in captions that featured both Finnish and English. They were present in summaries or direct translations of the "main point" of the post:

---

5 Maria also used a Swedish hashtag (#midsommar) once.

13 Uusi video YouTubella, linkki kanavalle profiilissa! Kerrohan myös et minkä meikkiasian kanssa sinä kamppailet, niin voin laittaa sen listan jatkoksi. […] *I uploaded a new video on my Finnish YouTube channel. This one is about eyebrows*. (AT "New video on YouTube, link to my channel in profile! Also, let me know which makeup issue you struggle with, so I'll add it to the list […] I uploaded a new video on my Finnish YouTube channel. This one is about eyebrows.") (Annie, May 2020).

In (13), Annie promotes a new video on her YouTube channel. In the Finnish version of the caption, she tells her followers where to find the video, and asks them to submit their video requests. The English translation is more compact. Its main function appears to be explaining the point of the post, so that followers who do not speak Finnish are not left wondering. While the caption also has community-specific functions – Finnish speakers are instructed to act in specific ways while English speakers are not – the purpose of the post is explained in both languages, unlike in examples (8) and (9), where Finnish was only used to provide additional information, without any summary or explanation of the surrounding English content. Some influencers in the data – Enni and Hanna in particular – often wrote long captions in one language, followed by a shorter translation that explained the content of the post in the other language. Interestingly, changes in such multilingual practices could be observed over time. In the next section, I will discuss the ways in which the influencers' multilingual practices appeared to have evolved between 2020 and 2023.

4.2.2 Changes between 2020 and 2023

Out of the six influencers studied, four featured English less prominently on their Instagram accounts in 2023 compared to 2020. Interestingly, Elliroosa and Maria, who otherwise favored Finnish consistently, began using English in one specific context: the newly launched Instagram Reels. Cece was the only influencer whose language choice stayed the same – she posted consistently in English in both 2020 and 2023.

Annie and Enni had posted predominantly in English in 2020, but in 2023, both had shifted to posting mostly in Finnish. In Annie's case, some stylistic code-switching remained, as well as occasional English hashtags, but she no longer provided instructional translations of the post content in English. Enni had used English extensively in 2020, only opting for Finnish in posts that were collaborations with Finnish brands and photographers or featured specifically Finnish products and topics. By 2023, Enni no longer prioritized English in the same way, although she did post two captions that were entirely in English during the observation period. For the most part, however, English was only present in occasional hashtags and in Reels videos that had an English audio track. Interestingly, Hanna's posting style already changed during 2020 – she went from posting predominantly in English to prioritizing Finnish as the year progressed (see also Limatius 2023). In 2023, Hanna posted mostly in Finnish, but with some stylistic English use and English hashtags, as well as English audio tracks in Reels.

Indeed, the launch of Reels appeared relevant in terms of multilingual practices, as influencers who otherwise prioritized Finnish started including some English elements through the introduction of Reels. In Maria's case, all her posts during the 2023 observation period were Reels. While she had included some English hashtags in her photo posts in 2020, all her captions had been in Finnish. In 2023, she used English audio tracks in some of her Reels, thus introducing new forms of multilingual content:

14 **Reel audio**: I got this little container to organize my craft stuff and I'm feeling so organized. Yay! **Captions superimposed on the video**: I got this little container to organize my cosmetics and I'm feeling so organized. **Post caption**: *Kosmetiikkakaaoksen järjestely on loppuelämän työmaa.* 🙆‍♀️ It is what it is. 😅 *#kaaos #kosmetiikka #kauneustubettaja #järjestely* (AT: "Organizing the cosmetics chaos is enough work for the rest of one's life. 🙆‍♀️ It is what it is. 😅 #chaos #cosmetics #beautyyoutuber #organizing"). (Maria, February 2023).

The example in (14) illustrates the multiple layers of meaning-making in a multilingual Reel post. Maria has used an English audio track where another person is talking about organizing their "craft stuff". She applies the audio to a new context, which is constructed multimodally. On the humorous video clip, Maria is shown throwing a small container towards a large, messy collection of cosmetics on a desk, and in the English captions superimposed over the video, the words "craft stuff" from the original audio have been replaced by "cosmetics". In the written caption of the Reel post, she has included a Finnish statement about the troubles of keeping organized, followed by stylistic code-switching with the English expression "it is what it is", and Finnish hashtags. While the video by itself is understandable to any Instagram user who speaks English, the Finnish caption and hashtags connect it to more local communities and identities.

Similar to Maria, Elliroosa's content mostly focused on Reels in 2023. While she had utilized English in hashtags and some comments in 2020, these practices were rare in 2023, only appearing in one post and one comment respectively. However, the advent of Reels had again brought with it some new multilingual practices. Similarly to Maria, Elliroosa utilized English audio tracks in her Reels. In addition, her Finnish speech that appeared as audio for some of the Reels also consistently featured English beauty terminology (e.g., "contour", "bronzer", "setting spray", "glowy"). While these terms have existing Finnish equivalents, as discussed in Section 4.2.1, this type of multilingual practice likely results from English terms becoming part of the shared linguistic repertoire of the beauty community, not from influencers' personal language attitudes.

5. Conclusion

As the analyses presented in this chapter illustrate, the multilingual practices of social media influencers are diverse and dynamic. Social media itself is never stagnant, but constantly evolving. Influencers, perpetually balancing on the line between personal and professional, need to be willing to adapt to new trends and algorithmic changes, and to develop their content, including linguistic practices, accordingly.

In the case of the present study, both the questionnaire and Instagram data indicate a shift towards prioritizing the use of the Finnish language over English among Finnish beauty influencers. An interesting exception to this trend were the Reels videos, where the use of English audio tracks introduced new ways of constructing multilingual, multimodal content. Such audio tracks can be viewed as a shared linguistic resource among Instagram users from various social and cultural backgrounds, taken from their original context, modified, and recontextualized to create new meanings in more specific, local communities (e.g., Leppänen *et al*. 2013). In terms of language use produced by the influencers themselves, however, English appeared to have taken a less central role by 2023. In their questionnaire responses, the influencers framed this change as a strategic move to improve brand relationships and audience reception in the local, Finnish context. However, the technological affordances of social media platforms, such as advances in automatic translation (e.g., Kytölä 2023: 185–6) also have their role to play, as illustrated the respondent in (6). In the future, the significance of such affordances in the context of influencer communication should be studied in more detail. As the accuracy of social media algorithms improves, influencers' content may reach more targeted audiences, and they can adjust their language practices accordingly.

Throughout the analyses of the two datasets, the ways influencers carefully think about their linguistic practices and use them professionally are highlighted. This is particularly interesting as the legitimacy of influencer work is often questioned in mainstream media discussions (e.g., Droz-dit-Busset 2022). The results of the present study illustrate influencers' strategic approach to communication – choosing the right language or combining elements from different languages in specific ways can thus be viewed as professional resources.

Even though commercial motivations were emphasized in the influencers' questionnaire responses, the effects of language ideologies were also present. Both English and Finnish use were discussed in terms of accessibility, and the influencers also talked about their personal language preferences. In the Instagram data, the presence of creative language play and the construction of identities through shared beauty community vocabulary demonstrate that multilingual practices are an important communicative resource for influencers. At the same time, the move towards prioritizing Finnish that was explicitly articulated in the questionnaire data and reflected in the Instagram data may also be connected to shifting language-ideological stances (e.g., Leppänen & Pahta 2012; Mortensen 2023). For example, based on the questionnaire responses, many influencers seemed to equate "Finnishness" with the Finnish language. However, more in-depth interviews with influencers are needed to study the role of language ideologies in influencer communication in more detail. Another theme that warrants further investigation is the central role of the audience. As influencers appear likely to change their linguistic practices based on audience reception, future studies could examine the language attitudes among these audiences. For example, do the audience members have ideological motivations for voting in favor of an influencer changing their bilingual social media content to monolingual?

This chapter has added a new perspective to research on influencer communication, which has previously mainly focused on the construction of authenticity and relatability. Through a discourse-centered online ethnography approach (Androutsopoulos 2008), the study has zoomed in on language choice and multilingual practices from two perspectives, combining questionnaire data with ethnographic observation of Instagram communication. However, the chapter is limited by its focus on a small group of influencers who operate within a specific genre: the Finnish online beauty community. In the future, quantitative studies focusing on larger datasets are needed to explore broader trends and changes in influencers' multilingual practices. In addition, comparing the linguistic practices of beauty influencers with another category of content creators may be relevant. Notably, all the participants of the present study were women in their 20s or 30s, and previous research has demonstrated some gender-based differences in language choice on social media (e.g., Coats 2019). As such, investigating language choice and multilingual practices within a genre that includes more male and/or non-binary influencers is another potential topic for a future study. However, despite its limitations, the present study sheds light on the language-related factors influencers need to consider when crafting their social media presence. In addition, it highlights the benefits of discourse-centered online ethnography as a methodological approach to investigating social media work.

Bibliography

Primary sources

Cece Bombshell. (@cece_bombshell). Instagram. https://www.instagram.com/cece_bombshell/
Hanna (@hannamariaw). Instagram. https://www.instagram.com/hannamariaw/
Heikka, Elliroosa (@elliroosa). Instagram. https://www.instagram.com/elliroosa/
Maria Glow (@mariaglow). Instagram. https://www.instagram.com/mariaglow1/
Muukka, Annie (@anniemuukka). Instagram. https://www.instagram.com/anniemuukka/
Remes, Enni (@ennireme5). Instagram. https://www.instagram.com/ennireme5/

Secondary sources

Abidin, Crystal 2016. 'Aren't these just young, rich, women doing vain things online?': Influencer selfies as subversive frivolity. *Social Media + Society* 2 (2). https://doi.org/10.1177/2056305116641342

Abidin, Crystal 2018. *Internet Celebrity: Understanding Fame Online*. Bingley: Emerald Publishing Limited.

Androutsopoulos, Jannis 2008. Potentials and limitations of discourse-centred online ethnography. *Language@Internet* 5, article 8. https://www.languageatinternet.org/articles/2008/1610

Androutsopoulos, Jannis 2014. Languaging when contexts collapse: Audience design in social networking. *Discourse, Context & Media* 4–5: 62–73. https://doi.org/10.1016/j.dcm.2014.08.006

Blackwood, Robert 2019. Language, images, and Paris Orly Airport on Instagram: Multilingual approaches to identity and self-representation on social media. *International Journal of Multilingualism* 16 (1): 7–24. https://doi.org/10.1080/14790718.2018.1500257

Blommaert, Jan 2010. *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press.

boyd, danah 2011. Social network sites as networked publics: Affordances, dynamics, and implications. *A Networked Self: Identity, Community and Culture on Social Network Sites*, ed. by Zizi Papacharissi. 39–58. New York and London: Routledge.

Braun, Virginia & Victoria Clarke 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3 (2): 77–101. https://doi.org/10.1191/1478088706qp063oa

Clausnitzer, J. 2023, 17 November. Total population of Finland from 2017 to 2022, by age group. *Statista*. https://www.statista.com/statistics/521152/population-of-finland-by-age/

Coats, Steven 2019. Language choice and gender in a Nordic social media corpus. *Nordic Journal of Linguistics* 42 (1): 31–55. https://doi.org/10.1017/S0332586519000039

Dencheva, Valentina 2023, 28 June. Distribution of Instagram influencers worldwide in 2022, by category. *Statista*. https://www.statista.com/statistics/1123051/instagram-influencers-share-world-category/

Droz-dit-Busset, Olivia 2022. "So-called influencers": Stancetaking and (de)legitimation in mediatized discourse about social media influencers. *Discourse, Context & Media* 49: 100629. https://doi.org/10.1016/j.dcm.2022.100629

Dubravac, Vildana & Nadja Skopljak 2020. Foreign and multilingual language play on social media as an identity marker. *Journal of Multicultural Discourses* 15 (1): 61–79. https://doi.org/10.1080/17447143.2019.1701678

Eckert, Penelope & Sally McConnell-Ginet 1992. Think practically and look locally: Language and gender as community- based practice. *Annual Review of Anthropology* 21: 461–90.

García, Ofelia & Li Wei 2014. *Translanguaging: Language, Bilingualism and Education*. Basingstoke and New York: Palgrave Macmillan. https://doi.org/10.1057/9781137385765

Georgalou, Mariza 2021. New Greek migrant (dis)identifications in social media: Evidence from a discourse-centred online ethnographic study. *Humanities and Social Sciences Communication* 8 (155). https://doi.org/10.1057/s41599-021-00831-9

Greiwe, Craig 2018, 27 July. The influencer's new clothes: Fake followers, engagement and accountability in influencer marketing. *Forbes*. https://www.forbes.com/sites/forbesagencycouncil/2018/07/27/the-influencers-new-clothes-fake-followers-engagement-and-accountability-in-influencer-marketing/

Hudders, Liselot & Steffi De Jans 2022. Gender effects in influencer marketing: an experimental study on the efficacy of endorsements by same- vs. other-gender social media influencers on Instagram. *International Journal of Advertising* 41 (1): 128–149. https://doi.org/10.1080/02650487.2021.1997455

Hurley, Zoe 2019. Imagined affordances of Instagram and fantastical authenticity of female Gulf-Arab social media influencers. *Social Media + Society* 5 (1):1–16.

Institute for the Languages in Finland. On language – Languages in Finland. https://www.kotus.fi/en/on_language/languages_of_finland. Accessed 28 Nov 2023.

Isotalus, Pekka 2023. Aristoteleesta somepersoonaan: Sosiaalisen median vaikuttajat vanhojen vaikuttamisen teorioiden näkökulmasta. *ProComma Academic: Vaikuttava viesintä*, ed. by Hanna Reininkainen & Salla-Maaria Laaksonen. 29–45. Helsinki: ProCom – Viestinnän ammattilaiset.

Jerslev, Anne 2016. In the time of the microcelebrity: Celebrification and the YouTuber Zoella. *International Journal of Communication* 10 (19): 5233–5251.

Kim, Bokyung, Seoyeon Hong, & Hyunmin Lee 2021. Brand communities on Instagram: Exploring Fortune 500 companies' Instagram communication practices. *International Journal of Strategic Communication* 15 (3): 177–192. https://doi.org/10.1080/1553118X.2020.1867556

Kováčová, Dominika 2021. Becoming #Instafamous: The analysis of (in)formality in self-presentation on Instagram. *Internet Pragmatics* 5 (1): 12–37. https://doi.org/10.1075/ip.00069.kov

Kytölä, Samu 2023. Metapragmatics of "bad" English in Finnish social media. *English in the Nordic Countries: Connections, Tensions, and Everyday Realities*, ed. by Elizabeth Peterson & Kristy Beers Fägersten. 185–203. London: Routledge.

Lee, Carmen 2016. *Multilingualism online*. London: Routledge. https://doi.org/10.4324/9781315707211

Lee, Carmen & Dennis Chau 2018. Language as pride, love, and hate: Archiving emotions through multilingual Instagram hashtags. *Discourse, Context & Media* 22: 21–29. https://doi.org/10.1016/j.dcm.2017.06.002

Leppänen, Sirpa & Päivi Pahta 2012. Finnish culture and language endangered – Language ideological debates on English in the Finnish press from 1995 to 2007. *Dangerous Multilingualism: Northern Perspectives on Order, Purity and Normality,* ed. by Jan Blommaert, Sirpa Leppänen, Päivi Pahta, & Tiina Räisänen. 142–175. Basingstoke and New York: Palgrave Macmillan.

Leppänen, Sirpa *et al*. 2013. Entextualization and resemiotization as resources for (dis)identification in social media. *Tilburg Papers in Culture Studies 57*. https://research.tilburguniversity.edu/en/publications/entextualization-and-resemiotization-as-resources-for-disidentifi

Limatius, Hanna 2023. Examining the multimodal and multilingual practices of Finnish social media influencers. *Multilingualism from Manuscript to 3D: Intersections of Modalities from Medieval to Modern Times*, ed. by Matylda Włodarczyk, Jukka Tyrkkö, & Elżbieta Adamczyk. New York: Routledge. https://doi.org/10.4324/9781003166634-11

Meer, Dorothee & Katharina Staubach 2020. Social media influencers' advertising targeted at teenagers: The multimodal constitution of credibility. *Visualizing Digital Discourse: Interactional, Institutional and Ideological Perspectives*, ed. by Crispin Thurlow, Christa Dürscheid, & Federica Diémoz. 245–270. Boston and Berlin: De Gruyter Mouton. https://doi.org/10.1515/9781501510113-013
Mohr, Susanne 2022. Investigating English in multilingual contexts online: Identity construction in geotagged Instagram data. *Frontiers in Communication* 6: 1–14. https://doi.org/10.3389/fcomm.2021.778050
Mortensen, Janus 2023. Beyond threat or opportunity: English and language ideological tensions in the Nordic countries. *English in the Nordic Countries: Connections, Tensions, and Everyday Realities*, ed. by Elizabeth Peterson & Kristy Beers Fägersten. 104–124. London: Routledge.
Pahta, Päivi, Janne Skaffari, & Laura Wright, eds. 2017. *Multilingual Practices in Language History: English and Beyond*. Boston: De Gruyter.
Pennycook, Alastair 2017. Translanguaging and semiotic assemblages. *International Journal of Multilingualism* 14 (3): 269–282. https://doi.org/10.1080/14790718.2017.1315810
Pérez-Sabater, Carmen & Ginette Maguelouk Moffo 2019. Managing identity in football communities on Facebook: Language preference and language mixing strategies. *Lingua* 225: 32–49. https://doi.org/10.1016/j.lingua.2019.04.003
Pöyry, Essi *et al*. 2019. A call for authenticity: audience responses to social media influencer endorsements in strategic communication. *International Journal of Strategic Communication* 13 (4): 336–351. https://doi.org/10.1080/1553118X.2019.1609965
Reinikainen, Hanna *et al*. 2020. 'You really are a great big sister' – parasocial relationships, credibility, and the moderating role of audience comments in influencer marketing. *Journal of Marketing Management* 36 (3–4): 279–298. https://doi.org/10.1080/0267257X.2019.1708781
Wiley, Danielle 2021, 8 March. What you need to know about working with nano influencers. *Forbes*. https://www.forbes.com/sites/forbesagencycouncil/2021/03/08/what-you-need-to-know-about-working-with-nano-influencers/
Zappavigna, Michele 2018. *Searchable Talk: Hashtags and Social Media Metadiscour*se. London: Bloomsbury Academic.
Zhou, Shuang *et al*. 2021. Social media influencers' narrative strategies to create eWOM: A theoretical contribution. *International Journal of Information Management* 59. https://doi.org/10.1016/j.ijinfomgt.2020.102293

Flagging multilingual features in post-Conquest manuscripts
Verbal and visual

Janne Skaffari
ORCID: 0000-0001-7349-9904

In written language as well as in spoken communication, code-switching is often accompanied by flagging. Directing attention to features which are important or potentially difficult to process, flags also appear in medieval manuscripts and may have a visual or a verbal form. Visual flagging may involve, for example, the use of red ink, underlining, or marginal manicules, while verbal flags include metalinguistic labels specifying the embedded language or the main language, and intratextual translations of other-language material. Based on data from a large number of manuscripts produced in England in the long twelfth century, this study examines both Latin code-switches in Early Middle English texts and English switches appearing in Latin manuscripts. It considers motivations for both code-switching and concomitant flagging, and outlines a tentative typology of verbal and visual flagging, which may also be applicable to other periods and language pairs.

Keywords: code-switching, flagging, medieval manuscripts, multimodality

https://doi.org/10.51814/ufy.1041.1455

1. Introduction[1]

As a multilingual practice, code-switching – the use of two or more languages or language varieties in the same instance of communication, be it spoken or written – has become a prominent research topic in English historical linguistics particularly in the last decade or so (see Schendl & Wright 2011; Pahta, Skaffari, & Wright 2018). Such research often addresses the syntax of multilingual sentences, the socio-pragmatics of historical code-switching and the ubiquitous problem of distinguishing single-word switches from lexical borrowings. There is also a multimodal dimension in written code-switching as the presence of another language may be visually highlighted in a text. In addition to visual *flagging*, some verbal commentary – essentially verbal flagging – may appear adjacent to switches as well. Until recently, such flags have seldom merited more than a fleeting mention in published research (e.g. Voigts 1996; Rogos-Hebda 2016; Skaffari 2016), but there is now growing interest in linking multilingual practices with their material and visual context (Kopaczyk 2023: 122). The associations between language and the visuality of script may also interest palaeographers (see Aiello 2021).

The linguistic starting point of the present chapter is code-switching (henceforth, CS), which has a fairly long history as an object of linguistic enquiry but which has not been applied to written materials for equally long (see, for example, Sebba (2012) for a discussion of CS in written data). Research on spoken CS, which continues to offer frameworks for examining written CS, has established that switching from one language to another may be accompanied by changes in pitch and speed, pauses, metalinguistic comments and also non-verbal behaviour (e.g. Auer 1999: 314; Poplack 2004: 593), although these are not always present (see Gumperz 1982: 59–60). Such prosodic and other cues have equivalents in the written medium: just as CAPITALS can be interpreted as shouting in today's digital communication, foreign or 'other-language' words are often italicised in print media to signal difference from the main language of the text and perhaps, sometimes, give

1 The idea for this study emerged while I was working for Professor Päivi Pahta's research project Multilingual Practices in the History of Written English, funded by the then Academy of Finland. Most of the manuscripts had been consulted during my own Academy-funded project Multilingualism in the Long Twelfth Century. In addition to the now Research Council of Finland, I wish to thank the Institute of Medieval and Early Modern Studies for a library fellowship at Durham University.

them a certain *je ne sais quoi*.[2] This type of typographical highlighting in printed material has a parallel in hand-written texts, which has been called *script-switching* by, for example, Gardner-Chloros (2009: 21); Kaislaniemi (2017) discusses this concept at length. Other means of marking code-switches include the use of red ink and underlining. All of these are what Machan (2011: 305) and Rogos-Hebda (2016: 38) call *bibliographic codes*, which at times match changes in the linguistic code of the text.

There are several reasons for the relative sparsity of CS research on the visual and other flags accompanying language mixing. The visual, extra-linguistic features may have seemed irrelevant to linguists interested in syntax or sociolinguistics, or the features have been inaccessible to linguists. Inaccessibility is related to the fact that larger-scale studies of historical CS have of necessity often relied on existing diachronic or historical corpora (a case in point is the seminal paper by Pahta & Nurmi 2006), which are typically based on editions and include little information about the layout and appearance of the original sources. Case studies of CS in individual manuscripts or early printed books allow a different approach: their authors may also observe the physical context, but they do not necessarily consider the visual changes at the points where another linguistic code appears. Although such switch-points sometimes have also verbal flags, explicitly naming the language to or from which the switch is made, these *metalinguistic* flags do not necessarily catch the researcher's eye, perhaps because the languages involved have been regarded as obvious for the contemporary producer or reader as well as the later researcher, or because the focus has lain on syntactic or sociolinguistic aspects of CS. Flagging does, however, help to structure the text and may index the producer's expectations of the recipient or attitude toward the text. It may be seen to constitute *metadiscourse*, which Hyland (2017: 17) defines as ways in which "we use language out of consideration for our readers or hearers based on our estimation of how best we can

2 Numerous CS studies call the codes involved *matrix* and *embedded* languages. As terms, these are most appropriate for syntactic research on intrasentential CS, but as the switches in my material are of various types, and as syntax is not in focus, I will refer to *main* and *other* languages, the latter term inspired by Auer (1999: 314) and, in particular, Poplack (2018). Embedding nonetheless remains a practical concept.

help them process and comprehend what we are saying". Facilitating comprehension is extremely relevant in multilingual texts.

In contrast to earlier studies, this chapter discusses both visual and verbal flags accompanying CS in medieval texts. It cannot aim at an exhaustive catalogue of flagging phenomena but explores the range of flags attached to both 'English in Latin' and 'Latin in English', as seen in manuscripts from the two centuries following the Norman Conquest in 1066, or the 'long twelfth century'. Paying attention to both visual and verbal flagging makes the approach richer. Like most research on historical CS, the chapter draws on both historically orientated work and studies of contemporary spoken CS, although concepts from the latter need modifications when researching historical "multilingual literacy practices" (Schipor 2018: 39–42).

This study begins with a brief introduction to multilingualism in post-Conquest England and the primary sources consulted (Section 2). The multilingual context of this period is what produced the CS and forms of flagging described in Section 3, which explores different types of linguistic and non-linguistic cues accompanying changes in code. These practices are reviewed and elaborated on in Section 4. The outcome of this discussion is a tentative typology of flagging, capable of accommodating examples sought from the manuscripts of the long twelfth century but potentially applicable to other periods and contexts as well. This list of flagging types is presented in Section 5. Finally, Section 6 summarises the key findings and provides an outlook for future research on CS and flagging.

2. Medieval multilingualism and post-Conquest texts

Medieval multilingualism is discernible in multiple ways: it includes both societal and individual bi- and multilingualism; it shows in different topic domains, such as religion, and genres, such as instructional writing; and it is evident in the intertextuality of writing, with material adopted and adapted from various sources, including translations especially from Latin. In England after the Norman Conquest, individuals with bilingual (English-French) skills must have made use of CS in speech, but it is noteworthy that in the written medium – in manuscripts produced in the country – it was typically English and Latin that appeared together within texts, not the two vernaculars, English and French.

Compared to the preceding Old English period with its rich tradition of vernacular writing, the post-Conquest period has seemed quite difficult for historical linguists because of its alleged lack of English data.[3] This, however, is in part a matter of perspective: while it is true that the number of texts containing original material composed at this time is not large in comparison with the preceding and following centuries, older vernacular texts were still copied and circulated, often with modifications reflecting language use at the time of copying rather than that of composition (see e.g. Faulkner 2012).

The best-known new English texts surviving from the period are the last annals of the *Anglo-Saxon Chronicle* (in the Peterborough manuscript) and *Poema Morale* from the twelfth century, and *Ancrene Wisse* and the Katherine Group from the early thirteenth. The copied material, which is more abundant, includes, for example, many homilies by Ælfric of Eynsham, reproduced and reused for at least two hundred years after his death (Swan 2000: 67). Excellent resources for locating also less known English texts – and small amounts of consecutive English surviving in Latin books – from this period are the catalogues by Laing (1993) and Da Rold *et al.* (2010–2013), which directed the present author to view the original manuscripts on site in the repositories that hold them. The original objective was to account for CS as a linguistic phenomenon in such post-Conquest manuscripts. At some of the repositories I also engaged in what Daniel Wakelin (2014) has called "do-it-yourself digitisation", photographing relevant folia of dozens of manuscripts. While some libraries have made many of their manuscripts available as digital images on their websites, the coverage is still far from comprehensive, and a neglect of the long twelfth century can sometimes be discerned. Visiting the libraries has therefore been a crucial part of the project, not least because of the visual aspect of CS, which literally caught my eye and subsequently led to a change in the focus of my research.

I have thus far viewed examples of CS in over 120 manuscripts from around 1200, nearly all of them *in situ*. They were identified with the help of the aforementioned catalogues. Of the repositories in London, Oxford and Cambridge, which hold approximately 90 per cent of the relevant items, the most important single repository is the British Library. The majority of the texts are literary in the broadest sense of the term. The domain of religion is

3 For a recent discussion of the "linguistic ecology" of the period, see Faulkner (2022: 68–101).

dominant: the mainly English material includes, for example, homilies, while the Latin sources are more varied, including also saints' lives, theological texts and history. Latin is typically the main language in the data; this also applies to two of thirds of the nineteen manuscripts cited below.

3. Flagging in post-Conquest material

Flagging multilingual features in written materials means indicating or highlighting the presence of another language in a text or book.[4] More broadly, flagging can be explained as directing "attention to a feature of the written or spoken text that is important or requires more processing from the reader/hearer" (Nurmi & Skaffari 2021: 501). Flagging therefore facilitates comprehension, which is particularly relevant in multilingual communication, as not every potential participant of the communicative event can be expected to know both or all of the languages in the text or conversation.

In this section on flags in post-Conquest manuscripts, visual flagging is discussed before verbal; the latter cannot be subsumed under the former, although both types of flagging are *visible* to the reader. Anything that the reader can discern on the page is indeed visible, but not all of it is primarily *visual*, in the sense of 'non-linguistic'.[5] What is visible is central to my study and therefore described carefully, while the actual content of the other-language units is not, so the code-switches are not consistently quoted or translated below. Moreover, CS may also be unmarked, or "smooth" (Poplack 2004: 593), but such flagless switches are outside the scope of this study.

Before focusing on the visual, it is perhaps necessary to adjust expectations about the material characteristics of the manuscripts. The English books of the period were not visually impressive: Treharne (2012: 97) points out that expensive *de luxe* copies were rare. Vernacular books were practical materials meant for reading and use, not for display. The range of visual flags is thus not extremely wide, but the technically less demanding verbal flags may appear in all kinds of material.

---

4 For a discussion of 'highlighting' and other related terms, see Varila *et al.* (2017: 11–13).

5 Understanding the content provided on the physical page may require both reading skills and visual or graphic literacy (cf. Ruokkeinen *et al.* 2024).

3.1. Visual flags

Visual flagging was executed in medieval manuscript production by scribes and rubricators who in order to emphasise a particular feature of the text drew attention to it by choosing a position, size, colour or style for it that differed from its surroundings (see Carroll *et al.* 2013: 57). A very typical example of this practice – unlikely to surprise anyone familiar with reading medieval manuscripts – is the use of red ink for Latin passages, while the vernacular text around them is in ordinary black or brown. Extremely clear examples of this can be seen in, for example, the late twelfth-century London, Lambeth Palace Library MS 487, in which Latin quotations stand out in the English-language homilies, sometimes half a dozen of them on a single page (see Figure 1). Similarly, Ælfric's Christmas homily in Cambridge, University Library MS Ii.I.33 begins on f. 29r with a Latin *rubric*; the term itself suggests red colour, as its etymon, the Latin *rubrīca*, meant 'red ochre' and hence 'chapter heading (written in red)' (*OED* s.v. *rubric*). In Cambridge, Trinity College, MS B.14.52, which contains the *Poema Morale* and the 'Trinity Homilies', the ten-line Latin passage in the English homily on f. 14r can be located with ease since, unlike the English, it is dotted with thirteen red *litterae notabiliores* (larger or otherwise visually outstanding capital letters; here, large capitals in red).

Flagging could also be done by underlining words or longer stretches of text in the other language. Red ink could be used for this purpose (see Stowe 34, discussed below), but the colour did not have to differ from the rest of the text on the page. Moreover, it was possible to draw a frame or a box around other-language units or their flags.[6]

As flagging, the use of red ink as well as underlining and boxing must have been more conspicuous and thus more effective than, for example, slight changes in graph size, if also potentially slower to execute. With reference to script choice or script-switching, it was customary to use Caroline script for Latin and Insular script for English (e.g. Treharne 2000: 25); Caroline minuscule had been imported from the Continent approximately one hundred years before the Norman Conquest (Parkes 2008: 110). Change of script according to the language is evident in, for example, the oldest manuscript in the present

6 For a less concrete use of frames as a concept in linguistic CS research, see Gardner-Chloros (2009: 75).

study, Cambridge, University Library, MS Hh.I.10, whose earliest items are from the eleventh century (Kato 2010). However, it seems that the difference between the languages in terms of "visual quality" was lost by the thirteenth century (Aiello 2021: 36). Even without shifts in script or graph size, other-language expressions often somewhat stand out on the page, as the distribution of letters inevitably differs between Latin and Old or Middle English. The greater frequency of abbreviation and suspension signs in embedded Latin may alone catch the reader's eye, as may the presence of thorns and eths (*þ*, *ð*) in embedded English. These features are not deliberate flagging, however, as they are part and parcel of writing in the selected languages.

Figure 1. London, Lambeth Palace Library MS 487, ff. 18v–19r. Image courtesy of Lambeth Palace Library.

It is worth noting that visual enhancement is at times a means for differentiating between rhetorical units (see Machan 2011: 3) in the text, rather than highlighting an other-language unit. The red heading mentioned above is a typical example of this, and relevant here are also quotations: the Latin within the vernacular text may not be originally by the English author but introduced from another source. There is thus a link between quoting and change of language, but all quotations may not be highlighted: a case in point is London, British Library, MS Stowe 34, in which red underlining consistently accompanies Latin quotations (with only a few exceptions, which may be oversights), but not quotations in English, which must, ultimately, be translations from Latin authorities. Multilingual practices are thus intertwined with intertextual ones in medieval writing; visual flagging may accompany either or neither, and it may at times be difficult to discern which purposes the same type of visual flag serves in different contexts.

The examples of the use of red ink above are of Latin appearing in dominantly English-language texts, but similar means of flagging were also available for marking English in Latin. For example, the Latin sermons in London, British Library, MS Harley 3823 contain a few English verses, not clearly visually flagged apart from f. 354v, on which two couplets appear with red underlining; however, variants of the same verses are not visually flagged in Durham, Cathedral Library, MS B.I.18. Another use of red is visible in Oxford, Bodleian Library, MS Laud Misc. 511 (example 1):

1 Angł maiden stod at welle and
wep weilawei · late cemet ẏe
lith of dai·(f. 110v)
'In English: A maiden stood at the well and wept Weilawei, late comes the light of day.' (transl. JS)

Here the two and a half lines of the English lyric, appearing in a Latin sermon, are not underlined but struck through with red ink (not reproduced in example 1, due to technical constraints). The intention cannot have been to remove the English: strikethroughs in red mark also other words and abbreviations without making them illegible, and the use of red and blue ink leaps to the eye across this page. The code-switch is introduced with *Angł*, a shortened form of the language label *anglice* 'in English', preceding what Wenzel (1986: 225–226) calls an "abrupt quotation" and a "mysterious snatch" of vernacular verse. The visual flag is thus accompanied by a metalinguistic one, a feature discussed below (Section 3.2).

As noted above, the post-Conquest manuscripts containing English texts are often quite plain and unadorned, with little or no colour or other costly visual features, although they contain Biblical quotations and Latin words and phrases, elsewhere often picked up by various means. In some cases, the only visual cues guiding the reader to a quotation or a code-switch are marginal manicules (hand-shaped pointers). In Durham, Cathedral Library MS B.I.18 (a Latin manuscript of *Summa de Vitiis*), manicules can be seen in the margins or between columns on many folia, two of them pointing to embedded English (ff. 37r and 85v), both code-switches preceded by a suspended form of *anglice*. Similarly, London, British Library, MS Cotton Claudius D vii (f. 184v) also displays multiple flags: in the left margin, a manicule points to a red box with the words "angł dictū" inside, and the Latin body text next to the box contains the *anglice* label and almost two lines of English. Without careful palaeographical analysis, it is difficult to estimate whether manicules and the like are contemporary with the text or represent annotations by "any of the users of the manuscript" (Schipor 2018: 13), be they subsequent readers or much later scholars; for example, the brace next to the English 'Bede's death song' in the Latin of Oxford, Bodleian Library, MS Digby 211 (f. 108r) is later than the text in the column (Aiello 2021: 52). The same may be true of some of the underlining or boxing in the manuscripts as well. Multiple layers of markings are nonetheless evidence of use and, importantly, indicative of readers' engagement with other-language material.

A further visible feature is the use of margins or flyleaves for incorporating more text, including additions in another language; a fair share of the limited English material in mainly Latin manuscripts is housed in the extremities of page and book. For example, the bottom margin on f. 25r in Oxford, Bodleian Library, MS Digby 45, contains two and a half lines of the English version of "Candet Nudatum Pectus": "Naked was hẏs wite brest", etc. In contrast, in Oxford, Bodleian Library, MS Rawlinson C 317, this Latin poem and a version of its English translation appear together in the body text on f. 89v, separated by a paraph touched with red.[7] Similarly, marginal and inter-

7 A Latin-and-English sequence also survives in Durham, Cathedral Library, MS A.III.12. This is the oldest of the English "Candets" (Hanna 2011: 189). The English lines always appear with Latin, according to Hanna (2011: 191), but Digby 45 is actually an exception to this (see also Aiello 2021: 57–59).

linear *glosses* – providing translations or explications – are loci for multilingual practices (see 3.2 for related phenomena).[8] Whether we encounter other-language passages in the margin or glosses between lines of text, these are flagging in verbal form but in visually marked positions on the page. Verbal flags may also be placed within the body text, as we shall see next.

3.2. Metalinguistic flags

Metalinguistic flags are verbal means of pointing to code-switched units or acknowledging their presence: they may name the main language, the other language, or both. As observed by Nurmi and Skaffari (2021: 508), who look at various types of flags for Latin in the history of English and discover similarities in practices from Old to Present-Day English, the languages involved can be identified directly by their names or indirectly, for example as ‘our language’. This, however, is not the only way to flag code-switches verbally.

Examples of label-type verbal flags can be found in many of the manuscripts consulted. In Latin texts, *anglice* – also mentioned by, for example, Voigts (1996: 818) and already seen above – appears shortly before the English response to a toast in London, British Library, MS Cotton Tiberius B xiii (f. 95v). In London, British Library, MS Cotton Faustina A v, ‘Bede’s death song’ included in the Latin text is poised between references to both languages on f. 43r: “hoc anglico carmine cōponens” and “ita latine sonat”. While it is less easy to find other-language labels in the English texts of the period, main-language labels can be discovered instead: on f. 3r in Cambridge, University Library, MS Ii.I.33, for example, a Latin quotation from the Genesis is linked to its English translation with ‘that is in English’, combining an “apposition marker” (Pahta & Nevanlinna 1997: 124) with a label for the target language of the translation (the main language of the text).

In addition to naming languages, verbal flagging can also be assumed to incorporate various forms of *reiteration*: the content of the switch may be repeated in the main language, quite faithfully or with an elaboration or an abridgement. The Cotton Faustina folio cited above is a case in point: the *anglico* label is followed by the English song or poem, and the *latine* label by

8 The best-known glossator of the latter part of the period must be the Tremulous Hand of Worcester (see e.g. Franzen 1991), whose Latin glosses survive in, for example, London, British Library, MS Cotton Otho C i vol. 2 (Laing 1993: 79).

reiteration – i.e., translation – in the main language of the text. Reiteration can be used in spoken language to clarify the meaning of the code-switch, or to emphasise a point (see e.g. Gumperz 1982: 78–79). As a feature of medieval bilingual texts, reiteration has been acknowledged at least since Voigts (1996: 816), but there has been little sustained interest in it. Importantly, Diller (1997–1998) has called this type of reiteration "support";[9] see Skaffari (2016) for other possible terms, including *intratextual translation*, and Nurmi and Skaffari (2021) for support as *mediation*. Both Diller (1997–1998) and Skaffari (2016) identify elaboration as the most frequent type of support in their respective primary sources from different centuries: the little Latin quoted from an authoritative source – not unlike what Gumperz (1982: 71) called "truncated, idiomatic stock phrases" – sufficed to establish the reliability of the message conveyed, whereas more space was given to support, which made the content intelligible to a wider audience. The need for this must have been particularly keenly felt when producing texts with an instructional purpose, such as much of religious writing; many of the texts cited here can be assigned to such domains. The reproduction of a relatively large range of homiletic texts in England at this time is a compelling cue to what many of the books may have been intended for: they aided religious instruction, for which expensive decoration was not deemed necessary but which would not have functioned equally well if reiteration in the vernacular had not supported the authoritative Latin.

As seen above, support was at times accompanied by labels for the source or target language, or both. In Oxford, Bodleian Library, MS Bodley 343, for example, the Latin term *avarus* mentioned in the English text on f. 166v is translated with OE *gytsere* 'avaricious person', followed by "on englisc". This type of metalinguistic flagging was not, however, always done, as can be seen in the preface to the Early Middle English *Ancrene Wisse*, in Cambridge, Corpus Christi College, MS 402 (example 2):

2 Þeos riwle is eauer inwið & rihteð þe heorte. Et hec ÷ caritas quam describit apłs. de corde puro et consciencia bona & fide non ficta.

9 A similar phenomenon has been described as "guidance added" in research on translating allusions (Leppihalme 1997: 82). Translation is crucial in support for CS.

> Þeos riwle is chearite of schir heorte & cleane inwit & treowe bileaue. (f. 1r)
> ‘This rule is always internal and directs the heart. *And this is the charity that the Apostle describes, ‘of a pure heart and a clear conscience and sincere faith’*. This rule is the charity of a pure heart and a clear conscience and true faith.’ (Millett 2009: 1; original italics)

Since *ap[osto]l[u]s* is mentioned as the source within the Latin sequence but not in English, the code-switch here is longer than its support. Similarly to the rest of the preface, example 2 is without metalinguistic labels for languages; as there is no striking language-related flagging of the visual kind either, the intratextual translation is the only identifiable flag. The Present-Day English rendition, however, is monolingual but flags visually both the embedded translation from Latin and the quotation within it – the former with italics, the latter with punctuation – to guide the reader through the multiple levels of the text.

4. Discussion

The previous section contained varied examples of flagging as metadiscoursal assistance to readers encountering a text with multilingual content; these were obviously just a small portion of what the post-Conquest manuscripts contained. We saw how the change of linguistic code was often accompanied by a change in the bibliographic code, made explicit by a language label or supported by an explanation. It is time to address the motivations behind CS on one hand and flagging on the other, and other factors contributing to the variation witnessed in the material.

The global motivation commonly identified by historical CS researchers for switching from English into Latin is the authority carried by the latter; as is well known, it was the international language of the Church and of learning. In contrast, it is harder to pinpoint an equally all-encompassing single motivation for the use of English in Latin texts, although instances of switching in this direction do survive in these manuscripts. Looking at English verses in somewhat later material, Wenzel (1974: 55) has suggested the desire for “memorable and appealing” phrasing as a motivation for their inclusion in the

Latin texts. English switches in the post-Conquest Latin material are words from mortals rather than the word of God; such switches into the language of the original speaker or context may have been considered more authentic or reliable as evidence, or they may have added a local touch to the content provided in Latin (cf. the English response in the Cotton Tiberius example above). Regardless of motivation, some of them have come to preserve snatches of vernacular verse, songs or proverbs which can no longer be accessed elsewhere (Wenzel 1986: e.g. 226). Such "scraps of English" (Laing 1993: 40 *et passim*) are therefore not unimportant.

In addition to verses and quotations, the switches to English may also be names, for example place names in boundary clauses – clauses specifying the geographical boundaries of a plot of land – included in Latin charters (see Schendl 2004). In my material, English boundary clauses appear within Latin in the bottom margins of ff. 328v–329r in Oxford, Bodleian Library, MS Bodley 297. It must have been more expedient to refer to places in the original language rather than with potentially less transparent Latin translations.

All of this suggests that as writers used the two different languages intentionally and for distinguishable – although not necessarily easily confirmable – purposes even in the same text, they were indeed able to tell them apart. That English and Latin appeared together therefore does not imply the emergence of a single, un-separable linguistic system. Towards the end of the Middle English period, the situation may have been different, as Trotter (e.g. 2011) has pointed out, but this applies only to some registers and to the use of the vernaculars, English and French, rather than English and Latin.[10]

As for flagging, there are multiple potential motivations for making multilingual features more noticeable in the text. The reason for reiterating content must have been the desire to ensure that the code-switch is understood. This clarifying function has been identified by many CS researchers (e.g. Gumperz 1982: 78), and it is also a feature of metadiscourse (Hyland 2017: 17). As CS phenomena may often be similar in writing and speech, we can review the findings from medieval manuscripts with the help of the classification of flagging patterns presented in the study of CS in spoken lingua franca English by Hynninen, Pietikäinen, and Vetchinnikova (2017: 103–109). Their term for clarification in its various forms is *explication*, but they also identify another

---

10 The thoroughly mixed code of business writing, involving also Latin, is a case in point (see e.g. Wright 2005).

pattern, *contextualisation*, which refers to mentioning where the switch comes from; this resembles the language labels quoted in 3.2. Sometimes the names of the cited sources or authorities are spelled out as well. Their third type, *hedging using pragmatic markers*, may seem less relevant, but it is explained as features slowing the discourse down at points needing more attention, that is, in the vicinity of CS.[11] In manuscripts, visual flagging may have served two purposes: highlighting the switched sequence – often probably because of its content or the value or origin thereof – on one hand, but possibly also providing a hedge on the other, alerting the reader approaching a change in the linguistic code. In the century or so after the Norman Conquest, when monasteries and other institutions were typically led by men of Norman descent, the presence of English in a Latin text may have been surprising and the language itself unintelligible to at least some of the potential readers, which cannot have been the case with Latin – the lingua franca of the learned – appearing in otherwise vernacular material. Latin must have been more acceptable, and as Hynninen, Pietikäinen, and Vetchinnikova (2017) point out, it is the less acceptable, more problematic linguistic choices that tend to be flagged.[12] However, we do find instances of visual flagging for both English-in-Latin and Latin-in-English, as section 3.1 indicated.

On another level, the use of flagging may be guided by genre conventions or the manuscript's visual programme, topics which unfortunately could not be discussed at length in connection with the examples above. Visually marking a different rhetorical unit, such as a heading or a quotation in Latin, is an example of such conventions; it also helps the reader to navigate on the page (e.g. Rogos-Hebda 2016: 40–41). Furthermore, the appearance of the manuscript may reflect the individual preferences of the producer or his commissioner or patron. It is therefore important to bear in mind that the flagger may not have been the original code-switcher, but a copyist – or a series of them – was involved in the process; it is, thus, also the scribe who "communicates" to the reader (Jucker & Pahta 2011: 3–4). Moreover, as a material feature of the manuscript, visual flagging was subject to the time and

11 The fourth and last type, request for help (Hynninen, Pietikäinen, & Vetchinnikova 2017: 105), is not relevant here.

12 Similarly, Trotter (2011: 365) argues that not flagging Middle English words in later Insular French texts indicates that choosing to use English equivalents of French lexical items was unconscious and unremarkable.

money available for producing a book – the more colourful the book, the more time-consuming and expensive it was to produce – and did not solely reflect attitudes towards either multilingual practices or intertextual contents.

5. Towards a typology

The types of flagging and the visible practices in executing them are summed up in Table 1 below, in an attempt to devise a tentative typology of flagging English-Latin CS.[13] A typology need not provide a frequency-based ranking of the categories or types but may, instead, account for the range of possibilities; a quantitative approach was not possible as word counts for most of the material were not available, and therefore frequency counts could not have been normalised. Quantitative and qualitative variation within the material must nonetheless be acknowledged: in one manuscript, several flags may appear together on the same page, while in another we may find a different array of flags, or none at all. To ensure that Table 1 is not a mere summary of the flags encountered in a handful of primary sources displaying an exceptional wealth of relevant features, the types were collected from across the manuscripts consulted.

Examples of the different types were collected from both the Latin and the English-language manuscripts of the period (cf. quotations and descriptions in Sections 3 and 4). It was rare that a particular type of flag could only be located in an English text but not in a Latin one, or vice versa; further searches in medieval manuscripts from England, perhaps beyond the temporal boundaries applied here, could certainly provide a fuller range of data. Moreover, as I originally did not focus on flagging but on CS as a linguistic phenomenon, visual and verbal flags may have been recorded less consistently for some of the manuscripts accessed early in the project.

Some of the visual and verbal types might be divided further. Of the former, the use of red ink, which is now represented by just two types (writing in red on one hand and other use of red ink on the other, such as underlining), could well be described more explicitly, red ink being a versatile means of highlighting parts of text. Of the verbal flags, more attention could be paid to different ways of referring to the source and/or target language, beyond

13 For a more complex model linking multilingual and multimodal practices, but not verbal flagging, see Kopazcyk (2023: 125).

"other metalinguistic labelling", which now covers labelling the main language (and thus the language of the translation) and both of the languages involved.

Table 1. Types of flagging in the material consulted.

Type of flagging	Type of execution
Visual flagging	Code-switched text in red
	CS marked in red otherwise
	CS marked in main ink
	CS in different script
	CS in different size
	Code-switched unit placed differently
	CS flagged in margin
Verbal flagging	Metalinguistic label for the embedded language
	Other metalinguistic labelling
	Glosses provided
	Support in main language provided
No flagging	(Unmarked, smooth CS)

As noted above, Table 1 does not reflect typicality or frequencies. Nonetheless, it was not possible to avoid making some observations about the preferred forms of flagging. Although the same methods were available for writing and flagging English and Latin, the two languages seem to have been treated somewhat differently. For example, it was more difficult to locate red ink associated with English in a Latin text than vice versa. Other means of flagging were used instead, perhaps most notably metalinguistic labels. Regardless of the main language of the text, English seems to receive a language label more often than Latin, possibly because the latter was the default language of writing and therefore an unmarked choice; in a similar vein, Faulkner (2022: 70) notes that "a linguistic code has a name only if it is an object of discourse". Switching to English in a Latin text may have been regarded as exceptional enough to merit a verbal flag (but not red ink, for example), and as the language skills of every potential reader of Latin texts could not have been expected to include the vernacular, it was useful to provide a translation as well as a metalinguistic label for the embedded English.

Another observation arising from a close examination of the data has to do with what is actually flagged. Visual flags appear with rhetorically distinct parts of the text, such as other-language quotations and headings, whereas verbal flagging is more clearly 'CS flagging'. The metalinguistic cues point to the code-switched units within the text itself, or to their translations, rather than to instances that represent different types or levels of discourse: the red of the quotation, for example, may be primarily a discourse-structuring device and only secondarily a language-marking one, while naming the language(s) points more directly to the presence of two linguistic codes. This underlines the relevance of considering metalinguistic labelling in CS research.

The typology, despite these caveats and concerns, may well be of use in research outside the rather narrow temporal limits defining the material studied here. Linguistic practices and communicative functions are known for their longevity. Recently, this has been discussed with reference to both support for CS in multilingual texts (Nurmi & Skaffari 2021) and, more broadly, to language practices transhistorically enduring changes in society and in technology (Evans & Tagg 2020). Flagging remains important: in 21st-century texts, often digitally produced, distributed and accessed, red ink is no longer relevant, but the function it had in the manuscript era still is.

6. Conclusion

The main findings of this study are that English-Latin CS observable in the long twelfth century is often although not consistently flagged and that flagging takes multiple forms, both visual and verbal, or multimodal and linguistic. This underlines – sometimes literally – the fact that the juxtaposed languages carried separate functions within the same text or manuscript. In its varied forms, flagging facilitated readers' encounters with and understanding of other-language passages.

Multilingual and flagging practices also communicate to us the context in which the texts were produced. Post-Conquest England was multilingual in writing and in speech, and the default language of its texts was Latin. In this de-vernacularised writing culture, snippets of English nonetheless appeared in supposedly monolingual texts composed in Latin, either in the margins or carefully placed within the Latin prose. Some new English texts were also produced, and those writing in English could hardly conceal or suppress the relevance of the prestigious lingua franca of Europe. As all readers or audiences did not have multilingual literacy skills, the supporting translations must have been useful in both Latin and English texts.

The present exploration, I hope, invites researchers working on historical CS to consider the material as well as the linguistic context of each instance of two or more languages interacting on the page and with the reader. The changes in bibliographic code that match switches from one linguistic code to another may serve multiple purposes, which have not yet been examined fully; they may highlight the content, origin or authority of the other-language sequence as well as code-switching itself, or they may potentially serve as a type of hedging. The verbal – metalinguistic and reiterative – flags also merit more attention, since they may suggest what text producers thought about the intelligibility and acceptability of their choices in the eyes of their intended audiences. Overall, it is the recipient's view on code-switching and other multilingual practices that we should also pay attention to in future: marginal notes in another language or manicules added onto the page remind us that the original code-switcher's output did not become interaction until a reader received and reacted to it.

