Writing History in Late Modern English

Explorations of the Coruña Corpus

Edited by
Isabel Moskowich, Begoña Crespo,
Luis Puente-Castelo and
Leida Maria Monaco

John Benjamins Publishing Company

Copyright 2019. John Benjamins Publishing Company. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 2/10/2023 3:18 AM via
AN: 2294299 ; Moskowich, Isabel, Universidade da Coruna, Crespo Garcia, Begona, Puente-Castelo, Luis Miguel, Monaco, Leida Maria.; Writing History in Late Modern English : Explorations of the Corua Corpus
Account: ns335141

Writing History in Late Modern English

Writing History in Late Modern English

Explorations of the Coruña Corpus

Edited by

Isabel Moskowich
Begoña Crespo
Luis Puente-Castelo
Leida Maria Monaco
University of A Coruña

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/z.225

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2019027979 (PRINT) / 2019027980 (E-BOOK)

ISBN 978 90 272 0424 0 (HB)
ISBN 978 90 272 6201 1 (E-BOOK)

Cover image: Tomb of Sir John Moore in the Jardín de San Carlos, A Coruña
© 2015, Merce https://www.flickr.com/photos/merceblanco/20668333459 CC BY-SA

© 2019 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · https://benjamins.com

Table of contents

About this book vii

Writing History in Late Modern English: Explorations of the Coruña Corpus – *A Preface* 1
Javier Pérez-Guerra

CHAPTER 1
A review of the development of historical writing and writers in English from 1700 to 1900 5
Elena Alfaya Lamas

CHAPTER 2
"There were always Indians passing to and fro": Notes on the representation of Native Americans in CHET documents 21
Marina Dossena

CHAPTER 3
An introduction to CHET, the *Corpus of History English Texts* 41
Isabel Moskowich

CHAPTER 4
Typical linguistic patterns of English History texts from the eighteenth to the nineteenth century: An information-theoretic approach 57
Stefania Degaetano-Ortlieb, Katrin Menzel and Elke Teich

CHAPTER 5
Exploring the narrative dimension in late Modern English History texts 83
Leida Maria Monaco

CHAPTER 6
Time and history: A preliminary approach to binomials in late Modern English Astronomy and History texts 103
Paloma Núñez-Pertejo

CHAPTER 7
"Were this eſtimation, however, to be depended on": Inversion conditionals as evidence of paradigmatic change in CHET **129**
Luis Puente-Castelo

CHAPTER 8
Modal verb categories in CHET **149**
Francisco Alonso-Almeida and Francisco J. Álvarez-Gil

CHAPTER 9
A corpus-based study of some certainty adverbs in the *Corpus of History English Texts* **167**
María José Esteve Ramos and Inés Lareo

CHAPTER 10
How intimate was the tone of female History writing in the Modern period? Evidence from the *Corpus of History English Texts* **185**
Begoña Crespo

CHAPTER 11
Neither *I* nor *we*: Inexplicit authorial voice in eighteenth century academic texts **215**
Margarita Mele-Marrero

CHAPTER 12
Do writers express the same attitude in historical genres? A contrastive analysis of attitude devices in the *Corpus of History English Texts* **235**
María Luisa Carrió-Pastor

CHAPTER 13
On cognitive complexity in scientific discourse: A corpus-based study on additive coherence relations **259**
Iria Bello

Index **277**

About this book

In the case of the preceding volumes about the *Coruña Corpus of English Scientific Writing* (Moskowich & Crespo, 2012 and Moskowich *et al.*, 2016), the corpora themselves were compiled and saved in a CD-ROM included with the book so that both the corpus and the works about it came out together.

On this occasion, however, this book reaches the public some time after the corpus it deals with has been completely finished. In fact, all the studies gathered here have been written using a version of the *Corpus of History English Texts* that is not the one available at present, but an earlier one (which we have called beta version) that, although very similar, may have some slight difference in word-count of some samples. The *Corpus of History English Texts* (CHET) has been published in open-access with a permanent URL by the University of A Coruña Press. To use CHET or any other corpus in the collection, please download both the corpus and the *Coruña Corpus Tool*, check the manual and follow the steps there to successfully load the corpus.