Bibliography

Primary sources

Cambridge, Corpus Christi College, MS 402
Cambridge, Trinity College, MS B.14.52
Cambridge, University Library, MS Hh.I.10
Cambridge, University Library, MS Ii.I.33
Durham, Cathedral Library, MS A.III.12
Durham, Cathedral Library, MS B.I.18
London, British Library, MS Cotton Claudius D vii
London, British Library, MS Cotton Faustina A v
London, British Library, MS Cotton Otho C i vol. 2
London, British Library, MS Cotton Tiberius B xiii
London, British Library, MS Harley 3823
London, British Library, MS Stowe 34
London, Lambeth Palace Library, MS 487
Oxford, Bodleian Library, MS Bodley 297
Oxford, Bodleian Library, MS Bodley 343
Oxford, Bodleian Library, MS Digby 45
Oxford, Bodleian Library, MS Digby 211

Oxford, Bodleian Library, MS Laud Misc. 511

Oxford, Bodleian Library, MS Rawlinson C 317

Secondary sources

Aiello, Matthew 2021. English vernacular script in the thirteenth century (c.1175–c.1325). *New Medieval Literatures 21*, ed. by Wendy Scase, Laura Ashe, Philip Knox & Kellie Robertson. 28–77. Cambridge: Boydell & Brewer. https://doi.org/10.1017/9781800101630.002

Auer, Peter 1999. From codeswitching via language mixing to fused lects: Towards a dynamic typology of bilingual speech. *The International Journal of Bilingualism* 3 (4): 309–332. https://doi.org/10.1177/13670069990030040101

Carroll, Ruth, Matti Peikola, Hanna Salmi, Mari-Liisa Varila, Janne Skaffari, & Risto Hiltunen 2013. Pragmatics on the page: Visual text in late medieval English books. *European Journal of English Studies* 17 (1): 54–71. http://doi.org/10.1080/13825577.2013.755006

Da Rold, Orietta, Takato Kato, Mary Swan, & Elaine Treharne (eds.) 2010–2013. *The Production and Use of English Manuscripts 1060–1220*. University of Leeds. Updated 2018, Stanford University. https://em1060.stanford.edu

Diller, Hans-Jürgen 1997–98. Code-switching in medieval English drama. *Comparative Drama* 31 (4): 506–537. https://doi.org/10.1353/cdr.1997.0016

Evans, Mel & Caroline Tagg 2020. Introducing transhistorical approaches to digital language practices. *Message and Medium. English Language Practices across Old and New Media*, ed. by Caroline Tagg & Mel Evans. 1–12. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110670837-001

Faulkner, Mark 2012 Archaism, belatedness and modernisation: 'Old' English in the twelfth century. *Review of English Studies* 63 (259): 179–203. https://doi.org/10.1093/res/hgr050

Faulkner, Mark 2022. *A New Literary History of the Long Twelfth Century: Language and Literature between Old and Middle English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009031844

Franzen, Christine 1991. *The Tremulous Hand of Worcester: A Study of Old English in the Thirteenth Century*. Oxford: Clarendon Press.

Gardner-Chloros, Penelope 2009. *Code-Switching*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511609787

Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511611834

Hanna, Ralph 2011. Editing 'Middle English lyrics': The case of Candet Nudatum Pectus. *Medium Ævum* LXXX (2): 189–200. https://doi.org/10.2307/43632870

Hyland, Ken 2017. Metadiscourse: What is it and where is it going? *Journal of Pragmatics* 113: 16–29. http://doi.org/10.1016/j.pragma.2017.03.007

Hynninen, Niina, Kaisa S. Pietikäinen, & Svetlana Vetchinnikova 2017. Multilingualism in English as a lingua franca: Flagging as an indicator of perceived acceptability and intelligibility. *Challenging the Myth of Monolingual Corpora*, ed. by Arja Nurmi, Tanja Rütten, & Päivi Pahta. 95–126. Leiden: Brill Rodopi. https://doi.org/10.1163/9789004276697_007

Jucker, Andreas H. & Päivi Pahta 2011. Communicating manuscripts: Authors, scribes, readers, listeners and communicating characters. *Communicating Early English Manuscripts*, ed. by Päivi Pahta & Andreas H. Jucker. 3–10. Cambridge: Cambridge University Press.
Kaislaniemi, Samuli 2017. Code-switching, script-switching and typeface-switching in Early Modern manuscript letters and printed tracts. *Verbal and Visual Communication in Early English Texts*, ed. by Matti Peikola, Aleksi Mäkilähde, Hanna Salmi, Mari-Liisa Varila, & Janne Skaffari. 165–200. Turnhout: Brepols. https://doi.org/10.1484/M.USML-EB.5.114135
Kato, Takako 2010. Cambridge, University Library, Hh.1.10. *The Production and Use of English Manuscripts 1060–1220*, ed. by Orietta Da Rold, Takako Kato, Mary Swan, & Elaine Treharne. University of Leicester. Updated 2018, Stanford University. https://em1060.stanford.edu/manuscript/cambridge-university-library-hh-1-10
Kopaczyk, Joanna 2023. The challenges of bringing together multilingualism and multimodality: Unpacking the Structural Model of Multilingual Practice. *Multilingualism from Manuscript to 3D: Intersections of Modalities from Medieval to Modern Times*, ed. by Matylda Włodarczyk, Jukka Tyrkkö, & Elżbieta Adamczyk. 119–138. London & New York: Routledge. https://doi.org/10.4324/9781003166634-7
Laing, Margaret 1993. *Catalogue of Sources for a Linguistic Atlas of Early Medieval English*. Cambridge: D. S. Brewer.
Leppihalme, Ritva 1997. *Culture Bumps. An Empirical Approach to the Translation of Allusions*. Clevedon: Multilingual Matters. https://doi.org/10.21832/9781800417908
Machan, Tim William 2011. The visual pragmatics of code-switching in Late Middle English Literature. *Code-Switching in Early English*, ed. by Herbert Schendl & Laura Wright. 303–333. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110253368.303
Millett, Bella (transl.) 2009. *Ancrene Wisse – Guide for Anchoresses: A Translation Based on Cambridge, Corpus Christi College, MS 402*. Exeter: University of Exeter Press.
Nurmi, Arja & Janne Skaffari 2021. Managing Latin: Support and intratextual translation as mediation strategies in the history of English. *Text & Talk* 41 (4): 493–513. https://doi.org/10.1515/text-2019-2011
OED = *Oxford English Dictionary* 2023. https://doi.org/10.1093/OED/6351789661
Pahta, Päivi & Saara Nevanlinna 1997. Re-phrasing in early English: The use of expository apposition with an explicit marker from 1350 to 1710. *English in Transition: Corpus-Based Studies in Linguistic Variation and Genre Styles*, ed. by Matti Rissanen, Merja Kytö, & Kirsi Heikkonen. 121–183. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110811148.121
Pahta, Päivi & Arja Nurmi 2006. Code-switching in the Helsinki Corpus: A thousand years of multilingual practices. *Medieval English and its Heritage: Structure, Meaning and Mechanisms of Change*, ed. by Nikolaus Ritt, Herbert Schendl, Christiane Dalton-Puffer, & Dieter Kastovsky. 203–220. Frankfurt: Peter Lang.
Pahta, Päivi, Janne Skaffari, & Laura Wright (eds.) 2018. *Multilingual Practices in Language History: English and Beyond* (Language Contact and Bilingualism 15). Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9781501504945
Parkes, M. B. 2008. Handwriting in English books. *The Cambridge History of the Book in Britain*. Volume II: 1100–1400, ed. by Nigel Morgan & Rodney M. Thomson. 110–135.

Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521782180.010

Poplack, Shana 2004. Code-Switching. *Sociolinguistics: An International Handbook of the Science of Language and Society*, vol. 1, 2nd ed., ed. by Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, & Peter Trudgill. 589–596. Berlin: De Gruyter. https://doi.org/10.1515/9783110141894.1.4.589

Poplack, Shana 2018. *Borrowing: Loanwords in the Speech Community and in the Grammar*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780190256388.001.0001

Rogos-Hebda, Justyna 2016. The visual text: Bibliographic codes as pragmatic markers on a manuscript page. *Studia Anglica Posnaniensia* 51 (3): 37–44. https://doi.org/10.1515/stap-2016-0013

Ruokkeinen, Sirkku, Aino Liira, Mari-Liisa Varila, Otso Norblad, & Matti Peikola 2024. Developing a classification model for graphic devices in early printed books. *Studia Neophilologica* 96: 69–93. https://doi.org/10.1080/00393274.2023.2265985

Schendl, Herbert 2004. Hec sunt prata to wassingwellan: Aspects of code-switching in Old English charters. *Vienna English Working Papers* 13 (2): 52–68.

Schendl, Herbert & Laura Wright (eds.) 2011. *Code-Switching in Early English*. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110253368

Schipor, Delia 2018. *A Study of Multilingualism in the Late Medieval Material of the Hampshire Record Office*. PhD Thesis UiS no. 407. Stavanger: University of Stavanger.

Sebba, Mark 2012. Researching and theorizing multilingual texts. *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, ed. by Mark Sebba, Shahrzad Mahootian, & Carla Jonsson. 1–26. New York & London: Routledge.

Skaffari, Janne 2016. Code-switching and vernacular support: An early Middle English case study. *Multilingua* 35 (2): 203–226. https://doi.org/10.1515/multi-2015-0033

Swan, Mary 2000. Ælfric's Catholic Homilies in the twelfth century. *Rewriting Old English in the Twelfth Century*, ed. by Mary Swan & Elaine M. Treharne. 62–82. Cambridge: Cambridge University Press.

Treharne, Elaine 2000. The production and script of manuscripts containing English religious texts in the first half of the twelfth century. *Rewriting Old English in the Twelfth Century*, ed. by Mary Swan & Elaine M. Treharne. 11–40. Cambridge: Cambridge University Press.

Treharne, Elaine 2012. *Living Through Conquest: The Politics of Early English, 1020–1220*. Oxford: Oxford University Press.

Trotter, David 2011. Intra-textual multilingualism and social/sociolinguistic variation in Anglo-Norman. *Conceptualizing Multilingualism in Medieval England, c. 800–c. 1250* (Studies in Early Middle Ages 27), ed. by Elizabeth M. Tyler. 357–368. Turnhout: Brepols. https://doi.org/10.1484/M.SEM-EB.4.8018

Varila, Mari-Liisa, Hanna Salmi, Aleksi Mäkilähde, Janne Skaffari, & Matti Peikola 2017. Disciplinary decoding: Towards understanding the language of visual and material features. *Verbal and Visual Communication in Early English Texts*, ed. by Matti Peikola, Aleksi Mäkilähde, Hanna Salmi, Mari-Liisa Varila, & Janne Skaffari. 1–20. Turnhout: Brepols. https://doi.org/10.1484/M.USML-EB.5.114128

Voigts, Linda Ehrsam 1996. What's the word? Bilingualism in late-medieval England. *Speculum* 71: 813–826. https://doi.org/10.2307/2865721

Wakelin, Daniel 2014 Framing, meaning and the matter of the page. Paper read at Linguistic Meets Book History: A Pragmatics on the Page symposium. University of Turku.
Wenzel, Siegfried 1974. Unrecorded Middle-English verses. *Anglia* 92: 55–78. https://doi.org/10.1515/angl.1974.1974.92.55
Wenzel, Siegfried 1986. *Preachers, Poets and the Early English Lyric*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400854141
Wright, Laura 2005. Medieval mixed-language business texts and the rise of Standard English. *Opening Windows on Texts and Discourses of the Past*, ed. by Janne Skaffari, Matti Peikola, Ruth Carroll, Risto Hiltunen, & Brita Wårvik. 381–399. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.134.29wri

“I rede þou lerne wel þis of me”
Advice giving in Middle English medical discourse

Irma Taavitsainen
ORCID: 0000-0001-9032-5930
Andreas H. Jucker
ORCID: 0000-0003-3495-2213

Advice giving has been one of the main concerns of instructional medical discourse in the history of English. These texts tell their readers how they should behave in order to preserve or regain their health. In this contribution, we focus on the Middle English period up to 1500 with a brief look at some earlier manifestations in Old English. Our data comes from the digital corpus of *Middle English Medical Texts* (MEMT). We approach the speech act of advice giving through a careful analysis of relevant meta-illocutionary expressions, i.e. terms that are used to either perform a speech act (performative uses) or to talk about them (narrative uses). In MEMT, only two lexical items are attested, *rede*, which goes back to Old English and occurs mainly in the older text tradition of remedy books, and *counsel*, which is more often attested in the specialized texts and in the surgical texts. The examples also show that advice giving in MEMT is not restricted to issues of preserving and regaining health but extends to courteous behavior in the interaction between the medical professionals and their patients.

https://doi.org/10.51814/ufy.1041.1456

Keywords: speech acts, advice, Middle English Medical Texts (MEMT), medical discourse, meta-illocutionary expressions

1. Introduction

Most medical texts from the early periods are instructive. They give readers advice on how to live in order to preserve their health or to regain it once they have lost it, and they also instruct how to prepare medicines. In a present-day context, advice giving has received a considerable amount of scholarly attention both in medical and in non-medical contexts, but studies on advice giving in the early periods are still lacking (but see Schrott 2014 for an investigation of advice giving in Old Spanish texts). This article sets out to fill this gap. It explores forms and functions of advice giving in medical discourse with a focus on Old English (OE) and Middle English (ME) texts, relating strategies of the past to modern practices as described in the recent literature. Different vernacular realizations of advice giving are analyzed in their socio-historical context with different types of readers and across time.

Advice giving is seen as part of a wider pragmatic space (Jucker & Taavitsainen 2000). We see speech acts as fuzzy entities and pragmatic variables (cf. Terkourafi 2011; Jucker & Taavitsainen 2012). Advice giving overlaps with such neighboring speech acts as directives, recommendations and warnings. They are all concerned with something that the addressee might or should do at some point in the future, but they differ in the extent to which the speaker exerts pressure on the addressee. Advice giving can also be conceptualized as being part of the pragmatic space of sharing and helping (Põldvere, De Felice, & Paradis 2022: 1). Here, advising is seen as a special case of informing, which involves the offering of help for the addressee and has a fit from me-to-you, i.e., from the speaker to the addressee who is supposed to benefit from the advice.

We deal with the special language of medicine of the early periods and our texts have not been assessed from the present angle before. Our data gathering is performed partly digitally with corpus linguistic methods and partly by qualitative reading.[1] We build on both electronic corpus searches of

1 Cf. the data collecting methods of modern studies, which are very different with questionnaires, interviews and Conversation Analytical methods.

lexical items and apply qualitative analyses with contextual assessments to detect alternative expressions and devices to enhance the advice giving. The main sources were the relevant dictionaries and the digital corpus of *Middle English Medical Texts* (MEMT 1375–1500), which covers the early vernacularization period.

We shall first explain our theoretical approach with meta-illocutionary uses of speech act verbs and corresponding nouns in focus, then go through the relevant lexical items and describe our data. In the empirical part, we proceed from OE to ME keeping our focus on meta-illocutionary uses of speech act verbs and nouns but outlining other detected functions as well. At the end, we summarize our observations about the variation and diachronic development of pertinent lexical items, but we also discuss an important change in interpersonal language use in advice giving that emerged with our surgical examples.

2. Advice giving and the meta-illocutionary lexicon

Giving advice, either solicited or unsolicited, appears to be a very common communicative need. People regularly ask each other and receive responses in the form of recommendations about what they should – or should not – do, and how they should – or should not – behave. It is not surprising, therefore, that advice giving has been researched from many different perspectives and in a large range of different contexts (see for instance the contributions in Limberg & Locher 2012 or in MacGeorge & Van Swol 2018, and for a recent overview see Põldvere, De Felice, & Paradis 2022). Advice giving belongs to Searle's (1976) large class of directives because it is concerned with things that are to be done by the addressee at some point in the future, but they differ from requests, commands and other more prototypical cases of directives in that the speaker may be indifferent as to whether the addressee performs the suggested act or not, but assumes that it might be in the best interest of the addressee to do so. There is a rich literature on advice giving in conversational settings. Adolphs (2008), for instance, investigates a range of speech acts, including suggestions, in the *Cambridge and Nottingham Corpus of Discourse in English* (CANCODE), and Põldvere, De Felice, and Paradis (2022) use the *London-Lund Corpus of spoken British English* to investigate advice giving in conversational data.

In medical settings, advice giving appears to be particularly prominent. People are concerned how to preserve or regain their health and for this they turn to other people whom they consider to be more knowledgeable in this respect. In the medical setting, they get such advice from medical professionals (see Locher 2006; Locher & Turnherr 2017; D'Angelo & D'Angelo 2018; Stivers *et al*. 2018). There is a clear distinction between an advice giver and an advice seeker or advice recipient with an assumed epistemic difference between the roles (DeCapua & Huber 1995). The advice giver is in possession of some knowledge or information that they believe would benefit the advice seeker or recipient who may not possess the same knowledge or information. In many settings, the distribution of roles is pre-assigned because the setting specifies who is supposed to be more knowledgeable and who is supposed to be in need of advice or instruction. Zhang and Hyland (2021), for instance, investigated advice giving in the context of theses supervisions where the roles between advice giver and advice seeker are clearly distributed. In medical settings, too, it is generally clear who is in the role of advice giver and who is in the role of advice seeker or advice recipient. The participants do not have to negotiate these roles. This stands in contrast to everyday conversations between equals, where such negotiations are common (see Adolphs 2008: 91).

The nature of our data makes it difficult to automatically retrieve instances of advice seeking and advice giving. Our method, therefore, is both eclectic and exploratory. In general, there are two ways of investigating speech acts in historical speech act analysis (see Jucker 2024a; 2024b: Chapter 3). The researcher can search for manifestations of a speech act by searching for elements that are known to occur in such manifestations. This knowledge may have been provided by earlier research or by manually searching carefully chosen sample texts. For our purposes, this method appears to be of limited value. Preliminary searches and our familiarity with the data suggest that advice giving was not sufficiently conventionalized for recurrent patterns to emerge. This is also true for other speech acts, such as apologies or requests. Recent research has shown that historically, their conventionalization and the development of typical expressions marking their illocutionary force, such as *sorry* and *please*, are relatively recent (see, for instance, Tieken-Boon van Ostade & Faya Cerqueiro 2007; Jucker 2018).

The second method consists of an analysis of meta-illocutionary expressions denoting a particular speech act. The terms "meta-illocutionary expression" and "meta-illocutionary lexicon" (i.e., the collection of all meta-illocutionary expressions) appear to have been introduced by Schneider (2017; 2021; 2022; see also Schoppa 2022). It is a more precise term than "metapragmatic expressions" (Jucker & Taavitsainen 2013: Chapter 6; Jucker & Taavitsainen 2014) because it focuses more specifically on expressions that are used to talk about speech acts. The meta-illocutionary lexicon provides a first-order perspective on what speakers or writers of a language find to be salient pragmatic acts. It comprises names for those pragmatic acts that they need to talk about. This is central for our study because we are going to explore the pragmatic act of advice giving mainly through the lens of the meta-illocutionary lexicon.

Schneider (2017; 2021; 2022) and Schoppa (2022) provide several categorizations of the functions of meta-illocutionary expressions. In everyday discourse, they can have a performative, reporting, problematizing or challenging function (Schneider 2017: 230). Schoppa (2022: 71) distinguishes six functions: performative, problematizing, reporting, clarifying, naming and commenting. The distinction between these types is not always easy, but in our data the focus is clearly on the performative function. Meta-illocutionary expressions are regularly used to perform the named speech act. It seems that in medieval medical texts the selection of additional functions is somewhat different from those mentioned in the literature for present-day contexts, but the main line is in accordance with previous research. It has been shown that in OE and ME speech acts, such as requests for instance, were regularly performed through explicit performatives (see, for instance, Bergner 1992; Kohnen 2000).

3. Early English medical texts

The first vernacular English medical texts were remedy books with recipes and prognostications that survive from the eleventh century. After the late OE period there is a gap of about two centuries before medical writings emerged in a larger scale with first vernacular translations, adaptations, and some more original compositions in the late fourteenth and fifteenth centuries.[2] ME medical texts can be grouped into three sets according to their underlying traditions (see Voigts 1984). Several surgical texts were rendered into English at this point. Some of them derive from university settings and belong to top medieval science with theoretical passages. Mostly they are, however, instructive and provide some very interesting examples of advice giving in the learned style and abound in features typical of the scholastic thought style (Taavitsainen & Pahta 1995; Taavitsainen 2002; Taavitsainen & Pahta 2004; Taavitsainen & Schneider 2018).

Another group consists of specialized texts in early translations of e.g., ancient authorities' learned texts, but texts pertaining to natural philosophy, reproduction and theoretical subjects do not contain advice giving as they are mostly expository and argumentative.[3] Some of the specialized texts are, however, more practical and some come close to the third category of remedy books. Middle English versions of Latin texts show difficulties in rendering abstract ideas in a language whose lexical or syntactic resources were not developed for expressing scientific ideas (see Pahta & Carrillo Linares 2006: 116). Alongside with the surgical and specialized texts, the third group consists of a large body of more practical works. This category provides direct continuation from the Old English leech books with traits of learned medicine. Some elements from classical sources are added and mixed with oral folk traditions such as *materia medica* with herbal lore, charms, and prognostications (see Figure 1). In addition, practical verses in rhymed couplets also belong to this group, forming the most popular layer of medieval medical texts. They were easy to memorize and repeat on suitable occasions.

2 There was a pan-European vernacularization boom as translations of Latin were made into several languages at about the same time, and knowledge of Arab learning spread to Europe (Siraisi 1990).

3 Few texts of the highest level were known in the last millennium. Pahta's edition of *De spermate* (1998) made a rare academic text on reproduction available to scholars. The identification of some important texts is fairly recent (for the Hippocratic commentary, see Tavormina 2006). Galen's texts have a complicated history (for details and editions, see Pahta *et al.* 2016).

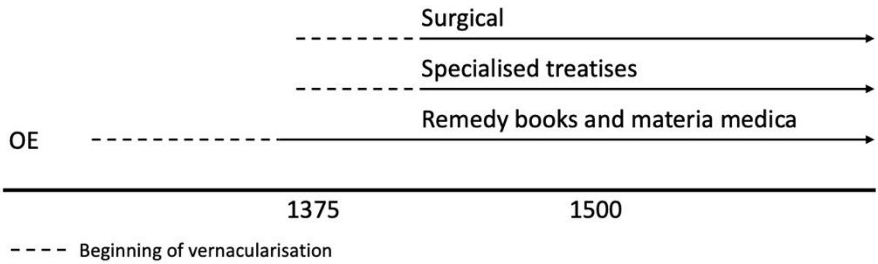


Figure 1. Chronology of three medical text traditions (modified from Taavitsainen, Pahta, & Mäkinen 2005).

Chronologically our data begins with some Old English examples. To detect manifestations of advice giving we consulted all available Old English scientific and medical sources and checked the Old English part of the *Helsinki Corpus*. We did not find very much, but a leech book dating from the eleventh century provides some examples pertinent to our research task. Next, we assessed the late medieval period with digital searches of the MEMT corpus. It covers the early vernacularization period 1375-1500 and contains all editions of medieval texts of the period known to us at the time the corpus was submitted for publication (Taavitsainen, Pahta, & Mäkinen 2005). Texts range from learned scholastic treatises to popular verse. Our searches yielded a fruitful harvest of meta-illocutionary uses of speech act verbs and nouns, and they also revealed some other interesting alternative ways of giving advice and devices of enhancing the advice.

The readers or listeners of early medical advice included a broad range of people. In addition to learned doctors and other practitioners of the medical profession like surgeons and barber-surgeons, readers included lay people, such as literate women and household members. Useful texts like health guides and regimen texts were often read aloud thus broadening the target audience even to the illiterate. The authors were a more homogeneous group, as it was customary for learned physicians, surgeons, and other medical authors to write for both specialists and lay people.

4. Meta-illocutionary expressions of advice giving in Middle English

In our search for a relevant set of meta-illocutionary expressions of advice giving in the early periods, we relied on the *Historical Thesaurus of the Oxford English Dictionary*, which provided a list of nouns and verbs (see below; cf. also Milfull (2004), who used the same source for lexical items on advice giving in Middle Scots). In addition, we relied on our own familiarity with vernacular medical texts and the recurrent patterns used in them, such as imperative forms of verbs, deontic collocations, and other modal expressions with a directive force. For instance, instruction in remedy books including recipes employ direct commands with imperative forms of verbs. They can be regarded as pertinent to our study as well, and although not at the heart of the present task, they are important in marking the neighboring speech act of giving directions. Furthermore, some patterns of strengthening the illocutionary force of the speech act also emerged.

According to the *Historical Thesaurus* of the *Oxford English Dictionary*, there are quite a number of nouns and verbs in the semantic field of advice giving. Table 1 provides an overview of all the items that go back to ME (i.e. first attested before 1500).

Table 1. Lexical items of advice giving c. 1100–1500.

Lemma	Date	Meaning
Nouns		
rede	Old English–	Counsel or advice given by one person to another. Also: †a piece of advice (obsolete).
rathe	c1175–1300	Counsel, advice. Also: a counsellor, a friend. Cf. rede, n.
counsel	?c1225–	Opinion as to what ought to be done given as the result of consultation; aid or instruction for directing the judgement; advice, direction.
governail	a1382–1500	Guidance, advice, counsel. Also: contrivance.
advice	c1390–	Opinion given or offered as to what action to take; counsel; recommendation.

advisement	1409–	Advice or guidance as to what action to take; counsel; (also) a piece of advice. Cf. advice, n. 2.
vising	c1480	Advice, counsel.
Transitive verbs		
rede	Old English-1650	To suggest (a course of action) by way of advice; to advise (a thing), to give (counsel). transitive. With simple object. Frequently with cognate.
rothe	c1175-1400	transitive. To counsel, advise (a person). Also reflexive and intransitive.
berede	a1225-1350	transitive. To advise, inform, counsel; to plan.
counsel	1297-	To give or offer counsel or advice to (a person); to advise.
inform	c1350	transitive. To give instruction to (a person, the mind, etc.); to educate, teach, train; (later more generally) to impart knowledge or learning to.
Intransitive verbs		
rede	Old English-1591	intransitive. To give advice or instruction. Obsolete.
rothe	c1175-1400	transitive. To counsel, advise (a person). Also reflexive and intransitive.
counsel	?c1430-1842	intransitive. To give or offer counsel or advice. With to: to give advice to a particular effect or purpose. Obsolete.
to give to rede	c1460-1540	to give to rede: to state by way of advice, to advise. Obsolete.
advise	1481-	intransitive. To give advice or counsel.

The OE repertoire is narrow with one item only, the precursor of *rede.* Its attestations in the early period are few, but increase in ME. The predominating item in MEMT is, however, *counsel*. We will focus on its meta-illocutionary use and different attestations, but we found examples of other uses as well that have not been mentioned in the literature before. According to OED *advise* came to the language in the fourteenth century, but it was not recorded in our data, which is limited to medical and scientific texts only.

In the earliest period,[4] we found both the masculine noun *rǣd* (also *a-rǣd*) ‘counsel or advice; what is advisable, benefit, advantage’ and the

4 We are grateful to Leena Kahlas-Tarkka for checking the relevant item (*rede*) in the OE Corpus, the OE part of the *Helsinki Corpus* and the relevant dictionaries, as well as for providing the examples with translations.

corresponding verb *rædan* with the meaning ‘to counsel, give advice; to ask advice; to deliberate; to suggest (a course of action)’ (s.v. Bosworth-Toller). Two originally distinct verbs seem to have coalesced in this verb form, and therefore a wide variety of meanings, including ‘to read’ have been recorded. First-hand evidence of the meaning ‘counsel’ is given by Ælfric in his *Glossary* (G1 B1.9.2 1) *consilium ræd*.

The wealth of religious writings, chronicles and laws in OE yield several instances of *ræd*, whereas the limited amount of scientific and medical writing does not give much evidence for the use of this word. The only relevant example for the medical or scientific register can be found in the *Leechbook*: “**Ræd** bið gif he nimð mealwan mid hire ciþum, seoþe on wætere, sele drincan.” ‘It is advisable to take mallow with its seeds, to boil in water and give to drink.’ (LcH II (2) B.21.2.1.2.2.)

Another pertinent example from the Old English period which also advises and exhorts to wise behavior has been recorded: “Ne acse þu nanre wicce **rædes**, se sech þu riht æt deaden: soþlice god ascuneð swylce þing.” ‘Do not ask for advice from a witch/magician but look for justice from those who are dead: truly God detests such things.’ (*Proverbs* 1(Cox) B7.1 Distics of Cato.)

5. Manifestations of advice giving in MEMT

In the MEMT corpus (1375-1500), the form *rede* is attested 378 times, but the great majority of instances denote either the verb ‘read’ or the color ‘red’.[5] Few examples in the meaning of ‘giving advice’ are found in a specialized text and the remedy-book tradition had some, but in general it is not very common. *Counsel* (with spelling variants) prevails with a total of 42 hits and occurs particularly frequently in surgical texts, but all categories showed some examples (see Table 2).

5 E.g. all 36 hits in *De spermate* at the learned end of the data referred to the color.

Table 2. Spread of *rede* and *counsel* (including spelling variants) as meta-illocutionary expressions in MEMT in performative and narrative functions.

Source file	***rede***		***counsel***		**Total**
	Perf.	**Narr.**	**Perf.**	**Narr.**	
Surgical texts					
arderne_fistula.rtf				6	6
chauliac_ulcers.rtf			1	5	6
chauliac_wounds.rtf				4	4
chirurgie_de_1392.rtf			1	1	2
lanfranc_chirurgia_magna_1.rtf				1	1
lanfranc_chirurgia_magna_2.rtf				1	1
mondeville_chirurgie.rtf				1	1
Specialized texts					
benvenutus_grassus.rtf	1		4		5
caxton_ars_moriendi.rtf				2	2
daniel_liber_uricrisiarum_2.rtf		1		1	2
torrella_tretece_of_the_pokkis.rtf				1	1
Remedy books					
a_tretys_of_diverse_herbis.rtf	1				1
astrological_compendium.rtf	2				2
bloodletting.rtf	2				2
caxton_gouernayle_of_helthe.rtf	1			1	2
de_caritate_the_priuyte_of_priuyteis.rtf				3	3
leechbook_1.rtf				1	1
practical_verse.rtf	2		1		3
quinte_essence.rtf				1	1
regimen_sanitatis.rtf				2	2
rupescissa_remedies.rtf		1			1
secreet_of_secreetes.rtf			1	2	3
sidrak_and_bokkus.rtf	1	1			2
when_the_mone_is_in_aries.rtf				1	1
Total	**10**	**3**	**8**	**34**	**55**

Table 2 shows an interesting difference between the two meta-illocutionary expressions. *Rede* is used more often in its performative function while *counsel* is used more often in its narrative function. The number of occurrences is small, but a chi-square test shows that the difference is significant at the .01 level (χ^2=15.1032, df=2). This may be meaningful in the sense that *rede* is the older form going back to Old English when the performative use of speech act verbs appears to have been more common (cf. Bergner 1992; Kohnen 2000). *Counsel*, on the other hand, is more frequent in the more recent text traditions of specialized texts and surgical texts. But at this point, this can be no more than speculative. We now turn to manifestations of these meta-illocutionary expressions in the three text traditions.

5.1. Surgical texts

Chirurgie de 1392 (with an alternative name *Compilacioun of sirurgie*) is not a direct translation but a more original compilation from various sources and personal observations by a London surgeon, possibly also physician, from the year 1392. Example (1) comes from a preamble to the third part of the text. It records an emphatic plea to the readers of the text in a word pair where the first component gives special coloring to the second (see Pahta & Nevanlinna 1997).

1 In þe þridde partie I haue putt an antidotarie of sirurgie wiþ watris, emplastris, poudris, oilis & oynementis, & oþere dyuers medicyns in general; wherfor **I praie & counseile ȝou** þat vsen þe worchinge in þis doctr=i=ne, contynue þa=t= ȝe ben gr=a=cious & helpinge to þe pore for goddis sake & to þe riche for a competent salarie. (chirurgie_de_1392.rtf)

The most famous medieval surgeon, Guy de Chauliac (c. 1300–1368), taught at the University of Montpellier and wrote in Latin. Several vernacular versions appeared in various languages, e.g., both full and incomplete versions of his *Chirurgia magna* and other works circulated widely in Middle English. The first scholarly editions date from the nineteenth century and the most recent came out some decades ago (see MEMT bibliography). Differences between them are found both in syntax and vocabulary. These texts contain several passages in the logocentric scholastic style e.g., in meta-illocutionary performative uses of the verb with the first-person pronoun, as in (2). In (3) the past tense verb occurs in the third person recording his past action; this is a clear case of narrative use (see Werlich 1982).

2 ffor þat in drawing out it is oft-tymeȝ ybroken, ȝe, bot of þis **I consaile þat** þe tenteȝ be bounden with a threde þat if þai cleued-to And were profounded to mych, þai may be
(chauliac_ulcers.rtf)

3 And if it war nede, be þer a litel y-added of opio or of mandrace, **as counsaileþ william**
(chauliac_wounds.rtf)

In (4) below the noun *counsaile* is used descriptively. The scholastic style favored impersonal expressions and the passive voice; sometimes the third person refers to the author and provides a distancing device (see Taavitsainen & Pahta 1998). In this early period the language of writing science had not conventionalized yet and there is a great deal of fluctuation.

4 if þe bone be corrupte, it is þe **counsaile** of hym for to kut þe flesh & for to vncouere þat bone as mych as …
(chauliac_wounds.rtf)

The same text shows another passage in (5) with similar scholastic features. The reference to G is likely to be an abbreviation for Galen, the number one authority of medieval medicine, and the accuracy of the reference indicates a learned provenance; Avicen and Albucasis were also frequently cited in scholastic writings.

5 Wondeȝ þat nedeþ sewyng aboue, which ar made bi al þe transuers of principale musculeȝ & þat smyteþ þe grete veyneȝ & arterieȝ & nerues & medulle, bringeþ in most perile. **ffirst counsaileþ G in primo 6.=i= amphorismorum**, þe vesic kutte or þe brayne or þe herte or diafragma or any of þe subtile entraleȝ or þe lyuer or þe wombe, it is mortale after þat of 5=ti= amphorismorum, Quibuscumque ydemata &c:. fforsoþ, if þe bone be corrupte, **it is þe**

counsaile of hym for to kut þe flesh & for to vncouere þat bone as mych as shale be possible, & þat with rasours & corrosyueȝ, as it shal be said with-inforþ of an vlcerate legge.
If, forsoþ, it be not possible to þe þat þu make or do inscisioun**,** **Auicen counsaileþ** wasshyngeȝ wiþ clisterieȝ **after þe maner of Albucasis**;
(chauliac_wounds.rtf)

John Arderne is one of the best-known figures of medieval medicine in England and, as was customary at the time, he wrote in Latin. He practised surgery first in Newark, Notts. from 1349 to 1370, and then in London. His works circulated in at least four different vernacular translations after 1400.

In the surgical tradition, *counsel* often occurs in the form *to ask counsel*, as in (6), (7) and (8) taken from *Fistula in ano.* All examples below are narrative and record past actions in the third person, and many of them come from case reports where the author boasts about his own surgical skills that surpassed all others'.

6 by The forsaid sir Adam, forsoth, suffrand fistulam in ano, made **for to aske counsel** at all the lecheȝ and cirurgienȝ that he myȝt fynde in Gascone, at Burdeux, at Briggerac, Tolows, and Neyrbon, (arderne_fistula.rtf)

7 Aftirward I cured sir Reynald Grey, lord of Wilton in Waleȝ and lord of Schirlond biside Chesterfelde, whiche **asked counsel** at the most famose leches of yngland, and none availed hym. Aftirward I cured sir Henry Blakborne, clerk, Tresorer . . .
(arderne_fistula.rtf)

8 in the ȝere of oure lord 1370, I come to london, and ther I cured Iohn Colyn , Mair of Northampton, that **asked counsel** at many lecheȝ. Aftirward I helid or cured Hew Denny, ffisshmanger of london, in Briggestrete; and William Polle (arderne_fistula.rtf)

5.2. Specialized texts

Example (9) comes from Benvenutus Grassus' text explaining the procedure of an eye operation. The meta-illocutionary expression explicitly flags the sentence as a piece of advice.

> **9** and þere sotylly I kut without any trouble of the tonycle. And þan **hade I rede** ij nedyls myghtyly knytt togeder …
> (benvenutus_grassus.rtf)

Benvenutus Grassus was probably an Italian thirteenth-century ophthalmologist. The text deals with eye diseases and injuries and circulated in various European vernaculars. The ME version represents an independent translation of the Latin original with added material from different sources and gives evidence of difficulties in rendering learned texts into the vernacular. On the basis of phrases like "Moreoure, quod Beneuucius, I wyl þat ȝe practysers" the text seems to be a paraphrase or an adaptation of the Latin original rather than a translation.

The phrase *I counsel you* occurs four times in this text and conveys a didactic tone in the style of writing scholastic science. The imperative form is also present in advice giving for playing safe with cures in order to keep a good name, as in (10).

> **10** **I counsel yow** that when ȝe se thys maner of pannycle not new but incarnate and harded vpon the tonycle of the eye, **take yt not** in cure, for ȝe may haue no worschype þerof but hurtyng your name and fame among the people.
> (benvenutus_grassus.rtf)

Another case in specialized treatises occurs in the vernacular translation of Henry Daniel's *Liber uricrisiarum* that dates from 1379 and achieved a wide circulation. It is based on Isaac Judaeus' *De urinis* and deals with diagnostics by urine and uroscopy. The mid-fifteenth century manuscript that was used as base for the edition belonged to a barber-surgeon of London, giving a reliable indication of its readers. In (11), the phrase *we taken counsell*, shows an authoritative use of *we*. The repetitive word pair is used strategically to interpret the word of foreign origin (*counsel*) with the native equivalent (*red*).

11 For when we will witten and knowen the state and the disposicioun of manys body wythinnin, and namely of þe reyns, **we taken counsel and red** at vryn. Also, vryn is seyde of this Latyn worde vrere, that is to seyn "brenyng, as fyre or hote thing dothe," for hit is be wey of kende hote and drye. Be reson of his dryehede he is desictatif, þat is forto sey, dryend;
(daniel_liber_uricrisiarum_2.rtf)

The third case in this category is found in Torrella's *Tretece of the pokkis*, a specialized text on syphilis from the very last years of the fifteenth century. In (12), the authority is ascribed to the centre of learned scholastic medicine in Paris and the noun *counsel* is used in a narrative clause.

12 The tretece of the pokkis: And the cure by **the nobull counsell of parris** A syrop mervelus & expert with weche inumerable hathe been cured and allso preservid from all maner of the skyn & of dolors of the pannicles lacertts & nervis.
(torrella_tretece_of_the_pokkis.rtf)

5.3. Remedy books

In this category, the *secreta secretum* tradition circulated widely in different versions (see the MEMT catalogue) and contains several instances of the verb *counsel*. Allegedly these texts contained Aristotle's teaching to Alexander, or as the preface expressed its provenence states "This booke …was sent fro þe grete philosophir Aristotill to the nobill prince King Alexander." Example (13) begins eloquently with direct address preceded by the vocative interjection *O!*

13 Qwerfor, o Alysaundyr, if þu vse this thow schalt vse none odyr, for þis suffysith. Nor þu schalt neuir blede nor be boystid but be þe **consel** of a man þat is experte in astronomye, and for þis cause þe profytabylnes of medycinal kunnyng is exaltyd in þat, and þe natural wylle is enclyned and claryfiyd in þat.
(de_caritate_the_priuyte_of_priuyteis.rtf)

A combination of a directive command in the imperative form and an *if*-clause of volition with a second explicit exhortation "obserue my counsel" is found in the same text, see (14).

14 Knowe þan þat þer be of metys sondry … and confortid to dygeste þe alymentys. **Gouerne** wele þan þi body if thow wult þat it be in hele, and **obserue my counsel**, for þat is to þe chef solas. (de_caritate_the_priuyte_of_priuyteis.rtf)

In (15), *counseile* is used in a performative function in advice giving for the well-being of both the body and the mind.

15 and this shall take the fleume from thy mouthe and stomac, and yiueth hete to the body, and driueth awey the wynde, and yiueth good sauoure**. Moreouer I counseile the** that thou comon often with the noble and wise men of thy reame, of alle suche maters as thou hast a-doo, and gouerne thaim graciously aftir the custumes. (secreet_of_secreetes.rtf)

In comparison, the other types of texts within the remedybook tradition favor the native *rede*. The simplest examples are found in practical verse. In (16), advice for the common cure of bloodletting carries an authoritative tone with the first person singular. The rhyme words *blede* / *I rede* served as a memory aid and *I rede* had the function of a standard line filler as well.

16 And fro scall withouten lesynge.
ij at the temples ther most **blede**
ffor Stoppynge and Akynge, **I Rede;**
(practical_verse.rtf)

Other instances can be found in herbal lore and in an astrological compendium, in (17) and (18).

17 Wha~ne it is yus to gedere browt **I rede** ye gaddererys for ȝete hem nowth, Y=e= tyme y=t= he on hy~ bere Fro alle perylys it wyl hy~

were
(a_tretys_of_diverse_herbis.rtf)

18 xxviij is noght gude, ne at the xxx; & er thu bled in tha that folus, **I rede** the a-bide: The iij day & the iiij, & the xviij als
(astrological_compendium.rtf)

A somewhat different instance of *rede* is found in the narrative function in a collection of recipes, in (19).

19 yit sum men mowe **rede** and inquere of þe heuenely cunnyng and bigynnyng of quynte essence þat I haue y-shewid here bifore.
(rupescissa_remedies.rtf)

Frequent examples of *rede* occur in a popular encyclopedia *Sidrak and Bokkus* (see example 20), written in a question-and-answer format in rhyming couplets as a dialogue between a Christian philosopher Sidrak and a heathen king Bokkus. Their discussion can be characterized as a brief digest of medieval learning in science, interspersed with theology and moral teaching.[6]

20 Þerfore, I seie, amys þei do Þat etith more þanne mister is to, Ouþer drinke, wheþer it be: **I rede þou lerne wel þis of me.**
(sidrak_and_bokkus.rtf)

Counsel is not common in the leech book category, but an emphatic example of the meta-illocutionary use occurs in an anonymous recipe collection with dietary advice, in (21). It employs a repetitive word pair with the verbs *admonish* and *counsel* that work into the same direction.

21 Jn all wise **J monyche and counsell** the that thow sett not lyttell by þis confection ffor I prevyd it full notable in many cawsis . . .
(leechbook_1.rtf)

Advice giving with *counsel* occurs in a lunary text, too. Astrology was an esteemed branch of science in the Middle Ages with the aim of indicating

6 We selected its emphatic phrase of enhancing advice as our motto for this article.

appropriate times for various actions in a broad range of texts from learned treatises to popular moonbooks (see Taavitsainen 1988). A particularly intriguing case of advice giving according to the moon is found in the *Guild-book of the Barber Surgeons of York*, a professional institutional book from 1486. It shows the fuzzy border of professional and popular; astrology was observed in medical treatments in all levels. In this lunary text, appropriate times are pointed out simply by *yt ys gude* and the opposite by *yt ys yll and perilous*, and the prognostications are given for all kinds of actions in a playful tone. Medical rules are concise and based on the *homo signorum* doctrine, e.g. in Aries it is forbidden "to do ought tyll a manys hede" and Lion gives comprehensive advice for the Canicular days (see Taavitsainen 1994). In (22), the advice for Aquarius achieves a jocular tone, which is rare in medieval medical texts.

22 Whene the Mone ys in the Watirwarde, that ys callyde Aquarius, **yt ys gude** to wake one the watters and to deyll wyth them, and to speke wyth frendis and to aske them of helpe and of **counsell**, and for to wede wyues, and to enture into religione, and for to founde houses, and for to by aritages, and for to enture into lordschyppes and newe possessions
(when_the_mone_is_in_aries.rtf)

6. Advice giving and courtesy

Several of the examples quoted in the previous section show the authors' concern for the interpersonal aspect of their texts. They give advice on how to preserve or regain one's health, but they also give professional advice on appropriate behavior in medical situations. In his surgical text, Lanfrank, for instance, advocates that advice should only be given if asked for, women should not be talked to foolishly, nor should the patient or his servants be criticized. Instead, the patient should be addressed courteously (example 23).

23 be he trewe, vnbeliche, & plesyngliche bere he him-silf to hise pacientis; speke he noon ribawdrie in þe sike mannis hous. Ȝeue he no **counseil**, buts he be axid; ne speke he wiþ no womman in folie in þe sik mannes hous; ne chide not wiþ þe sike man ne wiþ noon of

hise meyne, but **curteisli** speke to þe sijk man, and in almaner sijknes bihote him heele.
(lanfranc_chirurgia_magna_1.rtf)

Mondeville and John Arderne provide similar advice on good behavior in (24) and (25).

24 þe pacientis owen to obeien in alle maner þingis þa=t= ben parteynynge to þe cure of here syknesse ouþir maladie, & þei schal not wiþstonde **þe counseil** of her surgian in no wise /
(mondeville_chirurgie.rtf)

25 And it is seid in anoþer place, "Shrewed speche corrumpith gode maners." When seke men, forsoth, or any of tham bysyde comeþ to the leche to aske help or **counsel** of hym, be he noȝt to tham ouer felle ne ouer homely, but mene in beryng aftir the askyngis of the personeȝ; to som reuerently, to som comonly. ffor after wise men, Ouer moche homelynes bredeþ dispisyng.
(arderne_fistula.rtf)

In fact, courteous behavior is a recurrent topic in MEMT as (26) and (27) illustrate.

26 To euery tale, sonn, gyff not credence;
Be not hasty nor sodenly vengeabill,
To poure folke do no violence,
Curteys of langage; of fedyng mesurabyll,
On sondry metis not gredy at þe tabill,
In feding gentill; prudent in daliaunce;
To sey þe best sette al-wey þy plesaunce.
(practical_verse.rtf)

27 Haue the leche also clene handes and wele shapen naileȝ & clensed fro all blaknes and filthe. And be he **curtaise** at lordeȝ bordeȝ, and displese he noȝt in wordes or dedes to the gestes syttyng by;
(arderne_fistula.rtf)

These examples provide a clear link to the conduct book tradition starting in the Renaissance period, mostly in response to Italian courtesy and conduct books, which were widely translated and read in England. Paternoster (2022: Chapter 2), in her study of etiquette books in nineteenth-century Western cultures, provides an overview of the early beginnings of advice literature on good manners with a clear distinction between conduct books and courtesy books (etiquette books appeared much later in the nineteenth century). According to her classification, courtesy books deal with the ideal courtier and his behaviour. A particularly prominent example is Baldassare Castiglione's book, *Il libro del cortegiano* 'The Book of the Courtier', published in 1528 and soon translated into English and many other European languages. Conduct books, on the other hand, address a broader audience which also includes townspeople, not just courtiers. Paternoster mentions Giovanni Della Casa's *Galateo ovvero de' costume* (1558) as a pertinent example (see Culpeper 2017 on the reception of the *Galateo* in England). According to this distinction, the medical texts in MEMT are, of course, much closer to the conduct book tradition than to the courtesy book tradition. As pointed out above, they are addressed to a wide range of readers, including professionals and lay people, and indeed, Paternoster (2022: 39) links conduct books back to instructional literature in English going back to the thirteenth century.

The examples quoted above also provide a link to the use of the term *courtesy* in other Middle English texts, including fictional texts (Jucker 2020: Chapter 3). In these fictional texts, the term *curteisie* is used as a term to refer to a virtue or a praiseworthy moral ideal. It is used for aristocratic characters, but also with an ironic twist to members of the clergy or to lower class characters. In MEMT five examples of the term *curteis* (including spelling variants and other word forms) are attested, and they all occur in the context of advice giving.

7. Conclusion

The above survey of the uses of the meta-illocutionary expressions *rede* and *counsel* (and their spelling variants) in MEMT show the salience of these expressions in the context of Middle English advice-giving medical discourse. They were regularly used both to perform the speech act of advice giving and to talk about it. They are attested in 24 out of 172 source files of MEMT, but the examples show that there is a great deal of variation in how they are used across the different medical texts.

From a diachronic perspective, *rede* becomes rarer in time while *counsel* takes over. This is perhaps not very surprising as it is the expression *counsel* that survives into later stages of the English language up to Present-Day English. *Rede* with the meaning of 'giving advice', on the other hand, is now archaic and restricted to poetic and literary uses (OED *rede* n.). It may also be significant that the chronological difference coincides with the traditions of medical writing: the native *rede* is more common in remedy books, which have a tradition going back to OE, while the surgical texts and the specialized texts of the more recent learned tradition favor *counsel.*

In the previous section, we have shown that advice giving in Middle English medical discourse goes beyond recommendations for healthy living and ways of restoring one's health. At least some of these texts also provide counsel for the medical profession on how to behave in their interactions with patients and members of their households. As such, they provide a fascinating picture not only into suggested treatments, but also into the discursive interactions in medical contexts.

Bibliography

Corpora and dictionaries

Bosworth Toller's Anglo-Saxon Dictionary Online, https://bosworthtoller.com
Dictionary of Old English Corpus, https://doe.artsci.utoronto.ca/
Helsinki Corpus, www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus/
Historical Thesaurus of the Oxford English Dictionary, www.oed.com/thesaurus
MEMT. *Middle English Medical Texts* (2005). Compiled by Irma Taavitsainen, Päivi Pahta and Martti Mäkinen. With software by Raymond Hickey. Amsterdam: John Benjamins. CD-ROM.
OED = *Oxford English Dictionary*, www.oed.com/

References

Adolphs, Svenja 2008. *Corpus and Context. Investigating Pragmatic Functions in Spoken Discourse* (Studies in Corpus Linguistics 30). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/scl.30
Bergner, Heinz 1992. The pragmatics of medieval texts. *Cooperating with Written Texts: The Pragmatics and Comprehension of Written Texts*, ed. by Dieter Stein. 163–177. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110881196.163

Culpeper, Jonathan 2017. The influence of Italian manners on politeness in England, 1550–1620. *Journal of Historical Pragmatics* 18 (2): 195–213. https://doi.org/10.1075/jhp.00002.cul

D'Angelo, Jonathan & Anne-Lise D'Angelo 2018. Advice from healthcare professionals. *The Oxford Handbook of Advice*, ed. by Erina L. MacGeorge & Lyn M. Van Swol. 197–216. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190630188.013.10

DeCapua, Andrea & Lisa Huber 1995. 'If I were you ...': Advice in American English. *Multilingua* 14 (2): 117–132. https://doi.org/10.1515/mult.1995.14.2.117

Jucker, Andreas H. 2018. Apologies in the history of English: Evidence from the *Corpus of Historical American English* (COHA). *Corpus Pragmatics* 2 (4): 375–398. https://doi.org/10.1007/s41701-018-0038-y

Jucker, Andreas H. 2020. *Politeness in the History of English. From the Middle Ages to the Present Day*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108589147

Jucker, Andreas H. 2024a. "Is that a request or a command?" Speech act meta discourse and illocutionary indeterminacy. *Trust and Uncertainty: Perspectives from Linguistics and Literary Studies* (SPELL 44), ed. by Niklas Fischer, Steve Oswald, Kilian Schindler, & Julia Straub. Heidelberg: Winter.

Jucker, Andreas H. 2024b. *Speech Acts. Discursive, Multimodal, Diachronic*. Cambridge Elements in Pragmatics. Cambridge: Cambridge University Press.