We want to thank Benjamins for granting us permission for the open access re-issue of the *Corpus of English Texts on Astronomy* (CETA) and the *Corpus of English Philosophy Texts* (CEPhiT) as part of the Coruña Corpus open-access collection.

We would also like to thank the University of A Coruña (UDC) for its support and the use of infrastructures and very especially the people at the University Press as this was a new experience both for them and for the people in MuStE. The efforts of Dr Fernández-Ferreiro, director of the University publishing service, certainly produced a rewarding result. MuStE is indebted to him.

Our gratitude to the Spanish MINECO for the grants received inside their National Programme for Excelence in Scientific and Technical Research (grant numbers FFI2013-42215-P and FFI2016-75599-P). This partially covered the compilation of CHET and its presentation to the academic community.

The Editors
A Coruña, April 2019

CHET: https://ruc.udc.es/dspace/handle/2183/21849
The *Coruña Corpus of English Scientific Writing* collection: https://ruc.udc.es/dspace/handle/2183/21846

Writing History in Late Modern English

Explorations of the Coruña Corpus – *A Preface*

Javier Pérez-Guerra
Universidade de Vigo

Objectivity has traditionally been understood to constitute the main, defining, distinctive feature of historical writing. From a textual perspective, the purpose of ensuring objectivity is reflected in a number of conventions in the genre, such as the extensive use of notes and footnotes, copious reference to previous works, documentation, etc. However, the way in which objectivity has been materialised diachronically in historical writing, and thus its linguistic reflex, has been subject to variation and change over the course of earlier centuries. From the lively, individualistic, narrative-based historical literature of the Renaissance to the focus on sociological (and, later, socio-economic) issues in the Enlightenment; from story-telling in the past to pure explanations in the eighteenth century; from chronicles of "folly, irrationality, and religious superstition" (Baker, 2002: 187) to empirical factual investigation at the close of the nineteenth century, when historical knowledge was seen as having to be transformed into facts and quantitative data (Raphael, 2011: 109). In this vein, Baker (2002: 195) notes that for nineteenth-century historians "the study of history was (…) a kind of inductive science (…) that offered to show how peculiar circumstances of time and place determine the kind of law and government best fitted to a specific people"; that is, "[h]istory was still predominantly perceived as an art form, but the professionalizing historians also began pointing to history as something in between art and science" (Berger, 2011: 25).

It is in this context that the *Corpus of History English Texts* (henceforth CHET) provides the perfect textual data source for rigorous analyses of historical writing in 1700–1900 Britain. The choice of this timeframe, as Moskowich (2017: 85; also Isabel Moskowich, this volume) explains, is "motivated by the socio-historical context of scientific writing [and] the rise of the scientific method", which puts its own stamp and identity on historical writing. As Woolf (2012: 473) has observed, historical writing grew significantly in importance from the Late Middle Ages, when History was a "minor genre", to become "the most widely read of literary forms and the chosen subject of hundreds of writers" by around 1800.

https://doi.org/10.1075/z.225.pre

© 2019 John Benjamins Publishing Company

CHET, the most recent component of the ambitious Coruña Corpus project, comprises more than 400,000 words of eighteenth- and nineteenth-century texts. The (raw) texts are enriched with TEI markup, providing not only information about textual features but also metadata on the authors (sex, age, geographical provenance) and the texts (year of publication, sample chosen, reception). Since textual balance is key in corpus compilation, Moskowich and her team have selected two samples of approximately 10,000 words each per decade, these extracted from different sections of (mainly) first editions of original works.