Jucker, Andreas H. & Irma Taavitsainen 2000. Diachronic speech act analysis: Insults from flyting to flaming. *Journal of Historical Pragmatics* 1 (1): 67–95. https://doi.org/10.1075/jhp.1.1.07juc

Jucker, Andreas H. & Irma Taavitsainen 2012. Pragmatic variables. *The Handbook of Historical Sociolinguistics*, ed. by Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre. 293–306. Oxford: Blackwell. https://doi.org/10.1002/9781118257227.ch16

Jucker, Andreas H. & Irma Taavitsainen 2013. *English Historical Pragmatics*. Edinburgh Textbooks on the English Language. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9780748644704

Jucker, Andreas H. & Irma Taavitsainen 2014. Complimenting in the history of American English: A metacommunicative expression analysis. *Diachronic Corpus Pragmatics*. (Pragmatics & Beyond New Series 243), ed. by Irma Taavitsainen, Andreas H. Jucker, & Jukka Tuominen. 257–276. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.243.16juc

Kohnen, Thomas 2000. Explicit performatives in Old English: A corpus-based study of directives. *Journal of Historical Pragmatics* 1 (2): 301–321. https://doi.org/10.1075/jhp.1.2.07koh

Limberg, Holger & Miriam A. Locher (eds.) 2012. *Advice in Discourse* (Pragmatics & Beyond New Series 221). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.221

Locher, Miriam 2006. *Advice Online. Advice-giving in an American Internet Health Column* (Pragmatics & Beyond New Series 149). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.149

Locher, Miriam A. & Franziska Thurnherr 2017. Typing yourself healthy: Introduction to the special issue on language and health online. *Linguistik online* 87: 3–24. https://doi.org/10.13092/lo.87.4170

MacGeorge, Erina L. & Lyn M. Van Swol (eds.) 2018. *The Oxford Handbook of Advice*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190630188.001.0001

Milfull, Inge B. 2004. Advice: Performative expression in Middle Scots. *Anglistentag 2003. Proceedings*, ed. by Christoph Bode, Sebastian Domsch, & Hans Sauer. 57–70. Trier: Wissenschaftlicher Verlag.

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 & Maria José Carrillo Linares 2006. Translation strategies: *De spermate* and *De humana natura*. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS. R.14.52, Its Texts, Language, and Scribe* (Medieval and Renaissance Texts and Studies 292), ed. by M. Teresa Tavormina. 95–117. Tempe, Arizona ACMRS.

Pahta, Päivi, Turo Hiltunen, Ville Marttila, Maura Ratia, Carla Suhr, & Jukka Tyrkkö 2016. Communicating Galen's *Methodus metendi* in Middle and Early Modern English. *Communicating Early English Manuscripts*, ed. by Päivi Pahta & Andreas H. Jucker. 178–196. Cambridge: Cambridge University Press.

Pahta, Päivi & Saara Nevanlinna 1997. Re-phrasing in Early English: The use of expository apposition with an explicit marker from 1350 to 1710. *English in Transition. Corpus-based Studies in Linguistic Variation and Genre Styles*, ed. by Matti Rissanen, Merja Kytö, & Kirsi Heikkonen. 121–183. Berlin: De Gruyter. https://doi.org/10.1515/9783110811148.121

Paternoster, Annick 2022. *Historical Etiquette. Etiquette Books in Nineteenth-Century Western Cultures*. Cham: Palgrave-Macmillan. https://doi.org/10.1007/978-3-031-07578-0

Põldvere, Nele, Rachele De Felice, & Carita Paradis 2022. *Advice in Conversation. Corpus Pragmatics Meets Mixed Methods* (Cambridge Elements, Pragmatics). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009053617

Schneider, Klaus P. 2017. Is that a threat? Forms and functions of metapragmatic terms in English discourse. *AAA - Arbeiten aus Anglistik und Amerikanistik* 42 (2): 225–242.

Schneider, Klaus P. 2021. Notes on variational metapragmatics. *Journal of Pragmatics* 179: 12–18. https://doi.org/10.1016/j.pragma.2021.04.023

Schneider, Klaus P. 2022. Referring to speech acts in communication: Exploring meta-illocutionary expressions in ICE-Ireland. *Corpus Pragmatics* 6: 155–174. https://doi.org/10.1007/s41701-022-00123-w

Schoppa, Dominik Jan 2022. Conceptualizing illocutions in context: A variationist perspective on the meta-illocutionary lexicon. *Corpus Pragmatics* 6: 63–88. https://doi.org/10.1007/s41701-022-00115-w

Schrott, Angela 2014. A matter of tradition and good advice: Dialogue analysis and corpus pragmatics in Old Spanish texts. *Diachronic Corpus Pragmatics* (Pragmatics & Beyond New Series 243), ed. by Irma Taavitsainen, Andreas H. Jucker, & Jukka Tuominen. 303–327. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.243.18sch

Searle, John R. 1976. A classification of illocutionary acts. *Language in Society* 5: 1–23. https://doi.org/10.1017/S0047404500006837

Siraisi, Nancy G. 1990. *Medieval & Renaissance Medicine: An Introduction to Knowledge and Practice*. Chicago and London: University of Chicago Press. https://doi.org/10.7208/chicago/9780226761312.001.0001

Stivers, Tanya, John Heritage, Rebecca K. Barnes, Rose McCabe, Laura Thompson, & Merran Toerien 2018. Treatment recommendations as actions. *Health Communication* 33 (11): 1335–1344. https://doi.org/10.1080/10410236.2017.1350913

Taavitsainen, Irma (ed.) 1988. *Middle English Lunaries: A Study of the Genre*. Diss. Helsinki: Société Néophilologique. 121–183. Boston and Berlin: De Gruyter.

Taavitsainen, Irma 1994. A zodiacal lunary for medical professionals. *Popular and Practical Science of Medieval England*, ed. by Lister M. Matheson. 283–298. East Lansing, MI and Cambridge: Colleagues' Press/ Boydell and Brewer.

Taavitsainen, Irma 2002. Historical discourse analysis: Scientific language and changing thought-styles. *Sounds, Words, Texts and Change. Selected Papers from 11 ICHL, Santiago de Compostela, 7-11 September 2000*, ed. by Teresa Fanego, Belén Méndez-Naya, & Elena Seoane. 201–226. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.224.14taa

Taavitsainen, Irma & Päivi Pahta 1995. Scientific thought-styles in discourse structure: Changing patterns in a historical perspective. *Organization in Discourse*, ed. by Sanna-Kaisa Tanskanen, Brita Wårvik, & Risto Hiltunen. 519–529. Turku: Publications of the English Department, University of Turku.

Taavitsainen, Irma & Päivi Pahta 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

Taavitsainen, Irma & Päivi Pahta (eds.) 2004. *Medical and Scientific Writing in English 1375-1500.* Cambridge: Cambridge University Press.

Taavitsainen, Irma, Päivi Pahta, & Martti Mäkinen 2005. Introduction to MEMT. *Middle English Medical Texts*. CD ROM. Amsterdam and Philadelphia: John Benjamins.

Taavitsainen, Irma & Gerold Schneider 2018. Scholastic argumentation in early English medical writing and its afterlife: New corpus evidence. *From Data to Evidence*, ed. by Carla Suhr, Terttu Nevalainen, & Irma Taavitsainen. 191–221. Leiden and Boston: Brill. https://doi.org/10.1163/9789004390652_010

Tavormina, M. Teresa 2006. Commentary on the Hippocratic prognostication, Part 1. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS. R.14.52, Its Texts, Language, and Scribe*. Medieval and Renaissance Texts and Studies 292, ed. by M. Teresa Tavormina. 387–448. Tempe, Arizona ACMRS.

Terkourafi, Marina 2011. The pragmatic variable: Toward a procedural interpretation. *Language in Society* 40: 343–372. https://doi.org/10.1017/S0047404511000212

Tieken-Boon van Ostade, Ingrid & Fátima María Faya Cerqueiro 2007. Saying *please* in Late Modern English. *"Of Varying Language and Opposing Creed": New Insights into Late Modern English*, ed. by Javier Pérez Guerra, Dolores González Alvarez, Jorge L. Bueno Alonso, & Esperanza Rama Martínez. 421–444. Bern: Peter Lang.

Voigts, Linda E. 1984. Medical prose. *Middle English Prose: A Critical Guide to Major Authors and Genres*, ed. by A.S.G. Edwards. 315–335. New Brunswick N.J.: Rutgers University Press. https://doi.org/10.36019/9781978816954-017

Werlich, Egon 1982. *A Text Grammar of English*. Heidelberg: Quelle and Meyer.

Zhang, Yan (Olivia) & Ken Hyland 2021. Advice-giving, power and roles in theses supervisions. *Journal of Pragmatics* 172: 35–45. https://doi.org/10.1016/j.pragma.2020.11.002

Dental treatment and related vocabulary in Late Medieval England

Juhani Norri

The importance of the late medieval period in the shaping of English medical vocabulary is widely recognized. From circa 1375 onwards, the English language increasingly emerged from the shadow of Latin and French as the language of medical writing. Many of the terms used by the translators and authors of these works appear in an English-language text for the first time. This article first outlines what Late Middle English and pre-1550 Early Modern English medical works tell us about contemporary practitioners dealing with dental troubles. The causes to which such conditions were attributed and the methods of treatment applied are also discussed. The lexical analysis that follows addresses the vocabulary for teeth, sicknesses of teeth, medicinal preparations, and instruments. In each area, equivalent terms are frequent, the same concept carrying a variety of names. The English terms are often direct translations from Latin or French. The translators occasionally confuse similar words in the source language. Many of the terms reflect medieval theories about the causation of dental conditions.

Keywords: dental vocabulary, Middle English, Early Modern English, translation strategies

https://doi.org/10.51814/ufy.1041.1457

1. Introduction

Toothache can be an agonizing condition, difficult or impossible to alleviate without proper treatment of the underlying cause. The medieval surgeon Guy de Chauliac (1300?–1368?) well knew the pain that dental troubles inflicted upon his patients when he stated that “[i]t is schewed þat among þe passiouns of all þe body of þe whiche a man is ful litel ioyed, þe toþe akþe is most grevous” (Ogden 1971: 484; *schewed* ‘shown’). In churches, carvings dating from the Middle Ages show people whose jaws have been bandaged, most likely in an attempt to mitigate tooth pain (Kelly 2009: 49). Without anesthetics, x-rays, antibiotics, and many other relatively recent developments in medicine and dentistry, most of the patients must have faced a veritable ordeal. Contemporary medical works prescribe a whole variety of medicinal preparations against toothache, but the characterization “this is harde for to thole” accompanying one remedy (Ogden 1969: 17; *thole* ‘tolerate’) rather reminds one of the saying “desperate diseases must have desperate remedies”.

In medieval England, as during the Roman occupation of Britain, the regular consumption of coarse stone-ground bread had an abrasive effect on teeth, the skeletons excavated at some Roman burial sites showing badly worn-down teeth (Jackson 1988: 120; Cruse 2004: 182; Kelly 2009: 49). The poor level of oral hygiene also contributed to dental problems. The plentiful advice given in medieval “regimens of health” on how to take care of teeth and gums went largely unheeded, as did the pleas made by many medical authors (Demaitre 2013: 192). Not all things were worse than now, however. Sugar intake was less, with the result that tooth decay or caries was a rarer phenomenon, especially before the introduction of sugar into Europe by soldiers returning from the Crusades (Jackson 1988: 120; Cruse 2004: 182).

The present article surveys the discussions of teeth and their ailments in Late Middle English and Early Modern English medical manuscripts and books. More particularly, the period investigated ranges from circa 1375 to 1550. From about 1375 onwards, extensive English-language medical writings started to appear after a long dominance of Latin and French. As for the terminal date of the study, medical books printed before 1550 mostly reproduce material that was available as manuscripts in the earlier centuries, there being an upsurge in original material in the second half of the sixteenth century (Norri 1992: 53–55). The medical works analysed represent the main categories of Middle English medical writing first systematically distinguished by Voigts (1982), that is, academic treatises, surgical manuals,[1] and remedy-books. The tripartite classification is supported by linguistic evidence, as the medical writings in the three groups differ from each other with respect to lexical, grammatical, and stylistic features (Pahta & Taavitsainen 2004: 14). Besides observing what the Middle English and Early Modern English treatises examined can tell us about dental practitioners, theory, and treatment, the article also contains a brief outline of the contemporary terminology employed for teeth, their troubles, and ways of treating them.

2. Medieval practitioners treating dental conditions

Instructions for preparing medicines against toothache are found in each of the three varieties of medical writing examined. Physicians, surgeons, and various empirics took an interest in creating pastes, plasters, fumigations, and other forms of medication, to ease the plight of the patients. The late fifteenth-century banns (public proclamations) advertising the services of an unidentified itinerant doctor state that "he wil curyn ye toyȝthhache & helpyn ye pacyent be ye grace of god" (Voigts 2011: 265). More radical forms of treatment such as tooth extraction, cauterization, and bloodletting, were mostly practised by surgeons and barber-surgeons. Rawcliffe (1995: 125) cites a text from 1519 that lists "drawyng of tethe" among procedures "which restyth onely in manuall operacon, princypally with the handes of the werkman". Such manual skills were usually not taught in universities, where future physicians were expected to "speculate on the causes of illness, erect medical systems, and heal and cure the sick" (Bullough 1966: 109). In medieval times, neither Oxford nor Cambridge had surgery as part of their curricula. Instead, aspiring surgeons and barber-surgeons were trained through apprenticeship to an experienced practitioner. In England, the period of apprenticeship in the fourteenth and fifteenth centuries varied from five or six years to as many as twelve, the sons of master surgeons probably gaining their freedom sooner than other apprentices (Bullough 1959: 450–451; Ussery 1971: 12). Besides qualified practitioners, procedures like tooth extraction were also carried out by empirics, especially in rural areas and among poor people. Henry IV is

1 Often called just *surgeries* in studies of medieval medicine.

known to have awarded a pension to a Matthew Flynt, “tothdrawer” from London, for providing the poor with free dental care (Rawcliffe 1995: 144).

In Middle English and Early Modern English medical writings from the period 1375–1550, one is hard-pressed to find any references to the practitioners engaged in treating dental troubles. In the works studied, the above-mentioned Guy de Chauliac provides the most detailed description of the division of labour between the various groups of healers:

1 The firste lore is þat þise operaciouns or wirchynges ben particuler and moste approprede to barboures and to tothe drawers, and þerfore leches haue lefte þe forsaide werk to ham. It is siker forsoþe þat suche wirchers be dressede or gouernede by leches.

The secounde lore is þat it byhoueþ a leche counseilynge in suche þinges þat he knowe þe helpes of tieth. (Ogden 1971: 485; *lore* ‘teaching’, *wirchynges* ‘procedures’, *moste* ‘especially’, *approprede* ‘appropriate, proper’, *siker* ‘safe’, *wirchers* ‘practitioners’, *dressede* ‘guided, directed’)

Guy de Chauliac, considered one of the most eminent surgeons of the entire Middle Ages, would typically not undertake tooth extraction himself. Instead, the most dramatic and ultimate form of treatment would be left to barbers and tooth-drawers, who needed to be supervised by “leeches”, that is, physicians or surgeons (*Middle English Dictionary*[2] s.v. *leche* n.3, sense 1a). Procedures like blood-letting, cupping, purging, and cauterization, used against toothache like many other troubles, were carried out by both surgeons and barber-surgeons. The latter also had a role in preventive care and maintenance of dental hygiene, as appears from the following passage in Roesslin (1545: f. 147v):

2 To kepe and preserue the teathe cleane. Fyrst yf they be very yelow and fylthy or blackisshe, let a barbar skowre, rubbe & pyke them cleane and whyte.

2 Hans Kurath *et al.* eds. Hereafter *MED*. In the passage cited, *leche* translates the Latin word *medicus* (McVaugh 1997: 356).

3. Theories about dental conditions

The theoretical foundation of medieval medicine lay in the doctrine of the four humours going back to ancient Greek medical authors. The human body was thought to consist of blood, phlegm, bile (also called choler), and black bile (or melancholy). The proportions of these bodily fluids determined a person's physical and psychological disposition, as seen from the adjectives *sanguine*, *phlegmatic*, *choleric*, and *melancholy*, still used of mental traits (Siraisi 1990: 104–106). Sickness resulted from an imbalance of the humours, or from them becoming unstable or putrid. The following passages in the medical treatises studied illustrate the application of humoral theory to dental problems, in particular toothache and caries:

3 Sumtyme a mannes teeþ ben frete & ben holow, & þis comeþ of humours corrupt þat falliþ to þe teeþ. (Fleischhacker 1894: 265)

4 If the reume do destill into the cheekes and teeth, I haue proued that the iuice of ground iuy, and that herbe whiche we call mouse eare, taken within a quyll into the nosethrilles, oftentymes purgeth excedingely the reume and taketh away the ache of the teethe. (Elyot 1541: f. Y4r; *reume* 'watery bodily humour thought to drain from brain')

5 For the toth ake. The Causes. The synowes being very hote or cold or great quantity of humors, fallynge from the head to the gummes. (John XXI 1550?: f. F8v)

Rheum or corrupt humours distilling from the brain and trickling downwards was thus regarded as a common reason for aching or decaying teeth. But the source of trouble could also be located in the stomach. Demaitre (2013: 193) observes that toothache was attributed to "humors from the brain or vapors from the stomach". Both causes are mentioned in many of the works examined, among them the following two:

6 In a mannes teeþ ben diuerse greuauncis: ache, wormes, stynking, rotenes. Ache comeþ in diuerse maneris: Oþerwhilis, it comeþ of þe hede and of þe humour of þe heed þat falleþ doun into þe teeþ, and þan comenli þe ouer-teeþ aken. Oþerwhilis, it comeþ of þe stomake, and þen þe neþir teeþ ake. (Getz 1991: 91–92; *oþerwhilis* 'sometimes')

7 Tooþache comeþ somtyme of þe vice of þe stomake, somtyme of þe vice of þe brayne, whanne colde humours oþir hote rewmatik infectiþ þe synewis of þe teeþ and brediþ ache; of vice of þe stomake whanne it comeþ of hote humoures þat beþ in þe stomake and fumosite þerof resolueþ and todeliþ, and comeþ vpward, and bitiþ, and prickeþ, and greueþ þe synewis and rootes of þe teeþ, and brediþ þerinne ache and stenche also. (Seymour 1975: 370; *vice* 'sickness', *todeliþ* 'spreads')

Since ancient Egyptian medicine, dental pain and decay had also been attributed to worms infesting teeth and gums, a belief that lasted into the eighteenth century (Jackson 1988: 15–16; Cruse 2004: 183; Kelly 2009: 49). Cameron (1993: 12) cites an Old English passage in which the worms are said to fall from the teeth when subjected to therapeutic smoke, seen if a black sheet is placed below the mouth.[3] In his discussion of dental ailments, Andrew Boorde (1547: f. M1r) compares the size of a dental worm and an itch-mite infesting hands: "The worme is lytle greatter than ye worme in a mans hande". Humoral theory and the concept of the worm are sometimes combined in Middle English medical writings, as appears from Bernard of Gordon's statement that "humours falleþ to þe teeþ & rotiþ & takiþ hete & in þat manere þey may engendre wormys" in Bodleian Library, Ashmole manuscript[4] 1505, f. 128v. The same idea of harmful humours generating worms is presented by both Bartholomæus Anglicus and Gilbertus Anglicus:

8 Wormes brediþ in þe cheke teeþ of rotid humours þat beþ in þe holouȝnes þerof. (Seymour 1975: 371)

9 Oþerwhilis, þe ache comeþ of corrup humours þat ben at þe rote of þe teeþ...If it comeþ of a keen humour amonge þe teeþ, it makeþ wormes in þe teeþ, eiþir it makeþ þe teeþ holewe. (Getz 1991: 91; *oþerwhilis* 'sometimes', *keen* 'sharp')

Certain foods were considered bad for one's teeth. The list of nutriments to avoid was sometimes quite long, as seen from the passages below.

3 Some recipes mention worms falling into a dish of water held by the patient (Boorde 1547: f. M1r; Heinrich 1896: 212; Ogden 1969: 19).

4 Hereafter MS.

10 Vndirstonde þese ben þo þyngges þat wiche noyn þe teeþ: rotunhede of metis & of drynkes & ofte to be dronkun, salte fische, ofte to spewe, hony, mylke cruddus, datis, ficus, lekys, quiksyluere in what maner it be vsid, ceruse, vynegre, al maner sowre þyngges be þe froytis or oþer. (Cambridge, Gonville, & Caius MS 176/97, p. 206; *noyn* 'hurt', *metis* 'foods', *spewe* 'vomit', *cruddus* 'curds')

11 Meates whiche do hurte the tethe. Very hotte meates. Nuttes. Swete metes & drinkes. Radyshe rootes. Harde meates. Mylke. Bytter meates. Moche vomyte. Leekes. Fyshe fatte. Lymones. Colewortes. (Elyot 1541: f. D2v; *meates* 'foods')

12 This payne ... may come by drynkynge of hote wynes, eatynge of hoote spyces, or eatynge of hote aples, peares, and such lyke. (Boorde 1547: f. L4v)

It is interesting to note that the compiler of the treatise in the Gonville and Caius MS mentions quicksilver or mercury as a substance that is harmful for teeth. Since the 1830s (Porter 1996: 376), amalgam fillings containing mercury were widely used in dental practice, to be largely abandoned in recent decades after the potential health hazards of such fillings came to light. Medieval writers also state that frequent vomiting and consumption of sour fruits may jeopardize one's teeth, a view shared by modern dentists.[5] The order in which the foods were eaten was a further consideration. It was thought inadvisable to "ete cold mete after ryght hote or hoot after cold" (Dawson 1934: 26).[6]

Looseness of teeth, a common complication of periodontal disease, is another dental ailment mentioned in the works studied. The condition was attributed to flaccidness of the adjacent structures, as seen from the heading "How þe teþ buþ made meuable of feblenesse and loosyng of þe senewes" in an English version of Theodoric of Cervia's surgical manual (Cambridge, Magdalene College, Pepys Library MS 1661, p. 186).

In medieval medical writings, bad breath is associated with disturbances in a variety of different body parts, including teeth and gums, brain, stomach,

5 https://www.cosmeticdentistinbrisbane.com.au/2020/09/22/does-vomiting-cause-dental-problems/, https://hancockvillagedental.com/foods-that-irritate-your-teeth/.

6 "A sodeyn changing fro hote metis to colde" is one of the causes of toothache listed by Gilbertus Anglicus (Getz 1991: 91).

and lungs (Demaitre 2013: 192). That dental cavities may cause halitosis is recognized in some of the Middle English and Early Modern English treatises examined, among them the works by Bartholomæus Anglicus and Eucharius Roesslin:

13 Stenche of mouþe comeþ somtyme of corrupcioun of þe teeþ and of þe gomes. (Seymour 1975: 369)

14 Of stynkynge breth…If yt be engendryd in the mouth onlye, then most comunely it cummyth of sum rotten & corrupted holowe teth, whiche in this case must be plucked out. (Roesslin 1545: f. 148r)

It was sometimes thought possible to predict the outcome of a disease from the patient's teeth. In his discussion of a feverish patient, Bernard of Gordon tells his readers that "ȝiff hys teeþ be drye as it were tre, þat is euel signe & mortel" (Bodleian Library, Ashmole MS 1505, f. 127v).

4. Treatment of dental conditions

Medieval attempts to cure toothache and other dental afflictions were largely based on humoral theory (Kelly 2009: 49; Demaitre 2013: 192). The four humours were in various proportions associated with the four qualities of hot, cold, moist, and dry. If the causative humour was dominated by the hot quality, the medicinal preparation would primarily consist of ingredients regarded as cold. The effects of a moist humour were counterbalanced with dry medicines. Bloodletting, cupping, and purging were common procedures for evacuating excessive or troubled humours. The basic tenets of humoral theory and their importance for diagnosis and treatment can be seen in the following and many other passages in Middle English and Early Modern English medical writings:

15 If it [toothache] come of vicis of þe heed or of þe gomis, and þe cause come of hete, þat þou miȝt knowe bi reednes of þe place & bi þe hete þerof, & þan þou schalt lete him blood in þe heed veine, & þan þou schalt lete him blood in þe veine þat is vndir his tunge…If þe akynge come of coold humouris, þat þou miȝt knowe if þe place be not reed ne to-swollen, þan purge hi wiþ cochijs or wiþ pigra. (Fleischhacker

1894: 264; *vicis* 'flaws', *heed veine* 'cephalic vein', *cochijs*: 'pills or boluses for purging excess bodily humours, esp. from the head', *pigra* 'purgative based on aloes')

16 Oþerwhilis, þe ache comeþ of corrup humours þat ben at þe rote of þe teeþ. And if it comeþ of hete of þe heed, þe chekes ben to-swollen, and reed, and hoot. In ache þat comeþ of cold, þer is palenes of þe cheke and colde of þe teeþ. And if it comeþ of moistnes, þer is swelling and softenes of þe cheke, and if it comeþ of drines, þer is hardenes and litil swelling with þe ache. (Getz 1991: 92; *oþerwhilis* 'sometimes')

17 Fyrst purge the heed with pilles of cochie. And vse gargaryces. And yf it do come of any colde cause, chewe in the mouth diuers tymes the roote of horehounde (Boorde 1547: f. M1r; *pilles of cochie:* see *cochijs* in (15), *gargaryces* 'gargles')

Fumigations, as mentioned earlier, were used to expel the supposed worms in the teeth. In many recipes, seeds of henbane (*Hyoscyamus niger*) are placed on glowing coals or stones, the patient inhaling the smoke through a pipe.[7] Grieve (1982: 403) comments on the treatment still practised in some parts of Britain in the early twentieth century. She notes that fumes from henbane seeds are "dangerous, having caused convulsions and even insanity in some instances". Medieval physicians and surgeons were aware of such hazards, as appears from the warning "be-ware þat þu drawe nott thyn ande for strikynge of þe sauoire of þe sedis in-to thi throtte" in one remedybook (Ogden 1969: 19; *ande* 'breath').[8] Notes of caution also accompany the use of other poisonous plants in dental care, among them hellebore.[9] Platearius warns against ingesting the juice from the plant, stating that "it is to dreden þe þe iouce of ellebori come not within be cause þat it woll dissolue þe membris,

---

7 A picture of such treatment appears in a thirteenth-century Anglo-Norman translation of Roger of Parma's *Surgery* (Hunt 1992: 72–73).

8 The writings of Dioscorides, Celsus, and Pliny show that henbane was used to alleviate toothache in ancient Greece and Rome already. Like some medieval authors, Scribonius Largus believed that vapour from henbane seeds would expel the worms responsible for tooth decay (Jackson 1988: 121).

9 Hunt (1989: 106) identifies Medieval Latin *elleborus* in lists of plant-name synonyms with white hellebore (*Veratrum album*), black hellebore (*Helleborus niger*), and bear's-foot (*Helleborus foetidus*), all poisonous.

for why it is venemous & dystourblyng" (Cambridge, University Library MS Dd.10.44, f. 33r).

Some of the substances prescribed in the preparations against toothache and other dental troubles are prime examples of what has been termed the *Dreckapotheke*, "the use of remedies that are considered filthy, dirty or disgusting", with "excrement, urine, and any kind of dirt" as the basic ingredients (Maurer 2017: 1248). It is no wonder that sometimes the physician or healer was told not to reveal the true composition of the medicine to the patient, as in the first extract below.

18 ffor akyng of the holow toth. take a ravennys tord and put it in the holow toth and coloure hitt w^{t} the juse of peletre of Spayne that þe syke knaw it not ne wote not what it be (Dawson 1934: 24; *peletre of Spayne* 'pellitory of Spain, *Anacyclus Pyrethrum*', *wote* 'know')

19 Þe skyn of a serpent may be goode in þis cause, and ȝitt serpentys biþ venymous. I sey þat venymous bestis ne kepiþ noȝt here venym in þe skyn in þe superfluytees, as it schewyþ in þe torde of a grene euete þat is goode in þis caas. (Bodleian Library, Ashmole MS 1505, ff. 128v–129r; *superfluytees* 'body parts thought to develop from excess humours or fumes', *euete* 'newt')

20 Forto make yonge childrens tethe to growe hastely. Sethe hares braynes and anoynte therwith the gomes. (Cambridge, Trinity College MS O.8.35, f. 28r)

21 If the hollow toth be filled wyth crowes doung it breaketh the tothe and taketh awaye the payne. (John XXI 1550?: f. G3r)

In many of the recipes against toothache, the medicinal preparation is to be put in the ear or nostril rather than on the tooth, probably in an attempt to hold the medicine in place and thereby prolong the period of its effectiveness:

22 Item primerose put in thyn nose be þe hole syde abatyth peyne. (Schöffler 1919: 201)

23 Stamp garlec wiþ pepir and do þe iuse in þe ere of þe toþer side and it shalle help. (British Library,[10] Additional MS 34111, f. 123v)

10 Hereafter BL.

24 For þe towyth ache. Recipe þe ius of dayseys leuys & stampe them & weyt cotun ther in & put it in þe nose & draw your breyt yenward. (Bodleian Library, Ashmole MS 1389, p. 80)

Many different substances obtained from plants and animals, as seen from the recipes quoted, were used in the medicines against the agony of toothache. A badly damaged tooth could be destroyed or “broken” by filling the cavity with crow’s dung, or by touching it with the root of water crowfoot, *Ranunculus aquatilis* (John XXI 1550?: f. G2v). A passage titled “Ad trahendum dentem sine dolore” in Bodleian Library, Ashmole MS 1389, p. 23, informs the reader that the root of mulberry tree soaked in vinegar, dried, and powdered draws out the sick tooth without any pain. The latter recipe, with the addition of “piretrum” (pellitory), is also found in a Middle English translation of Lanfrank’s surgery, together with two other herbal preparations that help “to drawe out ony mannes tooþ wiþouten iren” (Fleischhacker 1894: 265).[11] Guy de Chauliac’s opinion of such promises of painless tooth extraction is outright scathing:

25 Many medecynes forsoþe ben putte of olde men, þe whiche drawe out tieth or liȝtne ham to drawynge out wiþ yren, as, þe mylke of tytymalle with pylettre and with þe roote of mulbery and of capparys and wiþ arsenyk and putte in þe roote of þe toþe, or þe strong water or þe grece of frogges of þe wode and of trees. Neuerþelatter þay ȝeue many byhestes and fewe werkes forsoþe. (Ogden 1971: 490–491; *tytymalle* ‘spurge’, *pylettre* ‘pellitory’, *capparys* ‘capers’, *strong water* ‘nitric acid or some other powerful solvent’, *byhestes* ‘promises’)

When the dental condition failed to respond to medicines, the pros and cons of treatment with instruments would be considered. One possibility was to apply a red-hot cautery to the hole in the tooth. The surgeon Roger of Parma describes the procedure as follows:

---

11 Another recipe “to plucke out a Toothe”, from circa 1570, appears in the remedybook edited by Macgill (1990: f. 160v): “Take the braynes of an haire and seeth it in redd wyne, and therwth annoynte the tooth that you will haue out, and it will fall out wthout payne.”

26 For the toth ache if it be causide of a false toth. Then take a smal iren crokyde at the ende and make it rede hote in the fire and make the pacient to gape. And sett the ende of the yren in the myddest of the hole of the toth & it is a souerayne medicyn. (BL Sloane MS 240, f. 14v)

John XXI (1550?: f. G6r) advises a form of cauterization that must have been difficult to put into practice: “Light a sharpe sticke of ashe and whyle it burneth put into the hollow toothe first filled with triacle.” In order to prevent damage to the adjoining tissues, some medieval writers recommend that the cautery should be applied through a metal pipe. Cambridge, Gonville and Caius MS 176/97, p. 206, instructs the medical practitioner to “tak two longe yren nedlis or of copre & firy hem & ley hem to þe teeþ þat ackyth þorwe a pype of yren”.[12] Instead of a metal instrument or *actual cautery*, a caustic substance or *potential cautery*[13] could be used. Kelly (2009: 50) mentions that medieval healers sometimes poured acid into the dental cavity to numb the ailing tooth (cf. “strong water” in (25)). Bernard of Gordon was well aware of the hazards involved in the application of potential as well as actual cauteries:

27 Make hym a cauterye wiþ arsenyk, vitriole & cantarides. Neuerþelattre þou most bewar in þese þinges þat no þynge ne passe in, for it wolde myche agreue. (Bodleian Library, Ashmole MS 1505, f. 128r; *cantarides* ‘broken dried bodies of the blister beetle’)

Tooth extraction was regarded as the ultimate remedy in dental troubles. Gilbertus Anglicus states that a tooth infested by worms must be pulled out if fumigation with henbane seeds fails to work (Getz 1991: 94). Roger of Parma would similarly resort to the procedure in severe cases of dental decay, noting that “ther is no medicyn so good for a roten toth as is pulling out þerof, for þat is most sekerest” (BL Sloane MS 240, f. 14v). Other writers were more sceptical, including Andrew Boorde, according to whom extracting one tooth may lead

12 Bernard of Gordon is another writer warning against harm to the nearby areas of the mouth: “In þe last make a cauterie in his tooþ wiþ a ȝerde of yre put in wiþ a canele for harmynge of placis þat biþ neiȝ” (Bodleian Library, Ashmole MS 1505, f. 128v; *ȝerde* ‘stick’, *canele* ‘pipe’).

13 *MED* s.v. *cauterie* n., sense 1a.

to a chain reaction: “beware of pullynge out any toth, for pull out one and pul out mo” (Boorde 1547: f. M1r). Platearius and Bernard of Gordon both allude to the possibility of the ache continuing or returning even after the procedure:

> **28** Ʒyf þis medicyn & oþer forseid help nat, þan þe best cure is þis: for to craftly drawyn out þe teithe with all his roottys, for why & þer be leuyd ony rote þe whiche be nat drawe out, þan þe ake is worse þan before. (Cambridge, University Library MS Dd.10.44, f. 33v)
>
> **29** Ʒiff þe akynge ne be noʒt acesyd hirwiþ vnrote hym. And þan þou most do þerto cautelys þat þou ne do awey noon but ʒif he be greuous oþer rotyd oþer yfrete, for þe akynge wole come aʒen þere þe tooþ was. Þerfor it is ful lytel helpe to drawe awey teeþ. (Bodleian Library, Ashmole MS 1505, f. 129v; *cautelys* ‘cautions’)

Platearius emphasizes the importance of carefully extracting all the roots of the rotten tooth. Modern dentists would agree with him in the matter, knowing that “leftover tooth fragments can cause dental infections and a myriad of other problems if they are not properly removed”.[14]

Prosthetics have a long history in dentistry. The Etruscans, since the seventh century BCE, are known to have employed bridgework in which the substitute tooth or series of teeth was riveted to a band of gold with loops at each end. The loops were then used to attach the dental bridge to the healthy teeth at either side of the gap (Jackson 1988: 119–120; Cruse 2004: 184). Archeological evidence from Roman cemeteries reveals that artificial teeth were prepared from pebbles and wrought iron as early as the first or second century CE (Cruse, ibid.). In dental repairs, the favourite medium among the Romans was gold. Loosened teeth, for example, were tied to the adjacent firm teeth by gold wire, the patient’s finances allowing (Jackson 1988: 120).

Kelly (2009: 50) notes that although both Egyptians and Romans had developed dental bridges and implants, “this did not seem to be popular during the Middle Ages”. The medieval medical treatises studied for the present article, with hardly any references to artificial teeth, lend support to that view. Guy de Chauliac stands alone in his description of the techniques of treating “a tooth þat is movede and made feble”. Such a tooth, in his opinion, should first

14 https://bestdentistinhouston.com/piece-of-tooth-left-after-dental-extraction/.

be fastened with a small gold chain. In the case of dental loss, the surgeon recommends preparing an artificial tooth from another person's tooth or from cow bone:

30 The whiche, if it helpe not, bynde ham with a softe smal cheyne of gold, as Albucasis techeþ. And if þai falle away, make hem of oþer menis teeþ or of a kowes bone, and bynde ham with a sleiȝte, and he may be serued with hem longe tyme. (Ogden 1971: 489; *sleiȝte* 'prudence')

Rather than with gold wire or chain, loose teeth are usually treated by applying various medicinal powders. The "pouder to conferme teþ þat ben lose" in Cambridge, Peterhouse MS 118, II, f. 106va, contains white coral, red coral, and mastic. The following are two more recipes for the ailment, the latter mentioning eating elecampane (*Inula helenium*) leaves as a potential remedy:

31 For teth þat stonden noȝt faste. Take hertis-hornys, and brenne hem, and do þe askys in a lynyn cloth, and leye it þer-to; and þei schull syttyn faste. (Müller 1929: 98)

32 To make a waggynge tethe fast. Take & eate the lewis of elena campana. (Bodleian Library, Ashmole MS 1405, f. 11v)

The lines of treatment so far presented could be supplemented or replaced by something very different. As Rawcliffe (1995: 95) states, charms thought to possess healing powers are found in medical works owned by laymen and professional healers alike. The medical manuscript compiled by Thomas Fayreford, a fifteenth-century medical practitioner well versed in academic medicine, contains some twenty charms, with tooth problems, childbirth, and epileptic seizures forming prominent concentrations (Jones 1998: 176). In the charms against dental troubles recorded by Fayreford, St Appollonia and St Nichasius are often invoked, and sometimes characters should be written on the patient's cheek (ibid.). St Appollonia also figures in the dental cures found in the remedybooks here examined. The story according to which she had her teeth removed and jaw broken by her torturers explains why she was thought sympathetic to those suffering from toothache

(Kelly 2009: 47). References to the saint's terrible fate appear in some of the charms, including the mentions that "Sancta Apolonia virgo fuit inclita, cuius pro Christi nomine dentes extracti fuerunt" (Müller 1929: 130), "tiranni eius dentes cum malleis ferreis fregerunt et in hoc tormonto orauit ad dominum" (Heinrich 1896: 148), and "beata appollonia martir & virgo amaram & horribilem dencium excussionem constanter sustinuit" (ibid., 221).[15]

Besides St Appollonia, other Christian figures make their appearance in passages dealing with dental afflictions. The "Carmen pro Dolore Dencium" in Heinrich (1896: 102–103) is based on the story of Jesus curing St Peter's toothache. A similar incident occurs between Jesus and St Abraham in another remedybook (Cambridge, Trinity College MS R.14.51, ff. 10r–10v):

> **33** A charme for tethe and þat a gode. Take vyne lefes and stampe hem wele and lay it to þe holere cheke. And charme hit with þis worde: Seynt Abraham lay on slepe, ora pro nobis, vpon þe Mounte of Olyuete, ora pro nobis. There come Oure Lord God, ora pro nobis. What doste þou? Slepeste þou, Abraham? Ora pro nobis. I ne slepe, I ne wake, ora pro nobis. My tethe I wepe for wrake, ora pro nobis. Ryse vp, said Jhesu, Abraham, and go with me, ora pro nobis. Schall not but gode falle vnto, ora pro nobis, neyþer to none oþere man, ora pro nobis, þat þis worde saye or thynke can, ora pro nobis. And þerto sey a Pater Noster and an Auee. (*wrake* 'pain, suffering')

The two Christian prayers mentioned in the charm should also be uttered when gathering sage leaves for a medicine alleviating toothache, the recipe in Ogden (1969: 18) beginning as follows:

> **34** To make a spyce for þe tothewerke: Tak lij leues of sawge gedirde on a foure sqwarede bedde on crose from on cornere to an oþer, and

15 English translations: "St Appollonia was a famous virgin whose teeth were extracted for the name of Christ", "tyrants crushed her teeth with iron hammers and in this torment she prayed to the Lord", "blessed Appollonia, martyr and virgin, steadfastly endured the bitter and terrible knocking out of her teeth". A version of the St Appollonia charm also appears in Ogden (1969: 18). I am grateful to University Lecturer Soili Hakulinen for help in translating the Latin passages into English.

saye at þe pullynge of euer ilk a lefe a Pater noster and Aue Maria... (*spyce* ‘medicine’, *bedde* ‘bed of herbs’)

Demaitre (2013: 192) states that in medieval times oral hygiene was largely neglected, “regardless of the extensive coverage in regimens of health and in spite of the emphatic pleas by nearly every author of a practica”. There are similar instructions in the Middle English and Early Modern English medical works studied. Gilbertus Anglicus, Theodoric of Cervia, and Andrew Boorde, among others, advise their readers on ways of maintaining good oral hygiene:

35 To drye þe teeþ aftir mete with a drye lynen clooþ is profitable, for þat shal clense hem, þat no mete cleue not, ne no corrupcion amonge þe teeþ to make hem roten. (Getz 1991: 97; *mete* ‘meal; food’)

36 Aȝens stynkyng of þe mouth & of teth or of gomys. Wasshe þi mouth thries a day with sauge and eysil. (Cambridge, Magdalene College, Pepys Library MS 1661, p. 13; *sauge* ‘sage’, *eysil* ‘vinegar’)

37 To mundyfy the teeth, wasshe them euery mornynge with colde water and a lytle roche alome. (Boorde 1547: M1r; *mundyfy* ‘cleanse’, *roche alome* ‘rock alum’)

Dental cosmetics is another issue addressed in medieval medical treatises. In particular, ways of whitening discoloured teeth are prescribed by many authors. Surgeon Lanfrank, for example, explains that “if a mannes teeþ ben blac, in þis maner þou schalt make hem whit” (Fleischhacker 1894: 265). In the remedybooks, plenty of recipes have titles such as “For ȝelwe & stincking teþ” (BL Sloane MS 3153, f. 46v), “For yelowe, blake or roten tothe or stynkyng” (Cambridge, Trinity College MS O.8.35, f. 27v), and “Forto make blake teþ wite and veyre” (BL Sloane MS 2457, f. 19v; *veyre* ‘fair’). Teeth that were too long were also considered unsightly. In such cases, Bernard of Gordon recommends something similar to modern odontoplasty, removal of small amounts of enamel. The procedure now involves little or no discomfort, the oral surgeon using a sanding disc or diamond bur to reduce the size of the tooth. Medieval patients were likely to experience a different level of anxiety altogether: “Whan teeþ wexiþ to longe oþer wexiþ to mych, strayne his tooþ &

sawe hym awey as myche as nediþ" (Bodleian Library, Ashmole MS 1505, f. 130r).[16]

5. Tooth-related vocabulary in Late Medieval England

The present section provides an overview of Middle English and pre-1550 Early Modern English terms for teeth, their ailments, the medicines employed, and the instruments wielded by the practitioners. The medical manuscripts and books examined were all part of the corpus for Norri's (2016) *Dictionary of Medical Vocabulary in English, 1375–1550*.[17] Except for direct quotations from medieval texts, the forms of the lexemes cited below are the same as the head-words in that two-volume work.

Details about the anatomy of teeth are rare in medical writings from the Middle Ages. Demaitre (2013: 192–193) comments on the paucity of such information, adding that even where more detailed descriptions are found, they are almost entirely based on Avicenna (980–1037). Demaitre cites a passage from Bernard of Gordon displaying the author's cursory regard for the anatomical specifics of teeth. The English translation in Bodleian Library, Ashmole MS 1505 (f. 129r) contains a similar vague reference to "anatomy" as the source for further detail:

> **38** In þe last entende þat þere biþ in a man 32 teeþ, but noȝt in alle men. And þe forteeþ biþ clepyd scindentes, þat biþ kuttynge. Þe greter teeþ biþ iclepyd molares. And biþ dyuersyd as þou myȝt se in anathomia. (*entende* 'understand', *clepyd* 'called')

Even in surgical treatises with separate sections on human anatomy, observations about teeth are mostly rather superficial. Lanfrank simply states that "in a mannes mouþ ben .xxxij. teeþ, & þerof sittiþ .xvj. in þe cheke boon, & summen han but .xxviij" (Fleischhacker 1894: 261; *cheke boon* 'jawbone'). William of Saliceto is equally succinct when he informs the reader that "in þat vppere chawl ben festned teþ 16 & in sume 14" (BL Additional MS 10440, f. 4v;

16 For reducing the size of teeth, Guy de Chauliac tells the practitioner to use a file: "If þe toþe were encresede ouer kynde, even it, and playn it sliely with a file, and move it nouȝt" (Ogden 1971: 490).

17 Hereafter *DMV*.

chawl 'jaw').[18] Henri de Mondeville is a third surgeon whose discussion of dental anatomy remains on a very general level:

> **39** The number of teþe ys dyuerse in diuerse men. And þerfor sum men were wunte to haue xxxij[ti] teþe, and sum men xxviij[ti]. And þer ben diuerse opinions wheþer þat þei be bones oþer none. (Cambridge, Peterhouse MS 118, I, f. 22vb; *wunte* 'accustomed')[19]

The surgeon with the richest vocabulary for teeth in the writings examined is Guy de Chauliac. The passage below is from one of the English translations of Chauliac's *magnum opus*. The anatomy of teeth and the relevant terminology are explained more fully than in the passages from the other surgical writers.

> **40** The tieþ forsoþe beeþ of þe kynde of bones, þogh þai be saide to haue felynge...Neuerþelatter þat is by resoun of some synowes goynge downe fro þe þridde payre to þe roote of ham. Þay beeþ at þe moste 32[ti], þat is to say 16 in eyþer iowe (þogh þat in some men is nouȝt ifounde but 28), þat is to say: þe 12 [read: 2] euene tieþ and 2 fourcornerde and 2 tuskes, 8 grynderes and 2 caysales. And þay haue rootes isette in the iowes: some oon, some two, some þre, some foure. (Ogden 1971: 45; *synowes* 'nerves')

In the extract, the names for specific teeth include *tusk* 'canine tooth', *grinder* 'molar tooth', and *caisal* 'wisdom tooth'. But Chauliac's influence on Middle English dental nomenclature went beyond those three words. The entire Latin text of his *Chirurgia* was translated into English at least three times over, and there is a fourth independent English rendering of the anatomical section (Book 1) of the work (Wallner 1991). The translators used different techniques when dealing with the Latin words and phrases. As seen

18 There are folios missing after f. 4, possibly containing a description of the lower jaw.

19 The passage continues with a discussion of the reasons why teeth were regarded as bones by some authors, but not by others. In the surgical treatise compiled by an anonymous London surgeon, the section on teeth reproduces what Henri de Mondeville wrote about the matter (Grothé 1982: f. 27rb).

from the passage below, the text in the Cambridge, Jesus College MS Q.G.23 (f. 36ra) is terminologically quite different from the manuscript edited by Ogden.

> **41** Þe teeþ forsoþ ben of þe kinde of bonis þoȝ þei be seide to haue in hem feling...Neuerþeles þat is bi cause of certein sinewis þat descenden doun fro þe 3 peire or couple of sinewis to þe rotis of hem. Þei ben forsoþe for þe more partie 32, .s. 16 in euereiþer chaule þoȝ in som be founde but 28 .s. 2 dualis or flat, 2 quadrupli & square & 2 canini or hounde teeþ & 8 grinders & 2 wong teeþ. And þei haue rotis fast in þe chaulis, som oon, som 2, som 3 or 4. (*sinewis* 'nerves', *chaule* 'jaw')

Corresponding to *tusk* in the Ogden edition, the above passage has the repetitive word pair "canini or hounde teeþ" (*DMV* s.vv. *canini*, *tooth / hound tooth*), and instead of the Latin adoption *caisales*, the reader finds a word going back to Old English, *wang tooth* (*DMV* s.v. *tooth / wang tooth*). Chauliac's Book 1 was affixed to at least two other surgeries, without any recognition of authorship, which further spread the dental terminology based on his *Chirurgia*. John of Bradmore's surgical manual called *Philomena* closely adheres to the Latin Chauliac, adopting the different names for teeth, including *molares*, untranslated:

> **42** As for þe moste parte comonly thes ar teth 32, þat ys to say 16 on eydyr syde or chape bon & in sum ar fownd but 28 teth. And in þe names of the teth ar thes, þat ys to say 2 duales most inwarde and next after 2 quadrupli and after them 2 canini & 8 molares & 2 caysales & þe teth hath rotes festynyd in þe chawelles. Sum teth haue on rote, sum 2, sum 3, sum 4. (BL Harley MS 1736, f. 20r; *chape bon* 'jawbone', *chawelles* 'jaws')[20]

Perhaps surprisingly, the most prolific source of different names for teeth is John Trevisa's translation of Bartholomæus Anglicus' encyclopedia *De proprietatibus rerum*, in which the medical sections largely stem from classical

20 The Latin terms are also kept as such in the Chauliac appropriation inserted into Hieronymus Braunschweig's surgical work, translated from Dutch (1525: f. B2rb).

authorities like Constantinus Africanus, Galen, Hippocrates, and Avicenna (Se Boyar 1920: 183). In Book 5, on anatomy, chapter 20 deals with teeth, with the Latin terms sometimes kept as such, at other times translated into English. For example, Latin *pares* and *quadrupli* occur as names for central and lateral incisor teeth, respectively, but the two general terms meaning ‘incisor tooth’ are Englished, appearing as *inkitter* (< Latin *incisor*) or *forekitter* (< Latin *praecisor*):

> **43** Constantinus seiþ þat a man bar xxxii. teeþ; sextene ben isette in þe chekebones; foure þerof ben isette in þe formest partye, and ben iclepid pares and quadrupli also. And þese teeþ ben brode and scharpe, and phisicians clepeþ hem ynkitters and forekitters, for þey ben able to kerue al þing. (Seymour 1975: 202; *chekebones* ‘jawbones’)

Christopher Langton uses yet another word for an incisor tooth, *divider*, so called “bycause they teare the meate” (Langton 1545?: f. D3v; *meate* ‘food’). In Middle English and Early Modern English medical writings, the same referent often carries a host of names in different treatises. The vocabulary for canine and molar teeth, like that for incisors, varies from writer to writer. The terms for canine teeth in *DMV* comprise *canine*, (plural) *canini*, *dog tooth*, *eye tooth*, *hound tooth*, *neck of a maid*, and *tusk*, while for the molars we find *axle tooth*, *cheek tooth*, *great tooth*, *grinder*, *gum* (short for *gum-tooth*), and *molar*. Besides the above-mentioned *caisal* and *wang tooth*, there was a third word denoting a wisdom tooth, *tooth of sapience*.[21] Some of the lexemes are only attested in a single text, one characteristic of the medical vocabulary of the period 1375–1550 being its impermanence (Norri 2004: 130). A certain amount of lexical turnover is to be expected at a time when the use of the vernacular is being pioneered even in the most learned kinds of medical writing.

In the discussions of dental troubles, *toothache* and its variant *aching of (the) teeth* are the most frequent terms. The painful affliction is also called *toothwark*, *gout of teeth*, and *gout in the teeth*. The two longer phrases express

21 Norri (1998: 173–182, 251–255, 285–288) discusses equivalent terms for body parts in academic treatises, surgical manuals, and remedybooks during the late Middle English and pre-1550 Early Modern English periods.

the idea that toothache was thought to be caused by morbid humours dropping on the teeth, Old French *goutte* and Latin *gutta* meaning ‘a drop’. The same notion underlies *flux of the teeth*, used in an English translation of Bernard of Gordon (Bodleian Library, Ashmole MS 1505, f. 130v) for the flow of watery matter from brain to teeth.

Tooth decay is another area with a clustering of lexemes. The list of the relevant headwords in *DMV* is long indeed: *breaking*, *corrosion*, *corruption*, *fretting*, *gnawing*, *holing* (cf. *hole* ‘cavity in tooth’), *perforation of teeth*, *piercing (of the teeth)*, *putrefaction*, *rottedness*, *rottenness (of teeth)*, *rotting (of teeth)*, *rottingness*, and *thirling* (cf. *thirl* ‘hole’) all occur in passages dealing with dental caries and cavity formation. A third lexical concentration occurs in the terminology for an unpleasant feeling in teeth attributed to sensitivity to cold or food containing acid. In modern dentistry, acids from everyday foods and drinks are known to be one reason for the development of cold-sensitive teeth. In his surgical treatise, Joannes de Vigo states that the sour things responsible for the condition may come from outside or inside the human body:

44 Congelation chaunceth to the teeth of outward or inwarde thynges. Of outwarde when a man eateth soure thynges. Of inwarde when sower vapours ascende from the stomacke. (Vigo 1543: f. Ff2rb)

Among other writers, Guy de Chauliac emphasizes the role of hot substances in the treatment of *congelation* of teeth:

45 Of girlynge and congelacioun of þe tieþ.[22] Holde hote wyne or aqua vite in the mouthe, or rubbe þe tieth wiþ rotede salte. Or lay þerto nottes or walnotes rostede al hote and suche oþer hetynge þinges, or propurly soche, as purseleyne, and chewe þe sede þerof. (Ogden 1971: 490; *purseleyne* ‘purslane, *Portulaca oleracea*’)

The different words and phrases for the condition listed in *DMV* include *chilling (of teeth)*, *congealing of teeth*, *congelatio*, and *congelation (of the teeth)*. To these can be added a mistranslation of Latin *congelatio* as *curding*

22 The Latin heading in McVaugh (1997: 360) reads “De stupore et congelacione dencium”.

("cruddynge") in Ogden (1971: 483). The translator has confused two meanings of the verb *congelare*, 'to become frozen, freeze' and 'to cause to solidify, coagulate'.[23]

Numbness of teeth is a fourth ailment for which the works studied employ a notable number of equivalent terms. Severe dental pain due to injury or tooth decay may eventually stop, the affected tooth becoming numb when the infection spreads to the pulp and regions around the roots.[24] Joannes de Vigo (1543: f. Ff2rb) cites less dramatic reasons for the loss of feeling, that is, holding cold liquids or narcotic medicines in the mouth. Medieval medical writers were familiar with the condition, variously calling it *astonying of teeth*, *dormitation*, *grilling of the teeth*, *sleeping*, *slumbering of teeth*, *stonying of teeth*, and *stupor*. The verbs *astonen* and *stonen* were both used in the senses 'stun', 'stupefy', and 'deprive of feeling' (*MED* s.vv. *astonen* v., *stonen* v.2). One of the meanings of the verb *grillen* was 'to shudder, quake, be afraid' (*MED* s.v., sense 1c). In (45), the tooth malady is compared to mental stupor.

From outside the four lexical concentrations discussed, mention can be made of the phrases for grinding of teeth or bruxism (*grinding of the teeth*, *grinting of the teeth*) and looseness of teeth (*shaking of the teeth*).

The pharmaceutical lexemes that appear in the instructions for dental care and cure are mostly general words like *ball* 'pill or bolus', *fumigation*, *ointment*, *paste*, *pill*, *plaster, powder*, or *pultes* 'poultice'. *Dentifrice* 'powder for cleaning teeth', from Latin *dentifricium*, is a rare instance of a term restricted to passages on dental matters. The three independent translations of Guy de Chauliac treat the original Latin word differently.[25] The translator of Cambridge, Jesus College MS Q.G.23 (f. 323ra) has taken over *dentifricium* in the form *dentifrice,* with a minor modification of the ending only. In the text edited by Ogden (1971: 490), *dentifricium* is rendered as "frotynge of tieth". The third translator combines the two approaches, producing the repetitive word pair "dentifriciez i. frotyng of teþe" (see the quotation from the New York Academy of Medicine MS 12 in *MED* s.v. *dentifricie*).

23 *Oxford Latin Dictionary* (ed. Glare) s.v. *congelō*. For further examples of mistranslations from Latin in Middle English medical writings, see Pahta and Carrillo Linares (2006: 111–115) and Norri (2017: 580–582, 596–597, 606–607).
24 https://southlanddentalcare.com/signs-you-need-emergency-dental-care/.
25 For the Latin passage, see McVaugh (1997: 360).

The different translational techniques are also seen in the handling of another Latin term employed by Chauliac, *resumptivum*, signifying a medicine for restoring a loose tooth.[26] The most thorough Englishing is again found in Ogden (1971: 489), where "restorynge medecynes" appears in the corresponding sentence. This time, the text in the Jesus College manuscript couples a Latin adoption and its English equivalent in the co-ordinated phrase "resumptifs or fastners" (f. 322rb). The translator of the New York Academy of Medicine manuscript adheres to the Latin original most closely, opting for "resumptiuez" (*MED* s.v. *resumptive*).[27]

In the terminology for instruments, the type of cautery most often mentioned in the context of teeth is a cautery with a narrow round point, easy to insert into a dental cavity. The Latin phrase *cauterium punctuale* is attested in the English works in its original form, slightly Anglicized (*punctual cautery*), or more fully translated (*pointed cautery*). The adjective may be used alone, as a noun: "The fourþe instrument is a punctale, hauynge a smal poynte and rounde" (Ogden 1971: 572).[28] Like cauteries, other instruments had a wider use than just dental operations. Several surgical manuscripts include a picture of a *levator*, also called *levatory* or *levor*, the lexical variants going back to Old French *levatoire*, *levier*, Anglo-Norman *lev(i)er*, and Latin *levator*, *levatorium*. Henri de Mondeville in fact had difficulties in coming up with an appropriate verbal explanation, confessing that "I maye nott well writte þe discripcion of a leuatorie but in þis maner", then adding a picture (Cambridge, Peterhouse MS 118, I, f. 91vb).[29] The instrument, used for extracting teeth as well as raising depressed portions of fractured bone, was either bifurcate (*biforked levatory*, *levatory with two branches*, *levor with two branches*, *two-forked levor*) or non-

26 McVaugh (1997: 359).

27 In their treatment of Latin *resumptivum*, the three translations conform to what Wallner (1991: 159) writes about them: "Whereas the New York MS closely adheres to its Latin antecent and the Paris MS attempts to render the text in idiomatic Middle English, the Jesus College version makes use of paraphrases and explanations." With *dentifricium*, however, the picture is slightly different.

28 Further examples of the nominal use (some ambiguous between a noun and an adjective) are given in *DMV* s.vv. *punctal* and *punctual*. For pictures of the *cauterium punctuale*, see Wallner (1965), McVaugh (1997: 418), and Cambridge, Jesus College MS Q.G.23, f. 372va.

29 Pictures of the instrument are also found in Wallner (1965), McVaugh (1997: 189), Cambridge, Jesus College MS Q.G.23, f. 185ra, and in John of Bradmore's *Philomena* in BL Harley MS 1736, f. 42v.

branching (*simple levatory*, *single levatory*, *simple levor*, *single levor*).[30] For tooth extraction, a barber-surgeon's or tooth-drawer's tool kit would also typically contain various types of forceps. Some of them had serrated jaws to enable a firmer grip of the tooth (*scotched pinsons*, *toothed pinsons*, *toothed tenacles*), while others had jaws forming a hollow (*hollow pinsons*, *hollow tenacles*, *cannulate tenacles*). Besides dental operations, both types of forceps were used to remove the shafts of arrows lodged in the body.

By the time our authors wrote their works, dental scrapers had a long history behind them, as shown by the pictures from Albucasis (936–1013) in Kelly (2009: 55). Guy de Chauliac writes about "hardenede filþes" in teeth, telling the reader to "schaue hem with schauynge knyfes and wiþ spatures" (Ogden 1971: 490; *spature* 'surgical chisel'). The three translations of Chauliac again manifest different translational approaches. The New York Academy of Medicine manuscript adopts the Latin *raspatorium* as *raspatoriez* (*MED* s.v. *raspatorie*), while the translator of the Jesus College text comes up with the partial translation *raspours* (f. 320va), from the verb *raspen* and the agentive suffix *-our* (*MED* s.vv). The terms for instruments for scraping a hard surface in the treatises studied further include *scraper*, *shaver*, and *raffe*, the last-mentioned of unknown origin.

6. Concluding remarks

The lot of anyone suffering from toothache or other dental troubles in medieval times was a sad one. The possibilities of removing the cause of the affliction were limited, and the aching would no doubt continue for an extended period of time in most cases, with numbness and loss of the tooth as the final outcome. If that happened, prosthetic repair would be very difficult and costly, involving the use of materials like gold chains. The remedies attempted reflect the severity of the pain felt by the patient. Dental cavities were cauterized with red-hot iron, acid was poured into them, and fumigations of seeds of poisonous plants were applied. God, Jesus, and saints are invoked in the many charms against toothache found in remedybooks. The theoretical framework of dentistry was largely derived from Arabic writers like Avicenna and Albucasis, who had flourished several centuries ago.

---

30 Cf. modern dental elevators, also called luxators.

On the more positive side of things, the importance of good oral hygiene was emphasized in contemporary medical writings, the authors advising their readers on the proper ways of cleansing one's teeth. The dangers that frequent vomiting and certain foods, especially acidic ones, posed for dental enamel were recognized quite correctly. Modern dentists would also agree with the exhortation to remove all the roots carefully when a tooth is extracted. Some writers candidly admit the difficulties of treating dental ailments, Guy de Chauliac even warning his readers against believing the empty promises of miscellaneous tooth-drawers.

The terminology surrounding teeth and their ailments presents a picture similar to earlier studies of Middle English and Early Modern English medical language. For a particular referent, there often exist several equivalent terms. This is seen in the lexical fields for teeth, dental conditions, medicinal preparations, and instruments alike. The proliferation of words and phrases sometimes stems from the same Latin work being translated several times over, each translator opting for their own solution in rendering the term in the source text. The three independent translations of Guy de Chauliac, in particular, had a significant role in the shaping of anatomical, including dental, terminology in late medieval England.

The present article is a general survey of dental treatment and related vocabulary in late medieval England. More detailed studies of anatomical, pathological, and therapeutical issues would no doubt further add to what we know about the development of this central area of medicine. In future lexical analyses, the origins, meanings, and uses of individual lexemes can be scrutinized more closely. It is also possible to widen the selection of works studied. The search string ".j1 dentristry" yields just three hits in the Voigts and Kurtz (2014) database,[31] but dental matters are often discussed in passages inside larger treatises on surgery or academic medicine. Nor should remedybooks, central for the study of charms and sources for lexemes not recorded elsewhere, be ignored in this context. More extensive studies of Latin and Arabic texts would be illuminating, as those works are ultimate sources for many of

31 In the search string, ".j1" refers to the "subject descriptor". The entry numbers for the three hits are 6274.00, 7165.00, and 7594.00. Interestingly, entry 7165.00 goes back to the tenth-century Haly Abbas, yet another indication of Arabic influence on medieval dentistry. The two other eVK2 entries refer to collections of medicinal recipes.

the ideas, and words, in the Middle English and Early Modern English medical writings.

Bibliography

For the texts included in *DMV*, the abbreviated titles used in the dictionary are given. In the case of manuscripts, whether edited or unedited, references are provided to the Voigts and Kurtz database (eVK2) entries that concern sections from the particular works. The information about early printed books includes the Pollard and Redgrave *Short-Title Catalogue* (*STC*) numbers.

Primary sources: manuscripts

BL Additional MS 10440: surgical treatise by William of Saliceto, anatomical section, on ff. 1r–17r (*DMV *Saliceto(1)*; eVK2 0914.00, 2243.00)

BL Additional MS 34111: remedybook beginning with the words “Experimenta dinamidiorum libri Galieni” on ff. 114v–167v (*DMV *ExperDinam*; eVK2 2628.00, 5834.00, 7932.00)

BL Harley MS 1736: surgical treatise, *Philomena*, by John of Bradmore on ff. 6r–185r (*DMV *BradPhilom*; eVK2 1409.00, 4353.00, 5460.50, 6214.00)

BL Sloane MS 240: surgical treatise by Roger of Parma on ff. 1r–137r (*DMV *RogerChir*; eVK2 2158.00, 4949.00, 6858.00, 7549.00)

BL Sloane MS 2457: remedybook on ff. 8v–32v (*DMV *RemSl.2457*; eVK2 0262.00, 3415.00, 3871.00)

BL Sloane MS 3153: remedybook on ff. 2r–97v (*DMV *RemSl.3153*; eVK2 0627.00, 1760.00, 2389.00, 2960.00, 5037.00, 5796.00, 6364.00, 6907.00)

Bodleian Library, Ashmole MS 1389: remedybook on pp. 1–256 (*DMV *RemAshm.1389*; eVK2 1078.00, 1085.00, 2657.00, 3678.00, 4275.00, 4292.00, 4453.00, 4475.00, 4541.00, 5576.00, 5977.00, 6206.00, 6720.00, 6754.00)

Bodleian Library, Ashmole MS 1405: remedybook on ff. 1r–69v (*DMV *RemAshm.1405*; eVK2 6040.00, 6106.00, 6204.50, 6514.00)

Bodleian Library, Ashmole MS 1505: academic treatise, *Lilium medicinae*, by Bernard of Gordon on ff. 4r–244v (*DMV *BernLilium*; eVK2 1595.00, 3857.00)

Cambridge, Gonville & Caius MS 176/97: academic treatise written by a so far unidentified “Austin” on pp. 37–228 (*DMV *Austin*; eVK2 1143.00, 1547.00, 1918.00, 7795.00)

Cambridge, Jesus College MS Q.G.23: surgical treatise by Guy de Chauliac on ff. 1r–408v (*DMV *Chauliac(2)*; eVK2 0411.00, 1836.00, 7644.00)

Cambridge, Magdalene College, Pepys Library MS 1661: surgical treatise by Theodoric of Cervia (attributed to both Theodoric and Lanfrank on p. 198), followed by a collection of medicinal recipes beginning with “a confeccion of Theodoricus sope”, on pp. 35–211 (*DMV *Theodoric(1)*; eVK2 3849.00, 6142.00, 6641.00)

Cambridge, Peterhouse MS 118, I: surgical treatise drawing on Henri de Mondeville and other sources on ff. 1r–169v (*DMV *Mondeville(a)*; eVK2 2253.00, 3580.00)

Cambridge, Peterhouse MS 118, II: antidotary of Henri de Mondeville's surgical treatise on ff. 1r–65v (*DMV *Mondeville(b)*; eVK2 0946.00, 2300.00, 6044.50) & a surgery based partly on Mondeville on ff. 82r–173v (eVK2 0493.00)
Cambridge, Trinity College MS O.8.35: remedybook on ff. 1r–124r (*DMV *RemTrin.O.8.35*; eVK2 0672.00, 1966.00, 2194.00, 3137.00, 3211.00, 3572.00, 4684.00, 6541.00, 6894.00, 7534.00, 8042.00, 8173.00)
Cambridge, Trinity College MS R.14.51: remedybook on ff. 1r–47r (*DMV *RemTrin.R.14.51*; eVK2 0353.00, 0608.00, 1045.00, 1333.00, 2679.00, 2837.00, 2905.00, 2906.00, 2995.00, 6925.00, 7651.00, 7689.00, 7833.00, 8153.00)
Cambridge, University Library MS Dd.10.44: academic treatise, *Practica brevis*, by Platearius on ff. 1r–100v (*DMV *PlatearPra*; eVK2 1536.00, 2502.00)
New York Academy of Medicine MS 12: surgical treatise by Guy de Chauliac on ff. 1r–181v (eVK2 1878.00, 1959.00, 3637.00, 7602.00)

Primary sources: editions and early printed books

Boorde, Andrew 1971 [1547]. *The Breuiary of Helthe*. A facsimile edition. Amsterdam: Theatrvm Orbis Terrarvm [London: W. Myddelton]. (*DMV BoordeBreu*; *STC* 3373.5)
Braunschweig, Hieronymus 1525. *The Noble Experyence of the Vertuous Handywarke of Surgeri*. London: P. Treueris. (*DMV BraunSurg*; *STC* 13434)
Dawson, Warren R. (ed.) 1934. *A Leechbook or Collection of Medical Recipes of the Fifteenth Century*. London: Macmillan. (*DMV RemLMSoc.136*; eVK2 2015.00, 2210.00, 2962.00, 3490.00, 6109.00, 7342.00)
Elyot, Sir Thomas 1937 [1541]. *The Castel of Helthe*. A facsimile edition. 1st ed. (now lost) 1534. New York: Scholars' Facsimiles and Reprints [London: T. Bertheleti]. (*DMV ElyotCastel*; *STC* 7644)
Fleischhacker, Robert von (ed.) 1894. *Lanfrank's "Science of Cirurgie."* Early English Text Society, Original Series 102. London: Kegan Paul, Trench, Trübner. (*DMV Lanfrank(1)* & *Lanfrank(2)*; eVK2 0835.00, 6802.00 & 0836.00, 6802.00)
Getz, Faye M. (ed.) 1991. *Healing and Society in Medieval England: A Middle English Translation of the Pharmaceutical Writings of Gilbertus Anglicus*. Wisconsin Publications in the History of Science and Medicine 8. Madison: University of Wisconsin Press. (*DMV GilbAngl(1)*; eVK2 0435.00)
Grothé, Richard (ed.) 1982. Le ms. Wellcome 564: deux traités de chirurgie en moyen-anglais. PhD dissertation, University of Montreal. (*DMV LondonSurg*; eVK2 0936.00, 2544.00, 6744.00, 6865.00)
Heinrich, Fritz (ed.) 1896. *Ein mittelenglisches Medizinbuch*. Halle: M. Niemeyer. (*DMV RemAdd.33996*; eVK2 2474.00, 6007.00)
John XXI (pope) 1550? *The Treasury of Healthe*. Translated from the Latin by Humfrey Lloyde, with additions. London: W. Coplande. (*DMV JohnXXITre*; *STC* 14651.5)
Langton, Christopher 1545? *An Introduction into Phisycke*. London: E. Whytchurche. (*DMV LangtIntr*; *STC* 15204)
Macgill, Elizabeth R. (ed.) 1990. *"This Booke of Sovereigne Medicines...Collected of Maister Doctour Ffecknam Late Abbott of Westmynster...": An Edition of Folger Ms. V.b.129 (ca. 1570)*. Ann Arbor: University of Michigan.
McVaugh, Michael R. (ed.) 1997. *Guigonis de Caulhiaco (Guy de Chauliac) Inventarium sive Chirurgia Magna*. Vol. 1, *Text*. Studies in Ancient Medicine 14 (1). Leiden: E. J. Brill. https://doi.org/10.1163/9789004377394

Müller, Gottfried (ed.) 1929. *Aus mittelenglischen Medizintexten die Prosarezepte des stockholmer Miszellankodex X.90*. Kölner anglistische Arbeiten 10. Leipzig: B. Tauchnitz. (*DMV StockhRec*; eVK2 1600.00, 5039.00, 5184.00, 5207.00, 5255.00, 5355.00, 5688.00, 5807.00, 6333.00, 6560.00)
Ogden, Margaret S. (ed.) 1969 [1938]. *The "Liber de Diversis Medicinis" in the Thornton Manuscript (MS. Lincoln Cathedral A.5.2).* Early English Text Society, Original Series 207. London: Oxford University Press. (*DMV LiverDivMd*; eVK2 1840.00, 2293.00, 6464.00, 7954.00)
Ogden, Margaret S. (ed.) 1971. *The Cyrurgie of Guy de Chauliac*. Vol. 1, *Text*. Early English Text Society, Original Series 265. London: Oxford University Press. (*DMV Chauliac(1)*; eVK2 0757.00, 1828.00, 3615.00, 7661.00)
Roesslin, Eucharius 1545. *The Byrth of Mankynde*. Corrected and augmented by Thomas Raynalde, physician. 1st ed., translated from the Latin by Richard Jonas, 1540. London: T. Ray[nalde]. (*DMV RoessByrth*; *STC* 21154)
Schöffler, Herbert 1919. *Beiträge zur mittelenglischen Medizinliteratur*. Sächsische Forschungsinstitute in Leipzig III, Anglistische Abteilung, Heft 1. Halle: M. Niemeyer. (*DMV BurgPract*; eVK2 0723.00)
Seymour, Michael C. *et al.* (eds.) 1975. *On the Properties of Things: John Trevisa's Translation of Bartholomæus Anglicus De Proprietatibus Rerum*. 2 vols. Oxford: Clarendon Press. https://doi.org/10.1093/actrade/9780198836834.book.1 (vol. 1), https://doi.org/10.1093/actrade/9780198837626.book.1 (vol. 2) (*DMV BarthPropr*; eVK2 1443.00)
Vigo, Joannes de 1968 [1543]. *The Most Excellent Workes of Chirurgerye*. Translated from the Latin by Bartholomew Traheron. A facsimile edition. Amsterdam: Theatrvm Orbis Terrarvm [London: E. Whitechurch]. (*DMV VigoChir*; *STC* 24720)

Secondary sources

Bullough, Vern L. 1959. Training of the nonuniversity-educated medical practitioners in the Later Middle Ages. *Journal of the History of Medicine and Allied Sciences* 14 (10): 446–458. https://doi.org/10.1093/jhmas/XIV.10.446
Bullough, Vern L. 1966. *The Development of Medicine as a Profession*. Basel: S. Karger.
Cameron, Malcolm L. 1993. *Anglo-Saxon Medicine*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511518706
Cruse, Audrey 2004. *Roman Medicine*. Stroud: Tempus.
Demaitre, Luke 2013. *Medieval Medicine: The Art of Healing, from Head to Toe*. Santa Barbara: Praeger. https://doi.org/10.5040/9798400684494
Glare, Peter G. W. (ed.) 2012. *Oxford Latin Dictionary*. 2nd ed. Oxford: Oxford University Press.
Grieve, Margaret 1982 [1931]. *A Modern Herbal*. 2 vols. New York: Dover Publications [Harcourt, Brace & Company].
Hunt, Tony 1989. *Plant Names of Medieval England*. Cambridge: D. S. Brewer.
Hunt, Tony 1992. *The Medieval Surgery*. Woodbridge: Boydell.
Jackson, Ralph 1988. *Doctors and Diseases in the Roman Empire*. London: British Museum Press.
Jones, Peter M. 1998. Thomas Fayreford: an English fifteenth-century medical practitioner. *Medicine from the Black Death to the French Disease*, ed. by Roger

French, Jon Arrizabalaga, Andrew Cunningham, & Luis García-Ballester. 156–183. Aldershot: Ashgate. https://doi.org/10.4324/9780429202995-8

Kelly, Kate 2009. *The History of Medicine: The Middle Ages, 500–1450*. New York: Facts on File.

Kurath, Hans *et al.* (eds.) 1952–2001. *Middle English Dictionary*. Ann Arbor: University of Michigan Press. Rev. online ed. https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary

Maurer, Petra 2017. Faeces and the old sole of a shoe: remedies of the *Dreckapotheke*. *Asiatische Studien* 71 (4): 1247–1292. https://doi.org/10.1515/asia-2017-0050

Norri, Juhani 1992. *Names of Sicknesses in English, 1400–1550: An Exploration of the Lexical Field*. Annales Academiæ Scientiarum Fennicæ, Dissertationes Humanarum Litterarum 63. Helsinki: Academia Scientiarum Fennica.

Norri, Juhani 1998. *Names of Body Parts in English, 1400–1550*. Annales Academiæ Scientiarum Fennicæ, Humaniora 291. Helsinki: Academia Scientiarum Fennica.

Norri, Juhani 2004. Entrances and exits in English medical vocabulary, 1400–1550. *Medical and Scientific Writing in Late Medieval English*, ed. by Irma Taavitsainen & Päivi Pahta. 100–143. Cambridge: Cambridge University Press.

Norri, Juhani 2016. *Dictionary of Medical Vocabulary in English, 1375–1550: Body Parts, Sicknessis, Instruments, and Medicinal Preparations*. 2 vols. London: Routledge. https://doi.org/10.4324/9781315577081

Norri, Juhani 2017. Translation from Latin and French as a source of new medical terms in Late Medieval England. *Romance Philology* 71 (2): 563–622. https://doi.org/10.1484/J.RPH.5.114789

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

Pahta, Päivi & María José Carrillo Linares 2006. Translation strategies: *De spermate* and *De humana natura*. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS R.14.52, Its Texts, Language, and Scribe* (Medieval and Renaissance Texts and Studies 292, 2 vols), ed. by M. Teresa Tavormina. 95–117. Tempe: Arizona Center for Medieval and Renaissance Studies.

Pollard, Alfred W. & Gilbert R. Redgrave 1986–1991 [1926]. *A Short-Title Catalogue of Books Printed in England, Scotland, and Ireland and of English Books Printed Abroad 1475–1640*. 2nd rev. ed. by William A. Jackson, Frederic S. Ferguson, & Katharine Pantzer. 3 vols. London: The Bibliographical Society. [Oxford: Oxford University Press].

Porter, Roy (ed.) 1996. *The Cambridge Illustrated History of Medicine*. Cambridge: Cambridge University Press.

Rawcliffe, Carole 1995. *Medicine & Society in Later Medieval England*. Phoenix Mill: Alan Sutton.

Se Boyar, Gerald E. 1920. Bartholomaeus Anglicus and his encyclopaedia. *Journal of English and Germanic Philology* 19 (2): 168–189.

Siraisi, Nancy G. 1990. *Medieval & Early Renaissance Medicine: An Introduction to Knowledge and Practice*. Chicago/London: University of Chicago Press. https://doi.org/10.7208/chicago/9780226761312.001.0001

Ussery, Huling E. 1971. *Chaucer's Physician: Medicine and Literature in Fourteenth-Century England* (Tulane Studies in English 19). New Orleans: Tulane University.
Voigts, Linda E. 1982. Editing Middle English medical texts: needs and issues. *Editing Texts in the History of Science and Medicine*, ed. by Trevor H. Levere. 39–68. New York/London: Garland.
Voigts, Linda E. 2011. Fifteenth-century English banns advertising the services of an itinerant doctor. *Between Text and Patient: The Medical Enterprise in Medieval & Early Modern Europe*, ed. by Florence E. Glaze & Brian K. Nance. 245–277. Florence: Sismel Edizioni del Galluzzo.
Voigts, Linda E. & Patricia D. Kurtz 2014. *Scientific and Medical Writings in Old and Middle English: An Electronic Reference*. Rev. ed. https://cctr1.umkc.edu/search. First published 2000 as a CD-ROM (Ann Arbor: University of Michigan Press).
Wallner, Björn 1965. Drawings of surgical instruments in MS. Bibl. Nat. Angl. 25. *English Studies* 46 (1–6): 182–186. https://doi.org/10.1080/00138386508597232
Wallner, Björn 1991. A newly discovered Guy de Chauliac manuscript. *Notes and Queries* 38 (2): 159.

Medical terminology of human reproduction in Late Medieval and Early Modern English

Alpo Honkapohja
ORCID: 0000-0002-9724-5180
Carla Suhr
ORCID: 0000-0002-5730-7180

In this paper, we compare the use of medical terminology related to human reproduction in a selection of texts in Middle English Medical Texts (MEMT, 2005) and Early Modern English Medical Texts (EMEMT, 2010). The vernacularisation of English medical writing in the medieval and early modern periods was accompanied by a huge influx of new medical terms, many of which were synonymous, as well as changes in medical knowledge and scientific thought-styles in general. We focus on three subsets of terms used to refer to the womb, the genitalia and the afterbirth, and conduct a corpus-assisted discourse analysis of the contexts and functions of the various terms in the subsets in order to assess motives for selecting particular terms over others. Finally, we compare our findings across the two corpora to determine any long-term trends or changes that can tell us how specialised medical terminology developed over time.

Keywords: medical terminology, corpus-assisted discourse analysis, diachrony, human reproduction

https://doi.org/10.51814/ufy.1041.1458

1. Introduction

The vernacularisation of science, particularly of medicine, in England has long been a vibrant field of study. The beginnings of this long process were first investigated, in a historical linguistic framework, in scholarly editions of individual treatises (see, for example, Pahta 1998) and in lexicographical research on the increasing medical lexicon (see, especially, Norri 2016, which is largely based on his extensive earlier work, and McConchie 1997). The compilation of corpora specialising in medical discourse from Late Middle English to Late Modern English (Middle English Medical Texts (MEMT, 2005), Early Modern English Medical Texts (EMEMT, 2010) and Late Modern English Medical Texts (LMEMT, 2019)) made it possible to study the development of medical discourse in a long diachrony. Thus, we know that, in the late medieval period, scholastic traditions were still strong, and that medical thinking about reproduction relied heavily on classical authorities. However, as vernacular writing was not yet fully standardised, the practices of medical writing could vary considerably (see, for example, the chapters in Taavitsainen & Pahta (eds.) 2004). We also know that medieval traditions continued into the early modern period, though new discoveries in anatomy and physiology as well as social changes in disciplinary practices also influenced medical texts (see the chapters in Taavitsainen & Pahta (eds.) 2011). With the aid of historical corpora, it is now possible to investigate the development of specialised terminology in a more nuanced way, by going beyond first attestations of terms to systematic analyses of their contexts of use.

In this paper, we will focus on the terminology used in one specific discipline within medicine, human reproduction, which was also affected by the more general developments in medical writing (Pahta & Ratia 2010: 94). Using a corpus-assisted discourse analysis, we investigate diachronically the use of terminology in Middle English (ME) and Early Modern English (EModE) texts specialising on this topic. First, we examine the range of terms used in these texts and how their repertoire changes over time. Second, we suggest reasons behind the selection of particular terms over others. This kind of microstudy goes beyond charting the appearance and disappearance of new terms by examining why some terms are preferred over others and how they are used in different contexts.

2. Background

Childbirth is a topic that has been handled in medical writing since the earliest extant texts. Nearly a fifth of the corpus attributed to Hippocrates and written in the fifth and fourth centuries B.C.E focuses on the female body (Green 2000: 5). The majority of ME texts on the topic are very much in the tradition of scholasticism, ultimately descending from Greco-Roman medical practice, especially as synthesised by Galen of Pergamon.

Late-antique and medieval gynaecological literature, spanning from fourth to fifteenth century, was vast, comprising hundreds of texts, including specialised treatises but also integrated into broader medical literature. Information on childbirth can be found in specialised treatises on the topic as well as in general collections on medicine and surgery (cf. Green 2000). Among these, the most influential was *Gynaecia* by Muscio based on the writings of Soranus of Ephesus, which played a crucial role in disseminating ancient medical knowledge throughout the Middle Ages (Green 1992: 55). Another very influential body of work were the writings attributed to Trota of Salerno, likely written in twelfth-century Italy, which gained widespread influence due to its compatibility with the prevailing Galenic framework.

Prior to 1375, Latin and Anglo-Norman French were the predominant languages of written communication, and even after that, Latin remained the lingua franca of scientific and institutional discourse. Medicine was at the spearhead of English vernacularisation, but Latin, French and later Greek and other languages continued to coexist in medical writing through the medieval and early modern periods (Pahta 2011: 117). According to Voigts (1996; see also Norri 2004: 101), by 1475 the full range of Latin registers and genres was available in vernacular translation; by 1700, English was the dominant language of medical writing (Pahta & Taavitsainen 2004: 5). The language shift was accompanied by a shift in scientific thought-styles: the scholastic emphasis on classical medical knowledge was gradually giving way to empiricism. Medieval traditions continued into the early modern period, though new discoveries in anatomy and physiology as well as social changes in disciplinary practices also influenced medical texts on reproduction (Pahta & Ratia 2010: 94). There was also an expansion in the readership of these texts.

The question of the readership of ME gynaecological literature – whether predominantly male or female – is problematic and remains a subject of debate. As Green (1992: 56) puts it, "the evidence, both textual and codicological, suggests something like a tug-of-war between men and women for possession of these texts" (see also Green 2000; 2005; Green & Mooney 2006). Nonetheless, there was an expansion in written culture in the course of the fifteenth and sixteenth centuries, which also made gynaecological books more readily available to practitioners of different levels of education as well as lay readers (Pahta & Taavitsainen 2010: 4–5).

The vernacularisation of medical writing that started in the mid-fourteenth century required new vernacular medical terms. Previously, the sixteenth century had been identified as the period of significant lexical expansion in scientific and medical terminology (see, for example, Görlach 1991), but Norri (2004: 101) argues that the foundations for the expansion were laid already in the period 1375–1500. According to him, about 11% of ME terms for body parts were in use already in Old English, and a handful of new terms appeared in the twelfth and thirteenth centuries, but it is in late ME that the lexicon explodes, which, of course, coincides with the period of significant increase in medical treatises written in the vernacular (Norri 2004: 109).

According to Norri (2004: 111–129), new medical words were borrowed mainly from Latin and French (or both), but a new term could also be coined by modifying non-medical words, often as metaphors; the term *mother*, for example, was extended to both the womb and the cerebral meninges as a reflection of their protective functions. *Mother* is also an example of semantic borrowing from Latin (*matrix* for 'womb' and *mater* for 'cerebral meninx'). Other ways of introducing new terms were by metonymy, for example by a part of something representing the whole, or simply by everyday words such as 'fit' being taken into a particular medical use. Affixation and compounding as well as noun phrases (for example 'bleeding of the nose') were also important ways to increase the medical lexicon. The end result is a vast number of new terms created in various ways. Comparisons of different translations of the same Latin originals reveal that translators used different techniques for dealing with Latin terms, indicating differences in the individual lexical inventories of translators (Norri 2004: 111–129). It has been suggested that Latin borrowings for technical terms often only had one meaning and were therefore less ambiguous than native words, which tended to have multiple meanings (Nevalainen 2006: 54).

Interestingly, Norri also notes that almost half of the medical terms for body parts and illnesses found in ME became obsolete before 1550 (2004: 130). This concerned mainly terms borrowed unchanged from Latin texts, many of which were often found in a single treatise and therefore must have reflected individual translation choices. However, many native coinages were also lost (Norri 2004: 131). The reasons for the disappearance of particular terms are hard to identify, but phonetic reasons and homonymy have been suggested for some terms, as well as a tendency to reduce the number of alternatives when there are multiple complete synonyms; tautological or unclear terms also tended to disappear from the lexicon (Norri 2004: 131–137). What is evident is that the late ME period was a period of great lexical expansion, but that many of the new terms were not permanent, thus leading to continued lexical expansion in the early modern period.

Medical terminology has been found to be one of the key contexts for code-switching (Pahta 2004; 2011). According to Pahta (2004: 82), in ME especially the more learned texts often made use of the terms in the source text by borrowing or code-switching.[1] It is evident that these code-switches were not always used to fill in lexical gaps, as they were often glossed with vernacular synonyms. In ME texts, foreign terms are commonly accompanied by explanations, but this practice seems to decrease in the early modern period (Pahta 2011: 124). Pahta (2004: 83) finds that code-switched terms "contribute to the precision and specificity required of scientific discourse", but that they are also used for stylistic, prestige or didactic purposes to signal group membership and familiarity with the language of the particular discourse community. It is important to note, though, that no clear-cut trends of any kind regarding code-switching of medical terms have been observed for ME or EModE medical texts (Pahta 2004: 97; Pahta 2011: 125).

1 The line between borrowing and code-switching is notoriously fuzzy, especially in historical texts (see Pahta 2004: 78–80). Pahta's studies (2004; 2011) on code-switching focus on switches that can be considered unambiguous on the basis of formal or contextual cues such as the retention of original inflections or explicit flags; Norri (2004) does the same when distinguishing between Latin and anglicised term. We follow the same practice for identifying code-switching.

Glossing medical terms, with and without code-switching, is one strategy for defining them. Other (sometimes overlapping) strategies for defining terms include explaining them when there were no English equivalents categorising them by listing their characteristics or symptoms; relating them to the symptoms or other body parts; and by explaining their origins (McConchie & Curzan 2011: 83–88). Of course, medical terms are often used without defining them. McConchie & Curzan (2011: 91) note that, in the early modern period, practically-oriented surgeons engage in defining less than the classically-trained physicians.

Previous research on medical terminology in ME and EModE has focused mostly on lexicographical studies that chart the repertoire of medical terms (see especially Norri 2016 and McConchie 1997), and studies on medical terms as loci of code-switching (see especially Pahta 2004; 2011). However, Sylwanowicz (2018) has investigated the use and distribution of medical terms used with reference to medical preparations in ME and EModE recipes. She found that medical terms referring to medical preparations were rather uncommon in recipes, and that the terms used were more often general terms rather than specific ones, possibly because the focus was on the ailment to be cured rather than the preparation used to cure it. Many of the terms were adopted from Latin or French, especially in the more learned compilations (Sylwanowicz 2018). In the present diachronic study, we focus on another domain of medicine, childbirth. We analyse in detail the contexts of use and functions of a small selection of synonymous code-switched, borrowed and native terms to see if there are changes in the way the terms are used across time. The material and analytical method are explained in the next section.

3. Material and method

The material for this study is drawn from two corpora, *Middle English Medical Texts* (MEMT, 2005) and *Early Modern English Medical Texts* (EMEMT, 2010). The domain-specific corpora were carefully compiled to accurately represent the totality of medical writing in English between the years 1375 and 1700.[2] Texts are included in their entirety or in 10,000-word samples usually taken

2 For the compilation criteria of the corpora, see https://varieng.helsinki.fi/CoRD/corpora/CEEM/compilation.html.

from the beginning of longer texts. It should be noted that the corpus sampling method inevitably means that we may miss some instances of the terms we investigate.

For this study, we narrowed down the selection into texts focusing on gynaecology and childbirth in the two corpora, which provides us with continuity in terms of topic and register (see Table 1). From MEMT, we selected four texts focusing on childbirth from the perspectives of gynaecology and embryology. The selection includes two longer treatises: *De spermate* and *Sykenesse of wymmen*. It also includes two shorter texts, *De spermate hominis*, a poem on embryology, as well as *Nature of Wommen*, a very concise English translation of a Latin adaptation of Muscio's influential *Gynaecia*. These texts span a range of dates in the century which saw the emergence of medical writing in English vernacular. *Nature of Wommen* dates from the 14th century (see Green 1992: 84–85). In contrast, *De spermate* survives in the voluminous Trinity College Cambridge R. 14.52, dating from 1458–1468. Tavormina (2006) describes the manuscript as "an unusually ambitious instance of vernacularisation of medical and scientific writings in late medieval England" (xii). Consequently, the ME subcorpus exhibits a continuity of almost hundred years. The total word count of the ME material is 14,740 words.

EMEMT contains a subcategory of texts dealing with midwifery and children's diseases. Of the ten texts in the subcategory, seven texts describe reproductive theory and anatomy, particular complications of reproductive organs, or practical information on the generation and birth of children. They take the form of treatises for practitioners, handbooks for midwives and even self-help books for women. The remaining three texts deal with children's diseases. Human reproduction is also discussed in some surgical texts (EMEMT category 5), but they were left out of the analysis as they did not contain significant numbers of the terms we focus on. With a word count of 102,923 words, the EModE material is much larger than the ME material.

Table 1. The MEMT and EMEMT texts selected for analysis. The three EMEMT texts focusing on children's diseases are marked with an asterisk.

MEMT text	**Edition**	**Word count**
De spermate	MS: Trinity College Cambridge R.14.52, ff. 28r–36v. Pahta, Päivi 1998. *Medieval Embryology in the Vernacular: The Case of* De spermate. Mémoires de la Société Néophilologique de Helsinki 53, 161–255. Helsinki: Société Néophilologique de Helsinki.	8,063
De spermate hominis	MS: National Library of Scotland Advocate's 23.7.11, ff. 89v–90r. Hargreaves, H. 1977. *De spermate hominis:* A Middle English poem on human embryology. *Mediaeval Studies* 39: 507.	345
Nature of Wommen	MS: BL Egerton 827, ff. 28v–30v. Green, Monica H. 1992. Obstetrical and gynecological texts in Middle English. *Studies in The Age of Chaucer* 14: 85–88.	1,011
Sekenesse of Wymmen 2	MS: Yale Medical Library 47, ff. 65v–71v. Transcribed from the manuscript by Francisco Alonso Almeida. By permission of Yale University, Harvey Cushing/John Hay Whitney Medical library.	5,321
EMEMT text	**Edition**	**Word count**
Roesslin	Roesslin, Eucharius. *The Byrth of Mankynde…* London: T[homas] R[aynald], 1540. STC 21153.	9,511
Phayer*	Phayer, Thomas. *The KEGIment of life, whervnto is added a treatyse of the Pestilence, with the booke of children newly corrected and enlarged by T. Phayer*. Facsimile edition 1969. Amsterdam/New York: Theatrum Orbis Terrarum and Da Capo Press. (Originally London: Edward Whitchurch, 1546). STC 11969.	10,260
Jorden	Jorden, Edward. *A BRIEFE DISCOVRSE OF A DISEASE CALLED THE Suffocation of the Mother…* London: Iohn Windet, 1603. STC 14790.	11,633
Guillemeau, *Child-birth*	Guillemeau, Jacques. *CHILD-BIRTH OR, THE HAPPY DELIVERIE OF VVOMEN… To which is*	10,593

	added, a Treatise of the diseases of Infants, and young Children: with the Cure of them… Facsimile edition 1972. Amsterdam/New York: Theatrum Orbis Terrarum and Da Capo Press. (Originally London: A. Hatfield, 1603). STC 12496.	
Guillemeau, *Nvrsing**	Guillemeau, Jacques. *CHILD-BIRTH OR, THE HAPPY DELIVERIE OF VVOMEN… To which is added, a Treatise of the diseases of Infants, and young Children: with the Cure of them…* Facsimile edition 1972. Amsterdam/New York: Theatrum Orbis Terrarum and Da Capo Press. (Originally London: A. Hatfield, 1603). STC 12496.	10,269
Sadler	Sadler, John. *THE SICKE VVOMANS PRIVATE LOOKING-GLASSE…* Facsimile edition 1977. Amsterdam/New York: Theatrum Orbis Terrarum and Da Capo Press. (Originally London: Printed by Anne Griffin, for Philemon Stephens, and Christopher Meridith, 1636). STC 21544.	10,680
Culpeper	Culpaper, Nicholas. *A DIRECTORY FOR MIDWIVES…* London: Printed by Peter Cole, 1651. Wing C7488.	9,355
Pemell*	Pemell, Robert. *De morbis puerorum, OR, A TREATISE OF The Diseases of Children…* London: printed by J. Legatt, for Philemon Stephens. Wing P1132.	10,129
Sharp	Sharp, Jane. *THE MIDWIVES BOOK…* London: Simon Miller, 1671. Wing S2969B.	10,025
Aristotle	*ARISTOTELES MASTER-PIECE Or The Secrets of Generation displayed in all the parts thereof…* London: Printed for J. How, 1684. Wing A3689.	10,468

We started by looking at the word lists of the corpora created by AntConc software (Anthony 2023) and identifying potential medical terms, which were then checked using KWIC (Key Word in Context) concordances. From the extensive list of concordances we selected three subsets of synonymous or near-synonymous terms that recur across multiple texts. Each subset contains both foreign-based and native terms. Though not all terms in each subset are found in all texts, at least some of them appear in ME as well as EModE. The three subsets of terms are as follows:

1. Terms referring to the womb: *womb, matrice/matrix, moder/mother* (excluding the sense of a parent), *uterus*
2. Terms referring to genitalia: *privy member/thing/place/passage/ entrance, private parts, privities, genitalia*
3. Terms referring to afterbirth: *afterbirth, afterburden, secondine/ secundine, second birth, latter birth*

The *Middle English Dictionary* (hence MED), *Oxford English Dictionary* (OED), Juhani Norri's *Dictionary of Medical Vocabulary in English, 1375–1550* (2016) and the *Historical Thesaurus of English* (HTE) were consulted for information on meanings, etymologies and first attestations and to find other potential terms for the subsets.

The KWIC concordances give us numerical data about the frequencies of each term in the two corpora. However, as the frequencies are often quite low and the sizes of the two data sets are very different, we use the quantitative data to compare frequencies of terms within each subset of terms in the two data sets. Diachronic comparisons of subsets of terms consider the relative ranking of the alternative terms as well as the appearance and disappearance of terms.

The quantitative analysis is used to support a qualitative discourse analysis, which focuses on the contexts of use and the functions of each term in the subsets. The terms were analysed by close reading of the entire passages in order to find out if we could assign a motive for selecting a particular term from the subset of alternatives. Possible motives could be related to context – for example, a theoretical versus practical passage, a definition of a term versus undefined use in running text, a translation versus original text, the anticipated readership – or function, for example as a discourse-organising heading. With this information, we can consider how vernacularisation and other contemporary developments in medical writing are reflected in the use of our three subsets of medical terms.

4. Results

We start by describing our findings on the three subsets of terms in ME, followed by EModE. The paper then discusses each term, presenting the token count, discussing the term's relevance in medical discourse and illustrating it by examples. Comparisons across time and the larger implications of our results are discussed in the final section of this paper.

4.1 Middle English

Table 2 shows the distribution of all terms across this subcorpus. The terms are divided into three groups, the womb, genitalia and afterbirth, although the semantic field of ME words for 'womb' partly overlaps with genitalia. Table 2 shows the raw token counts of these terms.

Table 2. Token counts for subsets 'womb', 'genitalia' and 'afterbirth' in MEMT.

Terms in subset 'womb'	**Instances**
womb	30
matrice/matrix	31
mother/moder	69
uterus	-
Terms in subset 'genitalia'	**Instances**
privities, privy part(s)/place/ passage/entrance, private parts	1
privy member	15
Terms in subset 'afterbirth'	**Instances**
secondine	5
second birth/latter birth	-
after-birth/after-burden	-

Moder

The most frequent word used in the context of childbirth is *moder* (which is attested with the spellings *moder*, *moders*, *modur*). Norri (2004: 116–117) classifies the term as metonymic, as the protective meaning of the word was extended to the female reproductive system, but also notes that the word is a semantic borrowing from Latin, which could stand both for *matrix* 'womb' and *mater* 'cerebral meninx'. He (2016: 697) cites 1425(1389) as the earliest attestation with the meaning 'womb'. The extension of the semantic field to both womb and genitalia is also evident from the MED definition.

> 6a.
> *Anat.* and *med.* The uterus of a woman, womb; also *fig.*; the reproductive organs or genitalia of a female animal; causes (passiounes) of the ~, diseases of the uterus; risinge of the ~, suffocacioun of the ~, a disturbance or disorder of the uterus.

A total of 147 instances of *moder* (74 in *Sekenesse of Wymmen*, 72 in *De spermate*, 1 in *De spermate hominis*) are found in the MEMT subcorpus. However, this also includes the meaning 'human parent', which is the exclusive one in *De spermate*. In this treatise, it often appears in discussions about the effects of the timing of conception and the humoral complexions of both parents, which affect the child. This usage is illustrated in example 1.

> **1** the chield forsoth be conceyved in the hour of sangwyn, the fader and moder in houres of blac coler, the more liefully he reteyneth the figure of fader and moder and of the nature of rede coler than of the nature of sanguyn, only he reteigneth the propre passioun from his bigynnyng, that is to say of bloode. (*De spermate*: f. 34r)

When all the tokens with reference to 'human parent' are excluded, *moder* still remains the most frequent childbirth term in the corpus. However, this usage is concentrated into a single text. All 69 tokens that refer to the female reproductive system are found in the *Sekenesse of Wymmen*, in which this meaning is the predominant one (there are also five tokens with identical spelling, in which the reference is to 'human parent'). The fact that the usage also encompasses genitalia is made explicit by a Latin gloss, given in example 2, which specifies that the Latin term for the 'mouth of the moder' is *vulva*.

> **2** Ffor aposteme yn þe moder.
> The posteme yn þe moder bredeþe yn dyuerse partyes oþer whiles in þe mouþe of þe moder, id est in vulua, oþer whyles with yn forther. (*Sekenesse of Wymmen*: f. 68v)

This example is also notable for another reason, as it contains a unique example within the present corpus of the Latin formula *.i.* 'i(d est)'", the earlier variant of *i.e.*, which is common in many other medieval medical translations" (Pahta & Carillo Linares 2006: 104). It is, however, the only instance of this formula in the present subcorpus. The terms which we included in the present study are, by and large, not glossed by equivalent Latin terms. The appearance of the bilingual *.i.* formula here stands as an exception rather than the norm.

Code-switching can, however, be found in textual organisation (cf. Pahta 2004: 90–91). Notably, the Latin word *matricis* (*matrix*, gen.) is used twice in section headings (which are in Latin) in the *Sekenesse of Wymmen*. It is in both cases referred to as *of þe moder* in the running text, as shown in example 3 (see also Pahta 2004: 91). Consequently, a comparison between the two longer texts, *De spermate* and *Sekenesse of Wymmen*, reveals two different translation approaches to the same subject matter.

> **3** Suffocacio matricis.
> Suffocacyon of þe moder is when a womans hert her lyghtes ben þrust to gedur be rysyng of þe moder so þat þei lese hir breþynge & men calle it a cardiacle, for it ys a greuaunce of þe hert. (*Sekenesse of Wymmen*: f. 65v)

As a result, an intriguing asymmetry emerges. While *moder* is the most frequent term for referencing the female reproductive system in the present corpus, its usage is confined to a single text: *Sekenesse of Wymmen*. In contrast, in *De spermate*, the word is used exclusively with the meaning 'human parent', while the reproductive system is denoted as *matris*. As both of these texts were copied in the fifteenth century, after the earliest attestation of *moder* meaning 'womb' (cf. Norri 2016: 697), this discrepancy stems from different translation strategies employed within each text. The variability in the translation strategies of ME medical translators is something which has also been noted by Norri (2004: 129) and Jones (1989). The task of creating whole new medical vocabulary in the vernacular was considerable.

The other two texts are considerably shorter. *Nature of Wommen* does not contain a single token of *moder*, whereas the fourth text, *De spermate hominis*, contains one instance of *moder* (as evidenced in example 4).

4 The womman resseiuithe þe sede of man anone Wiþin þe moder thei ben medillid togedre ful sone Þe tone is hote þe toþer is colde (*De spermate hominis*: f. 89v)

The word *moder* is used in a passage which explains how male and female seed commingle within the mother, which most likely refers to the female reproductive system. However, it could arguably also be understood as referring to the human parent.

Womb

The word 'womb' is found in all four of the texts (*De spermate*, 5 tokens; *De spermate hominis*, 1 token; *Nature or Wommen*, 3 tokens; *Sekenesse of Wymmen*, 21 tokens). This is a native Germanic word, which can be traced back to Old English *wamb*, *womb* and further to proto-Germanic **wambo*. It is the earliest word attested in the semantic fields included in the present study, as the metonymic extension/loan translation of *moder* to mean 'womb' seems to occur only in the fifteenth century. It is thus an example of what Norri (2004) means when he observes that the "Old English stratum contains a considerably larger number of terms for body parts (11 per cent) than for sicknesses" (108–109). HTE shows *womb* to be the only word in the category *female sex organs: womb* which is in continuous use from OE to the present day. Nonetheless, the semantic field of the term in ME was broader than its present-day meaning. This is evident from MED meaning 5a.

> 5a.
> (a) The human uterus, womb; also *fig.*; specif. the womb of the Virgin Mary; also, the vaginal canal, vagina [quot. a1450 *Diseases Women(2)*]; ~ gate, the vulva; fruit (wastme) of ~, offspring; (b) the uterus of an animal or a fowl; (c) *fig.* ?a progenitor; (d) ?transl. of L Alva, apparently the name of a star in the constellation Sagittarius; (e) in asseverations.

In the MEMT subcorpus, the term *womb* – in contrast to *moder* – is often used with meanings related to the development of the fetus in the womb. In the Pseudo-Galenic *De spermate* it is used referring to the development of the fetus discussed in the astro-medical, humoral theory framework, which is central to this work (example 5).

5 Therfor of so moche of his agilite he hath and of humours, whan he bigynneþ in the wombe of humours to transferre the propirtes of planetis and signes and to lotte the propirtes with variable workes. (*De spermate*: f. 32r)

In other instances, the reference is to the child moving in the womb (example 6).

6 It is to be wist and knowen that bi 3 monethis past in þe man and iiij in the womman bigynne the veynes to issue to the constitucioun and makyng of the nailes. From than bigynneþ the chield to moeve in the wombe, and after that menstr[{ua{]t of bloode, whiche is made cotidianly of mete and drynke in the bodie of the Moder (*De spermate*: f. 29r)

The word *womb* is also used as a collocate with child in *Sekenesse of Wymmen*, specifically in discussions of stillbirths, as illustrated in example 7. This example contains several of the terms we focus on in this study, showing their typical meanings in this translator's usage. The reference to *hir priuy member* is with respect to 'administering a treatment to external genitalia'. The reference to *womb* is with reference to a (dead) child. The term *secondine* is used to refer to the afterbirth and the term *moder* to the womb as well.