'From his own diary we learn': *Investigating the Corpus of History English Texts*, based on CHET, is a collection of scholarly studies dealing with different aspects and dimensions of the writing of History texts in the recent past. After a technical chapter providing some background about the discipline in the period, the trigger for compilation and the description of the corpus material (Isabel Moskowich), the volume contains outstanding explorations on the development of historical writing (Elena Alfaya), historiographic contributions (Marina Dossena) and analyses of linguistic patterns and constructions ranging from recurrent lexical combinations (Stephania de Gaetano-Ortlieb, Katrin Menzel & Elke Teich) to binomials (Paloma Núñez-Pertejo) and conditional structures (Luis Puente-Castelo). It also contains chapters looking at discourse-related strategies, such as the use of modals (Francisco Alonso & Francisco J. Álvarez-Gil), certainty adverbs (María José Esteve & Inés Lareo), authorial presence through the use of attitude devices (Carrió-Pastor), additive markers (Iria Bello), proxies for 'narrativity' from a multidimensional perspective (Leida Maria Monaco) and genre-specific modulating resources (Margarita Mele). Finally, the gender dimension is addressed in a chapter on female History writing (Begoña Crespo).

Rigour, meticulousness, thoroughness, attention to detail, exactness, exactitude, precision, strictness, and punctiliousness are all concepts which are central to historical writing(s). In a nutshell, objectivity, as Wright (2002: 9) puts it, "has often been a kind of 'noble dream' in historical thought", and this is well reflected in all the studies included in *'From his own diary we learn'*: *Investigating the Corpus of History English Texts*. Taking a variety of approaches, these contributions challenge the view that "British historical writing has no shape, that there exists no pattern beyond the randomness of spontaneous authorship and occasional moments of inadvertent intertextuality" (Bentley, 2011: 204).

References

Baker, Thomas N. (2002). National History in the Age of Michelet, Macaulay, and Bancroft. In Lloyd S. Kramer & Sarah C. Maza (Eds.), *A Companion to Western Historical Thought* (185–204). Oxford: Wiley-Blackwell.

Bentley, Michael. (2011). Shape and Pattern in British Historical Writing, 1815–1945. In Stuart Macintyre, Juan Maihuashca & Attila Pók (Eds.), *The Oxford History of Historical Writing. Volume 4: 1800–1945* (204–224). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199533091.003.0011

Berger, Stefan. (2011). The Invention of European National Traditions in European Romanticism. In Stuart Macintyre, Juan Maihuashca & Attila Pók (Eds.), *The Oxford History of Historical Writing. Volume 4: 1800–1945* (19–40). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199533091.003.0002

Moskowich, Isabel. (2017). Genre and Change in the Corpus of History English Texts. *Nordic Journal of English Studies*, 16(3), 84–106.

Raphael, Lutz. (2011). 'Experiments in Modernization': Social and Economic History in Europe and the United States, 1880–1940. In Stuart Macintyre, Juan Maihuashca & Attila Pók (Eds.), *The Oxford History of Historical Writing. Volume 4: 1800–1945* (97–114). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199533091.003.0006

Woolf, Daniel. (2012). Historical Writing in Britain from the Late Middle Ages to the Eve of Enlightenment. In José Rabasa, Masayuki Sato, Edoardo Tortarolo & Daniel Woolf (Eds.), *The Oxford History of Historical Writing. Volume 3: 1400–1800* (473–496). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199219179.003.0024

Wright, Johnson Kent. (2002). Historical Thought in the Era of the Enlightenment. In Lloyd S. Kramer & Sarah C. Maza (Eds.), *A Companion to Western Historical Thought* (123–142). Oxford: Wiley-Blackwell.

CHAPTER 1

A review of the development of historical writing and writers in English from 1700 to 1900

Elena Alfaya Lamas
Universidade da Coruña

This chapter is dedicated to the memory of Linda

> Here are presented the results of the enquiry carried out by Herodotus of Halicarnassus. The purpose is to prevent the traces of human events from being erased by time, and to preserve the fame of the important and remarkable achievements produced by both Greeks and non-Greeks.
>
> (*Herodotus, The Histories, 440 BC*)

1. Introduction

The establishment of History as an area of scientific knowledge dates back to the nineteenth century. Prior to this, the study of History was largely disdained in the universities and not seen as worthy of serious consideration. As a branch of knowledge it was regarded with certain contempt, one which was neither deep nor serious in its scope or potential.