7 Or elles temper borage with wyne & gif hir þat to drinke, or make a plastre of mugwort soden in watre from þe nauyll to hir priuy membre, ffor þat makeþe a woman sone to be delyuerd of child if it be ded in hir wombe and it draweþ oute þe secondyne. But let yt not ly to longe for yt woll drawe þe moder al oute. (*Sekenesse of Wymmen*: f. 70v)

The term *womb* is also often used to refer to sicknesses affecting the uterus, such as womb moving out of place (as in example 7). Additionally, in Old and Middle English the word could also be used to refer to the ‘stomach as the primary organ of digestion’. This is found once in *De spermate* (example 8).

> **8** Blac coler forsoth hath his power in the reynes, wherof rennyth bi al the body, of whos superfluite bien born passions in reynes, as the stone, and sperme and therof wastith, and ache in the wombe gendrith so that sumtyme a man may nat hold his vryne, and sumtyme he may nat sende it out. (*De spermate*: f. 30v)

There are also examples in which *womb* is clearly used to refer to ‘vagina’ or ‘the vaginal canal’ (cf. the MED definition).

> **9** Summe wyl fare so raþer þan somme but alle þo þat far so haue ouyr meche habundauns and we owen to loke to hem and teche hem gouernauns. Þei moste haue abstinens on day and hyr wombe moste be anoyntyd with good oyntement. (*Nature of Wommen*: f. 29v)

Example 9 from *Nature of Wommen* suggests applying a good ointment to the *wombe*, revealing a usage that seems to refer to the external entry part of the female reproductive system rather than the uterus.

Matris

The word *matris* occurs as 31 tokens with the spelling variants: *matryse*, *matryce*, *matrice*. The term shares a comparable semantic field as *moder* and *womb*, since it too can be used in a whole range of meanings within the female reproductive system. It differs from the previous two words by lacking the meaning ‘parent’ of *moder* (see example 1) as well as the meanings related to the digestive system of *womb* (example 8). The term can thus be said to be more purely in the medical register of gynaecology.

The word belongs to the group of words which entered the language in the ME period and which have dual origin in Latin and French/Anglo-Norman. Norri (2016) separates Latin (*matrix*) and Anglo-Norman/French forms (*matrice*) of the word into separate entries. The MED, on the other hand, groups both under the headword *mātrīs*. Interestingly, according to Norri (2016), the

Latin form is attested later (c. 1450) than the French form (c. 1375), which is supported by the present data.

All spellings in the present subcorpus seem to be derived from the French form with the exception of two tokens of the genitive *matricis* (from Lat. *matrix*, gen.), which are both found in Latin headings in *Sekenesse of Wymmen* (see example 3 above), and thus constitute code-switching used in textual organisation rather than borrowing. The French-influenced forms *matrice* and *matryce* are the preferred use in *De spermate* (example 10) and *Nature of Wommen* (example 11).

10 Forsoth if the matrice be long and strait, long and sclender shal the chield be. If it be short and straite, short and sclender shal the chield be. (*De spermate*: f. 29v)

11 Damyseles owen to be maydenes tyl þei haue purgyd and þe offycys of þe matryce mad clene. And þan þei arn clene of nature for to conceyue. (*Nature of Wommen*: f. 29r).

The use of *matryce* in the *Nature of Wommen* (example 11) is, actually, listed by Norri (2016) as the earliest attestation of the lexical item in English. It is found in a work with connections to continental France. The work is an abbreviated translation of a Latin Muscio-adaptation, which, according to Green (1992), likely “comes out of the university milieu of later-thirteenth-century Paris” (84–85). The only surviving copy of the Latin original survives in Paris, Bibliothèque Nationale, MS lat. 15081. Green notes that despite its short length, the ME summarising translation “encapsulates several major themes in mediaeval theories of female physiology and the processes of birth” (ibid.). It is a competent summary of a Latin text with French provenance.

If *Nature of Wommen* is indeed the earliest to borrow the term ‘matris’ to English, it is interesting to speculate about possible connection to continental French in this particular instance. As Norri (2004) notes, “[b]y the time our writers started their work, French influence on English medical vocabulary had already passed its peak” (112). Yet, this usage shows a sound-change typical of French (ibid: 111), in a translation of a work with likely French origin (albeit written in Latin). Nonetheless, with the datings, it has to be kept in mind that much material has been lost and remains unedited.

The distribution of forms of ‘matris' in the subcorpus, yet again, illustrates variability in translation strategies across the ME period. The two texts in which it is used are from opposite ends of the high period of vernacularisation, potentially separated by nearly a hundred years. The *Nature of Wommen* is dated to the fourteenth century by Green (1992) and to c. 1375 by Norri. The date of copying *De spermate* is as late as 1458–1468 (cf. Tavormina 2006). In contrast, *Sekenesse of Wymmen*, which originates in the first half of the fifteenth century, only uses *matrix* in Latin section headings and translates it to *moder* in the running text (see example 3 above).

Privy member, privy parties

The term ‘privy member' occurs 15 times in the ME corpus (spelling variants *pryuy*, *pryuey*, *priuy*, *priuey*, *membre*, *membris*). The form ‘privy parties' is attested by a single token in the *Nature of Wommen*.

The form ‘privy member' is restricted to a single text, the *Sekenesse of Wymmen*. It is a lengthy redaction of the gynaecological and obstetrical sections of Gilbertus Anglicus's *Compendium medicinae*, supplemented by material from other gynaecological treatises (see Green & Mooney 2006: 455–461). This was the most widely copied ME text of its type in the fifteenth century, and “more popular in England than even the Trotula texts” (Green 1992: 56). The work introduces itself as having been “composed so ‘that oo womman may help another in hir sikenes and nat discure hir privitees to … vncurteys men'” (Green & Mooney 2006: 455), yet it survives in manuscripts we know were copied and owned by men in the context of more general medical texts (ibid.).

The address to the reader and the use of euphemistic terms like ‘privy member' is typical of the so-called ‘Secrets of Women' tradition of gynaecological texts, which was prevalent in England and Germany in the later Middle Ages (cf. Green 2000: 11). The usage consists of translations from Latin expressions like *circa partes secretiores* ‘around more hidden parts' (Green 2000: 10). Green (2000: 13) also notes that the French tradition preferred terms were *marris* ‘womb', *naissance* ‘vagina [lit. ‘the birth']' and *flors* ‘menses [lit. ‘flowers']', which, while still euphemistic, were more direct than their English equivalents.

In *Sekenesse of Wymmen*, ‘privy member' is used to refer to female genitalia, when administering treatments (example 12, see also example 7 above).

In this case, the treatment is administered internally, as indicated by the preposition *yn to* ‘into’.

> **12** Poudre with oyle of pyliall & put yt in a smale soft lynnen cloþe and make yt of þe schappe of an ey. Þen put yt yn to hyr pryuy membre to let þe moder þat it fall not oute a ȝeyne. (*Sekenesse of Wymmen* 2: ff. 67v–68r)

Compared to the other terms used in the ME subcorpus, the semantic field of ‘privy member’ in the *Sekenesse of Wymmen* is well-defined. The term is exclusively used for the outer parts of female genitalia (the translator prefers *moder* for the internal parts, see examples 2 and 3). Nine out of fifteen tokens of *membre* have the feminine personal pronoun (*hir* or *hyr*) as a collocate (as in example 12). In the rest, the context makes it obvious the reference is to vulva or vagina. This usage contrasts with the considerable variety recorded elsewhere, including the EMEMT material in the present study (see below). Norri (2016) records several variants, including an expansion in meaning from vulva or anus attested ca. 1400 and penis in 1500. The single token of *preuy partees* in *Nature of Wommen* is found in an anatomical description of female genitalia, as consisting of *stonys* by the side of the *matryce* ‘by the which sperm is ministered by’ (example 13).

> **13** Off þe nature of wommen is for to speke. Fyrst we sal be gynne of þe preuy partees of hem þat is her stones be þe wheche spermat is mynystered by. Þe stonys sothly of wommen ben put on both partees of þe matryce and þo be mor round and smaler þan menys stones ben. (*Nature of Wommen* 2: f. 28v)

Jucker and Seiler (2023) have recently discussed the problematic word *queynte* appearing in Chaucer’s *Miller’s Tale*, arguing it to be a spelling variant of ‘cunt’ (58). Furthermore, based on the appearance of the word *cunt* in a range of genres, including glossaries and Lanfrank’s *Complete art of Surgery*, they argue that “ME *cunte* is used regularly in medical or scientific texts” (47). It has to be noted, however, that their analysis, which is based on surgical texts, contrasts with the absence of the word in the four ME gynaecological texts included in the present corpus. This suggests register variation within medical writing with surgical treatises opting for a different strategy, while the gynaecological treatises included in the present data are influenced by the 'Secrets of Women' tradition, in which the word is likely to have had some offensive connotations.

Secondine

The final anatomical word included in this selection is *secundine*. Norri (2016) defines it as "[f]etal membranes, placenta and umbilical cord; extruded from uterus after childbirth", and separates the French and Latin forms of the word (*secundine, secondine*). Similarly to *matris* (see above), the French form is older (c. 1375) than the Latin one (c. 1425). What is more, similarly to *matris*, the earliest attestation recorded by Norri is in the *Nature of Wommen* (example 14). HTE identifies *hala* and *cildhama* as earlier synonyms in the category 'placenta', but neither of them is used after Old English. Meanwhile, the term *afterbirth*, which is found in the EMEMT, was not attested before 1527. The OED lists *afterburden* as a possible ME alternative, but it is not attested in the MEMT material included in the present study either.

> **14** It falleth for a mydwyf to opyne þe mowth of þe matryse with hertys mary or with some oþir thyng þat is good þer fore þat it may lyghtly opene withoute querkenyng of þe schyld. Whan þe schyld is out it is for to gete out clene þe secundine. Þe secundine is a skyn þat is of þe veynes and of þe sennes of purpyl colour lyk to þe nauel of a schyld whan it is born. (*Nature of Wommen*: f. 30r)

The word *secundine* occurs only 5 times in the present subcorpus. It is found in the two texts quoted here. Four of the tokens are of French-influenced spellings (*secundine* 3, *secundyne* 1); one of a Latin variant (*secundina*). However, the Latin token is an instance of text-organising code-switching, since it occurs in a Latin heading (example 15).

This term differs from other ME words included in the present study by its clear explicit definition in both texts which use it. The term is not introduced by glosses. However, the Latin heading (example 15) makes the connection between Latin *secundina* and English *secondine* explicit. What is more, both the *Nature of Wommen* (example 14) and *Sekenesse of Wymmen* contain clear anatomical descriptions on how to identify it (example 15). The former also contains a reference to *a mydwyf*, who performs the delivery.

> **15** Secundina matricis. The secondyne ys a litull skynne þat is a boute þe childe in his moder wombe right as a þynn skynne, a cornell of a notte. And oþer whyles a woman ys delyuerd þer of when þat sche beryþe childe and oþer whiles yt a bydiþe still yn hir for febulnes of þe moder þat comeþe of mych fastynge, or of gret angur, or wrath, or of smytynge, or of longe flux of þe wombe, whech þinges often slethe a childe in þe moder wombe & þen þe moder delyuered hir of þe childe but þe secundyne a bydeþ styll. (*Sekenesse of Wymmen*: f. 71r)

4.2 Early Modern English

The selection of material for EModE is much larger than for ME, so the absolute numbers of instances in the two data sets are not comparable. What is evident from the quantitative analysis of terms in EMEMT (see Table 3) is that the repertoire of terms in each subset is larger than in MEMT. The qualitative analysis reveals that the distribution of terms within the data set is still uneven, with some texts dominating over others. There is also variation between authors in the selection and use of terms.

Womb, matrice/matrix, mother, uterus

Matrice/matrix, *mother* and *womb* are found in almost all of the texts, though in differing ratios. In the ten texts of the data set, the native Germanic *womb* is by far the most frequent choice in all but three texts. Eucharius Roesslin's *Byrth of Mankynde* (1540) and Jacques Guillemeau's *Child-birth* (1612) do not favour *womb* over *matrice/matrix* or *mother*. Roesslin uses *matrice* 21 times and *womb* 8 times (half of which are part of the construction 'mother's womb'), whereas Guillemeau uses the two terms more evenly, with 11 and 12 instances respectively. Both of these texts are translations, which probably contributes to the relatively higher frequency of *matrice/matrix* in comparison to other texts considered here. The third exception is Edward Jorden's *Suffocation of the Mother* (1603), a treatise that focuses on the condition in its title, so it is not surprising that the 'mother' mentioned in the condition's name is the most frequently (38 times) occurring term in his text. (Note that *moder*

Table 3. Token counts for subsets 'womb', 'genitalia' and 'afterbirth' in EMEMT.

Terms in subset 'womb'	**Instances**
womb	233
matrice/matrix	50
mother (excluding sense of parent)	54
uterus	6
Terms in subset 'genitalia'	**Instances**
privities	10
privy part(s)/place/passage/ entrance, private parts	20
genitals	9
Terms in subset 'afterbirth'	**Instances**
secondine/secundine	31
second birth/latter birth	9
after-birth/after-burden	5

rather than *matrice* is also used in the ME name for this condition, see example 3 above.) However, Jorden uses the Latinate term *matrix* 11 times and *womb* only 6 times, perhaps because the text is a learned treatise aimed at physicians rather than lay practitioners and therefore favours a more elevated style with Latin quotes and references as well as Latinate vocabulary.

There seem to be no distinct contexts for using *matrice/matrix* and *womb* in the EModE material. The two alternative terms are used quite interchangeably to discuss both theory and more practical issues such as the location and parts of the womb or various medical conditions pertaining to it. Though the two terms seem to be full synonyms, neither of them is in danger of disappearing, as often happened with new terms coined in late ME with full synonyms (see Norri 2004: 133–134). At the same time, the terms seem to have become restricted to the present-day meaning of ‘womb’, as there are no instances where the terms refer to the stomach or the vagina (cf. ME examples 8 and 9). In the following example 16 from John Sadler’s *Sicke VVomans Private Looking-glasse* (1636), both *womb* and *matrice/matrix* are used when advising on the procedure to cure a condition called “the descending or falling downe of the mother” (78).

> **16** When the inflation is dissipated; let the midwife annoint her hand with oyle of Mastick and reduce the wombe into its place. The matrice being put up, the situation of the patient must bee changed. (Sadler 1636: 82–83)

Sadler uses the term *matrice/matrix* alongside *womb* and *mother* more often than many of the other texts, but the example shows that he uses the three terms in similar contexts as full synonyms, and that he considers *matrice* an English word rather than an instance of code-switching. *Matrice/matrix* occurs as a Latin word in only three instances: in a Latin quote, in the Latin name of a medicine, and in a term specified as Latin.

Mother has the same meaning as *matrice/matrix* and *womb* (in addition to its meaning as “parent”), and it is occasionally used as an alternate for the other terms, but in the EModE texts it seems to have become more restricted in its context of use, thus avoiding extinction because of synonymity with the other two terms. This can be seen in the most common collocates of the term: “suffocation of the mother”, “fits of the mother” and “falling down of the mother”. The prevalence of these constructions is due in large part to Jorden’s text, which focuses on the suffocation of the mother and the fits that manifest the condition, and accounts for the majority of the tokens, but even in the other texts *mother* is rarely found outside these three constructions.

There are two particular discourse strategies found in historical medical discourse that encourage the co-occurrence of synonymous terms. The first one has to do with the structuring of texts in the classical order, according to which the exposition of an illness or body part begins by defining its name, often by glossing or providing a general explanation of the phenomenon. The definition is followed by descriptions of the causes, symptoms and cures for illnesses or the structure and functions of body parts.

Providing Latin, Greek or English equivalent terms for *womb* or particular illnesses related to the womb as a part of defining the term is not common in the material of this study, which aligns with the general findings of Pahta (2011). Only one author, Nicholas Culpeper, uses this kind of code-switching strategy systematically to start off the exposition on the anatomy and function of each male and female reproductive organ, including the womb. This is seen in example 17:

17 GAlen and Hippocrates, and most of the Greeks call the Womb μήτραν, and υστεραν, and some γαστέρα; and therefore the usual word the Septuagint gives for [great with child] is έν γαστρί εχειν, the Latins call it, Matrix, and Uterus; and we English, some follow the Latin word Matrix, but the only English word is, the Womb. (Culpeper 1651: 31)

Culpeper's decision to include all foreign terms for the English *womb* may have been motivated not only by the traditional discourse structure he followed, but also by his critical attitude towards the use of foreign terms by medical practitioners and a wish to reveal the secrets of literate terminology to his unlearned target audience. In his *Directory for Midwives* (1651: 6–7), he links the use of foreign terms with aims to increase one's status and to hide knowledge from commoners: "(we must borrow terms of other Nations that we may make the common people beleeve wonders, that so our selves may grow rich and proud, and keep folk in ignorance, though to their own undoing, and the loss of many a dear life)".

Only a few other authors engage in glossing as part of defining a medical term. Edward Jorden glosses the Latin terms for the condition known as suffocation of the mother at the beginning of his treatise-length description of the ailment (1603). His glosses also include *mother* as the name for the disorder: "the Mother, or the Suffocation of the Mother" (f. 5v). John Sadler describes the condition, too, in his manual for women (1636), but he omits the foreign terms from his definition, though he specifies that 'suffocation of the mother' is an English term and then provides an explanation of the term's origin (example 18). However, his description of the womb starts by indicating Greek terms (only) for the word and an explanation of the origin of the terms (example 19):

18 THis affect which simply considered is none, but the cause of an affect, is called in English the Suffocation of the Mother, not because the wombe is strangled, but for that it causeth the woman to bee choked. (Sadler 1636: 61)

19 First, touching the womb; of the Grecians it is called μύτρα the mother; or δυλφύς, saith Priscian, because it makes us all brothers. (Sadler 1636: p. 4)

The glosses of Culpeper, Jorden and Sadler ascribe the terms they define to a particular source language, perhaps as an appeal to authority. Jane Sharp's manual for midwives, *Midwives Book* (1671), also starts the description of the womb by providing the Latinate term together with the native word (example 20):

20 THe Matrix or Womb hath two parts, the great hollow part within, and the neck that leads to it, and it is a member made by Nature for propagation of children. (Sharp 1671: 34)

Contrary to the male authors, Sharp does not refer to *matrix* as a foreign term, and it is clear from her use of the term elsewhere in the text that the word is considered simply an English synonym for *womb*. Sharp's text is actually quite conspicuous in its absence of foreign terms, which are replaced by very concrete descriptions explained in layman's terms; she calls the womb, for instance, a "house for the infant to lie in" (35).

Another discourse strategy that promotes the co-occurrence of synonymous terms is the practice of using doublets as a signal of a more elaborate writing style. Like glossing, this practice is also connected to the more learned and traditional register of medical writing. In our sample of ten texts, only one text uses doublets consistently, and it is no surprise that it is the oldest of the texts. The translation of Eucharius Roesslin's *Byrth of Mankynde* is from 1540, but the original German was published in 1513. Roesslin regularly uses word pairs, usually consisting of a Latinate term and an English synonym, as seen in example 21, where *matrice* co-occurs with *womb*.

21 And sometyme the vesyke or bladder/ or other intralles beyng about the matrice or wombe be also apostumat & blystered/ whiche beynge

> greaued/ the matrice or wombe lykewyse is greaued with them/ and that hyndereth greatly the deliueraunce. Also sometyme in the fundament are emerrodes or pyles and other pusshes/ chappynges or chynnes which cause greate payne/ also hardnes and difficulte or byndynge of the belly/ which thinges for the grefe and payne that ensueth of them causeth the woman to haue lyttell power to help herselfe in her labor. (Roesslin 1540: fol. 14v)

Roesslin also uses the doublet ‘matrice or mother’ elsewhere in his text. What is notable in the above example is that not all doublets consist of medical terms: alongside ‘vesyke or bladder’ and ‘matrice or womb’ we see the more general doublets ‘hardnes and difficulte’ and ‘grefe and payne’. Thus it seems that in Roesslin’s case, doublets seem to be a stylistic choice rather than a gloss.

The final term in this subset is *uterus*, which, according to the OED, is first attested as an English word only in 1615. All 6 instances of this term appear in just two of our texts, and they are always flagged as Latin rather than English: 5 tokens appear in multi-word Latin terms or quotes, and the only single-word token is explicitly mentioned to be Latin (see example 17). It is therefore clear that *uterus* is not yet a part of the lexicons of the writers in our data sets. However, the other terms in this subset, *womb, matrice/matrix* and *mother*, are used synonymously, and the Latin-based *matrice/matrix* is considered a part of the English lexicon.

Privities, privy parts/passage/place/entrance, private parts, genitals

The ME *Sekenesse of Wymmen* refers to (female) genitalia with the term *privy member*, but in the EModE material there are no instances of this particular term. Instead, genitals are referred to with the terms *privities* (10 instances) and various constructions with the modifier *privy*: *privy parts* (11 instances), *privy passage* (6 instances), *privy place, privy entrance* and *private parts* (1 instance each). According to the OED, *privities* and *privy place* were in use already in ME, but *privy parts* and *private parts* only enter the English lexicon in 1533 and 1623, respectively. *Privy passage* and *privy entrance* are not found in the dictionary at all. *Genitals* (9 instances), the term borrowed from Anglo-Norman or French and ultimately deriving from classical Latin, was in use in English by the end of the fourteenth century.

All of the constructions formed from *private* or *privy* build on the idea of privacy. As such, they continue the 'secrets of women' tradition that stems from Late Medieval Latin treatises in England and Germany (see above and Green 2000: 11). Though the ME *privities* and EModE arrival *privy parts* appear most frequently in the texts analysed here (10 and 11 times, respectively), and both of them in four different texts, the number of variant forms suggests that the exact term is fluid as long as it includes the aspect of privacy. Even potentially idiosyncratic constructions like *privy entrance* and *privy passage* – the two terms not found in the MED or OED – are clear enough synonyms for *privy parts* or *privities* on the basis of context, even though *entrance* and *passage* are rather more specific in meaning than *parts*. The *privy* + noun construction is evidently quite productive, though not all constructions become established in the lexicon; Norri has noted this phenomenon already for ME (2004: 126–127). Honkapohja (2022: 81–82), on the other hand, notes that many EModE manuscript copies of John of Burgundy's medieval plague tract changed a reference to 'privy thing' either to 'privates' or 'privy members', suggesting that there may have been minute variations in this construction, which, with time, could become more or less acceptable in the register of medical writing.

As the rather low number of instances (39) shows, references to genitalia with these terms are not very common in the EModE material. Almost all of the references are found in just four texts: 13 instances in Roesslin's *Byrth of Mankynde* (1540), 7 instances in Sadler's *Sicke VVomans Private Looking-glasse* (1636), 8 instances in Culpeper's *Directory for Midwives* (1651), and 7 instances in Sharp's *Midwives Book* (1671).

The references to genitalia are mostly found in two contexts: to introduce anatomical descriptions of the reproductive organs, often as headings, as in example 22, or in discussions of treatments (example 23) or symptoms of ailments:

22 CHAP. 1. Of the Privy passage.
IN this I shal consider but these Seven following Parts. (Culpeper 1651: 27)

23 And when ye are thus bathed or washed/ then shall it be very conuenient for you to annoynte with the foresayde greses and oyles youre backe/ belly/ nauell/ sydes/ and such places as are nere to the

> preuye partes. Farthermore it shall be greatelye profitable for her to conueye inwarde in to the preuye parte these foresayde oyles or greses with a sponge or other thynge made for the purpose/ (Roesslin 1540: f. 19v)

The examples show clearly that *privy passage* and *preuye parte(s)* refer to a group of organs connected to reproduction, including both external and internal genitalia. According to Culpeper, the privy passage consists of seven parts (the lips, the nymphae or wings, the clitoris, the urinary tract, the caruncula myrtiformis, the hymen, and the neck of the womb), which he then proceeds to describe in detail. Culpeper thus effectively defines the term by relating it to other body parts, both the ones that are included under the term and the ones excluded from it, such as the womb. Culpeper's example is followed by Sharp, whose text is based on Culpeper. Roesslin's instructions about the use of oils to ease childbirth refers to both external and internal use of oils 'nere to the preuye partes' and 'inwarde in to the preuye parte', which also implies that the term is used collectively, though rather less exactly than in the first context. The rather vague use of *private/privy* terms found in Roesslin and also in Sadler when discussing treatments for ailments or symptoms of ailments suggests that they wished to retain an aspect of modesty or secrecy to female genitalia.

With one exception, the *privy/private* terms are never glossed. The one exception is the term *privy entrance*, only used once by Jane Sharp, which she glosses as 'or the neck of the womb', indicating a more restricted meaning than *privy part* for the term. The Latinate term *genitals*, which could have been a potential gloss for these terms, turns out not to be an exact match, which is probably why neither Culpeper nor Sharp uses it.

The Latinate term *genitals* (9 instances in total) is used in three texts, and it is never glossed. Culpeper uses the term exclusively as an umbrella term to cover all reproductive organs, male or female; for him, *genitals* seems to be a broader term than *privy parts*. In this broader sense, *genitals* is found in a text-organising flow-chart and the headings that introduce the sections where male and female reproductive organs are described in order. This practice bears a resemblance to the text-organising function of code-switched headings in ME medical texts (Pahta 2004: 90–92); the practice decreases before disappearing in the early modern period (Pahta 2011: 129). In Sadler, *genitals* is used twice in a very general sense for external genitalia, as in example 24:

> **24** The cause of this affect is suppression of the months, repletion of the whole body, immoderate use of Venus, often handling of the genitalls, difficult childe-birth, vehement agitation of the body, falls, blowes; to which also may bee added the use of sharp pessaries, whereby not seldome the wombe is inflamed. (Sadler 1636: 86–87)

Genitals – and on one occasion an adjectival construction *genital parts* – is also found in a similar broad sense in *Aristotle's Master-piece* (1684) to refer to both male and female reproductive organs as a whole. In this general sense, *genitals* could be understood to be a synonym for *private/privy* terms, but with a sense of clinical detachment rather than secrecy or modesty. Nonetheless, it is clear that in this subset of terms, the terms formed with the anglicised French *privy* and various nouns prevail over pure Latin or French borrowings.

Secundine, afterbirth

As is the case with *womb* and *privy parts*, the EModE texts make use of a range of synonyms for *afterbirth*. However, in the case of *afterbirth*, the Latin/French borrowing *secundine* is used more often (31 instances in three texts) than the various terms formed using the English words *birth* and *burden.* One possible reason for this is that the English terms are actually the newer terms: *afterbirth*, first used in 1527 probably as a variant of the fifteenth-century term *afterburden*, is used only a total of 5 times in two texts. *Second birth* is not recognised in MED or OED – though MED does record an example of *second burden* – but Norri (1998: 394) has identified a single instance of *second birth* in an early fifteenth-century surgical text. However, 8 instances of the term are found in Roesslin's 1540 text; he also has a single instance of the term *latter birth*, which is also not recognised by MED or OED. This is the only one of the three subsets of terms where a term borrowed from Latin or French dominates over native or partially native-formed terms.

The novelty of the English-based terms is perhaps evident in the fact that all instances of *afterbirth* and *afterburden* appear only as glosses of foreign terms or the anglicised term *secundine*. Roesslin glosses *secundine* with the native terms, whereas Culpeper glosses Greek and Latin terms with their three English terms (example 25):

> **25** THe Greeks call this τά δεύτερα, and τά ὑστερα; the Latins imitating them, call them Secundas, and Secundinas, and our Women, the Secundine, After-birth, and after-burden. (Culpeper 1651: 51)

Culpeper's glosses are followed by detailed descriptions and functions of the four parts of the *secundine* (the placenta and three different membranes, each of which only have a name borrowed from Latin). In one of these descriptions, *secundine* is glossed again with *afterbirth*. *Secundine* is used a further four times without glossing. Sadler, who glosses the term *womb* with its Greek equivalents, uses *secundine* twice without glossing or defining the term in any way.

Roesslin's use of terminology for *afterbirth* is very interesting, as he exhibits a lot of variation in his terms. He uses the term *secundine* 21 times, and in 7 instances he combines it with one or two other terms. His proclivity for couplets as a stylistic choice was noted with the term *womb*, and this may be an explanation for *secundine* as well. In example 26, the couplet/gloss is accompanied by a definition (note also the doublet "wrapped and conteyned"):

> **26** Farthermore when the secondine or seconde byrthe (in the which the byrth is wrapped and conteyned) doth ones appeare/ then maye ye knowe that the labor is at hande/ (Roesslin 1540: f. 21v)

Second birth is a term that seems to have been coined by Roesslin (or his translator), as it is not found in any other text and it is not found in MED or OED. The term is also used alone in the text, without any glossing. The single instance of *latter birth* is also unique to Roesslin; it is combined with *secundine*. Both instances of *after byrthe* are combined with *secondyne* and another term (*second birth* and *caul*) in a single passage (example 27):

> **27** Farther ye muste vnderstande that there be thre coueres or caules in the whiche the byrthe is contayned and lapped: of the whiche the one compassyth & embrasyth rownd aboute the byrthe/ and the

> other two caules also: and it is called the secondyne seconde byrth or the after byrthe: the whiche defendeth the byrthe frome noysum and yll humours encreasynge in the matryce after conception by retensyon of the flowres otherwyse wonte to passe and yssue furthe ones in the monethe/ the whiche yll humours yf they sholde touche or come nere to the byrth wolde greatlye peryshe and hurte the same. But after the deliueraunce of the pryncipall byrth these humours also with the foresayd caule or secondyne yssue furthe/ and is called the after byrthe. (Roesslin 1540: f. 12v)

In Roesslin's definition above, *secundine* seems to consist of just the first or outermost of the three membranes that surround the foetus in the womb – elsewhere he glosses *secundine* specifically as the *first caul* – whereas Culpeper includes the placenta and the two inner membranes as part of *secundine.* Culpeper's description of the four parts of the *secundine* indicates that medical authorities were not in agreement of the exact makeup of afterbirth, which is perhaps why Roesslin completely omits a description of the placenta from his text. In any case, it seems that medical terminology referring to the afterbirth seems to be still in flux in the early modern period.

5. Discussion & conclusion

Our study of ME and EModE medical terms related to the womb, genitalia and afterbirth has shown that the English lexicon of the time was very productive and had a high tolerance for synonyms derived from both native and foreign terms. Many new medical terms may have multiple meanings, which sometimes overlap, exhibiting origins by different word-formation strategies. Writers of vernacular texts made full use of the range of alternative terms, but especially in the earlier period, individual preferences can be seen in the choices of terms.

In ME, many of the terms in the three subsets cover a broader semantic field, and consequently they can refer to the uterus as well as to parts of or the entire internal and external genitalia or reproductive organs. In some cases, a term could refer to other parts of the body, too: *womb* could even refer to the navel, abdomen, bowels or stomach, cerebral ventricles or ventricles of the heart (Norri 2016: cf. *womb*). Not all of the terms in the subsets were in use in

ME, but nonetheless there are a number of potential alternatives in each subset. In the texts of this study, the selection and usage of terms in each subset appears to depend on translation strategies and to be specific to each text. Especially in the subset of terms referring to the womb, native terms seem to be preferred over the Latin/French borrowed term. This variability is particularly noticeable in the two longer texts.

By the early modern period, the ME terms have become more restricted to genitalia and/or reproductive organs, but they still exhibit some variation in their exact meanings, as exemplified by the differing understandings of the term *secundine*. *Mother* seems to have become the most restricted of the terms referring to *womb*, as it is usually found in very specific constructions that refer to particular ailments. But it is clear that many synonymous ME terms survive into the early modern period, contrary to what one could assume on the basis of general terminological developments (see Norri 2004).

New terms also appear, with especially the *privy* + noun construction showing continued productivity in the vein of the medieval "secrets of women" tradition, but some of them seem to be idiosyncratic coinages that are not adopted more largely into the lexicon. The newly borrowed Latinate term *genitalia* seems to be used in slightly different contexts to avoid overlapping with the *privy/private* terms. It is only in the subset of terms referring to afterbirth where the Latin borrowing is clearly preferred over native or anglicised terms, which is contrary to the trend noted by Sylwanowicz (2018) for terms referring to medicinal preparations in recipes. All in all, this study has shown that the medical terminology pertaining to our three subsets are still in flux both in terms of selections of terms as well as their meanings, and that with the exception of terms for afterbirth, borrowing terms from Latin was not the only or even preferred solution. Though much of this variation can be explained by the process of coining and establishing new medical terms necessary for vernacular writing, it may also reflect incomplete knowledge of human reproductive anatomy and reproductive processes.

When it comes to definitions of terminology, the ME texts mostly do not explicitly discuss and define the anatomical terms included in this study, contrary to Pahta's findings (2004: 83). The Latin gloss *vulua* for 'mouth of the mother' is a rare exception of a gloss (see example 2) and the only occurrence of the bilingual .i. glossing formula in the present corpus. Glossing (foreign or English) womb-terms in the EModE material is also not common, which matches with the observations by Pahta (2011: 124). Also McConchie & Curzan (2011: 77) note that glossing as a way of defining terms started to shift towards the modern 'lexical definition' in the early modern period. Definitions and/or explanations are found in only a few texts that continued the classical tradition for structuring treatises.

Our data set consisting of selected texts from the MEMT and EMEMT corpora has been quite limited in this study, and especially in the case of the EModE material it was evident that the corpus text samples were not always optimal for our purposes. Thus, extending the data set to full texts – and to texts outside EMEMT and available in the *Early English Books Online* database – would give us a more comprehensive view of the range of alternate terms and their different contexts of use. In the case of ME, extending the data set in any meaningful way is not possible due to the limited amount of material in existence.

In our study of ME and EModE gynaecological texts we have focused on three subsets of terms referring to the womb, to genitalia and to the afterbirth. Our aim was to identify particular contexts and functions that could explain the selection of individual terms from the subsets and to find out if there are any long-term diachronic trends. While we can see some long-term continuities and changes in the selection of terms, we could not identify any particular trends for the motives for these selections. Individual translators and authors made their choices on the basis of their individual preferences, which likely had to do with not only their personal lexicons but also with their expectations of their audiences, genre conventions and positioning amongst contemporary medical practitioners.

Bibliography

Anthony, Laurence 2023. AntConc (Version 4.2.4) Computer Software. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software.

Görlach, Manfred 1991. *Introduction to Early Modern English*. Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9781139166010

Green, Monica H. 1992. Obstetrical and gynecological texts in Middle English. *Studies in The Age of Chaucer* 14: 85–88. http://doi.org/10.1353/sac.1992.0002

Green, Monica H. 2000. From "diseases of women" to "secrets of women": The transformation of gynecological literature in the Later Middle Ages. *Journal of Medieval and Early Modern Studies* 30 (1): 5–39. http://doi.org/10.1215/10829636-30-1-5

Green, Monica H. 2005. Bodies, gender, health, disease: recent work on medieval women's medicine. *Sexuality and Culture in Medieval and Renaissance Europe*, ed. by Philip M. Soergel. 1–44. New York: AMS Press.
Green, Monica H. & Linne R. Mooney 2006. The sickness of women. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS R.14.52, Its Texts, Language, and Scribe* (Medieval and Renaissance Texts and Studies 292, 2 vols.), ed. by M. Teresa Tavormina. 455–568. Tempe: Arizona Center for Medieval and Renaissance Studies.
Hargreaves, H. 1977. *De spermate hominis:* A Middle English poem on human embryology. *Mediaeval Studies* 39: 507. http://doi.org/10.1484/J.MS.2.306846
Honkapohja, Alpo 2022. Tracing the Early Modern John of Burgundy. *Genre in English Medical Writing, 1500–1820*, ed. by Carla Suhr, Turo Hiltunen, & Irma Taavitsainen. 68–88. Cambridge: Cambridge University Press. http://doi.org/10.1017/9781009105347.006
HTE = The Historical Thesaurus of English. Available online at https://ht.ac.uk/ and https://www.oed.com/thesaurus.
Jones, Peter Murray 1989. Four Middle English translations of John of Arderne. *Latin and Vernacular: Studies in Late-Medieval Texts and Manuscripts*, ed. by A. J. Minnis. 61–89. Cambridge: Brewer.
Jucker, Andreas H. & Annina Seiler 2023. Translating Middle English (im)politeness: The case of Geoffrey Chaucer's *Miller's Tale*. *The Chaucer Review* 58 (1): 35–59. http://doi.org/10.5325/chaucerrev.58.1.0035
McConchie, Rod 1997. *Lexicography and Physicke: The Record of Sixteenth-century English Medical Terminology*. Oxford: Clarendon Press.
McConchie, Rod & Anne Curzan 2011. Defining in Early Modern English medical texts. *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen & Päivi Pahta. 74–93. Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511921193.006
MED = Middle English Dictionary. Available online at https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary.
OED = Oxford English Dictionary. Available online at https://www.oed.com/.
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 2004. Code-switching in medieval medical writing. *Medical and Scientific Writing in Late Medieval English*, ed. by Irma Taavitsainen & Päivi Pahta. 73–99. Cambridge: Cambridge University Press.
Pahta, Päivi 2011. Code-switching in Early Modern English medical writing. *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen & Päivi Pahta. 115–134. Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9780511921193.008
Pahta, Päivi & Maria Jose Carrillo Linares 2006. Translation strategies: *De spermate* and *De humana natura*. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS R.14.52, Its Texts, Language, and Scribe* (Medieval and Renaissance Texts and Studies 292, 2 vols.), ed. by M. Teresa Tavormina. 95–117. Tempe: Arizona Center for Medieval and Renaissance Studies.
Pahta, Päivi & Maura Ratia 2010. Category 2. Treatises on specific topics. *Early Modern English Medical Texts: Corpus Description and Studies*, ed. by Irma Taavitsainen & Päivi Pahta. 73–99. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/z.160.04cat2

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

Pahta, Päivi & Irma Taavitsainen 2010. An interdisciplinary approach to medical writing in Early Modern English. *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen & Päivi Pahta. 1–8. Cambridge: Cambridge University Press.

Nevalainen, Terttu 2006. *An Introduction to Early Modern English*. Edinburgh: Edinburgh University Press.

Norri, Juhani 2004. Entrance and exits in English medical vocabulary, 1400–1550. *Medical and Scientific Writing in Late Medieval English*, ed. by Irma Taavitsainen & Päivi Pahta. 100–143. Cambridge: Cambridge University Press.

Norri, Juhani 2016. *Dictionary of Medical Vocabulary in English, 1375–1550*. 2 vols. Abingdon: Ashgate. http://doi.org/10.4324/9781315577081

Roesslin, Eucharius 1545. *The Byrth of Mankynde, Otherwyse Named the Womans Booke Newly Set furth, Corrected and Augmented…By Thomas Raynold Phisition*. London: By Thomas Ray[nalde]. STC 21154.

Sylwanowicz, Marta 2018. Middle and Early Modern English medical recipes: Some notes on specialised terminology. *Anglica* 27 (2): 89–101. https://doi.org/10.7311/0860-5734.27.2.05

Taavitsainen, Irma & Päivi Pahta (eds.) 2004. *Medical and Scientific Writing in Late Medieval English*. Cambridge: Cambridge University Press.

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

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

Taavitsainen, Irma, Päivi Pahta, Turo Hiltunen, Martti Mäkinen, Ville Marttila, Maura Ratia, Carla Suhr, & Jukka Tyrkkö (eds.) 2010. *Early Modern English Medical Texts (EMEMT)*. CD-ROM. Amsterdam: John Benjamins.

Taavitsainen, Irma, Päivi Pahta, & Martti Mäkinen (eds.) 2005. *Middle English Medical Texts (MEMT)*. CD-ROM. Amsterdam: John Benjamins.

Tavormina, M. Teresa 2006. Preface. *Sex, Aging, & Death in a Medieval Medical Compendium: Trinity College Cambridge MS R.14.52, Its Texts, Language, and Scribe* (Medieval and Renaissance Texts and Studies 292, 2 vols.), ed. by M. Teresa Tavormina. xii–xxiii. Tempe: Arizona Center for Medieval and Renaissance Studies.

Voigts, Linda 1996. ‘What’s the word?’ Bilingualism in late-medieval England. *Speculum* 71: 813–826. http://doi.org/10.2307/2865721

“We could scarce distinguish one from another” Towards a phraseological perspective on modal auxiliaries in three categories of Late Modern English medical writing

Turo Hiltunen
ORCID: 0000-0003-1417-2703
Martti Mäkinen
ORCID: 0000-0002-6806-4217

The present study investigates the patterns of use of modal auxiliaries in three categories of medical writing in the eighteenth century. The data is drawn from the corpus of Late Modern English Medical Texts (LMEMT), and it is approached through phraseological register analysis. The study focuses on the distribution of modal auxiliaries in medical texts, per main verb, in active constructions, and on the division of auxiliaries between epistemic and root modalities (deontic and dynamic) in recipe texts, surgical texts, and medical periodicals. The study looks into certainty of knowledge and the normativity of medical writing as reflected in the use of modal auxiliaries. It is assumed that the three medical text categories attest to differing patterns of use of modal auxiliaries and that the variation is indicative of the communicative purpose of each category.

https://doi.org/10.51814/ufy.1041.1459

Keywords: modality, auxiliaries, medical writing, Late Modern English, phraseology, register analysis

1. Introduction

Modality is an inherent element in all scientific discourse, be it spoken or written. Among other things, different types of modality employed in text inform the audience of the author's level of certainty on the knowledge they disseminate, and of the levels of normativity in the discipline at hand in general. Different linguistic means are used to modify the information, answering questions such as, "What is true?", "What is plausible?" (epistemic modality), "What is necessary?", "What is permissible?" (root / deontic modality), and "What is possible or customary?", "What is the intention?" (root / dynamic modality) (Palmer 1990: 5–8; Coates 1995: 55; Marín-Arrese 2009: 30, 34).

Modality is often marked grammatically through the use of modal auxiliaries, and many studies on scientific and medical discourse have accordingly focused on them (e.g. Vihla 1999; Taavitsainen 2001). However, as far as historical medical writing is concerned, beyond studies investigating epistemic space in certain medical genres (Whitt 2023), modality in medical texts from Late Modern English (LModE) has received surprisingly little attention. In addition, we also argue that previous research focusing on this register has not sufficiently considered the contributions from phraseological and constructionist approaches, which posit that language users do not necessarily depart from the meaning of individual modal verbs when constructing their utterances, but rather make choices to use strings consisting of modal verbs *and* other linguistic material, which have precise and predictable interpretations that suit their communicative purposes (e.g. Sinclair 2004; Cappelle and Depraetere 2016).[1] The present study aims to take the first steps to address these issues in the context of 18th-century medical writing, using the corpus of *Late Modern English Medical Texts* (LMEMT) as the data. The perspective adopted here is thus an amalgam of phraseological research and register analysis: our main research question is: *How does the use of modal auxiliaries,*

1 We concur with McEnery and Hardie (2012: 210–11) in seeing "neo-Firthian" corpus linguistics and Construction Grammar as convergent approaches and treat them as compatible in the context of the present study, despite their differences.

in particular as regards phraseological patterning, reflect the purposes of writing of each sub-register of Late Modern English medical writing?

To achieve this, we will study the distribution of modal auxiliaries in medical texts, per main verb, in active constructions, and the division of auxiliaries between epistemic and root modalities in three LModE sub-registers, i.e. *Recipe Collections*, *Surgical texts*, and *Scientific periodicals* (on these categories, see further Taavitsainen & Hiltunen 2019: 279, 299, 317). Our focus lies on how the certainty of knowledge and norms are expressed in the medical sub-registers through the use of modal auxiliaries. We expect to find differing usage patterns of modals, and we believe that the observed variation correlates with the communicative purposes of the sub-registers.

The phraseological analysis is centred on the colligation between modal auxiliaries and lexical verbs, and the analytical measures employed are ATTRACTION and RELIANCE (Schmid 2000), which have been successfully used in the study of scientific genres in earlier work (e.g. Hiltunen 2021). Along with shedding light on the usage patterns characteristic of individual sub-registers, our chosen approach highlights a more general point in the study of modality: due to the polysemous nature of modals, the scrutiny of their immediate textual cotext is crucial to determining their meaning in use, which is the sum of the phraseological configuration in its entirety, including (but not limited to) the modal verb and the verb it modifies (cf. Coates 1995: 56; Dehouck & Denis 2023). The second important aim of the article is to test the applicability of our methodology in the study of modality, hence the exploratory nature of the study. This study will motivate the phraseological approach adopted and show its uses in the study of modal auxiliaries and modality in general.

2. Literature review

2.1 Modality in historical / modern medical and scientific writing

Modality is a well-researched phenomenon, and there are several relevant studies that converge with the present study. Vihla (1999), a corpus study of present-day medical writing, is possibly the closest match to our choice of approach, and Whitt (2023) uses data from LMEMT as we do. Studies on earlier or later phases of medical writing that focus on modality (even if not always on modal auxiliaries) include Whitt (2016a) (evidentiality in EModE medicine), Alonso-Almeida and Álvarez-Gil (2020) (modals in EmodE medical recipes), Alonso-Almeida (2015a) (modality and evidentiality in modern medical research papers), Alonso-Almeida (2015b) (epistemic modality and evidentiality in Early Modern English and Spanish medical writing), Kopaczyk (2013) (formulaic discourse in EModE medicine), Hiltunen and Tyrkkö (2011) (discursive practices in EModE medicine); Gray, Biber, and Hiltunen (2011) (stance expressions in EModE medicine), and Mäkinen (2022) (persuasive means in EModE medical recipes). Other studies on historical scientific registers outside the medical field that have informed the present study are Álvarez-Gil and García Alonso (2019) (*I think* in LModE scientific texts), Crespo-García (2011) (markers of persuasion in 18th century philosophical texts), and Whitt (2016b) (evidentiality and epistemology in EModE scientific discourse). Even if we draw on the same or similar pools of data or share the principles of semantic-pragmatic analysis of modals with many of the cited studies, our approach departs from them by letting the data-driven quantitative phraseological analysis inform our selection of items for qualitative analysis.

2.2 Modality and modal auxiliaries

Modality provides the means through which one can either assert or deny the possibility of utterances, or the permissibility and necessity of actions (Palmer 1990: 6). Modality in language can be expressed through many grammatical and lexical categories, e.g. adverbs, adjectives, full verbs, and modal auxiliary verbs. In this study, we focus on one facet of modality, modal auxiliaries, and we acknowledge that a quantitative approach alone will not fully capture the complexity of the phenomenon studied.

In the identification of modal auxiliaries, we lean on the formal criteria of the 'NICE' list of verb properties (Huddleston 1976: 333): modal auxiliaries 1) can be inverted with the subject, 2) can form a negative form with *-n't*, 3) can retrieve the meaning of a preceding lexical verb in an elliptical expression ('code'), and 4) can be used emphatically (Palmer 1990: 3–4). Any adjustments to the set of modals identified with the 'NICE' criteria will be discussed later in section 3.2.

As modality is a highly complex phenomenon, and therefore also widely studied, it is natural that the attempts to describe modality have created several, somewhat differing terminological domains, each expressing their theoretical underpinnings.[2] The current study adopts the general division of modality into epistemic and non-epistemic (root) modalities (there is a general agreement of this division, cf. Palmer 1990: 5–8; Sweetser 1990: 50–51; Coates 1995: 55–56; Warchał 2015: 74), and acknowledges deontic and dynamic modalities as the sub-categories of the latter. In the analysis of modals in medical registers, we occasionally resort to more fine-grained terminology, but overall, our aim is to describe the differences in modalities between the registers by allocating the modal auxiliary + main verb pairs to the three main classes of modality.

To define the different modalities we work with, EPISTEMIC MODALITY expresses how the speaker judges the truth-value of a claim or a proposition in terms of necessity, probability, or possibility (Palmer 1990: 2, 7; Sweetser 1990: 49). DEONTIC MODALITY refers to the necessity or obligation of future actions expressed by the speaker or author (Hoye 1997: 43). Finally, DYNAMIC MODALITY is related to the ability or potential of a person or a thing, and as such does not necessarily require attitude-expressing subjectivity from the speaker (Palmer 1990: 7, 10).

2.3 From frequency counts to phraseology

As was already noted, the present study sets out from the position that speakers create meanings not word by word, but in multi-word phrases: linguistic items sharing the same cotext are interdependent upon one another (Sinclair 2008: 409). In the context of modality this means that modality-expressing categories interact with the cotext: the sense of the modified proposition cannot, therefore, be derived from the mere presence or absence of a modality-expressing item. Rather, the modified sense is the sum total of all the linguistic categories of the proposition in interaction (Depraetere, Cappelle, & Hilpert 2023: 4–5). As the interaction of linguistic categories defines the sense of a modal expression, a natural conclusion is that modals are inherently polysemous: they can lend themselves to creating both epistemic and root senses. This versatility of modals also allows for effective and flexible ways of meaning making. Our data-driven approach focusing on identifying (on

---

2 For a lucid review of different theories of modality, see Warchał (2015: 71–92).

formal grounds) modal auxiliaries and the main verbs modified by them is designed to free us from pre-conceptions sometimes attached to modals.[3] The method allows us to observe the linguistic performance without interfering layers of interpretation.

3. Data and methods

3.1 Corpus

The data for this study comes from the *Late Modern English Medical Writing* corpus (LMEMT, Taavitsainen *et al.*, 2019), a representative corpus covering all the major registers in printed medical texts 1700–1800, which was POS-tagged for this study using CLAWS5 (C7 tagset). The corpus is divided into seven text categories, which represent different kinds of medical writing based on such factors as purpose and tradition of writing (see Taavitsainen & Hiltunen (2019) for more information on these categories). Of these, we chose three categories for analysis in this study: *Recipe collections* (REC), *Surgical and anatomical texts* (SUR), and *Scientific periodicals* (SC-PER). These categories were chosen as they represent maximally different types of writing:

- *Recipe collections* focus on the preparation and administration of remedies.
- *Surgical and anatomical texts* provide instructions on how to perform surgical operations.
- *Scientific periodicals*, which encompass two publications: *Philosophical Transactions of the Royal Society* (PT) and *Edinburgh Medical Journal*,

3 For example, even if the modal auxiliary *can* is most often associated with dynamic modality, it is also found in phrases that mitigate the asserted characteristics or ability (Carrió Pastor 2012: 123), as in *The weather in June **can** be nice*. The proposition can be paraphrased as *(In my experience,) on **some** days in June, the weather is nice*. The surface level of the proposition states the characteristics of weather in June; however, as the paraphrase involving a subjective element is possible, *can* may also introduce a sense of epistemic modality in the proposition, based on the speaker's prior knowledge.

represent the early stages in the history of the empirical research report.[4]

Given these differences and their concomitant effects at the textual level, which have been well-documented in earlier research, it is thus reasonable to expect that insofar as they give rise to differences in the use of modals (as predicted by our hypothesis), a corpus-based exploration will provide us data with which these can be described in detail.

3.2 Method of analysis

We begin the study with an exploratory approach assessing the frequencies of individual modal auxiliaries across the three categories of writing, following the approach used in several previous studies (e.g. Vihla 1999; Millar 2009; Leech 2014). This approach essentially consists of determining the frequencies of individual modal auxiliaries and comparing them across corpora, either synchronically (Vihla) or diachronically (Leech, Millar). As our primary focus is on sub-register differences, we do not consider possible diachronic change in the current study but adopt a synchronic perspective, treating the 18th century as a temporally undivided whole.

A limitation of this approach is that it does not take full account of the functions of clauses where modals occur, and hence the sense coefficient of the auxiliary and verb is lost. As an illustration of this, in (1) and (2) we see the modal auxiliary *must* used in an epistemic and dynamic clause, respectively. These senses can only be distinguished after analysing the clause in which the phrase SUBJECT + AUXILIARY VERB + VERB is embedded.

1 By a further Disquisition into this matter, we find that **the Doses must not only be** greater where the thickness of Blood is greater;

4 Note that *General periodicals* were not included in the analysis due to their communicative purpose that differs from that of *Scientific periodicals*. The category *General periodicals* in LMEMT, which encompasses articles from *The Gentleman's Magazine*, aims to represent how health matters were communicated to the general public, and was compiled following a different procedure (see Taavitsainen & Hiltunen 2019).

but that they must be encreas'd […] (REC, Cockburn, *The Practice of Purging and Vomiting Medicines*, 1705)

2 **You must put** two ounces of White Sugar-candy into each Bottle, and let the Water drop on it; […] (REC, Kettilby, *A Collection of Receipts in Cookery Physick and Surgery*, 1714)

In (1), *must* refers to the explanation (*disquisition*), making the sense expressed a logical conclusion of it. In other words, it is a clear example of epistemic necessity.[5] In (2), the same auxiliary conveys neutral or facultative obligation which makes the example a case of deontic modality, almost like prescription (Silk 2022: 204), again something that is made apparent only by the study of the cotext.[6] This being the case, quantitative analysis based on raw frequencies is problematic for two reasons, which are related to each other: first, the counts of individual wordforms may be altogether misleading, as they conflate different senses and different types of modality, as is evident in the case of *must*. Relatedly, they also fail to highlight what the characteristic uses of modals are in different textual contexts, yet it is precisely this type of information that would be relevant for describing how the use of modals is influenced by the conventions of registers and (sub-)registers.

For this reason, along with simple frequency analysis, the present study also looks into the phraseological patterning of modals.[7] The reasons for preferring a phraseological approach can be derived from a commonly agreed-on premise underlying much of previous work in phraseology and

5 Lyons (1977: 792) and Palmer (1990: 7) note that in natural language use epistemic modality is always to some extent subjective, and hence any logical conclusion is based on inference by the speaker: the emphasis is more on the plausibility than the verifiability of the proposition. Therefore, even if (1) belongs strictly speaking to alethic modalities – logically, the proposition is necessarily true (see Wright 1951; Palmer 1990: 6) – in literature on modality it is customary to discuss epistemic necessity in cases like this.

6 Regarding (2), one could also argue that it is "common sense" necessity that makes it a case of deontic modality (Álvarez-Gil & García Alonso 2019: 47). The ambiguity in modals illustrated by (1) and (2) can be solved through the study of the cotext; the pragmatics of aux + main verb phrases imply the sense potential of modal auxiliaries and motivates the study of modals in interaction with their cotext (Sweetser 1990: 49; Leclercq 2023: 60).

7 The analysis thus incorporates both 'text-linguistic' and 'variationist' perspectives, using the terms introduced in Biber (2012).

corpus linguistics: rather than selecting words individually, language users want to express complete meanings and choose the words to match those intended meanings (e.g. Sinclair 2004). It follows that to understand the specific modal meaning of an utterance, it is often necessary to analyse it in its entirety, rather than just the specific modal auxiliary (or another item expressing modality). This idea has been recently restated in the framework of Construction Grammar by Depraetere, Cappelle, and Hilpert (2023: 4–5): the study of modals should be extended to the adjacent lexical items and syntactic categories, combining the consideration of semantics with pragmatics.

An obvious challenge with this approach is that it is much more difficult to automate the analysis of complete utterances than count individual words.[8] Indeed, qualitative close reading might be the only way to ensure the correctness of such an analysis. For this reason, we opt for an approach that employs a semi-automatic method to look at phraseological features in the co-text of modal auxiliaries, focusing mainly on the **verb** with which the modal occurs. This analysis is complemented with a qualitative appraisal of the **subject** of the clause within which the modal verb is embedded. While these features are clearly not the only ones that have the potential to contribute to meaning, we argue that they are central in terms of clausal meaning and its relationship with modality, as they typically indicate what action was performed (verb), and by whom (subject). The influence of other possible factors (e.g. the object of transitive verbs) was left for further study.

We retrieved all the instances of modal auxiliaries from the three categories of *LMEMT* described above (using the tag **VM**). This query matched the following twelve wordforms:

1. *can*
2. *could*
3. *dare*
4. *may*
5. *might*
6. *must*
7. *need*
8. *ought*
9. *shall*
10. *should*
11. *will*
12. *would*

8 We would argue that this statement still applies to most corpus linguistic research scenarios, despite major advances that have recently taken place in natural language understanding, in particular after the introduction of large language models (LLMs) (cf. Yu *et al.*, 2024).

Of these, *dare* and *need* were highly infrequent and were excluded from further attention, as was the semi-modal *ought to*. On the auxiliary *will*, earlier literature is somewhat divided: it is formally a modal auxiliary but expresses modality only to some extent. This argument is often supported by the modal's participation in periphrastic future constructions, such as *It will rain tomorrow*, or the association of the modal with "high probability" and hence with "less-than-factual" situations (see the discussion in Depraetere and Cappelle 2023: 16–19). Nevertheless, the formal features and the 'NICE' criteria (see Section 2.2 above) define *will* as a modal auxiliary (Palmer 1990: 160), justifying its inclusion in the analysis.

After removing false positives from the dataset (caused by tagging errors), we manually identified for each line in the concordance the lexical verb with which the auxiliary is associated, as well as the subject of the clause where the VP is found. With verbs, it should be emphasised that we do not attempt to provide a comprehensive appraisal of the lexicogrammatical preferences of modals. Rather, our aim in this part of the analysis is a more exploratory one: we intend to identify MODAL AUXILIARY + VERB pairings that (1) stand out in terms of frequencies of use, and (2) are relevant to describing discourse meanings in that they can be linked to specific functions characteristic of early medical writing.

To investigate the co-occurrence tendencies between specific auxiliaries and lexical verbs across the registers in focus, we determined two metrics related to the modals and the verbs that accompany them: *attraction* and *reliance* (Schmid 2000; Hiltunen 2021). ATTRACTION refers to the degree to which a modal auxiliary 'attracts' a specific verb, and it is calculated by dividing the frequency of the modal auxiliary by the total frequency of the MODAL AUXILIARY + VERB pairs in each category. The RELIANCE value, in contrast, represents to what extent the occurrence of the modal depends on a particular verb, and it is obtained by dividing the frequency of the modal co-occurring with a specific verb by the aggregate frequency of the modal in the category. This approach is convenient, as it enables visual exploration of the data in a two-dimensional space, and as such, makes it possible to gauge the relative contributions of different types of word associations across text categories in an intuitive manner, and also avoid missing important information in the data (see Küchenhoff & Schmid 2015: 546). Given the exploratory nature of the analysis of verbal collocations, it was deemed that these advantages outweigh the documented shortcomings of the chosen measures as compared to other ways of measuring lexicogrammatical attraction,[9] including the fact that since we consider two measures, it is not possible to obtain a single ranking of the colligations (Schmid & Küchenhoff 2013: 550–51). Following the exploratory quantitative analysis, we qualitatively examined the context of the main MODAL AUXILIARY + VERB pairings to identify and describe their local textual functions in the three categories.[10]

Regarding the qualitative appraisal of subjects, we classified them inductively into a small number of categories in a process that was informed by close reading, previous familiarity with the texts, and related work attempting to categorise grammatical subjects in academic writing (e.g. MacDonald 1992; Thomas & Hawes 1994).[11]

4. Results

4.1 Overall frequencies of modals across categories

The frequencies of modal verbs across the three categories in focus are provided as a treemap in Figure 1. The treemap provides us with counts of modal auxiliaries in the three categories studied, and graphically represents (a) the counts of each modal with respect to the other modals in the three categories, and (b) the sum total of modals per category as compared to the other categories. As can be seen by comparing the relative sizes of the rectangles, texts in SC-PER boast the most modals and in REC the least. In addition, the parts of the graph corresponding both to REC and SUR begin in the same way – *may*, *will*, *must* – which could be indicative of broader similarities in the functions

9 It is not possible to discuss in detail the relative merits of different association measures here but see Schmid and Küchenhoff (2013), Gries (2015), and Küchenhoff and Schmid (2015).

10 Given that our main focus is on the sequence MODAL AUXILIARY + VERB, passive clauses were left out of this part of the analysis.

11 Subjects were first identified based on their position relative to the modal auxiliary, but manual verification was necessary, as they were frequently either located somewhere else than in L1 position in the concordance or implied in elliptical phrases. We also manually verified the accuracy of the identification of main verbs and active/passive (the latter was made difficult by the variable spelling of regular past tense suffixes (-*’d* vs. -*ed*).

of modals in these two categories. In this respect, the category of SC-PER stands apart from the other two, as the frequency-ranked list begins with *may*, *could*, *will*.

However, as previously indicated, the top-down analysis illustrated in Figure 1 does not directly tell us what types of modality are conveyed by the modals in different sub-registers, and to investigate that question, we now turn to phraseological analysis. Nevertheless, Figure 1 does guide our investigation: the modals selected for the ATTRACTION and RELIANCE analysis are the top three modal auxiliaries of each sub-register.

4.2 Core modal meanings in *Recipe collections*

To identify potentially interesting MODAL AUXILIARY + VERB pairings that could be linked to text functions characteristic of the three sub-registers of Late Modern English medical writing in focus, we determined the values of ATTRACTION and RELIANCE for each verb that co-occurred with any of the nine modal auxiliaries – *can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, and *would* – and plotted them on a two-dimensional coordinate space. Based on visual exploration of attraction/reliance values, we can identify several discursively relevant pairings. We illustrate the complete procedure of employing visual exploration to arrive at a list of core MODAL AUXILIARY + VERB pairings for the first verb in focus – *will* in *Recipe collections* – whereas the visualisations are omitted from the treatment of remaining pairings for reasons of space.

will

We shall first explore the category of *Recipe collections*, focusing on *will*, *may*, and *must*. Figure 2 plots the values of ATTRACTION and RELIANCE for all the verbs in the category that co-occur with *will* and have the raw frequency of four or larger.

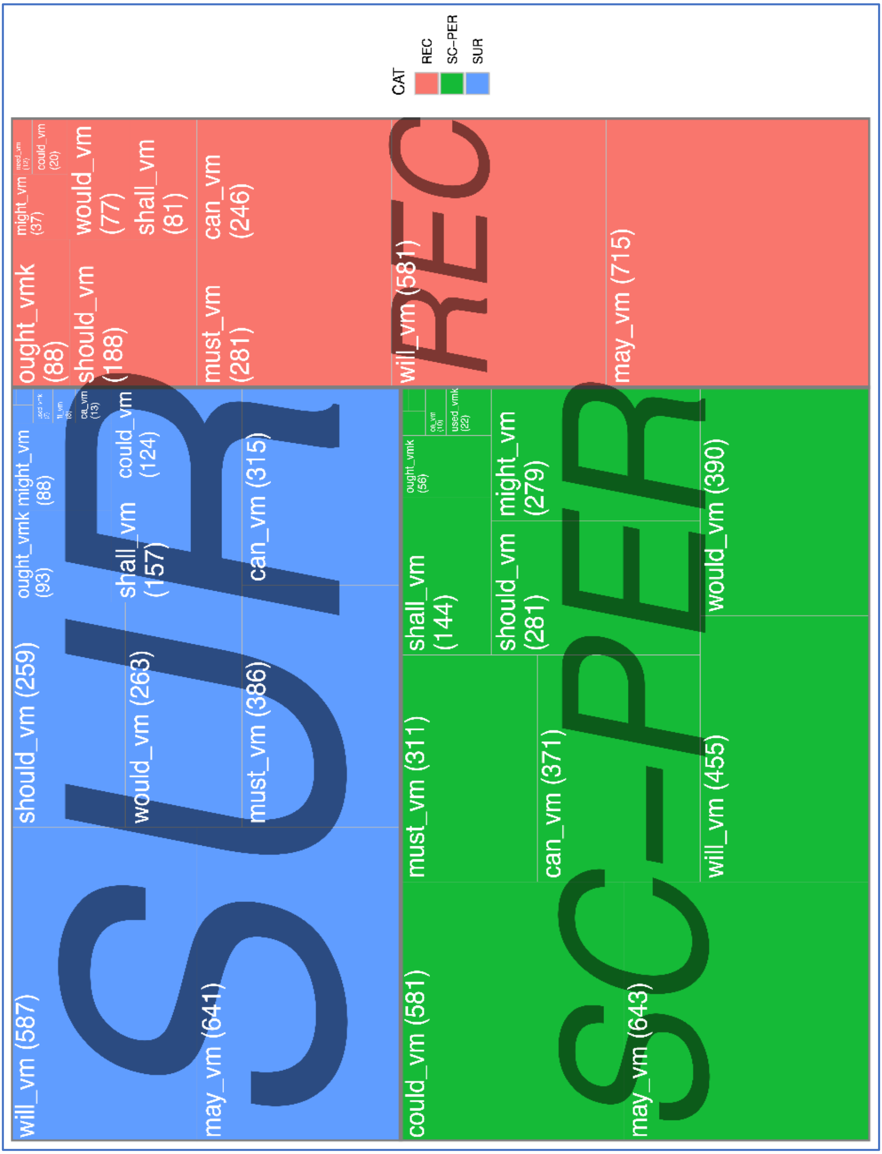


Figure 1. A treemap of modal auxiliaries and their frequencies across the three categories in focus.

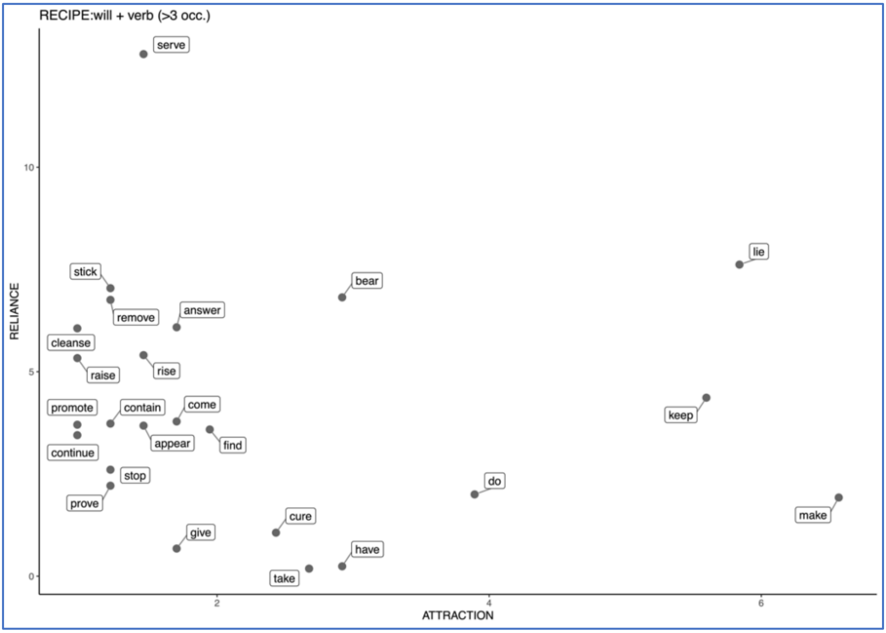


Figure 2. Attraction and reliance for verbs co-occurring with the modal auxiliary *will* in Recipe collections.

As can be seen, the top verb in terms of ATTRACTION is *make* (6.58), typically occurring in statements commenting on the effect of adding an ingredient, or the effect of the finalised produce (example 3). However, as *make* is a fairly frequent verb across the board, its usage in this category is not particularly dependent on this modal auxiliary, and this is reflected in its relatively low RELIANCE value of 1.92 (i.e. the occurrences of the *will* + *make* account for less than two per cent of the total frequency of *make* in *Recipe collections*). By contrast, *serve*, is only ranked on the thirteenth place in terms of ATTRACTION (1.46), but the instances of this verb are much more reliant on the use of *will*, such that *will* + *serve* comprise nearly 13 % of these instances. In terms of text function, *serve* is used for statements about the usability and effectiveness of the substance discussed, but also about the preservation of the substance (example 4).

3 …and adding the Sugar after it has stood in a gentle Warmth (about as much as that the Sun gives) for some weeks, it **will make** a beautiful-colour'd and grateful Tincture … (REC, Quincy, *Pharmacopoeia Officinalis Extemporanea*, 1718)

4 This **will serve** very well till a proper ointment can be prepared. (REC, Cole, *The Lady's Complete Guide*, 1788)

We could thus argue that in terms of discourse function, the use of the modal auxiliary *will* is contingent on both *make* and *serve*, but the nature of the two-way lexicogrammatical relationship is different in each case: *make* is highly attracted to but only moderately reliant on *will*, whereas *serve* is less strongly attracted by *will*, but relies more on it for its use. As our aim is to describe MODAL AUXILIARY + VERB pairings that are both lexicogrammatically salient and discursively key, we used a rough heuristic to identify three types of pairings (listed here in the order of importance), for which

1. both attraction and reliance are high (or relatively high), or
2. reliance is high but attraction is relatively low (or low), or
3. attraction is high but reliance is relatively low (or low).

The pairings we identified for *will* in *Recipe collections* are listed in Table 1.[12] It can be seen that all of them are linked to core textual functions which are present in the category: comments on the effect of ingredients, instruction on preparing, preserving, and administering medical substances, how effective they are and what effect they have in the treatment of a condition. In the most common cases, a statement on the future effect and preservability of medicines is made with the phrase *will* + *do/keep/serve*. The subject is often a medicine or an ingredient, and the verb phrase expresses something about the characteristic properties of the subject (Salkie 2010: 192): *in all the cases when*

12 For reasons of space, example phrases have been abbreviated and the names of the source texts have been omitted in the tables.

X is used, the result will be Y. As the subject in these cases is non-sentient, any type of modality involving subjectivity must be that of the author. If there is an authorial comment, it is the confidence in the presented proposition or the weak commitment to the predictability of the future event which would count among EPISTEMIC POSSIBILITY (Coates 1983: 179). Therefore, it is the temporal / predictive use of *will* that seems to characterize recipe texts.

Table 3. Auxiliary + Verb pairings and their textual functions for *will* in Recipe collections.

Auxiliary	Verb	Main textual function	Example
will	*make*	Effect of an ingredient / of the medicine	(3)
will	*serve*	Effect / preservability of the medicine	(4)
will	*keep*	Preservability of the medicine	*it **will keep** many years*
will	*bear*	Instructions on dosage / administration	*[take every hour] if the stomach **will bear** it*
will	*lie*	Instructions on dosage	*[take] as much as **will lie** on a shilling*
will	*do*	Effect of an ingredient / of the medicine; process information	*[it dries,] which it **will do** in a day*

In the few cases with a personal subject (all of them with the verb *do*) we can observe a deontic use of *will*, as in (5):

5 But observe, if the pain is violent, you **will do** right first to give the purging apozem to promote stools … (REC, Smith, *The Family Physician*, 1761)

Example 5 evaluates the alternative actions of the reader, and it can be paraphrased as *if X, you should Y*. Therefore, the example can be analysed as a case of DEONTIC NECESSITY. Such cases of *will* + *do* are but few, and they are all from expository passages, outside the instructional recipe texts. It can therefore be constituted that they are not necessarily characteristic uses of *will* in recipe texts, as such expository passage may occur in other medical sub-registers as well.

may and *must*

The pairings identified for *may* and *must*, listed in Table 2, are linked to statements of dosage and administration, efficacy of medicines, and regimens of health. Statements containing *may add* are found in recipes proper (and not in the adjacent commentary), in the part that describes the preparation of the medicine and the order of adding the ingredients. The phrases attest to ROOT POSSIBILITY, more specifically opportunity (see Depraetere & Cappelle 2023: 32).[13] Examples of *may require* are all cases of EPISTEMIC POSSIBILITY, as they can be paraphrased as *it is possible* ***that*** *X requires* (see Depraetere & Cappelle 2023: 39). The function of sentences with *may require* is likewise limited to instructions on dosage and administration of medicines. *May serve* denotes ROOT POSSIBILITY (opportunity), but unlike the phrase *may add* described earlier, it does not occur in the actual recipes to make medicines, but in discussions of their efficacy. The phrase *may take* is most often found with dosage instructions, and they are usually restricted by some condition, often by the age of the patient. The type of modality is, therefore, ROOT POSSIBILITY, more precisely situation permissibility (Depraetere & Cappelle 2023: 33).

The modal *must* seems not to be a feature of recipes proper, but it typically occurs in the adjacent prose, often in regimen-like instructional passages. The examples of *must* listed in Table 2 are all cases of ROOT NECESSITY (deontic necessity).[14] However, category-specific generalisations need to be made with caution, as the overall frequencies of individual AUXILIARY + VERB pairs are fairly low and sometimes the instances occur in just a couple of texts.[15]

---

13 The possibility is noted by a subject-external source, in this case the author of the text, but there is no consideration of the likelihood of the situation, hence it is a case of concessive use of *may*. The scope of the modal is narrow (it applies only to the verb and not to the whole situation), and there is no potential barrier nor condition that would cancel the action (i.e. the author does not list conditions under which the possibility would not stand).

14 Using the terminology of Depraetere and Cappelle (2023: 23–24), all examples are cases of objective deontic modality, or obligation where the deontic source is an authority, either named or implied, not the speaker/author of the text.

15 For example, of the 13 cases of *must take*, eight are from Smith's *The Compleat Housewife* (1728). The four cases of *must go* are all from regimens of health and from two texts (1788 Cole: *The Ladys Complete Guide* and 1795 Taylor: *Mrs Taylors Family Companion*) that seem to be (at least partly) copies of the same.

Table 2. Auxiliary + Verb pairings and their textual functions for *may* and *must* in *Recipe collections*.

Auxiliary	Verb	Main textual function	Example
may	*add*	Preparation	*you **may add** two spoonfuls of cinnamon water*
may	*require*	Dosage	*to be repeated as occasion **may require***
may	*serve*	Efficacy considerations	*it **may** well **serve** as a substitute*
may	*take*	Dosage	*a robust person **may take** a spoonful*
must	*take*	Dosage / Regimen of health	*older children **must take** more*
must	*go*	Regimen of health	*the patient **must go** into a cold bath*
must	*eat*	Regimen of health	*you **must eat** the following diet*
must	*anoint*	Dosage / Administration	*you **must anoint** the reins of the back*

4.3 Core modal meanings in *Surgical texts*

will

The auxiliary *will* likewise co-occurs with a number of verbs in *Surgical texts*, of which the key ones are listed in Table 3. *Will make* is a prominent word pair, but in this category it is not linked to the efficacy of substances or end products as in *Recipe collections*, but to instruction on how to instrumentally treat specific pathological conditions (6) or exposition based on observations on human physiology.

6 …; it [=*the bandage*] **will make** a pressure on the intestine while down, which may prove very pernicious, and has often been the sole occasion of a gangrene … (SUR, Pott, *A Treatise on Ruptures*, 1756)

Other verbs co-occurring with *will* are also linked to common textual characteristics of surgical writing. For example, the phrase *will find* is typically used in abstract sense, about finding an argument or proving it, or (to a lesser extent) to instruct a medical practitioner (e.g. *surgeon*, *lithotomist*) how to locate something in the human body (e.g. a stone in the urinary tract). Phrases *will come*, *will slip* and *will produce* describe physiological processes like the appearance of pain, and relatedly, *will happen* comments on the probability of the occurrence that is being discussed.

The phrase *will make* attests to the author or speaker's confidence in the proposition, often based on first-hand experience. The VP is found in passages of EPISTEMIC POSSIBILITY and NECESSITY, sometimes supported by modally harmonious epistemic adverbs or adjectives, e.g. *will **certainly** make him fearful*[16] and *will thereby make it **necessary***,[17] respectively. In the cotexts described above, *will find* almost exclusively communicates EPISTEMIC POSSIBILITY, and *will admit* ROOT/DYNAMIC POSSIBILITY. Verb pairs *will come* and *will slip* attest to EPISTEMIC POSSIBILITY, with a more or less certain outcome: the likely sense of the modified VP can be elicited from the adverbs often following the VP (*always, sometimes, easily, most probably*). The remaining verbs, *appear, happen* and *produce*, belong to the realms of EPISTEMIC POSSIBILITY and ROOT/DYNAMIC POSSIBILITY, and they are often further supported by adverbs such as *always, almost infallibly, frequently, most probably, seldom* and *sometimes*.

Table 3. Auxiliary + Verb pairings and their textual functions for *will* in *Surgical texts*.

Auxiliary	Verb	Main textual function	Example
will	*make*	Procedural instruction	(6)
will	*find*	Procedural instruction	*he **will find** it*
will	*admit*	Procedural instruction / exposition	*[do X] if the part **will admit** it*
will	*come*	Physiological description	***will come** out*
will	*slip*	Physiological description	*it **will slip** through*
will	*produce*	Physiological description / causes of symptoms	***will** produce **pain***
will	*appear*	Likelihood statement (‘to become evident’)	*it **will appear** that X*
will	*happen*	Likelihood statement	***will** frequently / seldom **happen***

16 SUR, Taylor: *An Account of the Mechanism of the Eye*, 1727
17 SUR, Pott: *A Treatise on Ruptures*, 1756

may and *must*

Table 4 lists the core word pairs for auxiliaries *may* and *must* and describes their textual functions.

Table 4. Auxiliary + Verb pairings and their textual functions for *may* and *must* in *Surgical texts*.

Auxiliary	Verb	Main textual function	Example
may	*observe*	Collaborative reasoning	(7)
may	*happen*	Likelihood statement	*it **may happen** that X*
may	*seem*	Likelihood statement	*X **may seem** strange*
may	*appear*	Likelihood statement	*it **may appear** to X*
must	*have*	Various	(8)
must	*take*	Procedural instruction	(9)

Interestingly, the functions assigned to these word pairs stand in stark contrast with those identified for *Recipe collections* in the previous section. As can be seen, instances with *may* are linked with cognitive processes and logical reasoning. With phrases like *may observe*, the subject is either *I* or inclusive *we* (7), where the reader is implicitly invited to participate in the cognitive process of working out the relative merits of different lines of treatment. The use of the auxiliary in these cases looks like permissive *may*, and as such it would point towards ROOT/DEONTIC MODALITY; nevertheless, the analysis of the cotext shows that *may observe* in surgical texts communicates EPISTEMIC POSSIBILITY. *Happen*, *seem* and *appear* comment on the likelihood of some proposition discussed in the text.

7 Mary Langley had the disorder in all seven weeks, and magnesia and nitre were joined with hemlock; and, upon the whole we **may observe**, that there are only seven cases out of the twenty, in which hemlock can possibly have all the credit of the cure, because other medicines were joined along with it; …(SUR, Kirkland, *An Essay On*

The Methods Of Suppressing Haemorrhages From Divided Arteries, 1763)

What is also interesting is that *must* co-occurring with *take* almost exclusively exhibits ROOT NECESSITY, as in example (8), whereas the instances with *must have* (excluding *have* as auxiliary) attest to EPISTEMIC and ROOT NECESSITY in equal numbers, as exemplified in (9) (epistemic) and (10) (root).

8 The knife he **must take** in his fingers, like a pen, and lean the little finger of his right hand on the cheek-bone of the patient; (SUR, Bischoff, *A Treatise On The Extraction Of The Cataract*, 1793)

9 After every capital operation we should give our utmost attention to the application of the compresses and bandage; …as long continued pain **must have** destructive consequences; (SUR, Gooch, *Cases And Practical Remarks In Surgery*, 1758)

10 It **must have** a gentle hold of the upper part of the Os Pubis or Sharebone; but take care it is not placed too high or too low: […] (SUR, Brand, *The True Method Of Reducing Ruptures*, 1771)

4.4 Core modal meanings in *Scientific periodicals*

will

The pairings identified for *will* in *Scientific periodicals* are found in Table 5.

The fact that *appear* stands out as a key colligate of *will* owes much to examples like (11). Here, the author confidently expresses a proposition – *blood globules are not filled with elastic fluid* – using the epistemic verb *appear* as an indicator of near-categorical certainty (cf. Vihla 1999: 56) accompanied by a cataphoric deictic expression *following*, which indicates where the evidence for the proposition is found. Both *will answer* and *will succeed* contribute towards expressing a claim with a high degree of certainty, occasionally further emphasized by a comparative (*it will **better** answer*) or a highly inclusive pronoun (*it will answer **every** purpose*), or at times downtoned by an adverb like *sometimes*, lessening the speaker's / author's commitment to the proposition. All the cases of *appear, answer,* and *succeed* fall under

EPISTEMIC POSSIBILITY / NECESSITY.[18] *Will do* occurs variously in constructions where the main contribution to the meaning comes from the direct object (e.g. *honour / justice / favour*), not the verb itself. *Will give* is found in expressions that communicate the author's confidence in the likelihood of future events. Both *will do* and *will give* mostly communicate EPISTEMIC POSSIBILITY.

11 And that they [=*blood globules*] are not fill'd with any sort of Elastick Fluid, **will appear** from the following Experiment. (SC-PER, PT Vol. 30, pp. 1000–1014, 1717–1719)

Table 5. Auxiliary + Verb pairings and their textual functions for *will* in *Scientific periodicals*.

Auxiliary	Verb	Main textual function	Example
will	*appear*	Introduction of evidence	(11)
will	*answer*	Prediction of accomplishment / outcome	***will*** *(not)* ***answer*** *the purpose / end*
will	*succeed*	Prediction of accomplishment / outcome	***will*** *(sometimes)* ***succeed***
will	*give*	Declaration	*I* ***will give*** *my account*
will	*do*	Phrasal	***will do*** *me the justice/honour*

may

Table 6 shows that as far as its verbal colligates are concerned, *may* predictably expresses different shades of meanings related to POSSIBILITY.

May have (and *may perform*) are found in statements like (12) where the proposition about the world is presented as a possible scenario (the person in question **possibly** contracting the plague), not a definite outcome.

12 …the case is different in Italy, and in the south of France; to which countries a ship with a fair wind **may perform** a voyage in eight days from the Levant; during which time a person **may have** the

18 The evidence-introducing cases of *appear* could also be identified as attestations of EVIDENTIAL MODALITY, but in that case *appear* should be interpreted as a verb of perception (Vihla 1999: 30).

plague about him, without being confin'd to his bed; of which there are many instances. (SC-PER, PT Vol. 47, pp. 0514–0516, 1751)

Table 6. Auxiliary + Verb pairings and their textual functions for *may* in *Scientific periodicals*.

Auxiliary	Verb	Main textual function	Example
may	*have*	Possibility statement	(12)
may	*seem*	Projection	*it **may seem** (ADJ) that*
may	*serve*	Projection	*this **may serve** to point out X*
may	*conclude*	Projection	*we **may conclude** that*
may	*happen*	Projection	*it **may** easily **happen** that*

A similar epistemic qualification is present in instances with the other colligates listed in the table, which moreover are specifically related to projection (Halliday & Matthiessen 2013: 515–19).[19] Thus, in (13), *may* mitigates the force of the claim being expressed in a simple catenative construction involving *serve* (compare with *This account **serves** to convince …*), whereas *may seem* in (14) has a different function, that of projecting an adversative clause. In other words, the author uses *may seem* to present a conclusion that is possible (i.e *that the present observations clash with some previous ones*) but not one which is ultimately adopted; instead, the actual claim is presented in the following sentence (*But […] they will be found […] to confirm them …*).

19 The verbs in Table 6 occur in main clauses which project some kind of "stance" or "attitudinal meaning" on the following nominal complement, which may either be a noun phrase or nominal content clause. The stance in these clauses is partially constructed through the choice of the verb and can be further modified with the help of modal auxiliaries like may (cf. Gray, Biber, & Hiltunen 2011). The specific semantic dimensions involved in the constructed stance are dependent on the particularities of the verb.

13 This Account **may serve** to convince those who are of Opinion that Boys are conceived on the right Side, and the Girls on the left; … (SC-PER, PT Vol. 44, pp. 0617–0621, 1746)[20]

14 These observations **may seem** at first view to clash with and contradict those I have related: But, upon closer consideration, they will be found in reality to confirm them, … (SC-PER, PT Vol. 49, pp. 0254–0264, 1755)

The examples of *may* in SC-PER and the related discussion has constituted the use of the modal to be related to possibility (ROOT POSSIBILITY in the case of *have* and *happen*, and EPISTEMIC POSSIBILITY with *seem, serve* and *conclude*).[21] It is worth emphasising here that the specification of the modal meaning of an utterance is not limited to auxiliaries. For example, in (14), the verb *seem* is used epistemically: rather than being directly related to visual perception, it describes "a cognitive state concerning what is subjectively probably conditional on evidence" (Brogaard 2015: 13). Elsewhere in the corpus, it is not unusual to find occurrences of modally harmonious configurations involving *may* and adverbs (e.g. *we may* ***reasonably*** *conclude*, *it may* ***easily*** *happen*). All this underlines the usefulness of looking at the phraseology of modal expressions and not merely their frequencies when describing how they are used in discourse.

could

Finally, the core word pairs with their characteristic textual functions for *could*, the auxiliary that was established as having a comparatively high frequency in this very category, are summarised in Table 7.

20 The idea of the left-right polarity of the body and its influence on sex determination of the child at conception was first recorded in ancient Greece (Pahta 1998: 41), and it might appear surprising that it is still found in medical writing in the 18th century. This idea was, in fact, even more persistent as it was perpetuated in medical discourse until at least the early 20th century.

21 Nurmi (2009: 338) has pointed out the rise of the epistemic sense of *may* in personal letters, effected by "educated high-ranking men" from the 16th century onwards. Personal letters and the different modalities of *may* provide, therefore, interesting points of comparison to scientific periodicals, which originally evolved from the register of letters.

Table 7. Auxiliary + Verb pairings and their textual functions for *could* in *Scientific periodicals*.

Auxiliary	Verb	Main textual function	Example
could	*discover*	Acquiring / failing to acquire knowledge	*I **could** (not) **discover** that X*
could	*perceive*	Acquiring / failing to acquire knowledge	*I **could** (not) **perceive** X*
could	*distinguish*	Acquiring / failing to acquire knowledge	*I **could** (not) **distinguish** X*
could	*get*	Success / failure in a procedure	*I **could** (not) **get** it out*

Three of the colligates – *discover*, *perceive*, and *distinguish* – are related to the domains of cognition and perception, and at first glance it would seem reasonable to hypothesise that authors use them to report how a specific piece of information was obtained. While such reports do occur (15), upon closer inspection this turns out to be only a partial explanation, as in addition we encounter a sizable sample of instances where these verbs occur in statements with negative polarity (in fact, only 1 out of 5 instances of *could discover / perceive / distinguish* reports a positive outcome). Using such statements, the authors relate their inability to perceive something, as illustrated in (16). Interestingly, this description bears a striking similarity with Sinclair's well-known description of the phrase *naked eye* as a partial unit of meaning, which collocates/colligates with modals and negative polarity items, and which is associated with the semantic prosody of 'difficulty' (Sinclair 2004: 30–35). Hence, *could* mostly communicates ROOT POSSIBILITY (or, perhaps more accurately, "root impossibility"). The pair *could get* likewise frequently occurs in statements with negative polarity, but with a slightly different function: they most often report a medical procedure that was attempted but turned out to be more or less unsuccessful.

15 …and the Remainder, where there had been no Adhesion, (as I **could perceive** from the smooth Surface of the Pleura) was torn away by Piece … (SC-PER, PT Vol. 41, pp. 0623–0624, 1739–1741)

16 …we **could** <u>scarce</u> **distinguish** one from another. (SC-PER, EMJ1 Vol. 4, pp. 0242–0244, 1747)

5. Discussion

The present study has followed an exploratory approach to modality in LModE medical writing, combining ideas from descriptive corpus linguistics, phraseology/constructionist approaches to grammar, and register analysis. The findings of the study have identified several meaningful connections between specific phraseological uses of modal verbs (operationalised as recurrent and salient pairings of AUXILIARIES and LEXICAL VERBS) and the communicative purposes of the three sub-registers of 18th-century medical writing. By doing so, the analysis lends support to the idea that the approach we have adopted is a viable one for the analysis of modality, which is sensitive to the syntagmatic connections of modal verbs and the discourse contexts within which individual instances are embedded, and feasible to implement on mid-sized specialised corpora, such as the individual text categories of the *Late Modern English Medical Texts* (LMEMT).

Some of the obtained results were entirely in accordance with the expectations. For example, it is unsurprising that modal expressions in *Recipe collections* are embedded in discourse where ways of preparing and administering medical substances are expounded, given that this is a major concern in much of the writing in this category. Here, our analyses acted as a confirmation of, and partially an elaboration on, informal hypotheses gleaned from earlier research and familiarity with recipes and related texts from earlier periods in the history of the register. Alongside this, the analysis also identified phraseological tendencies that were, perhaps, less obvious at first glance, but on closer inspection turned out to be entirely congruent with the rhetorical purposes of the texts. For example, the fact that *could* emerges as a salient modal auxiliary particularly in *Scientific periodicals* is due to a complex configuration of phraseological variables including the choice of the lexical verb and its associated semantics, the subject, and polarity, giving rise to recurrent variants of the phrasal template 1SG/PL + *could* + NEG + VERB + NP/*that*-clause.

While the present analysis has taken the first step towards the identification and description of these configurations, a comprehensive analysis clearly necessitates a study with wider scope that covers not only the remaining sub-registers of medical writing from the period, but also a wider range of phrase frames at a higher level of granularity. In particular, the qualitative appraisal of the role of clausal subjects suggests that their systematic incorporation into the framework of quantitative analysis emerges as a high priority for future work. Similarly, we have deliberately excluded from the analysis both passive constructions (e.g. *it* ***may*** *be* ***reckoned***) and other expressions that are clearly formulaic (e.g. *as hot as you* ***can bear*** *it* or *as plain as* ***could be***) for reasons of space; assessing their frequency and discourse function is likely to offer further insights into the use of modals.

References

Alonso-Almeida, Francisco 2015a. On the mitigating function of modality and evidentiality. Evidence from English and Spanish medical research papers. *Intercultural Pragmatics* 12 (1): 33–57. https://doi.org/10.1515/ip-2015-0002

Alonso-Almeida, Francisco 2015b. The functions of *seem* and *parecer* in early medical writing. *Discourse Studies* 17 (2): 121–140. https://doi.org/10.1177/1461445614564517

Alonso-Almeida, Francisco & Francisco J. Álvarez-Gil 2020. 'So that it may reach to the jugular.' Modal verbs in Early Modern English recipes. *Studia Neofilologiczne* 16: 61–88.

Álvarez-Gil, Francisco J. & Natalia García Alonso 2019. Analysis of the linguistic expression *I think* in a corpus of Late Modern English scientific texts. *Revista de Lenguas Para Fines Específicos* 25 (2): 44–57. https://ojsspdc.ulpgc.es/ojs/index.php/LFE/article/view/1149

Biber, Douglas 2012. Register as a predictor of linguistic variation. *Corpus Linguistics & Linguistic Theory* 8 (1): 9–37.

Brogaard, Berit 2015. Perceptual reports. *The Oxford Handbook of Philosophy of Perception*, ed. by Mohan Matthen. 237–253 Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199600472.013.005

Cappelle, Bert & Ilse Depraetere 2016. Short-circuited interpretations of modal verb constructions: some evidence from the Simpsons. *Constructions and Frames* 8 (1): 7–39.

Carrió Pastor, María Luisa 2012. A contrastive analysis of epistemic modality in scientific English. *Revista de Lenguas Para Fines Específicos* 18: 115–132.

Coates, Jennifer 1983. *The Semantics of Modal Auxiliaries*. London: Croom Helm.

Coates, Jennifer 1995. The expression of root and epistemic possibility in English. *Modality in Grammar and Discourse*, ed. by Joan L. Bybee & Suzanne Fleischman. 55–66. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.32.04coa

Crespo-García, Begoña 2011. Persuasion markers and ideology in eighteenth century philosophy texts (CEPhiT). *Revista de Lenguas Para Fines Específicos* 17: 199–228.

Dehouck, Mathieu & Pascal Denis 2023. Revisiting modal sense classification with contextual word embeddings. *Models of Modals. From Pragmatics and Corpus Linguistics to Machine Learning*, ed. by Ilse Depraetere *et al*. 225–253. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110734157-009

Depraetere, Ilse & Bert Cappelle 2023. English modals: an outline of their forms, meanings and uses. *Models of Modals. From Pragmatics and Corpus Linguistics to Machine Learning*, ed. by Ilse Depraetere *et al*. 14–59. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110734157-002

Depraetere, Ilse, Bert Cappelle, & Martin Hilpert 2023. Introduction. *Models of Modals. From Pragmatics and Corpus Linguistics to Machine Learning*, ed. by Ilse Depraetere *et al*. 1–13. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110734157-001
Gray, Bethany, Douglas Biber, & Turo Hiltunen 2011. The expression of stance in early publications of the philosophical transactions and other contemporary medical prose: innovations in a pioneering discourse. *Medical Writing in Early Modern English*, ed. by Irma Taavitsainen & Päivi Pahta. 221–247. Cambridge: Cambridge University Press.
Gries, Stefan Th. 2015. More (old and new) misunderstandings of collostructional analysis: on Schmid and Küchenhoff (2013). *Cognitive Linguistics* 26 (3): 505–536. https://doi.org/10.1515/cog-2014-0092
Halliday, M.A.K. & Christian M.I.M. Matthiessen 2013. *Halliday's Introduction to Functional Grammar.* 4th edition. London: Routledge.
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 & Jukka Tyrkkö 2011. Verbs of knowing: discursive practices in early modern vernacular medicine. *Medical Writing in Early Modern English* (Studies in English Language), ed. by Irma Taavitsainen & Päivi Pahta. 44–73. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.005
Hoye, Leo 1997. *Adverbs and Modality in English*. London and New York: Longman.
Huddleston, Rodney 1976. Some theoretical issues in the description of the English verb. *Lingua* 40: 331–383. https://api.semanticscholar.org/CorpusID:170479872
Kopaczyk, Joanna 2013. Formulaic discourse across Early Modern English medical genres: investigating shared lexical bundles. *Meaning in the History of English. Words and Texts in Context* (Studies in Language Companion Series), ed. by Andreas Jucker, Daniela Landert, Annina Seiler, & Nicole Studer-Joho. 257–300. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.148
Küchenhoff, Helmut & Hans-Jörg Schmid 2015. Reply to more (old and new) misunderstandings of collostructional analysis: on Schmid & Küchenhoff by Stefan Th. Gries. *Cognitive Linguistics* 26 (3): 537–547. https://doi.org/10.1515/cog-2015-0053
Leclercq, Benoît 2023. Modality revisited: combining insights from construction grammar and relevance theory. *Models of Modals. From Pragmatics and Corpus Linguistics to Machine Learning*, ed. by Ilse Depraetere *et al.* 60–92. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110734157-003
Leech, Geoffrey 2014. Where have all the modals gone? an essay on the declining frequency of core modal auxiliaries in recent standard English. *English Modality: Core, Periphery and Evidentiality,* ed. by Juana I. Marín-Arrese, Marta Carretero, Jorge Arús Hita, & Johan van der Auwera. 95–116. Berlin and Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110286328.95
Lyons, John 1977. *Semantics*. Cambridge: Cambridge University Press.
MacDonald, Susan Peck 1992. A method for analyzing sentence-level differences in disciplinary knowledge making. *Written Communication* 9 (4): 533–569.
Mäkinen, Martti 2022. Persuasion in Early Modern English medical recipes. *Genre in English Medical Writing, 1500–1820: Sociocultural Contexts of Production and Use* (Studies in English Language), ed. by Irma Taavitsainen, Turo Hiltunen, Jeremy J.

Smith, & Carla Suhr. 254–277. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009105347.016
Marín-Arrese, Juana I. 2009. Effective Vs. epistemic stance, and subjectivity/intersubjectivity in political discourse. A case study. *Studies on English Modality: In Honour of Frank Palmer* (Linguistics Insights: Studies in Language and Communication), ed. by A. Tsangalidis & R. Facchinetti. 23–52. Bern: Peter Lang.
McEnery, Tony & Andrew Hardie 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
Millar, Neil 2009. Modal verbs in TIME: Frequency changes 1923–2006. *International Journal of Corpus Linguistics* 14 (2): 191–220.
Nurmi, Arja 2009. May: the social history of an auxiliary. *Corpora: Pragmatics and Discourse: Papers from the 29th International Conference on English Language Research on Computerized Corpora (ICAME 29)*, ed. by Andreas Jucker, Daniel Schreier, & Marianne Hundt. 321–42. Brill.
Pahta, Päivi 1998. *Medieval Embryology in the Vernacular: The Case of* De spermate. Helsinki: Société Néophilologique.
Palmer, Frank Robert 1990. *Modality and the English Modals*. Linguistics Library. Longman.
Salkie, Raphael 2010. Will: tense or modal or both? *English Language and Linguistics* 14 (2): 187–215. https://doi.org/10.1017/S1360674310000055
Schmid, Hans-Jörg 2000. *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin: Mouton de Gruyter.
Schmid, Hans-Jörg & Helmut Küchenhoff 2013. Collostructional analysis and other ways of measuring lexicogrammatical attraction: theoretical premises, practical problems and cognitive underpinnings. *Cognitive Linguistics* 24 (3): 531–577. https://doi.org/10.1515/cog-2013-0018
Silk, A. 2022. Weak and strong necessity modals: on linguistic means of expressing 'a primitive concept OUGHT'. *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*, ed. by B. Dunaway & D. Plunkett. 203–245. Ann Arbor: Michigan Publishing. https://doi.org/10.3998/mpub.9948199
Sinclair, John 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
Sinclair, John 2008. The phrase, the whole phrase, and nothing but the phrase. *Phraseology: An Interdisciplinary Perspective*, ed. by Fanny Meunier & Sylviane Granger. 407–410. Amsterdam: John Benjamins.
Sweetser, Eve 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure* (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.
Taavitsainen, Irma 2001. Evidentiality and scientific thought-styles: English medical writing in Late Middle English and Early Modern English. *Modality in Specialized Texts: Selected Papers of the 1st CERLIS Conference*, ed. by Maurizio Gotti & Marina Dossena. 21–52. Bern: Peter Lang.
Taavitsainen, Irma & Turo Hiltunen (eds.) 2019. *Late Modern English Medical Texts: Writing Medicine in the Eighteenth Century*. Amsterdam: John Benjamins.
Taavitsainen, Irma, Turo Hiltunen, Anu Lehto, Ville Marttila, Päivi Pahta, Maura Ratia, Carla Suhr, & Jukka Tyrkkö (eds.) 2019. *Late Modern English Medical Texts (LMEMT)*. CD-ROM. Amsterdam: John Benjamins.

Thomas, Sarah & Thomas P. Hawes 1994. Reporting verbs in medical journal articles. *English for Specific Purposes* 13 (2): 129–148.
Vihla, Minna 1999. *Medical Writing: Modality in Focus*. Amsterdam: Rodopi.
Warchał, Krystyna 2015. *Certainty and Doubt in Academic Discourse: Epistemic Modality Markers in English and Polish Linguistics Articles*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
Whitt, Richard 2016a. Evidentiality in Early Modern English medical treatises (1500–1700). *Journal of Historical Sociolinguistics* 2 (2): 235–63. https://doi.org/10.1515/jhsl-2016-0014
Whitt, Richard 2016b. Using corpora to track changing thought styles: evidentiality, epistemology, and Early Modern English and German scientific discourse. *Kalbotyra* 69: 267–293. https://doi.org/10.15388/Klbt.2016.10376
Whitt, Richard 2023. Epistemic space and key concepts in early and late modern medical discourse: an exploration of two genres. *English Language & Linguistics* 27 (2): 241–69.
Wright, Georg Henrik von 1951. *An Essay in Modal Logic*. Amsterdam: North-Holland.
Yu, Danni, Luyang Li, Hang Su, & Matteo Fuoli 2024. Assessing the potential of LLM-assisted annotation for corpus-based pragmatics and discourse analysis. *International Journal of Corpus Linguistics*. https://doi.org/10.1075/ijcl.23087.yu

“Teens like science, not science class, study finds”
Clausal evidential parentheticals in Contemporary American English

María José López-Couso
ORCID: 0000-0003-3543-7408
Belén Méndez-Naya
ORCID: 0000-0003-2238-2364

The present paper focuses on a type of evidential parenthetical featuring a noun phrase subject with the noun *study* combined with a verb phrase, as in *Teens like science, not science class, study finds*. These parentheticals are recorded from the early 1940s in American English but become frequent only in the 1990s. Using data from the *Corpus of Contemporary American English* (COCA), we show that *study*-parentheticals are attested in two different patterns, which differ as regards the relative order of subject and verb, the non-inverted pattern being the most frequent in the data. Different variants have been identified for each pattern, depending on the complexity of the subject noun phrase (presence/absence of determiners and modifiers). Our analysis reveals that these structures are closely associated with the written medium, in particular with popular science journalistic texts. We argue that *study*-parentheticals show incipient grammaticalization and that the different parenthetical variants can be classified as constructional theticals (Kaltenböck, Heine, & Kuteva 2011), with one particular sequence, *study finds*, coming

https://doi.org/10.51814/ufy.1041.1460

close to a formulaic thetical which has become a staple in pop-sci headlinese.

Keywords: parentheticals, evidentiality, grammaticalization, journalistic texts, headlinese, American English.

1. Introduction

Evidentiality is a semantic notion concerned with the indication of one's source of information. In languages such as Tibetan or Quechua evidentiality is an obligatory grammatical category (Aikhenvald 2004: 3), while others make use of various types of evidential strategies, including both lexical means and grammaticalized expressions (Aikhenvald 2004).[1]

English lacks a fully grammaticalized evidential system but possesses various strategies to convey epistemic and evidential meanings, ranging from the most grammaticalized ones (e.g., modal verbs) to others showing lower degrees of grammaticalization, such as parentheticals (e.g., *it seems*) and adverbs (e.g., *perhaps*) (cf. Chafe 1986: 261; Mélac 2022: 340). Over the last few years, we have devoted some space to clausal parenthetical constructions in both historical and contemporary English, with a focus on various patterns used with epistemic and/or evidential value. In this chapter we turn our attention to a further evidential parenthetical type, which has become relatively frequent in the last decades. The examples in (1) and (2) below illustrate this recent pattern, which features a third person subject in the form of a phrase that contains a noun denoting an examination or an investigation, such as *study* or *research*, accompanied by a verb. The analysis of these structures shows their close association with a specific genre, namely the press.

1 This research is funded by the Spanish Ministry of Science and Innovation (National Programme for Excellence in Scientific and Technical Research; grant PID2020-114604GB-I00) and the Regional Government of Galicia, *Consellería de Cultura, Educación, Formación Profesional e Universidade*, grant ED431B 2023/03. We also thank an anonymous reviewer and the editors of this volume for their insightful comments. The usual disclaimers apply. The chapter is dedicated to Päivi Pahta, friend and colleague, whose work has always been an inspiration for us.

1 This targeted ridicule, *studies show*, can have lasting effects. (COCA, MAG, 2015, Newsweek)

2 Total or near-total blockage of two or perhaps one major artery may warrant angioplasty. Those noninvasive steps can frequently relieve angina and lengthen life, *research suggests*. (COCA, MAG, 2007, Consumer Reports)

The chapter is organized in the following way. Section 2 sets the scene by providing a characterization of clausal parentheticals; special attention is paid here to their syntactic and prosodic independence, their positional mobility, and their semantico-pragmatic features. Section 3, in turn, is concerned with methodological issues, including the materials used for the analysis (Section 3.1) and the searches run in the *Corpus of Contemporary American English* (COCA; Davies 2008–) (Section 3.2). The analysis of the data is the focus of Section 4, where we describe the parenthetical patterns identified in the corpus, their variants, and their distribution in the short diachrony (Section 4.1), the behaviour of our parentheticals as regards the position they occupy in relation to their host clauses (Section 4.2), the verbal predicates which occur in the different parenthetical variants (Section 4.3), and the distribution of the parentheticals under examination across genres (Section 4.4). Section 5 is devoted to the discussion of how the data described in Section 4 fit in with accounts on the origin and development of clausal parentheticals, special attention being drawn to several formal indicators of grammaticalization. Finally, Section 6 offers a brief conclusion and some suggestions for further research.