In this chapter, we will offer an account of the development of History writing in the late Modern English period, from 1700 to 1900, two hundred years in which a vast amount of such writing was produced in Britain. Our aim is to review the discipline History, the process of change and development of English historical writing, and to consider some representative examples of eighteenth and nineteenth-century History writers from CHET, the *Corpus of History English Texts*. The chapter is arranged as follows. In Section 2 we present a conceptual frame of reference for the description of the concepts of *History, Historiography and Philosophy*, this followed by the establishment and evolution of the current academic discipline of History and the profession of historian. Section 3 presents an outline and brief History of the discipline itself, from antiquity to the nineteenth century, focusing on general

https://doi.org/10.1075/z.225.01alf

© 2019 John Benjamins Publishing Company

tendencies for writers of History and paying special attention to the 1700–1900 period. We follow this with a review of the main historical societies and institutions in Britain at the time. In Section 4 we present some concluding remarks.

This chapter, then, will set the scene for further chapters in the current book, describing History in its social context during the late Modern English period. It seeks to offer the reader a sense of the development, engagement and concerns of English historical writing and writers in this period and covered by CHET, and thus to serve as a framework for the other chapters herein.

2. Conceptual frame of reference

It is in the late Modern period when a division between the concepts of *History, Philosophy of History* and *Historiography* begins to emerge. In previous periods there is no clear difference between these terms, and it is not possible to refer to *History* as a discipline or a distinct branch of knowledge. In the late Modern English period we begin to discern it within the areas of rhetoric and theological studies. For this reason, it is important to establish a conceptual framework in which key concepts are defined and inter-related, and which thus provides a recognizable identity for the writing of History. However, we must bear in mind that even today there are many different approaches to this academic subject.

History, viewed in a general way, can be seen as the written or oral narration of past events and movements related to humans and their habitat. Archival records and inscriptions in old buildings are the main sources for historians in reconstructing the past. History, then, is “the discipline that studies the chronological record of events based on a critical examination of source materials” (Vann, 2018).

History was previously not considered an academic discipline *per se*, and as such was not conducted on the lines of the scientific method. It was between the mid-eighteenth and the mid-nineteenth centuries that it began to develop a methodology of its own, when the predilection for an approach on the lines of Voltaire, an intellectual *Philosophical History*, came to dominate in England (O’Brien, 2001). Indeed, we can say that in Britain, *History* and *historian* are concepts that belong firmly to the nineteenth century. *Philosophy* is today defined as “the rational, abstract, and methodical consideration of reality as a whole or of fundamental dimensions of human existence and experience” (Vann, 2018). However, the *Encyclopaedia Britannica* adds that *Philosophy* cannot be defined in a definitive or exhaustive way. Meanwhile, the term *Philosophy of History* refers to the “philosophical examination, study, and theorizing about the past, including substantial/speculative philosophy of History but also issues like contingency and necessity in History” (Tucker, 2009: xii). Thus, philosophers of History conduct philosophical examinations of the discipline. This term was "used first by Abbé Bazin in the title of his book *La philosophie de l'histoire* published in Amsterdam in 1765" (Tucker, 2009: 26). *Philosophy of Historiography*, in turn, refers to "philosophical examination, study and theorizing about historiography, about what historians write, and its relation to the evidence, […]" (Tucker, 2009: xii) and thus examines, in a philosophical way, the epistemology of what we know about *History*, including issues such as accuracy, objectivity and skills used to describe the past, looking for evidence in the methods and concepts used by historians and analysing these.

During the eighteenth century, theology and natural sciences were considered to be part of Philosophy, and History was not a distinct discipline. Indeed, a diversity of what we now consider independent disciplines were subsumed in the eighteenth-century notion of *Philosophy*, "and many of the divisions among what is and what is not philosophy, that we take for granted – i.e. between philosophy and different areas of empirical inquiry and theology – were in the process of developing or had not yet taken shape" (Garrett, 2014). In the eighteenth century, Philosophy in its broadest sense comprised much of what related to the experience of being human, the term itself dating back to c. 1325 (Moskowich, 2016).

Currently, *Historiography* is understood as the inquiry into History as a branch of science. It is a relatively new area of historical studies whose scope is the study of documents as evidence of the past. It is "what historians write, about past events, about History" (Tucker, 2009: xii). It is, then, the History of *History* itself, and deals with the History of the different evolutionary phases of historical writing. As Arthur Marwick (2001) observes, Historiography is the History of historical thought, a definition usefully expanded by Vann (2018):

> The writing of History, especially the writing of History based on the critical examination of sources, the selection of particular details from the authentic materials in those sources, and the synthesis of those details into a narrative that stands the test of critical examination. The term historiography also refers to the theory and History of historical writing.

This idea of Historiography is quite recent in the West. The methods used nowadays by historiographers emerged and were gradually developed during the late eighteenth century, mainly in Germany and France. The concept of Historiography as an activity in fact dates back to the late-eighteenth and mid-nineteenth centuries. As Zammito (2009: 65) notes:

> Historiography established itself as a discipline between the late eighteenth and the mid-nineteenth century, it struggled explicitly to free itself from philosophy of history. Historiography achieved its "paradigmatic" form in the mid-nineteenth century, […] as explicitly anti-philosophical.

Prior to the eighteenth century, Historiography was almost never considered a part of science or education and did not form part of university curricula. It was in the nineteenth century that a distinction between Historiography and Philosophy developed. Whereas in the early and mid-eighteenth century there is no distinction between History, Historiography and Philosophy, with historical practice and writing disparate and sometimes amateurish, the transition to the nineteenth century is a time of intellectual ferment, and marks the beginning of a change. History becomes a professional practice (see Moskowich, this volume) and an independent discipline at the same time that universities begin to consider it a discipline. Thus, in 1895 Oxford created the postgraduate research degree of Bachelor of Letters or *Baccalaureatus Litterarum*, including a compulsory thesis dealing with some historical events or processes, something that clearly benefited the rise of History as a profession in Britain.

It is perhaps worth noting that there is a difference between the discipline, that is, a single branch of knowledge of which all historians are representatives, and the modern profession, referring to many different occupations in which historians develop their work. Every historian contributes to the discipline but not every historian is part in the professional field itself. As James M. Banner argues, "A discipline is a domain of knowledge, a capacious province of inquiry […]. It is a universe of thought, not, like a profession, an arena of action" (2011: 3). This is an important distinction, because when the professional work of writers of History was recognised, male historians had the opportunity to become "professional historians", previously having been regarded as amateurs. Nevertheless, women at the time were not allowed to access professional posts, and as a consequence their writings continued to be considered of an amateur nature. Professions, distinct from disciplines, are occupations to which access may or not be allowed in that they are controlled by the dominant groups in society. During the eighteenth century Historiography was dominated by men, and thus women's historiography was bypassed almost by default. In the nineteenth century the academic discipline was professionalised on a wholly male footing, and hence women remained relegated from the professional world, which restricted them to the "amateur historian" condition, despite their competence in the field. By the end of the century the profession of historian had become identified with the academic discipline, which entailed that male historians were also academics, an insurmountable obstacle that overshadowed female historians as independent amateurs. This being so, the genres that women tended to focus on included travel, biographies and culture. However, female historians made notable efforts to become professionals, writing systematic, well-argued and detailed formal pieces of historical writing. Such efforts, but also the fact that they remained relegated from the professional ranks, can be inferred not only from the field of History itself, but from the male-female representation of History writers in CHET. Bonnie Smith (2000) explains that during the eighteenth century, women considered most of their History writings "superficial" in that they would write about social issues, whereas men tended to write about political and economic events. Many men held the same views about female History writers. Indeed, the German Leopold von Ranke (1795–1886) and the French Gabriel Monod (1844–1912) were motivated to professionalise the discipline partly by their despise for female History writers, considering the output of women here to be un-scientific and sub-standard (Alvarado, 2012).