2. Setting the scene: Clausal parentheticals

Parentheticals are words, phrases, or clauses which are independent from their host or anchor both syntactically and prosodically, being set off from the rest of the utterance by pauses (and, typically, by commas in writing). The information conveyed by parentheticals is optional (Boye & Harder 2021) and their meaning is non-restrictive (Kaltenböck, Heine, & Kuteva 2011: 857).

Parentheticals may take diverse forms and show different degrees of syntactic complexity. The categories that can occur parenthetically range from clauses (e.g., *I think*, *as you say*, *what is more*) and phrases of various types (e.g., *between you and me*) to adverbs (e.g., *frankly*), interjections (e.g., *damn*), and discourse markers (e.g., *like*) (Kaltenböck 2007: 29–31; Dehé 2009; Kaltenböck, Heine, & Kuteva 2011: 856, 876). In this chapter we are concerned with a specific type of clausal parentheticals, namely those structures resembling a matrix clause (Quirk *et al.* 1985: 1112) which also have "a non-parenthetical use in which they take a declarative content clause as complement" (Huddleston & Pullum *et al.* 2002: 895). A paradigmatic instance of parentheticals of this kind is *I think*, exemplified in (3a) and (4a), while (3b) and (4b) illustrate the corresponding matrix clause counterparts.

3 a. <u>Doing the show from Philadelphia during the Pennsylvania primaries was,</u> *I think*, <u>hard enough</u>. (COCA, SPOK, 2011, NPR_FreshAir)
b. I think (that) doing the show from Philadelphia during the Pennsylvania primaries was hard enough.

4 a. <u>She would count as a groupie</u>, *I think*. (COCA, MAG, 2015, Esquire)
b. I think (that) she would count as a groupie.

Despite the similarities between the (a) and (b) sequences in (3) and (4), the two constructions show some crucial differences. For example, from a syntactic point of view, *I think* in (3a) and (4a) functions as an adverbial, more specifically, as a disjunct (Quirk *et al.* 1985: 612 ff.), while the rest of the utterance is the main clause. In (3b) and (4b), by contrast, *I think* is the matrix of the subordinate clause that follows, which functions as a complement of the predicate *think*. Another relevant difference between the (a) and the (b) counterparts in (3) and (4) concerns argument structure: while in the sequences in (b) the verb *think* takes a clausal complement, in the (a) examples *think* lacks the object argument, which suggests that parentheticals like *I think* are syntactically defective; hence the terms 'gap-containing parenthetical clauses' (Kaltenböck 2007: 41) or 'reduced parenthetical clauses' (Schneider 2007: 76).

The (a) and (b) sequences in (3) and (4) also differ pragmatically. Thus, in (3a) and (4a) the underlined statements constitute the primary focus of discourse, that is, they convey the most salient information, while *I think* is secondary. For instance, (4a) does not describe an act of thinking; it is rather about her being a groupie. The discourse secondariness of *I think* can be probed by addressability tests (Boye & Harder 2007: 581–585): only discourse-primary elements can be taken up by question tags, as in (5a), or can be addressed by *really*-queries, as shown in (5b):

5 a. She would count as a groupie, I think, wouldn't she?
b. Really? (i.e., does she really count as a groupie?)

In complementation structures, on the other hand, it is the matrix clause *I think* that is discourse-primary (see Boye & Harder 2007; 2012). This becomes evident in cases where the matrix *I think* is questioned (6a) or is addressed by *really* (6b).

6 a. Well, *what do you think*? *I think* that I'm done looking for answers around here, *that's what I think*. (COCA, 2012, TV, Crazy)
b. Really? (i.e., do you really think that?)

Deriving from their syntactic independence, parenthetical clauses such as *I think* enjoy flexibility when it comes to positional mobility: they may occur before their host, in medial position, or after their host. The literature on parentheticals (see, e.g., Heine & Kaltenböck 2021: 11 for a recent example)[2] has often noted the controversial status of clause-initial clauses of the type shown in (7), which can be indeterminate between a matrix clause reading (taking a zero clause as complement) and a parenthetical interpretation:

7 MATTHEWS: Yeah, I agree with you, *I think* the moment was important. *I think* he was a little angry, little fatigued. (COCA, SPOK, 2009, NBC Matthews)

In most cases, however, this ambiguity cannot be easily solved; it is only when the sequence is followed by a non-declarative complement, as the interrogative clause in (8) below, or when some hesitation sound (e.g., *uh*, *uhm*) or other fillers intervene, as in (9), that the parenthetical analysis seems to be preferrable (see Kaltenböck 2007: 45; Brinton 2008: 12).

---

2 Heine and Kaltenböck (2021) use the label 'thetical' instead of parenthetical. The term thetical was first used in Kaltenböck, Heine, and Kuteva (2011) precisely on account of the fact that not all parentheticals occur in interpolated position (2011: 856).

8 *I mean*, can you think of any other situation, Pop, when a man gets so close to a woman except when he is actually making love to her? (1981 Rendell, *The Best Man to Die* [BNC]; from Brinton 2008: 12)

9 Gloria: (Chuckles) Well, *I think, um*, Candy's assets will work out very nicely with the trust fund set. (SOAP, 2010, Young and Restless)

In terms of their semantics, the most common parenthetical clauses are those expressing the speaker's evaluation of the proposition encoded by the host (e.g., *I think*, *I guess*, etc.). Parentheticals of this kind are known as 'comment clauses' (cf. Quirk *et al.* 1985: 1112–1118; Brinton 2008). Another common type of clausal parenthetical is that featuring predicates entailing some sort of communication. These parentheticals have sometimes been referred to as 'reporting clauses' (e.g., *he said*, *it is said*) (Kaltenböck, Heine, & Kuteva 2011: 856; see also Vandelanotte 2009: chapter 8; Cichosz 2018: 183). An example of a reporting clause is provided in (10).[3]

10 Deep below*, it is rumored,* lies Saddam's main bunker, (COCA, 1996, MAG, Saddam's Inferno)

Among the wide range of clausal parentheticals available in English, the present chapter is concerned with a type of reporting clauses for which we use the umbrella term *study*-parentheticals. As mentioned above, this parenthetical type features a third person inanimate noun phrase as subject, in contrast to paradigmatic parentheticals, such as *I think* or *you know*, which contain a first- or second-person subject pronoun. The nouns occurring in the construction typically denote some kind of investigation or examination. Besides *study* (example (1) above) and *research* (example (2)), representative illustrations are the nouns *report*, *analysis*, *results,* and *survey*, among others. For our purposes in this chapter, however, we focus solely on the noun *study*, not only because it is the most frequent item occurring in the construction in our preliminary searches, but also because it is the most neutral and general noun in the set from a semantic point of view.

3 Quirk *et al.* (1985: 1114) would subsume the two groups of clausal parenthetical constructions discussed here under 'comment clauses'.

3. Methodology

3.1. Materials

As stated in Section 1, our study is based on data retrieved from the *Corpus of Contemporary American English* (COCA), which contains over one billion words of text, with 25 million words allotted for each year between 1990 and 2019. Due to its sheer size, COCA is ideal to examine low-frequency constructions, such as *study*-parentheticals, in the short diachrony. Moreover, sample searches in the *Corpus of Historical American English* (COHA; Davies 2010) demonstrated the recency of this parenthetical type: the earliest unambiguous examples of the construction are attested in the early 1940s (see example (11)), but they are highly sporadic in the data. It is only from the 1990s onwards that *study*-parentheticals increase in frequency in the COHA material, that is, precisely in the timespan represented in COCA.

11 High-pressure salesmanship and instalment methods have taken millions from the working class annually, *the studies showed*, with 70 percent of all policies lapsing and another 20 percent terminating by surrender. (COHA, 1942, MAG, NewRepublic)

The current version of COCA contains eight different genres, almost equal in terms of size: spoken texts, fiction, popular magazines, newspapers, academic texts, and (after the March 2020 update of the corpus) TV and movie subtitles, blogs, and other webpages. While all other text categories contain material from each year from 1990 to 2019, web pages and blogs were collected in October 2012 and cannot therefore be used for a diachronic study. Rather, they represent a synchronic picture of the language of those genres in that particular year. For our analysis we consider all eight genres in COCA, except when a diachronic perspective is adopted (cf. Sections 4.1 and 4.2 below); in such cases blogs and other web pages are excluded from the tally.

Whenever it was deemed necessary for the sake of illustration, the material from COCA was complemented with examples taken from observation (mostly from the authors' reading of the press and the internet) and with examples retrieved from random searches in various other sources.

3.2. Searches

As mentioned in Section 2, parentheticals may occur in three different positions in relation to their host clause: initially, medially, or finally. However, given the difficulties of teasing out initial parentheticals and matrix clauses in complementation structures (see Section 2 above), the searches carried out in COCA were intended to yield unambiguous examples of parentheticals, that is, those occurring either in medial or in final position. Instances of the type illustrated in (12) were therefore left aside. Although the default interpretation for *most studies suggest* in this example is that of a matrix clause followed by an object clause introduced by the complementizer zero (*mortality continues to decrease*…), in the absence of clear prosodic clues in examples such as these the parenthetical reading cannot totally be discarded.

12 *Most studies suggest* mortality continues to decrease -- or does not increase -- in the 40 -- 70 ng/mL range, but a few studies have shown an increased mortality within this range (COCA, 2012, WEB, hsph.harvard.edu)

The corpus searches, which were conducted in October–November 2022, are reflected in Table 1, together with the target sequences they sought to retrieve. We made use of the wildcard <*> for optional elements in the search strings and of punctuation marks (commas and periods) as indications of pauses to identify parentheticals in medial or in final position. The concordance lists obtained from the searches in Table 1 had to be further refined manually in order to remove false positives. An example is given in (13). As seen here, the search string (*empirical studies reveal*) appears between commas, but is clearly a matrix clause in a complementation structure, taking a dependent *that*-clause as object:

13 Regrettably*, empirical studies reveal*, however, *that* authoritarians are frequently enemies of freedom, antidemocratic, antiequality, highly prejudiced, mean-spirited, power hungry, Machiavellian and amoral. (COCA, 2016, NEWS, Chicago Sunday Times)

Table 1. Searches run in the corpus.

Search string	Target
, (*) (*) [study] (*) VERB , , (*) (*) [study] (*) VERB .	Examples in medial or in final position, where the nominal head *study* appears in the singular or in the plural and is followed by any simple form of the verb.[4] The nominal head may occur on its own or be preceded by a determiner, by a modifier, or by both a determiner and a modifier. The searches also allow for the presence of one intervening word (e.g., an adverb) between the noun phrase and the verb
, VERB (*) (*) [study] , , VERB (*) (*) [study] .	Same as above, but with inversion of noun phrase and verb

The searches also yielded examples of initial structures which, as discussed in Section 2, could be ambiguous between a matrix and a parenthetical interpretation. Consider in this respect (14) and (15) below. In these examples the existence of punctuation marks after the search strings could be interpreted as an indication of prosody, thus lending support to the analysis of the noun phrase + verb sequence in such instances as an initial parenthetical. However, given that ambiguity of initial strings cannot always be solved, we decided to leave all such cases out of the analysis.

14 And, *studies suggest*, overburdened kids are also at risk of muscle strain, fatigue, and back problems later in life. (COCA, 2005, MAG, Prevention)

---

4 Since COCA does not allow automatic searches longer than five words unless a premium license is purchased, we decided to leave out complex verb phrases, such as the present perfect in (i) and (ii) below. Due to the restriction to five words (punctuation marks count as a word), it is not possible to run automatic searches with one or two wildcards in front of the noun (e.g., < , * * [study] _vh _v?n .>).

(i) Fungi can be used to treat a violin to make it sound like a rare Stradivarius, *a study has found*. (COCA, 2012, BLOG, bussorah.wordpress.com)

(ii) Women from low socioeconomic backgrounds are 25 per cent more likely to suffer a heart attack than disadvantaged men, *a major new study has found*. (COCA, 2017, MAG, Medical Xpress)

15 If all consumers became as informed as medical experts, *the study concluded*, national headache-remedy brands would see their sales cut in half. (COCA, 2014, MAG, Atlantic)

After filtering the corpus examples in accordance with the criteria discussed in the preceding paragraphs, the number of valid instances amounted to 974.

4. Analysis of the data

This section is devoted to the analysis of the data from COCA, paying attention to the following issues: (i) the patterns in which the evidential parenthetical construction with the noun *study* is attested in the material (Section 4.1), (ii) the position the parentheticals occupy in relation to their host clauses (Section 4.2), (iii) the different predicates occurring in the parenthetical construction (Section 4.3), and (iv) the distribution of the parentheticals under analysis across genres (Section 4.4).

4.1. Parenthetical patterns

Two major patterns are attested in the COCA material, taking into account the relative position of subject and verb: one without inversion and another with subject-verb inversion. For these two patterns a number of variants have been identified in the data, depending on whether the subject noun phrase features determiners and/or modifiers or, on the contrary, the head noun *study* stands alone. Table 2 shows the distribution of patterns and variants of *study*-parentheticals in COCA.

Table 2. Parenthetical patterns with the noun *study* and their variants in COCA.

Pattern	Variants		Nº of tokens (%)	Total
Non-inversion	**I**	[study] VERB	145 (14.9%)	930 (95.5%)
	II	* [study] VERB	565 (58.0%)	
	III	* * [study] VERB	220 (22.6%)	
Inversion	**IV**	VERB * [study]	13 (1.3%)	44 (4.5%)
	V	VERB * *[study]	31 (3.2%)	
Total				974

As the data in Table 2 show, the most recurrent pattern in our American English material is the one without inversion, which accounts for over 95% of the instances. The pattern with inversion, on the other hand, is attested in less than 5% of the cases. In what follows, the variants of these two patterns are illustrated and described in some detail.

Variant I ([study] VERB), which represents 14.9% of all cases, is the least syntactically complex of the five sequences included in Table 2. As shown in examples (16) and (17), it features the noun *study* occurring on its own (with no determiners or modifiers) immediately followed by the verbal predicate, which appears almost invariably in the present tense in our data (141 out of 145 examples; 97.2%). The noun can be inflected for the singular (Variant I.a; 24 exx.), as in (16), or for the plural (Variant I.b; 121 exx.), as in (17).

16 The AJC headline on Monday read, "HPV shots don't make girls promiscuous, *study says*." (COCA, 2012, WEB, clatl.com)

17 Four to eight people grieve intensely for each suicide, *studies show*. (COCA, 2013, NEWS, Pittsburg Post Gazette)

Variant II (* [study] VERB) is the most recurrent sequence in our American English material, accounting for more than half of the instances (58%). Variant II shows a wider range of variability than Variant I, both in the noun phrase and in the verb phrase. Concerning the latter, in contrast to Variant I, the ratio of preterite tense forms to present tense ones is here 2:1 (68.7% vs. 31.3%). As regards the preverbal noun phrase, the noun *study* is preceded by either a determiner (II.a) or a modifier (II.b). Variant II.a is by far the most common, amounting to 540 instances, which represent 95.6% of the total in Variant II. The list of determiners available here is quite large, comprising (in order of decreasing frequency) articles (*a*, *the*; 92.8%), quantifiers (*many*, *most*, *one*, *several*, *some*, *another*, *other*; 4.3%), demonstratives (*that*, *this*, *these*; 1.8%), and possessives (*her*, *his*, *our*; 1.1%). In this set of prenominal elements, the definite article *the* stands out as the preferred choice (475 examples; 88% of the instances belonging to this variant), followed by the indefinite article *a* (26 occurrences; 4.8%); the remaining determiners in the set represent minor options.

Example (18) serves as an illustration of Variant II.a with the definite article *the*, whereas (19)–(22) show the use of other determiners preceding the noun *study*.

18 About a third of all strawberry samples had at least 10 pesticides, *the study found*. (COCA, 2018, NEWS, Chicago Sunday Times)

19 A brief workplace intervention using mindfulness techniques has a positive effect on an individual's sleep quality and sleep, processes that are vital for daily recovery from the demands of work, *a study finds*. (COCA, 2015, ACAD, OccupationalHealth)

20 Not only the quality, but the number of evaluations matter. The more the evaluations, *many studies show*, the more accurately they average out. (COCA, 1994, NEWS, Washington Post)

21 Nor, *this study reveals*, did wartime integration involve the two nations' citizenries as a whole. (COCA, 2004, ACAD, American Review of Canadian Studies)

22 Children are being exposed to unacceptable levels of the chemical, *our study found*. (COCA, 1999, MAG, Consumer Reports)

As mentioned above, in Variant II.b the prenominal slot is occupied by a modifier, rather than by a determiner. Only 25 examples out of 565 of occurrences of Variant II in the data belong to this type (4.4% of the total of instances). These premodifiers add descriptive information on the head noun *study*, restricting or qualifying it. Premodifying items in this pattern include several of the types of premodification identified by Quirk *et al.* (1985: 1321–1337), specifically adjectives, nouns, and genitives. The latter two types are only sporadically attested in our COCA material: 6 (24%) and 4 occurrences (16%) of modification by nouns (both common and proper nouns) and by genitives, respectively. Examples (23)–(24) illustrate the use of nouns as premodifiers of *study*, while (25) shows a genitive in premodifying function.

23 "That's right. With so many people dying from overwork, I thought we should do our part to stay healthy. Having a pet around the office helps you to relax. It improves office efficiency, *government studies show*," Nakamura says. (COCA, 1996, FIC, Audrey Hepburn's Neck)

24 New car quality, in fact, is at an all-time high, but complaints about electronic accessories have stalled automakers' efforts to improve their quality scores, *J. D. Power studies show*. (COCA, 2003, NEWS, USA Today)

25 Personality traits such as being receptive to the unfamiliar may directly explain why life for some centenarians is still as enjoyable and independent as it was in their younger years, *Poon's study finds*. (COCA, 2011, MAG, Psychology Today)

As expected, most of the premodifiers in our data are adjectives (15 out of 25 examples; 60%): *academic* (2), *behavioural*, *independent*, *latest*, *long-term*, *military*, *new* (4), *recent* (3), and *secret*. Note the overall positive semantic prosody resulting from the combination of the noun *study* with positive evaluative adjectives such as those conveying the idea of novelty (e.g., *latest*, *new*, *recent*), as in (26), or of prestige (e.g., *academic* in (27)).[5]

26 Those warmer air temperatures are significantly boosting soil temperatures in many regions, *new studies show*. (COCA, 2007, MAG, Science News)

27 But the majority of mergers**,** *academic studies show***,** fail to deliver for shareholders (COCA, 2005, SPOK, CNN_Next)

This also applies to Variant III (* * [study] VERB; 22.6% of the total), where the preverbal noun phrase contains both a determiner and a modifier. The most frequent sequence in this variant is that which combines the indefinite article *a* and the adjectival premodifier *new* (i.e., *a new study*), as in (28), which is found in COCA in 132 instances (60% of the total of occurrences of Variant III in the data). In addition to *new*, the premodifiers attested here mostly convey the ideas of novelty and recency (e.g., *recent*, *first-of-its-kind*) or of an authoritative study (e.g., *Harvard*, *MIT*, *Oxford*, *NASA*). On other occasions, they indicate the source (e.g., *federal*, *Canadian*, *Japanese*) or the date (e.g.,

5 Positive evaluative adjectives modifying the noun *study* can be seen as 'epistemic extensions' (Aikhenvald 2004: 6), with the speaker underlining the validity of their statement by referring to a reliable or trustworthy source. We thank an anonymous reviewer for this insightful suggestion.

2012, *2013*) of the study being referred to in the parenthetical. See (29)–(31) below as illustrations.

28 The action of a gene called *ATDC* is required for the development of pancreatic cancer, *a new study finds.* (COCA, 2019, MAG, Medical Xpress)

29 Fair-haired folks and redheads have a higher risk of Parkinson's than those with dark hair, *a Harvard study found*. (COCA, 2013, MAG, Men's Health)

30 It also provides motivational messages -- which, *a British study says*, may increase your chances of being smoke-free at the 6-month mark. (COCA, 2012, MAG, Men's Health)

31 Over forty percent of them, *a 2005 study showed*, descend from just four "founding mothers" having Middle-Eastern-profile mitochondrial DNA. (COCA, 2008, ACAD, Commentary)

In contrast to the three parenthetical variants just discussed, Variant IV (VERB * [study]; 1.3% of the total) and Variant V (VERB * * [study]; 3.2%) illustrate so-called 'quotative inversion' (see, among others, Collins & Branigan 1997; Cichosz 2018), a type of inversion which is particularly frequent in written registers (Joseph 2000: 312). In Variant IV (shown in (32)) the noun phrase contains only a determiner (mostly the definite article *the*, which occurs in 10 of the 13 instances of this type; 76.9%) or the modifier *recent* preceding the noun *study* (33). By contrast, in Variant V the noun phrase contains both a determiner and a modifier as prenominal elements. In the latter type, the most recurrent combination is that of the indefinite article *a* with an adjective, mostly *new*. Example (34) illustrates the sequence *a new study* following the verb, which occurs 12 times in the COCA material (38.7% of the examples belonging to Variant V), while (35) provides an example of the combination *a*/*one*/*two recent*, which is attested in the data on 7 occasions.

32 The patient purchased the medicine over the internet, *said the study*. (COCA, 2018, NEWS, USAToday)

33 Diners using big forks ate less of a * large serving than those using small ones, *says recent study*. (COCA, 2015, MAG, Saturday Evening Post)

34 After a hectic day - when the lure of the drive-thru is most magnetic - your metabolism tanks, *says* a new *study*. (COCA, 2014, MAG, Prevention)

35 "What are the scientists saying? Gatorade works," one print ad says. (Sports drinks don't improve athletic performance, *concludes* *a recent* *study*, but do replenish needed minerals). (COCA, 1990, NEWS, USA Today)

Interestingly, the vast majority of examples with inversion in COCA (36 out of 44) are attested in formal written genres (academic texts, newspapers, and, most notably, popular magazines). In addition to the potential correlation between Variants IV and V and a higher degree of formality, inversion can also be explained by the principle of end-weight. By way of illustration, consider (36) and (37) below, which show a relative clause (Quirk *et al.* 1985: 1244 ff.) and an *-ed* participle clause (Quirk *et al.* 1985: 1264–1265), respectively, as postmodifiers of the noun *study*.

36 Almost 11 million of these uninsured Americans are children age 18 or younger, *said the study*, which was conducted by the department's Agency for Health Care Policy and Research. (COCA, 1997, MAG, San Francisco Chronicle)

37 The drop-off in ordering screenings could be linked to "decision fatigue," the mental burn-out that makes it harder for people to think through big decisions the farther they get into their day, *said the study*, published Friday in JAMA Network Open. (COCA, 2019, MAG, MarketWatch)

In order to provide a snapshot of the diachronic development of the five parenthetical variants described in the preceding paragraphs, Table 3 shows their distribution across time from the 1990s to the year 2019. For practical purposes, we have subdivided COCA into three subperiods, each covering a decade. Since, as mentioned in Section 3.1 above, the material for blogs and other webpages in COCA was collected for just one year (2012), examples belonging to these two genres have been excluded from the count.

Table 3. Distribution of parentheticals with the noun *study* across time in COCA (excluding blogs and other webpages).

	Variant I	Variant II	Variant III	Variant IV	Variant V	Total
1990–1999	36 (26.3%)	114 (27.5%)	28 (17.9%)	3 (27.3%)	5 (20%)	186 (25%)
2000–2009	52 (37.9%)	107 (25.8%)	28 (17.9%)	2 (18.2%)	10 (40%)	199 (26.7%)
2010–2019	49 (35.8%)	194 (46.7%)	100 (64.2%)	6 (54.5%)	10 (40%)	359 (48.3%)
Total	137	415	156	11	25	744

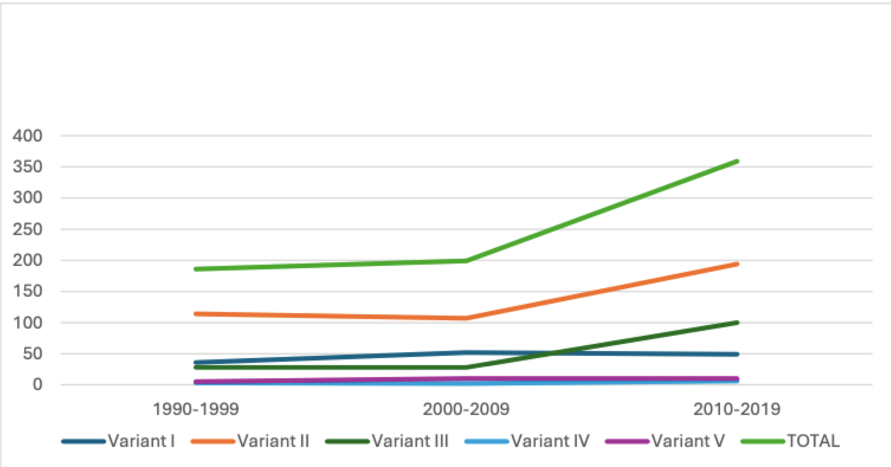


Figure 1. Parenthetical variants with the noun *study* across time in COCA (excluding blogs and other webpages).

As seen in Table 3, all five variants are attested in the three subperiods in COCA. The table also shows that parentheticals with the noun *study* have been on the rise over the last 30 years in American English. Note that the number of the constructions under analysis almost doubles from subperiod 1 (186; 25%) to subperiod 3 (359; 48.3%). Figure 1 shows that such an increase is especially noticeable from the 2000s to the 2010s (light green line) and particularly marked for Variants II and III (orange and dark grey lines).

4.2. Position of *study*-parentheticals

The distribution of our corpus instances (excluding material from the web and blogs) according to position is provided in Table 4.[6] As shown here, medial position occurs in just one sixth of the instances (125 exx., 16.8%).[7] On the other hand, final position is overall the preferred choice for the parenthetical constructions under analysis (619 exx., representing 83.2% of the total of examples),[8] in line with Biber *et al.*'s (1991: 923) findings for reporting clauses in fiction and news. Final-position parentheticals also predominate over those in medial position in all three individual subperiods in COCA. Note the slight decrease of medial position from subperiod 1 (22.6%) to subperiod 3 (13.4%).

Table 4. Position of parentheticals with the noun *study* across time in COCA (excluding blogs and other webpages).

	Medial position	**Final position**	**Total**
Subperiod 1: 1990–1999	42 (22.6%)	144 (77.4%)	186
Subperiod 2: 2000–2009	35 (17.6%)	164 (82.4%)	199
Subperiod 3: 2010–2019	48 (13.4%)	311 (86.6%)	359
Total	125 (16.8%)	619 (83.2%)	744

Final position is attested for all the parenthetical patterns described in Section 4.1. Examples (16), (22), (29), and (34) above show Variants I-IV, while (38) below illustrates Variant V.

38 How to make a food instantly higher in fiber, lower in fat, and less caloric? Slap an "organic" sticker on it, *says a Cornell study*. (COCA, 2011, MAG, Good Housekeeping)

---

6 As discussed in Section 3.2, initial sequences have been excluded from the analysis given their ambiguous status between matrix and parenthetical clauses.

7 In this respect, the behaviour of *study*-parentheticals differs from that of standard reporting clauses, for which medial position is not so marginal (Quirk *et al.* 1985: 1022).

8 Final position is also the unmarked position for *study*-parentheticals occurring in the web and in blogs in COCA: 114 exx. (89.8) and 93 exx. (90.3%), respectively.

In turn, medial position is restricted in our data to examples without inversion (Variants I-III). Our parentheticals in medial position occur mostly between the subject and the verb (80 exx., representing 64% of the relevant instances), as in examples (30) and (31) above. This is precisely one of the "weak spots" identified by Kaltenböck (2007: 43) in which parentheticals can easily be inserted. The second most common pattern is that illustrated in (39), where the parenthetical intervenes between an adjunct and the subject. Although parentheticals most frequently occur at the boundaries of phrases, they can sporadically be found in positions in which they "cut across phrasal constituents" (Kaltenböck, Heine, & Kuteva 2011: 869). Such disfavoured positions (cf. Kaltenböck 2007: 42 ff.; Brinton 2008: 8 and references therein) include, for instance, between an adjective and its complement, as in (40).

39 During pregnancy, *our studies suggest*, the trophoblast functions as a guest conductor. (COCA, 2007, MAG, Natural History)

40 They're also more likely, *studies show*, to drive up an exit ramp and into a head-on collision. (COCA, 2007, NEWS, Houston)

4.3. Predicates across parenthetical variants

A look at the verbal predicates occurring in the parentheticals under analysis also affords interesting insights into the nature of the construction. Table 5 provides the list of predicates attested in each of the five variants discussed in Section 4.1 above. As can be gathered from this list, all the verbs identified in the COCA material convey, in different ways, the source of information; in other words, they express evidential meaning.[9] Moreover, all these verbs belong to particular predicate types in Noonan's (1985) and Rudanko's (1989) well-known typologies of complement-taking predicates:

---

9 *Study*-parentheticals, with their evidential function, are connected with one of the types of reporting clauses studied by Vandelanotte (2009), namely his 'Subjectified Distancing Indirect Speech or Thought' (S-DIST), which are characteristic of news reports (Vandelanotte 2009: 325).

(i) Utterance predicates (Noonan 1985: 110–113), i.e., those which denote a transfer of information, describing the manner of transfer; e.g., *say*, *suggest*, *state*, *report*, etc.

(ii) Predicates of knowledge and acquisition of knowledge (Noonan 1985: 118–119), which describe the state or the manner of acquisition; e.g., *find*, *read*, *conclude*.

(iii) Demonstration predicates (Rudanko 1989: 77); e.g., *show*, *reveal*.

Table 5. Predicates attested in COCA in the five parenthetical variants with the noun *study*.

Variant I	**Variant II**		**Variant III**	**Variant IV**	**Variant V**
find (17) *indicate* (4) *reveal* (1) *say* (12) *show* (94) *suggest* (17)	*add* (1) *argue* (4) *assert* (2) *calculate* (1) *claim* (2) *conclude* (40) *confirm* (1) *contend* (1) *estimate* (2) *explain* (1) *find* (201) *indicate* (6) *maintain* (2)	*note* (14) *observe* (1) *predict* (3) *read* (4) *recommend* (1) *report* (19) *reveal* (13) *say* (176) *show* (45) *state* (2) *suggest* (20) *warn* (3)	*conclude* (8) *contend* (1) *declare* (1) *find* (115) *indicate* (7) *predict* (1) *report* (3) *reveal* (1) *say* (33) *show* (9) *state* (1) *suggest* (40)	*conclude* (1) *praise* (1) *report* (2) *say* (8) *suggest* (1)	*conclude* (1) *find* (7) *note* (1) *report* (4) *say* (14) *show* (1) *suggest* (3)

Interestingly, six of these predicates appear repeatedly in the non-inverted parenthetical pattern (Variants I, II, and III), namely *find*, *indicate*, *reveal*, *say*, *show*, and *suggest*.[10] Most importantly, these six predicates represent 86.8% of all occurrences of the parentheticals in the data, with *find* (340 exx.; 34.9%), *say* (243 exx.; 24.9%), and *show* (149 exx.; 15.3%) taking the lead. These three predicates are, semantically speaking, the most general in their respective categories. Thus, for example, *say* is the hyperonym of utterance verbs and *show* is the hyperonym of predicates of demonstration. Note also that the six predicates are the only ones attested in our Variant I, which

10 Some of these predicates (*find*, *say*, *show*, and *suggest*) also occur sporadically in the pattern with inversion (Variants IV and V).

contains the lowest number of types. By contrast, Variants II and III show a much wider range of types, including predicates with a more specific kind of meaning. For example, within the group of utterance predicates, together with the hyperonym *say*, we find *read*, *claim*, *explain*, *warn*, etc., i.e., predicates which also convey the manner of information. The predicate *say* is also the most frequent in the inversion pattern (8 and 14 exx. in Variants IV and V, respectively). The dominance of *say* here suggests the existence of a link with quotative inversion (cf. Section 4.1 above).

4.4. Distribution of *study*-parentheticals across genres

As mentioned in Section 3, COCA contains material belonging to eight different genres, nearly evenly divided in size. Table 6 shows that *study*-parentheticals are attested in all text categories represented in the corpus, though with marked differences among them.

Table 6. Distribution of *study*-parentheticals across genres in COCA.

Genre	Nº of tokens (%)
ACAD	45 (4.6%)
BLOGS	103 (10.6%)
FIC	3 (0.3%)
MAG	331 (34%)
MOV	1 (0.1%)
NEWS	355 (36.4%)
SPOK	9 (0.9%)
WEB	127 (13.1%)

Perhaps the most salient feature of the distribution of the parentheticals at issue here is their strong correlation with the written language, in particular with the press category, which comprises popular magazines (MAG) and newspapers (NEWS). Note that over 70% of the relevant corpus instances belong to these two genres (34% and 36.4%, respectively). Academic texts (ACAD) show a lower proportion of the parenthetical constructions under analysis (4.6%). On the other hand, *study*-parentheticals are far less common in those genres with a higher degree of speech-likeness, such as blogs (10.6%), and are practically non-existent in those categories which can be taken to represent speech, such as spoken (0.9%), fiction (0.3%), and TV and movie subtitles (0.1%). The strong connection of our parentheticals with the press category and, therefore, with the written language brings them close to Kaltenböck, Heine, and Kuteva's (2011: 856) 'reporting parentheticals' and Vandelanotte's (2009) S-DIST clauses, in contrast with other types of parentheticals, which are considered a typical feature of conversation and of written texts which resemble speech (Thompson & Mulac 1991; Biber *et al.* 1999: 197, 862–865, 983; López-Couso & Méndez-Naya 2014).

Interestingly, within the press category in COCA, our examples typically occur in texts which can be described as popular science (pop-science or popsci), a written genre concerned with the interpretation of science intended for a general audience which has experienced a boost over the last few years. Pop-science, which can be produced by either scientists or non-scientists, represents a bridge between scientific literature and journalism and aims at capturing the methods and accuracy of science, while making it more accessible to the general public. By contrast, academic writing is written by scientists and addressed to their peers. Even though popular science shares some linguistic features with academic writing (e.g., the use of epistemic modals indicating different degrees of likelihood; see Biber & Conrad 2009), in pop-science reference to the authorial sources can be vaguer than in academic texts, where precise mention of the sources is mandatory. As Chafe (1986: 268) puts it, "[i]n academic writing knowledge obtained through language is indicated with the formal device of citing a reference or personal communication". The difference between these two ways of conveying evidential meaning in scientific literature and in pop-science is illustrated in examples (41) and (42). While in (41) the evidential source (*Hilpert 2013: 133*) is expressed in its "most precise and deliberate form" (Chafe 1986: 269), in (42) the reference is undoubtedly more imprecise, though the use of the premodifier *Harvard*, a very prestigious academic institution, in the noun phrase *a Harvard study* serves to lend credibility to the statement made in the host clause, providing the source of evidence with a high degree of authority (see Section 4.1 above).

41 *Hilpert (2013: 133)* suggests that one of the Baayen's hapax-based measures of productivity, namely expanding productivity, could be used. (Hilpert 2017: 58)

42 A man with a good shot at getting a girlfriend may live three months longer than men who outnumber local women, *a Harvard study finds*. (COCA, 2011, MAG, PsychToday)

The growing importance of popular science among the general public and the strong connection of our *study*-parentheticals with pop-science become manifest in the existence of a popsci webpage precisely entitled *StudyFinds* (https://studyfinds.org/), of *The New York Times Magazine* popsci column *Studies show*. *What medical research really says* (https://www.nytimes.com/column/studies-show), and of the Wiley podcast called *This Study Shows* (https://audioboom.com/channels/5003898-this-study-shows). The latter is presented in the following way: “The only way for research to change the world is if the world knows and cares about it. In This Study Shows, a podcast from Wiley, we explore how to connect research with emotions and experiences and transform the way science is shared”.

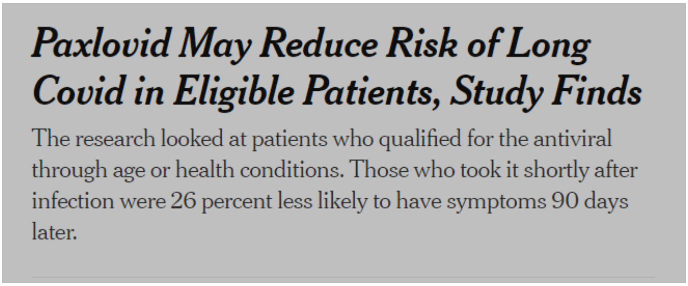
Another interesting feature of the parentheticals under examination is that they commonly occur in headlines, as shown in (43) and (44), where we find an explicit reference to a headline (see also example (16) above).

43 Greenland Ice Melt Reaching a Tipping Point, *Study Says*. (COCA, 2019, NEWS, Minneapolis Star Tribune)

44 But the Times apparently wasn’t interested in the headline, “*Bartenders Get Flu and Recover, Study Says*.” Such a headline would also not go far in justifying a smoking ban. (COCA, 2001, MAG, ConsumResrch)

Outside COCA, examples of this kind are not difficult to find, as shown by (45) and (46), taken from *The New York Times* (7 November 2022) and from *The Guardian* (20 September 2023), respectively.

45



Paxlovid May Reduce Risk of Long Covid in Eligible Patients, Study Finds

The research looked at patients who qualified for the antiviral through age or health conditions. Those who took it shortly after infection were 26 percent less likely to have symptoms 90 days later.

46

Fathers have ‘unique effect’ on children’s educational outcomes, study finds

Research claims children whose fathers read and play with them see a ‘small but significant’ increase in their educational attainment

Interestingly, it is in headlines where Variant I is normally found. Headlines serve as clickbaits, to spur the reader’s curiosity which is later quenched in the main body of the text. As illustrated in the examples below, when the topic advanced in the headline is taken up in the initial part of the article itself, the more complex Variant II (example (47) below) or Variant III sequences (examples (48) and (49)) are typically attested.

47 ‘Gigantic’ power of meat industry blocking green alternatives, *study finds*
[…] The “gigantic” power of the meat and dairy industries in the EU and US is blocking the development of the greener alternatives needed to tackle the climate crisis, *a study has found*. (*The Guardian*, 18 August 2023) [headline: Variant I.a; body of the article: Variant II.a]

48 Gulf Stream could collapse as early as 2025, *study suggests*
[…] The Gulf Stream system could collapse as soon as 2025, *a new study suggests*. (*The Guardian*, 25 July 2023) [headline: Variant I.a; body of the article: Variant III]

49 People who took the antiviral drug Paxlovid within a few days after being infected with the coronavirus were less likely to be experiencing long Covid several months later, *a large new study found*. (*The New York Times*, 7 November 2022) [corresponding to the headline in (45) above]

Note that in the body of the article itself the source of information can also be conveyed by means of an adjunct (50) or by a main clause (51).

50 Erythritol, an ingredient in stevia, linked to heart attack and stroke, *study finds*
A sugar replacement called erythritol — used to add bulk or sweeten stevia, monkfruit and keto reduced-sugar products — has been linked to blood clotting, stroke, heart attack and death, <u>according to a study</u>**.** (https://edition.cnn.com/2023/02/27/health/zero-calorie-sweetener-heart-attack-stroke-wellness/index.html)

51 Humans 'may need more sleep in winter', *study finds*. <u>Research shows</u> people get more deep REM sleep than in summer, and may need to adjust habits to season. (*The Guardian*, 17 February 2023)

As is well known, headlinese is characteristically condensed and can be regarded as "stylistically 'deviant'" (Molek-Kozakowska 2017: 897).[11] One of its most salient linguistic features is its fragmentary nature, as evinced by, for instance, the lack of determiners in cases where they would normally be required. As Moncomble (2018) puts it, absence of determiners is "part and parcel of the production of headlines". In fact, all our examples of Variant I.a (that featuring the singular noun *study* on its own) occur in headlines. Absence of the determiner is also attested in the inverted pattern, specifically in

11 On the pragmatics of headlines, see the 2024 special issue of *Journal of Pragmatics*, edited by Rita Finkbeiner.

Variant IV (VERB * [study]), but only in one example in the data (see example (33); *says recent study*), which is also found in a headline.

Another feature of headline language is the predominant use of the simple present to refer to past events, where a present perfect or even a preterite would be preferred in ordinary language. This is clearly seen in example (47) above (headline: *finds* vs. text: *has found*). Note that, as mentioned in Section 4.1, Variant I almost invariably shows a present tense verb phrase.

Despite the close association of Variant I with headlines, the structure may also occur in the body of the text, as illustrated in example (52). Such examples, however, feature the noun *study* in the plural (i.e., they correspond to Variant I.b). By contrast, all instances of singular *study* (Variant I.a) occur in headlines.

52 # For those born in that first wave and now entering adulthood, it's a tough, uncertain future. Some, like Abenes, will go to college or find jobs and eventually move out on their own. # But most will not, *studies show*. Most will continue to live at home and will, at best, find part-time, minimum-wage work -- or no work at all. (COCA, BLOG, 2012, specialedpost.com)

5. Discussion

As may have become apparent from the examples provided in the preceding section, *study*-parentheticals have a clear evidential function, that is, they convey the source of information the speaker/writer has for their statement. Of the three major categories of evidentials, namely direct perception, inference, and hearsay (Mélac 2022: 333), *study*-parentheticals present the information in the proposition as hearsay, so that the speaker's commitment is not really at stake. Hearsay evidentials of this kind fit in well with pop-sci texts, which portray the evidential source as serious and reliable, though not in such a precise manner as in academic writing. This is specially the case in examples containing modifiers referring to prestigious institutions (e.g., *Oxford*, *NASA*) or indicating that the study is up-to-date (e.g., *new*, *recent*) (see Section 4.4 above).

As far as their development is concerned, *study*-parentheticals seem to illustrate a similar pathway to that described in the literature for other clausal parentheticals. The emergence of parenthetical clauses has been explained by invoking several processes of change, among them grammaticalization, pragmaticalization, lexicalization, and co-optation (see Heine & Kaltenböck 2021: 4–7 for a good summary of the different proposals). Of these processes, grammaticalization stands out as the most widespread account and is the one we will adopt in the discussion that follows.

In a highly influential paper, Thompson and Mulac (1991) proposed the so-called matrix clause hypothesis, according to which epistemic parentheticals of the type *I think* originate in matrix clauses taking *that*/zero-clauses as complements. The increase in frequency of the complementizer zero is regarded as the trigger for the syntactic reanalysis of the matrix as a parenthetical, which, once reanalysed, can also occur in medial and in final position. A related grammaticalization account is the one posited by Boye and Harder (2007; 2012), who consider not only the changes in syntactic function, but also in discourse prominence, arguing that the development from a matrix clause to a parenthetical clause requires a previous usage reanalysis in which the dependent clause acquires primary discourse status (see Section 2). In fact, in line with Boye and Harder (2007), Kearns' (2007) study of *I think* suggests that semantico-pragmatic changes are precisely the catalyst for structural changes.

Indeed, the development of clausal parentheticals illustrates semantic, pragmatic, and structural indicators of grammaticalization. Among the former, clausal parentheticals show semantic bleaching and non-addressability (Boye & Harder 2007), together with the acquisition of (inter)subjective functions. As far as structural concomitants of grammaticalization are concerned, they typically show decategorialization and fixation (see López-Couso & Méndez-Naya 2014 and 2021 on *like*-parentheticals and on *chances are*- and *odds are*-parentheticals, respectively).

In an account compatible with grammaticalization, Heine and Kaltenböck (2021: 5) suggest that the development of (paren)theticals primarily involves a process of co-optation, whereby "main clauses without their complements […] are taken out of sentence grammar and transferred to the metatextual level of discourse processing to give rise to comment clauses and discourse markers in general". Co-opted elements can conventionalize and grammaticalize over time, giving rise to different types of (paren)theticals (Heine & Kaltenböck 2021: 7–8): (i) instantaneous theticals, which can be created spontaneously, maybe not to be uttered again (cf. also Quirk *et al.* 1985: 1114 on “freely constructed” comment clauses); (ii) constructional theticals, which have “a schematic format and function”, while involving a certain amount of variability (comment clauses are examples of this parenthetical type); and (iii) formulaic theticals, which are invariable (e.g., discourse markers, interjections, politeness formulae, etc.).

In our view, the data discussed in Section 4 suggest that *study*-parentheticals are moving along a grammaticalization pathway, as is also the case with other reporting clauses found in news reports (see Vandelanotte 2009: 325, 329). The three major parenthetical variants (that is, those illustrating the non-inverted pattern) can be placed along a cline of complexity which represents increasing grammaticalization, in line with the general pathway attested for other parentheticals. As shown in Figure 2, Variant III, which enjoys the highest degree of variability, would be placed at the leftmost end of the cline, while Variant I.a, the least syntactically complex, would be located at the right end of the continuum. Variant II.a is taken to be more constrained than Variant II.b, since the prenominal position is taken by a grammatical word (a determiner, mostly the definite article), while in Variant II.b various types of open-class modifiers (adjectives, nouns, etc.) can occupy the prenominal slot.

Higher degree of complexity **Lower degree of fixation**								**Lower degree of complexity** **Higher degree of fixation**
Variant III *a new/recent study finds*	>	Variant II.b *recent studies find*	>	Variant II.a *the study finds*	>	Variant I.b *studies find*	>	Variant I.a *study finds*

Figure 2. Complexity/fixation cline of *study*-parentheticals.

The growing grammaticalization of Variant I ([study] VERB) is supported by a number of features:

(i) The restricted range of predicates available in this variant (*find*, *indicate*, *reveal*, *say*, *show*, and *suggest*), all of which are fairly general from a semantic point of view, with *find*, *say*, and *show* (the hyperonyms of their respective semantic classes, namely knowledge and acquisition of knowledge, utterance, and demonstration) standing out in terms of frequency. Variants II and III, by contrast, feature a larger number of types, including verbs with more specific meanings conveying the manner of transfer of information, of knowledge acquisition, or of demonstration (see Section 4.3). Moreover, the range of predicates occurring in Variant I.a is even more constrained than in Variant I.b, thus pointing at specialization (Hopper 1991: 22). Only four predicates occur in Variant I.a (*find*, *say*, *show*, and *suggest*), with *find* present in over half of the tokens (13/24 exx.).

(ii) The almost complete restriction of the verb phrase to the present tense form (see Section 4.1), thus suggesting morphosyntactic fixation. The extensive literature on parentheticals has proved that the more grammaticalized a given parenthetical type, the more likely for it to show a verb phrase inflected for the present tense (Brinton 2008: 2; van Bogaert 2010: 401 and the references therein).

(iii) The disallowance of insertion of intervening material (e.g., an adverb), while this is perfectly possible in Variants II and III, as shown in (53), an example of Variant II. This points at an increasing bonding between the noun phrase and the verb phrase in Variant I. Increasing bonding is a well-known feature of grammaticalization, namely fusion (Brinton & Traugott 2005: 27).

53 According to the CDC study, the most common wound after head injuries involved upper limb fracture, followed by lower limb fractures. About 14 rides out of every 100,000 lead to injuries, *the study reportedly found*. (COCA, 2019, MAG, Gizmodo)

(iv) The absence of a determiner in Variant I.a, with the noun *study* in the singular, could be suggestive of decategorialization (Hopper 1991: 22, 30 ff.), in as much as the parenthetical type *study finds* lacks a determiner which would otherwise be mandatory, making the noun phrase somewhat deviant and defective. However, the omission of the determiner here could also be interpreted as genre-licensed (headlinese) (see Section 4.4 above).

In view of the above, it can safely be concluded that the parenthetical *study finds* is on its way towards becoming a formulaic phrase, with a growing restriction to just one predicate (*find*) (Hopper 1991: 22, 25). In Kaltenböck, Heine, and Kuteva's (2011) classification of thetical elements, Variants II and III would qualify as constructional theticals, while Variant I.a is moving towards becoming a formulaic thetical in a particular type of discourse.

6. Concluding remarks

This chapter has examined the development and use of an evidential strategy in the recent history of American English, whereby clausal parenthetical structures containing a noun phrase with *study* as head in combination with a verb phrase are employed to indicate hearsay, thus illustrating a type of reporting clause. The data from COCA (1990–2019) and complementary material have allowed us to identify a number of interesting trends:

(i) While s*tudy*-parentheticals occur sporadically from the 1940s, they only gain momentum from the 1990s, experiencing a marked increase in frequency over the last 30 years.

(ii) Although final position is by far the preferred choice for *study*-parentheticals, the construction is also attested medially, mostly at the boundaries of phrases (e.g., between the subject and the verb), but also, though only occasionally, across phrasal constituents (e.g., between an adjective and its complement).

(iii) In contrast to other parenthetical types, the *study*-parentheticals examined here are closely associated with the written mode of expression, in particular with journalistic discourse (the text-categories NEWS and MAG in COCA). More specifically, these parentheticals frequently occur in texts which can be characterized as pop-science. In texts of this kind *study*-parentheticals present the hearsay evidence as coming from a reliable, authoritative source, even though in a less precise manner than that required in academic writing.

(iv) *Study*-parentheticals show a certain degree of variability, appearing in the COCA material in two patterns and five different variants. The pattern featuring quotative inversion (e.g., *says a* (*new*) *study*) is only sporadically found. In the non-inverted pattern, the most frequent in the data is Variant II (* [study] VERB), in particular that showing a determiner in prenominal position (Variant II.a), followed by Variant III (* * [study] VERB) and Variant I ([study] VERB), in order of decreasing frequency.

(v) The major three variants in which *study*-parentheticals are attested in COCA can be placed along a cline of complexity/fixation, with Variant I.a, with the singular noun *study*, on its way to becoming a fossilized expression: (a) the predicates in Variant I.a are limited to just four verbs, namely *find*, *say*, *show*, and *suggest*, with *find* occurring in over 50% of the relevant cases (specialization); (b) the verb is almost invariably marked for the present tense (fixation); and (c) intervening material is disallowed between the noun and the verb (fusion). Finally, the lack of a determiner in Variant I.a, exclusively found in headlines in our data, would be a clear sign of decategorialization if this variant were ever attested outside headlines.

(vi) In Kaltenböck, Heine, and Kuteva's (2011) taxonomy, Variants I, II, and III would qualify as constructional theticals, with the parenthetical *study finds* coming close to a formulaic thetical which has become a staple in pop-sci headlinese, as testified by the existence of a website called *StudyFinds* (see Section 4.4 above).

The evidence presented in this chapter suggests that *study*-parentheticals represent an appealing but understudied type of reporting parenthetical construction, which clearly deserves investigation. One of the possible avenues for further research implies extending the analysis to other semantically related nouns, including *research*, *report*, or *survey*, as well as to animate nouns such as *expert*, *scientist*, or *researcher* in order to check whether animacy plays a role in the behaviour of these evidential parentheticals.

Bibliography

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
Biber, Douglas & Susan Conrad 2009. *Register, Genre and Style*. Cambridge: Cambridge University Press.
Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
Boye, Kasper & Peter Harder 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language* 31 (3): 569–606. https://doi.org/10.1075/sl.31.3.03boy
Boye, Kasper & Peter Harder 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88 (1): 1–44. https://doi.org/10.1353/lan.2012.0020
Boye, Kasper & Peter Harder 2021. Complement-taking predicates, parentheticals and grammaticalization. *Language Sciences* 88: 101416. https://doi.org/10.1016/j.langsci.2021.101416
Brinton, Laurel J. 2008. *The Comment Clause in English. Syntactic Origins and Pragmatic Development*. Cambridge: Cambridge University Press.
Brinton, Laurel J. & Elizabeth C. Traugott 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.
Chafe, Wallace 1986. Evidentiality in English conversation and academic writing. *Evidentiality: The Linguistic Coding of Epistemology*, ed. by Wallace Chafe & Johanna Nichols. 261–272. Norwood, NJ: Ablex.
Cichosz, Anna 2018. Parenthetical reporting clauses in the history of English. The development of quotative inversion. *English Language and Linguistics* 23 (1): 183–214. https://doi.org/10.1017/s1360674317000594
Collins, Chris & Phil Branigan 1997. Quotative inversion. *Natural Language and Linguistic Theory* 15 (1): 1–41. https://doi.org/10.1023/A:1005722729974
Davies, Mark 2008–. *COCA: Corpus of Contemporary American English*. Available online at https://www.english-corpora.org/coca/.
Davies, Mark 2010–. *The Corpus of Historical American English (COHA): 400 million words, 1810–2009*. Available online at https://www.english-corpora.org/coha/.
Davies, Mark. 2011–. SOAP: *Corpus of American Soap Operas*. Available online at https://www.english-corpora.org/soap/.
Dehé, Nicole 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45 (3): 569–615. https://doi.org/10.1017/s002222670999003x
Finkbeiner, Rita (ed.) 2024. *The Pragmatics of Headlines*. Special issue of *Journal of Pragmatics*.
Heine, Bernd & Gunther Kaltenböck 2021. From clause to discourse marker: On the development of comment clauses. *Language Sciences* 87: 101400. https://doi.org/10.1016/j.langsci.2021.101400
Hilpert, Martin 2017. Frequencies in diachronic corpora and knowledge of language. *The Changing English Language. Psycholinguistic Perspectives*, ed. by Marianne Hundt, Sandra Mollin, & Simone E. Pfenninger. 49–68. Cambridge: Cambridge University Press.
Hopper, Paul J. 1991. On some principles of grammaticalization. *Approaches to Grammaticalization*. Vol. I, ed. by Elizabeth C. Traugott & Bernd Heine. 17–35. Amsterdam: John Benjamins.

Huddleston, Rodney & Geoffrey Pullum *et al.* 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
Joseph, Brian D. 2000. Textual authenticity: Evidence from medieval Greek. *Textual Parameters in Older Languages,* ed. by Susan C. Herring, Pieter van Reenen, & Lene Schøsler. 309–330. Amsterdam: John Benjamins.
Kaltenböck, Gunther 2007. Spoken parentheticals in English: A taxonomy. *Parentheticals*, ed. by Nicole Dehé & Yordanka Kavalova. 25–52. Amsterdam: John Benjamins.
Kaltenböck, Gunther, Bernd Heine, & Tania Kuteva 2011. On thetical grammar. *Studies in Language* 35 (4): 848–893. https://doi.org/10.1075/sl.35.4.03kal
Kearns, Kate 2007. Epistemic verbs and zero complementizer. *English Language and Linguistics* 11 (3): 475–505. https://doi.org/10.1017/s1360674307002353
López-Couso, María José & Belén Méndez-Naya 2014. From clause to pragmatic marker: A study of the development of *like* parentheticals in American English. *Journal of Historical Pragmatics* 15 (1): 66–91. https://doi.org/10.1075/jhp.15.1.03lop
López-Couso, María José & Belén Méndez-Naya 2021. From complementizing to modifying status: On the grammaticalization of the complement-taking-predicate-clauses *chances are* and *odds are*. *Language Sciences* 88: 101422. Special Issue *Complement and Parenthetical Constructions: Theory and Description*, ed. by An Van linden, Lieselotte Brems, & Lieven Vandelanotte. https://doi.org/10.1016/j.langsci.2021.101422
Mélac, Eric 2022. The grammaticalization of evidentiality in English. *English Language and Linguistics* 26 (2): 307–329. https://doi.org/10.1017/s1360674321000101
Molek-Kozakowska, Katarzyna 2017. Stylistic analysis of headlines in science journalism: A case study of *New Scientist*. *Public Understanding of Science* 26 (8): 894–907. https://doi.org/10.1177/0963662516637321
Moncomble, Florent 2018. The deviant syntax of headlinese and its role in the pragmatics of headlines. *e-REA* 15 (2). https://doi.org/10.4000/erea.6124
Noonan, Michael 1985. Complementation. *Language Typology and Syntactic Description*. Vol. 2, ed. by Timothy Shopen. 42–140. Cambridge: Cambridge University Press.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
Rudanko, Juhani 1989. *Complementation and Case Grammar*. State University of New York Press: Albany.
Schneider, Stefan 2007. *Reduced Parenthetical Clauses as Mitigators. A Corpus Study of Spoken French, Italian and Spanish.* Amsterdam: John Benjamins.
Thompson, Sandra & Anthony Mulac 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. *Approaches to Grammaticalization*. Vol. II, ed. by Elizabeth C. Traugott & Bernd Heine. 313–339. Amsterdam: John Benjamins.
Van Bogaert, Julie 2010. A constructional taxonomy of *I think* and related expressions: Accounting for the variability of complement-taking mental predicates. *English Language and Linguistics* 14 (3): 399–427. https://doi.org/10.1017/S1360674310000134
Vandelanotte, Lieven 2009. *Speech and Thought Representation in English: A Cognitive-Functional Approach*. Berlin and New York: Mouton de Gruyter.

Intraindividual register variation in Early Modern English

Terttu Nevalainen
ORCID: 0000-0003-3088-4903

What is known as the "third wave" of variationist sociolinguistics focuses on the individual. In broad terms, individual variation comprises free variation and linguistically conditioned variation, on the one hand, and socially motivated variation, on the other (Werth *et al*. 2021). Applying these concepts to 17th-century English, this chapter analyses socio-pragmatically motivated intraindividual register variation in two case studies, one focusing on a startling departure from the expected norms of address term use, and the other on the writer's variation in expressing degrees of emotional involvement. Both studies draw on a corpus of personal letters that provides rich metadata on the individual writers; in the second case, the findings are compared with the writer's published work. Further consideration is given to corpus linguistic methods for retrieving intraindividual variation. Integrating the three waves of variationist sociolinguistics, access to a comparative baseline of the language community at large provides a means of assessing the multi-layered social evaluation of variation in individual usage.

Keywords: corpus linguistics, historical sociolinguistics, personal correspondence, Elizabeth Stuart, Thomas Browne

https://doi.org/10.51814/ufy.1041.1461

1. Introduction

The variationist approach to sociolinguistics typically conceptualizes linguistic variation and change at three levels: the macro-level of the community, the meso-level of social networks, and the micro-level of the individual (e.g., Auer & Hinskens 2005; Nevalainen 2015: 249–250). Penelope Eckert (2012; 2019) identifies these foci in developmental terms as the first, second, and third wave of variationist sociolinguistics, tracing the transition of the field from the first-wave demographic determinants of linguistic variation to the third-wave conceptualization of linguistic variation in terms of social practices.

Foregrounding the complementarity of the three waves as being “part of the same ocean”, Schilling (2018: 340) remarks that language users can make social meaning by conveying both group associational and interactional meanings. Apart from purely linguistically conditioned variation (Werth *et al.* 2021), socially motivated variation can merge public demographic information such as the speaker’s regional dialect and their momentary context-bound practices that shape their language in interaction. Ultimately, Eckert (2019: 751) argues, sociolinguistic variation ranges from social category membership to momentary affective states.

The concept of *register* is instrumental in approaching and accounting for this wide range of potential sociolinguistic variation. As registers are defined as alternate ways of “saying ‘the same’ thing” considered appropriate to particular contexts of usage (Silverstein 2003: 212), socio-situational contexts provide a framework for recognizing register-specific patterns of language use – and deviations from them. Registers can be realized on different levels of specificity, ranging, for example, from differences between fictional prose and verse (Busse 2002) to variation based on the writer–recipient relationship and the purpose of writing in personal correspondence, both subjects addressed in the rich tradition of letter-writing manuals (Bannet 2008).

As intraindividual register variation springs from diverse contexts and motivations it moves on a scale from reactive to creative. Examples of the reactive kind would include the use of address terms determined by the social roles and power relations of the interlocutors in highly formal, status-marked contexts, such as court trials, with departures from these commonplaces giving rise to register interpretations of impoliteness or disrespect (Kryk-Kastovsky 2006; Nevalainen 1994). At the creative end of the scale may come the use of the second-person pronoun *thou* that, instead of reinforcing the writer's critical comments, as it would do in early modern courtroom discourse, can serve to convey affection and banter among social intimates (Nevala 2018).

Highlighting its range of motivations, this chapter analyses intraindividual linguistic variation in two case studies of 17th-century English. The first one focuses on Elizabeth Stuart's startling departure from the expected norms of address term use in her personal correspondence, and the second discusses Sir Thomas Browne's variation in conveying degrees of emotional involvement among intimates and in public discourse. In both studies the observed linguistic variation is viewed as the writer's means of expressing their assessment of the appropriate register in a given communicative context.

Past socio-historical and pragmatic work on intraindividual variation is reviewed in Section 2, and Section 3 considers the material and methodological underpinnings of the field and the approaches adopted in this chapter; their details are discussed in Section 4. The two case studies present intraindividual variation of very different kinds: remarkable deviation from expected norms, on the one hand (4.1), and variation in encoding emotional involvement, on the other (4.2). Section 5 assesses intraindividual register variation and its possible social motives in the light of the research presented.

2. Past work and approaches

It is a commonly held view that intraindividual variation has been neglected in historical sociolinguistics (Schiegg & Huber 2023: 2). In practice the situation is much less bleak, but as past studies on the topic are typically informed by various philological, literary, sociolinguistic, and pragmatic frameworks, they would not have caught the attention of those departing from a three-wave variationist perspective (Bell 2016: 402–403). The following brief review of past work illustrates the material and methodological versatility of the research tradition and the range of the kinds of linguistic variation observed from syntax to spelling and from code-switching to ventriloquizing.

Focusing on a single individual, Raumolin-Brunberg (1991) explored intra-writer variation in the published and private writings of Sir Thomas More (1478–1535), including history, polemics, devotional prose, and public and private correspondence. The linguistic focus of the quantitative study was his noun-phrase use. On the basis of their NP characteristics More's texts could be classified into three main groups, with groups (1) and (3) representing opposing tendencies: (1) history writing and private letters (high NP frequency, low NP coverage, short and simple NPs), (2) polemics and semiofficial letters, and (3) prayers and official letters (high NP coverage, long NPs, N heads with branching structures). The author concluded that these differences rarely followed genre divisions but could instead be accounted for in terms of multiple dimensions of variation (1991: 281).

Dramatists provide a popular source for the study of intra-writer variation. Ulrich Busse (2002) analysed *you* vs. *thou* variation in Shakespeare's works, in both drama and poetry, and constructed a typology of "*thou*fulness" by creating "a link between a statistically more or less probable form (*thou*) and its stylistic value as the marked term of the dyad" (2002: 288). He observed that *thou* dominated in verse and *you* in prose, which confirmed Shakespeare's poetry as an elevated register compared to his drama. The distinction also cut across drama in that History Plays and Tragedy emerged as the most "*thou*ful" plays, while *you* prevailed in Comedies.

Nurmi and Pahta (2004) analysed code-switching, "the use of more than one language in the course of a single communicative episode", in correspondence in 1410–1550 and combined demographic and interactional perspectives in their work. At the community level, the authors found that nearly one in five of the writers studied code-switched in their letters, men more often than women. Code-switching was linked to the writer's social rank: members of the clergy and other writers with university education switched to Latin, while switching to French occurred in merchant letters. The study reported individual variation and micro-level interactional functions of code-switching within social groups. These intraindividual switches could create intimacy and establish group-membership but also occasionally served as a form of exclusion, confirming that "the explanations for variation do not lie exclusively with demographic variables" (Nurmi & Pahta 2004: 434, 448).

Hernández-Campoy, Conde-Silvestre, and García Vidal (2019) investigated spelling variation in the late 15th-century Paston Letters, concentrating on the replacement of <þ> by the incoming <th> among male writers. The process was a change from above in social terms, associated with the formality of the context and the high social rank of the addressee. The change was evidenced in intraindividual variation in that the writers showed upward accommodation to <th> when writing to recipients highest up on the social scale (nobility) and to professionals, but downward accommodation to <þ> when writing to social equals (lower gentry).

Nevalainen (2009) reported a unique instance of register switching where Lady Katherine Paston produced an imitation of the spoken words of a third party, as it were, by "ventriloquizing" them (cf. Tannen 2001). Conveying family news to her son William, who was studying away from home in the 1620s, she recounted an incident with little Philip, whose hands had been bitten by William's ['his master's'] three puppies. She reproduced Philip's words twice, transmitting to William a close imitation of the child's speech: "his Nastas goggs did dite ha hans", and again: "nasta lillo goggs did dite his hans" (1).

1 he cam with tears in his eies to me and sayd: his Nastas goggs did dite ha hans: I was desirous onc more to heer his pretty playnt: he cried nasta lillo goggs did dite his hans: your 3. pup*i*s weare so bowld with him: (CEEC, Katherine Paston to William Paston, 1627?; PastonK 99)

Some more expected register variation emerged when Lady Katherine's "*thou*ful" letters to William, characterized by terms of endearment, praise, and humor, were compared with her adult-directed language in the letters she wrote to her male relatives. For example, she always addressed them with the second-person pronoun *you*.

Showing that methodological and conceptual pluralism continue to mark the field, the papers presented in Schiegg and Huber (2023) include, for example, patterns of variation in discourse-ending formulae, self-corrections as stylistic choices in a draft letter, and spelling variation of Middle English scribes. In her contribution, Iyeiri (2023) regards intra-text variation as an instance of intra-writer variation in Middle English manuscripts. Viewed from the perspective of texts, her analysis might be extended to variation in the philological principles of editing manuscripts.[1]

---

1 Pahta (1998: 151) puts the text editor's role in a nutshell: "Decisions on editorial methods reflect on both the substance of the text and on the representation of its linguistic realisation." Now twenty-six years on, the research community thanks Päivi not only for her insightful early editorial work but also for her leading-edge scholarship that ranges from historical multilingualism and socio-pragmatics to

3. Material and methods

As the studies discussed illustrate, corpus linguistics provides the foundation especially for quantitative studies in intraindividual variation. Discussing the complex relationship between social roles and language practices in the study of Late Modern English, Pahta *et al*. (2010: 9) endorse the use of corpora and corpus methodology for two main reasons: corpus data and techniques make it possible "to observe differences in usage between data sets, as well as to reveal lexical and grammatical characteristics of the text that are not immediately observable to the reader but still form an important part of the style of the text". Systematic text comparisons provide a good heuristic tool, they argue, because "linguistic features indexing identity and social roles may encompass a variety of linguistic phenomena […], but it is difficult to know beforehand which will be relevant in a given situation".

This is also at issue with intraindividual register variation more broadly, especially when it operates on a wider set of register-defining phenomena than single linguistic features. Community-level studies of corpora based on ego-documents such as personal letters provide the basis for comparing individual practices and help detect outliers whose usage deviates significantly from the average community or group-level norms.[2] Investigating the use of address terms in the *Corpus of Early English Correspondence* (CEEC), Nevalainen and Raumolin-Brunberg (1995) also paid attention to correspondents whose usage departed from the attested norms. One of them was Elizabeth Stuart, Queen of Bohemia, who was selected as the subject of the first case study in this chapter. The outlier approach to individual variation was also explored in the studies based on the CEEC discussed in Nevalainen, Palander-Collin, and Säily (2018). This comparison revealed Sir Thomas Browne's variable use of linguistic features, which contrasted the letters he wrote to his son and daughter. He is the subject of the second case study in this chapter. Both case studies draw on the *Corpus of Early English Correspondence*. Sir Thomas Browne's letters are, moreover, compared with his published work,

---

the special language of medical writing, and from Middle English to the use of English in Finland today. Long may that work continue.

2 Community-level register variation has also been studied in multigenre corpora, for example, by means of a multidimensional analysis of the co-occurrence of a large number of linguistic features. Douglas Biber has applied this MD-method to a variety of corpora, including one of 18th century English (Biber 2001); cf. 4.2, below.

and where possible, Elizabeth Stuart's personal letters are quoted from a recent diplomatic edition.

Both quantitative and qualitative methods are used in these case studies. Furthermore, corpus data can be explored by means of diverse visualization techniques at various stages of the study from data exploration to interpreting the (socio)linguistic patterns that emerge from the analysis of the data (Siirtola *et al.* 2011; Sönning & Schützler 2023). A simple word-cloud heuristic is used in Section 4.2, where Thomas Browne's private and public writings are compared, showing that corpus-based distant reading combined with close reading can help discover and contextualize fine-tuned intraindividual register variation.

4. Two case studies

The studies to be discussed follow the basic sociolinguistic tenet that individuals vary their language use according to the communicative situation. In essence, descriptions of community-level sociolinguistic variation therefore rest on individual variation. Discussing women's letter writing in the 16th century, Daybell (2010: 67) concludes that "[f]actors such as rank, social status, local and wider influence, and the nature of the correspondence had as much, if not more, impact than gender on the tone of the letter". All these various factors are illustrated in the correspondence of Elizabeth Stuart, Queen of Bohemia.

4.1. Elizabeth Stuart: Departing from expected norms

Elizabeth Stuart (1596–1662), the only surviving daughter of King James VI of Scotland and I of England, was married to Fredrick V (1596–1632), the Protestant ruler of the Palatinate, who accepted the throne of Bohemia in 1619. However, the following year he was deposed and expelled as part of the conflict between Catholics and Protestants that escalated into the Thirty Years' War (1618–1648). As the Spanish occupation of the Palatinate made them stateless, Fredrick and Elizabeth and their large family lived in exile in the Netherlands, unsuccessfully trying to regain their territory in the Palatinate. In England, Elizabeth came to be known as a champion of the Protestant faith, and due to her short reign in Bohemia, she was dubbed as the "Winter

Queen". Asch (2016) notes that she had "a considerable talent to inspire admiration, loyalty, and love in Germany as well as England", which earned her the title of the "Queen of Hearts" (cf. Akkerman 2021).

Much of the work that Elizabeth undertook to promote the Palatine cause was carried out by means of correspondence. The forms of address in her letters varied, as would be expected, according to the recipient. For example, she addressed her male relatives partly in formal terms "Sire" or "Most gratious soveraygn and deare father" (her father, King James I), with routine affection "My only dear brother" (her brother, Charles I), or simply "Sonne" (her son, Charles Louis, the Elector Palatine after the death of his father Fredrick V). The courtiers, diplomats, and noblemen she corresponded with were typically addressed as "My Lord". But Elizabeth also showed creative language use and remarkable intraindividual register variation when, writing to James Hay, Earl of Carlisle, in 1630, she addressed the older courtier and diplomat as "Thou ugly, filthy camel's face". This letter, dated the 12th of June 1630, runs as follows (Green 1857: 482–483):

2 Thou ugly filthy camel's face,
You chid me once for not writing to you; now I have my revenge and more justly chide you; for not having heard from you so long as I fear you have forgot to write. I have charged this fat fellow to tell you all this, and that I cannot forget your villainy. He can inform (you) how all things are here, and what they say to the peace with Spain; and though I confess I am not much rejoiced at it, yet I am so confident of my dear brother's love, and the promise he hath made me, not to forsake our cause, that it troubles me the less. I must desire your sweet face to continue your help to us, in this business which concerns me so near; and in spite of you, I am ever constantly

Your most affectionat frend
Elizabeth[3]

3 Transcribed from the subscription in Elizabeth's own hand appended to Green's modern-spelling edition of the letter, which is not included in Nadine Akkerman's excellent original-spelling edition, as a rule cited here. Mary Anne Everett Green was Elizabeth's first biographer and one of the most prolific editors of the Calendar of State Papers (Akkerman 2021: 4–5).

The recipient of the letter, James Hay (c. 1580–1636), the 1st Earl of Carlisle, shown in Figure 1, was a courtier and favourite of both James I and Charles I, and undertook various diplomatic missions for them both. He had arrived in England with King James and, although he successfully blended into his new surroundings, Schreiber (2008) comments that he was "a singularly ineffective diplomat", unable, for example, "to negotiate the return of the palatine family to its German territory".

Not having received the word from Carlisle she had been expecting, Elizabeth has cause to chide him in (2) just as he had done on an earlier occasion when she had failed to write to him. She retaliates by addressing him with the second-person singular pronoun *thou*, which could give rise to various readings from affection and intimacy to condescension and disrespect. Combined with the following "ugly filthy camel's face", the salutation would strongly suggest the latter interpretation.

Figure 1. James Hay, 1st Earl of Carlisle (National Galleries of Scotland; CC BY-NC).

However, contrary to what the impolite term might lead one to assume, the beginning of the letter could also be read as friendly banter rather than a personal insult, as Elizabeth chides Carlisle for the "villainy" of failing to write to her, greeting him with the nickname she also employs elsewhere in her correspondence with Carlisle.[4] But as she goes on to discuss the treaty with Spain that was made on terms unfavorable to her husband, the salutation could preface her frustration not only at Carlisle's failure to write to her but also, perhaps more seriously, to help ensure a positive outcome in the Spanish affair.

Mary Anne Everett Green (1857: 482), Elizabeth's first biographer, takes the letter as evidence of the patience with which Elizabeth resigned herself to her circumstances. This is borne out in the second half of the letter that records Elizabeth's personal reaction to the political situation. However, as she also needs to be able to rely on Carlisle's continued help in the future, she changes her initial "ugly, filthy Camel's face" to "your sweet face" at the end of the letter, signing off as "Your affectinat frend". This expression, with some modification of the adjective, was Elizabeth's routine subscription formula, followed by her signature. As Nevala (2009: 89) comments, in personal correspondence "friend" does not necessarily imply a close friendship; in fact, acquaintances and more distant recipients are referred to as "friends" more often than intimates. After her singular creativity in the salutation, Elizabeth returns to a reactive, formulaic expression of goodwill at the end.

Although "ugly filthy camel's face" does not seem to occur in the salutation of Elizabeth's other surviving letters, she uses it in the body of several extant letters addressed to Carlisle. An earlier letter, dated the 18 March 1630, gives more background to the treaty with Spain and also to Elizabeth's level of (in)formality with Carlisle. Here the salutation is the expected "My Lord" and, unlike in (2), Elizabeth also has reason to thank Carlisle for having received his letter. She takes up the treaty with Spain towards the end of her letter in (3), expressing her reliance on both King Charles and Carlisle being able to further the pro-Palatine cause (Akkerman, 2015: 800–801). The letter closes

4 In line with the tone of the beginning of the letter, Elizabeth also mentions "this fat fellow", Sir Henry Vane, who was to deliver her message to Carlisle. Akkerman (2015: 804) mentions that she often used the adjective *fat* to describe both Vane and Sir Thomas Roe.

with Elizabeth’s continued appreciation of Carlisle, and ends in her appealing to him by using his nickname “honest worthy camels face”:

> **3** I heere the treatie with Spaine goeth one, you know my minde, that I ame still incredulous that they will doe anie thing, except it be vppon dishonnourable conditions, but I ame confident that the king my g deare Brother will not suffer anie such thing to be ~~though~~t spoken ~~one~~, for and that you will doe all you can in it, for I haue euer found you so true a frend to me in all occasions as binds me euer to be honest worth>y< camels face
>
> Your constant affectionat
> Frend
> Elizabeth

Elizabeth also repeats the nickname in other letters, both before and after the conclusion of the Spanish treaty. On 29 May 1629, she finished her letter to Carlisle: “I will loose no meanes whereby I may shew your ouglie Camels face that I ame Your most constant frend Elizabeth” (Akkerman 2015: 745). As late as in the mid-1630s, she concluded her letter using the nickname and assuring her reliance on Carlisle’s good judgement:[5]

> **4** I ame so confident of your loue, and you so much power with me, as for your sake I shall euer esteeme those you recommend to me beleeue this worthie Camels face of her that is euer
>
> Your most constant frend
> Elizabeth

We may conclude that the startling sobriquet was a peculiar form of endearment that Elizabeth only applied to Carlisle. She could use it to evoke different readings in the salutation of the letter and in the closing address at the end. Used in the salutation with *thou* in (2), it could be interpreted as banter with some critical undertones in that particular context, whereas at the end it routinely confirmed the good relations between the writer and the recipient of the letter. There is no way of knowing how and when the peculiar

---

5 Akkerman (2011: 372) dates the letter between the end of November 1633 and 23 November 1635.

usage came about, but Akkerman (2021: 9) mentions that Elizabeth, like her parents, was in the habit of using nicknames as terms of endearment. For example, she also addressed Sir Thomas Rowe, her closest confident at the royal court, not only as "My Lord" and "Good S^r^ Thomas Roe" but also as "Honest Thom Roe" and even "Honest fatte Thom" (cf. Akkerman 2015: 804).

In their community-level analysis of the use of address terms in the 16th and 17th centuries, Nevalainen and Raumolin-Brunberg (1995: 589) conclude that positive orientation is available to social superiors regardless of social distance and that the range of positive address forms such as nicknames is practically open-ended. It is, however, noteworthy that their mutual selection is only licensed by the close distance of the correspondents. This was particularly true of members of the royal family, whose informal terms of address could never be reciprocated by their subjects.

4.2. Thomas Browne: Varying degrees of involvement

As discussed in 4.1, in personal correspondence register variation is observed in address terms, which are usually in harmony with the register choices in the body of the letter. Other possible cues to register variation include linguistic features undergoing change over time, as incoming and outgoing features are often associated with different social meanings. Analysing outliers, individuals who deviated from the community norm, proved a useful discovery method for intraindividual variation in a study of the tail end of the diffusion of verbal *-s* in the course of the 17th and 18th centuries (Nevalainen 2018).

Figure 2 shows that the verb *have* (both main verb and auxiliary) was much slower in showing the incoming form than the bulk of other verbs in the CEEC: in the first half of the 17th century, when the other verbs had already passed the 80%-mark, in *has -s* had reached the frequency of 20% and was only nearing 50% in the second half of the century. Variation could naturally be found at the level of individual writers. One of those who used the outgoing *hath* form more frequently than most writers in the twenty-year period from 1680 to 1699 was Sir Thomas Browne, who also showed intraindividual variation by using the incoming variant *has* a few times in his private correspondence.

Sir Thomas Browne (1605–1682), was a medical doctor, polymath, and a notable author who was knighted for his merits by Charles II in 1671. Born in London, he came from the merchant ranks: his father was a liveryman of the

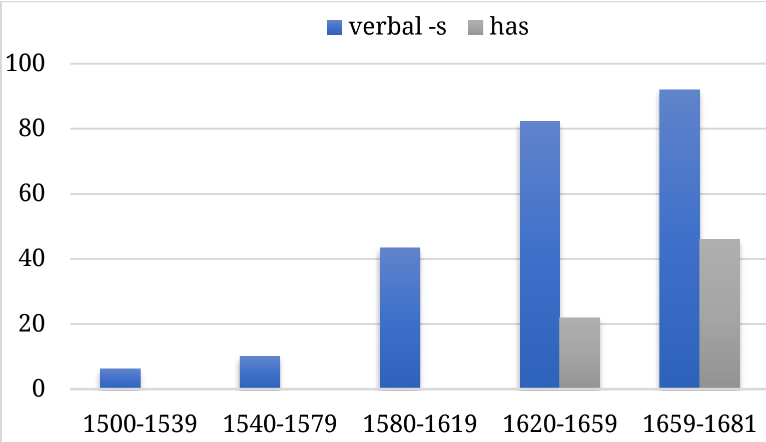


Figure 2. The spread of verbal -*s* (%) in the language community (modified from Nevalainen 2018).

Mercers' Company, and after his death, Thomas's education became primarily the responsibility of his uncle, who was a grocer. Browne studied in Oxford and on the continent specialized in medicine in Montpellier, Padua, and Leiden. He set up his practice in Norwich in 1637, where he married Dorothy Mileham (1621–1685) in 1641 (Barbour 2013; Robbins 2008).

Much of Browne's preserved private correspondence is addressed to his sons, Edward (1644–1708) and Tom (1647–1667?), with a few letters to his daughter Elizabeth (b. 1648). His eldest son Edward followed in his father's footsteps as a physician. He lived and practised in London but also travelled widely abroad, as did Tom, who died young. After her marriage to George Lyttelton in 1680, Elizabeth moved to live in Guernsey. In the CEEC, *has*, the incoming third-person singular variant of *have* only appeared in Browne's letters to Elizabeth ("Betty"), while *hath* was the sole form found in letters addressed to his other correspondents, including his son Edward. The passages quoted in (5), addressed to Betty, and (6), addressed to Edward are dated the same day, 6 June 1681, and largely contain the same family news. They provide a clear case of Browne's register variation shown not only in his use of different forms of *have* (shown in bold) but also in other linguistic features.

5 Dearest Betty,
Thy letters are still our best divertion, to hear you and all that belong to you gat so well to Portsmouth was very wellcom to us, yr thoughts for us are Equald with ours for you. I am sure their Passes not a day that we are not severall times talking of you. Poor Tomeys Cough **have** brought him in to a great heat, but I hope it will not be so bad as that feavor was wch you were so helpfull to him; his stomack very bad, we are this after noon goeing to bath [him] by his own desire. Our Tommy **has** had a grievous Cof and feavor, yr sister Frank **has** dun more for him then I could have thought; he was bled and bathd and I bless God he had got down amongst us again and is much delighted with yr Letters. (CEEC, Sir Thomas Browne to his daughter Mrs. Elizabeth Lyttelton, 6 June 1681; Browne, 194)[6]

6 D. S.,
Mr Deane Astley went to London on this day was sevenight, and sayd hee would call upon you. […] This day, God bee thancked, wee had a fine showre of raine; the spouts of our howse have not Runne for 8 or 9 weekes before. I had a cough for 6 weekes very feirce in the night and it held mee till within these 12 or 14 dayes, most persons in my howse had it or have it except my wife. Frank **hath** it and it **hath** been with hooping and vomiting, butt is persua[dable] to take litle and will not abstaine from going to morning and evening prayers which wee daylie have at our owne parish church. Tommy **hath** had it with some hooping and vomiting, butt now vomits butt seldome, butt sleepes prettie well in the night and at any time when hee lyeth downe in the day; hee **hath** been very hot and so that hee beggd to bleed a litle and to goe into a *balneum dulce* which [he] had used in a sicknesse before. (CEEC, Sir Thomas Browne to his son Dr. Edward Browne, 6 June 1681; Browne, 192–193)

Word clouds may be used to explore the lexical variation that characterizes the two passages, which can be suggestive of subregisters within the

---

6 The letter to Elizabeth is here quoted in its extant, full form, but the longer one to Edward is shortened in the middle to make the passages comparable in length and content.

broader private correspondence register.[7] Figures 3 and 4 include grammatical words, which point to certain differences in the structure and orientation of the letters.

Figure 3. Word cloud of letter (5) to Betty.

Figure 4. Word cloud of letter (6) to Edward.

7 Drawn using https://wordart.com/.

In the letter to Betty (5) the pronoun *you* stands out, suggesting orientation towards the recipient, whereas in the letter to Edward (6) the pronoun *it* is more prominent, pointing to a focus on things, in this case the particular disease that several family members had suffered from. This subject was of particular interest to Edward not only as a physician but also because the illness had affected his son Tommy, who was staying with his grandparents and is referred to as *hee* in the topic position in the sentence in (6). Tommy is also the referent of *him* in the letter to Betty (5), suggesting the semantic roles of both patient and beneficiary, foregrounding the care the other family members provided in restoring him back to health.

Both passages show frequent use of *and* and *to*. In both, the prominence of *to* arises from grammatical constructions that require the use of the preposition or particle. The conjunction *and* dominates in the letter to Edward. Most of the time Browne employs it as a clause-linking device as he describes how each family member had suffered from the illness. No similar structure marking with *and* is found in the letter to Betty.

The lexical choices in the two letters reveal further subregister differences. Betty is addressed as *Dearest Betty* and her letters (*thy letters*) are mentioned with particular appreciation at the beginning. Edward, in turn, is greeted casually with *D.S.*, an abbreviation of *Dear*(*est*) *Son*, but the letter opens with a piece of news of Mr. Deane Astley's London visit. The two letters also differ in some of their third-party references, which are more involved in the letter to Betty: *our Tommy* and *yr sister Frank* in (5) as opposed to plain *Tommy* or *Tom* and *Frank* in (6), where the first-person possessive *my* appears in *my house* and, interestingly, *my wife*, with reference to Edward's mother, Dorothy Browne. In this passage, which is concerned with the writer's own illness, the letter recipient is phased out.

That Browne's use of *has* and *hath* correlates with his other lexical choices can be seen in his use of involvement features such as first- and second-person pronouns, private verbs, and degree adverbs and adjectives (cf. Biber 2001). The differences are not absolute but stand out when comparing the latter halves of the letters to Betty in (7) and (8), and to Edward in (9) and (10).

Apart from including the East Anglian dialectal third-person suffixless form *have*, (7) contains *I*, *you*, and *we*, the verb *hope*, the intensifier *very* and two instances of *so,* as well as several evaluative adjectives (*bad*, *great*, *helpful,* *poor*). With two instances of *has*, (8) presents first- and second-person pronouns and determiners, the verbs *bless* and *think*, the degree words *more* and *much*, and the adjectives *delighted* and *grievous*.

7 **Poor** Tomeys Cough **have** brought him in to a **great** heat, but **I hope** it will not be **so bad** as that feavor was wch **you** were **so helpfull** to him; his stomack **very bad**, **we** are this after noon goeing to bath [{him{] by his own desire.

8 (**Our** Tommy **has** had a **grievous** Cof and feavor, **yr** sister Frank **has** dun **more** for him then **I could have thought**; he was bled and bathd and **I bless God** he had got down amongst **us** again and is **much delighted** with **yr** Letters.

By contrast, the contexts of *hath* in the letter to Edward in (9) and (10) lack first- and second-person pronouns except in the aside mentioning the daily prayers the family members (*wee*) have in their (*our*) parish church (9). Compared to the letter Browne wrote to Betty, the lack of private verbs and the scarcity of degree adverbs and adjectives are noticeable with only two instances of *very*, one earlier in (6) where Browne talks about his own cough, and the other in (10) where he describes Tommy's fever. The other intensifier he has is the vague – boosting or downtoning – use of *pretty* in *pretty well*. He also downplays the severity of Tommy's illness in expressions such as *some whooping*, *but seldom*, and *bleed a little* (10).

9 (Frank **hath** it and it **hath** been with hooping and vomiting, butt is persua[dable] to take **litle** and will not abstaine from going to morning and evening prayers which **wee** daylie have at **our** owne parish church.

10 (Tommy **hath** had it with **some** hooping and vomiting, butt now vomits **butt seldome**, butt sleepes **prettie well** in the night and at any time when hee lyeth downe in the day; hee **hath** been **very** hot and so that hee beggd to bleed **a litle** and to goe into a *balneum dulce* which [he] had used in a sicknesse before.

Comparing the descriptions of illness in the two letters shows the writer's varying degrees of emotional involvement and detachment: the Betty letter discloses the severity of Tommy's illness as well as Browne's hopes and fears concerning it, whereas the letter to Edward gives a more detached account of the disease that both his sister Frank and son Tommy have. There may have been several reasons for these subregister choices. The writer may not have wanted Edward to be unduly worried about his son, who was on the mend. Equally likely, the more businesslike register may have been his usual mode of writing to his son, with whom he was in the habit of communicating on a wide variety of scientific topics.

Comparing the two letters further, Browne's subregister choices could also be interpreted as reflections of his understanding of the social roles and identities of his children: Betty's role as a carer appears in (5), as does that of her sister Frank. Neither of them is mentioned in that role in Browne's letter to his son, whom he rather approaches collegially as a medical practitioner, explicitly signalled by the use of the Latin phrase *balneum dulce* ('sweet bath').

Thomas Browne's ability to enact his social role as a scholar and medical practitioner is amply demonstrated in his published work. He is particularly well known for a book that is regarded as a precursor of popular scientific writing entitled *Pseudodoxia Epidemica or, Enquiries into Very Many Received Tenents and Commonly Presumed Truths* (1646). In it, Browne aimed to promote reason and the scientific method and to rectify common misconceptions. The publication ran into six editions between 1646 and 1672 and was translated into several languages, including Latin. Browne noted in the preface that the text was intended for the Gentry, specifically to "the knowing and leading part of Learning". Testifying to his linguistic sensitivity, he acknowledged the difficulty that general readers might find with his language as "the quality of the Subject will sometimes carry us into expressions beyond meere English apprehensions" (Robbins 2008).

Thomas Browne's published work bears witness to his lexical creativity, being credited as the 73rd most cited source in the *Oxford English Dictionary.* In all, 761 quotations are found to provide first evidence of a word, and 1,558 first evidence of a particular meaning.[8] Typically of a scientific or medical nature, Browne's contributions include words such as *ambidextrous*, *analogous*, *ascetic*, *carnivorous*, *coexistence*, *compensate*, *cryptography*, *cylindrical*,

8 OED3, last accessed 14 July 2023.

disruption, *electricity*, *ferocious*, *hallucination*, *herbaceous*, *insecurity*, *locomotion*, *polarity*, *precocious*, *pubescent*, *suicide*, *therapeutic*, and *veterinarian*. He borrowed or derived the words and their senses from sources where most of them were used in professional circles in their Latin, Greek, or French forms. He was probably also an early adopter of words that were already in use in English at the time.

The passage from the preface to *Pseudodoxia Epidemica* in (11) illustrates the register that Browne adopted when addressing his reader. The conservative third-person singular present-tense suffix is -*th* not only in *doth* and *hath*, as in the letter to Edward, but also in lexical verbs (*leadeth*, *requireth*). The other items in boldface, *medicall*, *Vroscopy*, and *paradoxologie* (*OED*: 'The action or habit of indulging in or propounding paradoxes'), are listed by the *OED* as lexical innovations first attested in English in this text.[9]

11 Some consideration we hope from the course of our Profession, which though it **leadeth** us into many truths that passe undiscerned by others, yet **doth** it disturbe their communications, and much interrupt the office of our pens in their well intended transmissions: and therefore surely in this worke attempts will exceed performances: it being composed by snatches of time, as **medicall** vacations, and the fruitlesse importunity of **Vroscopy** would permit us. And therefore also perhaps it **hath** not found that regular and constant stile, those infallible experiments, and those assured determinations, which the subject sometime **requireth**, and might be expected from others, whose quiet doors and unmolested hours afford no such distractions. Although whoever shall indifferently perpend the exceeding difficulty, which either the obscurity of the subject, or unavoidable **paradoxologie** must often put upon the Attemptor, will easily discerne, a worke of this nature is not to bee performed upon one legge, and should smell of oyle if duly and deservedly handled. (*Pseudodoxia Epidemica* (1646), Part I, *To the Reader*)

---

9 However, some earlier 17th-century attestations of *medical* are found in Early English Books Online (https://earlyprint.org/lab/).

Like in the letter to Edward, the word cloud in Figure 5 highlights Browne's use of *and* in this passage. Here it connects not only clauses but also phrases, increasing the structural complexity of the text at the phrase level. The word cloud also suggests other structural differences between the printed text and the private letters. The most notable one is the prominence of the definite article *the* and the preposition *of* in Figure 5, which point to frequent use of nouns in the passage, an impression reinforced by the relative pronoun *which* and the pronoun *it*. Close reading confirms the "nouniness" of the passage, reminiscent of Raumolin-Brunberg's (1991: 281) heavy NP category 3 (high NP coverage, long NPs, noun heads with branching structures) in Thomas More's writings. An interesting metatextual comment appears in Browne's use of *us* and *our* in (9) with reference to himself as a medical professional who suffers from a lack of time to put his findings in writing in a "regular and constant stile".

Figure 5. Word cloud of the passage in (11) from *Pseudodoxia Epidemica*.

The conclusion to be drawn from the word-cloud comparisons is that despite some apparent similarities, such as the frequency of the conjunction *and*, the published text differs substantially from Browne's private letters. Especially the letter to Betty has little in common with *Pseudodoxia* either in terms of structure or orientation.

5. Discussion and conclusion

Several approaches can be adopted to account for intraindividual variation and register choices in socio-pragmatic terms. One is the individual's understanding of their social roles and those of their interlocutors, and the ways in which this understanding can be harnessed in social meaning making. Pahta *et al.* (2010: 1) stress the negotiability of social roles in interaction:

> Particular social roles can imply particular linguistic choices that are appropriate to enact those roles, but at the same time individuals can make linguistic choices and mobilize parts of their linguistic repertoire to index, negotiate and construct their social roles.

The social role aspect provides one possible approach to Sir Thomas Browne's (sub)register differences when reporting basically the same news to his son and daughter. We may assume that his experience as a widely published scholar, a practicing medical doctor, and a father of a large family, must have contributed to his susceptibility to a wide range of register variation. His ability to use different registers may therefore be detected in the way he fine-tuned the two letters to correspond to their recipients' social worlds and his own role in them, whereas his public role as a scholar addressing a wide readership made further demands and set different constraints on his register choices in print.

Another, related basis for interpretation is the degree of intimacy *vs.* distance that marks the relations of the interlocutors. It can be judged on external criteria such as family membership, but power differences that affect the degree of intimacy and emotional involvement are found both within the family and outside it. In personal correspondence these distinctions are in part subject to traditional genre conventions and more marked in earlier periods than in more recent times (Nevalainen 2001; Nevalainen & Raumolin-Brunberg 1995). However, just like the expression of social roles, the writer can monitor the intimacy *vs.* distance axis for various effects, as could be seen in the case of Elizabeth Stuart, Queen of Bohemia, who deftly used her striking nickname for the Earl of Carlisle in letter salutation to let him understand her affective state and reaction to his inaction (cf. Eckert 2019: 751).

As shown by the studies discussed in this chapter, personal letters continue to provide a rich source for research into intraindividual variation.

One way of broadening the scope of the field would be to have access to corpora that consist of different genres and registers authored by one and the same individual. This principle was applied here in the comparison of Sir Thomas Browne's private and public writings. In general, including extensive metadata on the writers' backgrounds in historical corpora would enable the exploration of the possible social correlates of intraindividual variation in their multilayered contexts. Here visualization techniques could also help to discover "characteristics of the text that are not immediately observable to the reader" (Pahta *et al*. 2010: 9).

Bibliography

Akkerman, Nadine 2021. *Elizabeth Stuart, Queen of Hearts*. Oxford: Oxford University Press.

Akkerman, Nadine (ed.) 2015. *The Correspondence of Elizabeth Stuart, Queen of Bohemia: 1603-1631*. Vol. I. Oxford: Oxford University Press.

Akkerman, Nadine (ed.) 2011. *The Correspondence of Elizabeth Stuart, Queen of Bohemia: 1632-1642*. Vol. II. Oxford: Oxford University Press.

Asch, Ronald G. 2016. Elizabeth, Princess [Elizabeth Stuart] (1596–1662), queen of Bohemia and electress palatine, consort of Frederick V. *Oxford Dictionary of National Biography.* Retrieved 11 Dec. 2023.

Auer, Peter & Frans Hinskens 2005. The role of interpersonal accommodation in a theory of language change. *Dialect Change: Convergence and Divergence in European Languages*, ed. by Peter Auer, Frans Hinskens, & Paul Kerswill. 335–357. Cambridge: Cambridge University Press.

Bannet, Eve Tavor (ed.) 2008. *British and American Letter Manuals 1680–1810*. London: Pickering and Chatto Publishers.

Barbour, Reid 2013. *Sir Thomas Browne: A Life*. Oxford: Oxford University Press.

Bell, Allan 2016. Succeeding waves: Seeking sociolinguistic theory for the twenty-first century. *Sociolinguistics: Theoretical Debates*, ed. by Nikolas Coupland. 391–414. Cambridge: Cambridge University Press.

Biber, Douglas 2001. Dimensions of variation among 18th-century speech-based and written registers. *Towards a History of English as a History of Genres* (Anglistische Forschungen 298), ed. by Hans-Jürgen Diller & Manfred Görlach. 89–109. Heidelberg: Universitätsverlag C. Winter.

Browne, Thomas 1646. *Pseudodoxia Epidemica, or, Enquiries into Very Many Received Tenents and Commonly Presumed Truths by Thomas Browne*. London: Printed by T.H. for E. Dod. http://name.umdl.umich.edu/A29861.0001.001

Busse, Ulrich 2002. *Linguistic Variation in the Shakespeare Corpus: Morpho-syntactic Variability of Second-Person Pronouns* (Pragmatics & Beyond New Series 106). Amsterdam & Philadelphia: John Benjamins. https://www.jbe-platform.com/content/books/9789027296191

CEEC = *Corpus of Early English Correspondence* (1402–1800). 1998–2006. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, Samuli Kaislaniemi, Jukka Keränen, Mikko Laitinen, Minna Nevala, Arja Nurmi, Minna Palander-Collin, Tanja Säily, & Anni Sairio. University of Helsinki. https://varieng.helsinki.fi/CoRD/corpora/CEEC/

Daybell, James 2010. Gender, obedience, and authority in sixteenth-century women's letters. *The Sixteenth Century Journal* 41 (1): 49–67. http://www.jstor.org/stable/27867637

Eckert, Penelope 2012. Three waves of variation study: The emergence of meaning in the study of variation. *Annual Review of Anthropology* 41: 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828

Eckert, Penelope 2019. The limits of meaning: Social indexicality, variation, and the cline of interiority. *Language* 95 (4): 751–776. https://doi.org/10.1353/lan.2019.0072

Green, Mary Anne Everett 1857. *Lives of the Princesses of England, from the Norman Conquest*. Vol. 5. London: Longman, Brown, Green, Longman & Roberts.

Hernández-Campoy, Juan M., Juan Camilo Conde-Silvestre, & Tamara García Vidal 2019. Tracing patterns of intra-speaker variation in early English correspondence: A change from above in the *Paston Letters*. *Studia Anglica Posnaniensia* 54 (s1): 287–314. https://doi.org/10.2478/stap-2019-0014

Iyeiri, Yoko 2023. Intra-text variation as a case of intra-writer variation: Middle English scribal behaviours, with a focus on the spelling variation of woman in MS Pepys 2125. *Intra-Writer Variation in Historical Sociolinguistics* (Historical Sociolinguistics 5), ed. by Markus Schiegg & Judith Huber. 473–490. Oxford & New York: Peter Lang Publishing.

Kryk-Kastovsky, Barbara 2006. Impoliteness in Early Modern English courtroom discourse. *Journal of Historical Pragmatics* 7 (2): 213–243. https://doi.org/10.1075/jhp.7.2.04kry

Nevala, Minna 2009. *Friends* will be "friends"? The sociopragmatics of referential terms in early English letters. *Corpora: Pragmatics and Discourse. Papers from the 29th International Conference on English Language Research on Computerized Corpora*, ed. by Andreas H. Jucker, Daniel Schrier, & Marianne Hund. 81–101. Amsterdam & New York: Rodopi.

Nevala, Minna 2018. "Ungenteel" and "rude"? On the use of *thou* in the eighteenth century. *Patterns of Change in 18th-century English: A Sociolinguistic Approach* (Advances in Historical Sociolinguistics 8), ed. by Terttu Nevalainen, Minna Palander-Collin, & Tanja Säily. 77–95. Amsterdam: John Benjamins. https://doi.org/10.1075/ahs.8.06nev

Nevalainen, Terttu 1994. Ladies and Gentlemen: The generalization of titles in Early Modern English. *English Historical Linguistics 1992* (Current Issues in Linguistic Theory 113), ed. by Francisco Fernández, Miguel Fuster, & Juan José Calvo. 317–327. Amsterdam & Philadelphia: John Benjamins.

Nevalainen, Terttu 2001. Continental conventions in early English correspondence. *Towards a History of English as a History of Genres* (Anglistische Forschungen 298), ed. by Hans-Jürgen Diller & Manfred Görlach. 203–224. Heidelberg: Universitätsverlag C. Winter.

Nevalainen, Terttu 2009. Grasshoppers and blind beetles: Caregiver language in Early Modern English correspondence. *The Language of Daily Life in England (1450–1800)*, ed. by Arja Nurmi, Minna Nevala, & Minna Palander-Collin. 137–164. Amsterdam & Philadelphia: John Benjamins.

Nevalainen, Terttu 2015. What are historical sociolinguistics? *Journal of Historical Sociolinguistics* 1 (2): 243–269. https://doi.org/10.1515/jhsl-2015-0014
Nevalainen, Terttu 2018. Going to completion: The diffusion of verbal *-s*. *Patterns of Change in 18th-century English: A Sociolinguistic Approach* (Advances in Historical Sociolinguistics 8), ed. by Terttu Nevalainen, Minna Palander-Collin, & Tanja Säily. 97–116. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/ahs.8.07nev
Nevalainen, Terttu, Minna Palander-Collin, & Tanja Säily (eds.) 2018. *Patterns of Change in 18th-century English: A Sociolinguistic Approach* (Advances in Historical Sociolinguistics 8). Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/ahs.8
Nevalainen, Terttu & Helena Raumolin-Brunberg 1995. Constraints on politeness: The pragmatics of address formulae in early English correspondence. *Historical Pragmatics: Pragmatic Developments in the History of English* (Pragmatics and Beyond New Series 35), ed. by Andreas Jucker. 541–601. Amsterdam & Philadelphia: John Benjamins.
Nurmi, Arja & Päivi Pahta 2004. Social stratification and patterns of code-switching in early English letters. *Multilingua* 23 (4): 417–456. https://doi.org/10.1515/mult.2004.23.4.417
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, Minna Palander-Collin, Minna Nevala, & Arja Nurmi 2010. Language practices in the construction of social roles in Late Modern English. *Social Roles and Language Practices in Late Modern English* (Pragmatics & Beyond 195), ed. by Päivi Pahta, Minna Nevala, Arja Nurmi, & Minna Palander-Collin. 1–27. Amsterdam & Philadelphia: John Benjamins.
Raumolin-Brunberg, Helena 1991. *The Noun Phrase in Early Sixteenth-century English: A Study Based on Sir Thomas More's Writings* (Mémoires de la Société Néophilologique de Helsinki 50). Helsinki: Société Néophilologique.
Robbins, R.H. 2008. Browne, Sir Thomas (1605–1682), physician and author. *Oxford Dictionary of National Biography*. Retrieved 16 Dec. 2023.
Schiegg, Markus & Judith Huber (eds.) 2023. *Intra-Writer Variation in Historical Sociolinguistics* (Historical Sociolinguistics 5). Oxford & New York: Peter Lang Publishing.
Schilling, Natalie 2018. Investigating stylistic variation. *The Handbook of Language Variation and Change*, ed. by J. K. Chambers & Natalie Schilling. 327–349. 2nd edition. Hoboken, NJ: Wiley Blackwell.
Schreiber, Roy E. 2008. Hay, James, first earl of Carlisle (c. 1580–1636). *Oxford Dictionary of National Biography*. Retrieved 12 Dec. 2023.
Siirtola, Harri, Terttu Nevalainen, Tanja Säily, & Kari-Jouko Räihä 2011. Visualisation of text corpora: a case study of the PCEEC. *How to Deal with Data: Problems and Approaches to the Investigation of the English Language over Time and Space* (Studies in Variation, Contacts and Change in English 7), ed. by Terttu Nevalainen & Susan M. Fitzmaurice. Helsinki: VARIENG. https://varieng.helsinki.fi/series/volumes/07/siirtola_et_al/

The studies in this volume cover various aspects of the study of multilingualism and language variation in English across genres and registers, as well as across centuries.

Voir aussi www.ufy.fi
ISSN 2984-0961 (online)
ISSN 0355-0192 (print)

MODERN LANGUAGE SOCIETY OF HELSINKI
SOCIÉTÉ NÉOPHILOLOGIQUE DE HELSINKI
NEUPHILOLOGISCHER VEREIN IN HELSINKI