Many academic disciplines other than History were professionalised in this way in the mid- or late-nineteenth century, so in this sense History was no different from many academic disciplines at the time. Also at the time many European universities were secularised, the curricula were changed, and new disciplines were included so as to complete and enhance traditional teaching. History progressively established itself as an academic discipline in Europe, not least as a result of the efforts of scholars working at the University of Göttingen (founded in 1737), who dedicated much of their time and resources to History. Thus research began to expand to other universities, and a number of academic journals and scientific academies emerged, as we will see in Section 3. The University of Manchester was the first to follow the Göttingen model and professionalise the discipline by means of postgraduate studies.

According to Hesketh (2011), History became a professionalised discipline between the 1860s and the 1890s in Victorian England under the influence of Natural Philosophy and the German historians. In 1900 it had already become a settled academic discipline, with the introduction of stable skills, publishing standards and principles of exactness and scholarly precision.

3. General outline of the discipline and its main representatives

This section presents a contextualised overview of History from its earliest inception to 1900. It aims to describe the diachronic development of the discipline over a considerable time-span, with the late Modern period as our principal focus.

History begins with the earliest extant writings of the ancient literate civilizations: the Greek, the Mesoamericans and the Chinese cultures. And although Western History provides ample examples of writers and chronicles that predate the Greek Herodotus (484–425 BC), he is known as the "Father of History". His work *The Histories* has served as a basis for the emergence and establishment of History as a branch of knowledge due to the nature of the inquiries he made in his treatment of the historical subjects that he dealt with therein. The models and systems used in past cultures were often very different to current ones, and Historiography is the branch of study which considers the development of these. The concept of *Universal History* became widely accepted in the first century BC, although it had previously been used with the Conquests of Alexander III of Macedonia (356–323 BC). The Romans inherited their Historiography from the Greeks, but being eminently practical in so many things, they were interested more in History than Philosophy (Sreedharan, 2000). The rise of Christianity goes hand in hand with the decline of Historiography in Europe, and from the fifth century secular Historiography dies out, with History coming to be almost exclusively written about Catholic monks and other members of clergy, an approach that lasted some eight centuries. It is during these dark ages for culture and History writing that the Venerable Bede, so-called "Father of English History", stands out in England, earning enduring fame for his historical writing *Historia Ecclesiastica.* It is said to be the best History of its time and the only substantial source of information of its kind for the years 597–731. Written in England in 731, many future generations of historians would derive their own writings from this text. During the Middle Ages most writings on History were again focused on sacred matters. The thirteenth century then saw Richard Hoveden's *The Annals of English History*. Annals are accounts of a year in list form, and themselves emerged from English religious scriptoria. Missionaries had spread the practice of writing annals in the continent from the seventh century.

During the Renaissance, society and culture are reshaped and transformed in Europe. Within this long period, broadly taken to extend between the fourteenth and seventeenth centuries, History continues to be approached mainly from a religious point of view. Likewise, Europeans begin to discover Arabic translations of philosophical Greek texts, and these would subsequently come to influence European History writers. In all these earlier periods, we do not find professional historians. Rather, they were amateurs who worked in other spheres and from time to time dedicated their efforts to the writing of History. From 1400 to c.1550, four main genres were adopted in such writing: clerical historiography, chivalric and militaristic historical writing, biographies of notable people, and urban chronicles (Woolf, 2012).

From the sixteenth to the eighteenth centuries, both English and Scottish writers of History adopted continental Humanism. Indeed, "It is also now widely recognized that the Renaissance […] elevated Scottish historiography onto a new intellectual level" (Allan, 2012: 498). The Scottish Hector Boece (1465–1536) published the biography of the bishops of Murthlack and Aberdeen in 1522, *Vitae Episcoporum Murthlacensium et Aberdonensium*, and also a History of the Scots, *Historia Gentis Scotorum* (1527). George Buchanan (1506–1582) also wrote a History of Scotland, *History of Scotland, Rerum Scoticarum Historia* (1582), in which a section is dedicated to the History of Hector Boece.

During the seventeenth century knowledge was mainly focused on natural science, with History still labelled as a "dubious science", as stated by the French René Descartes (1596–1650). However, Philosophy made some contributions to History and Historiography, and in the mid-eighteenth century a vogue for a Voltaire-inspired *Philosophical History* took hold in England.[1]

The end of the Renaissance is particularly associated with (i) the profound shift from manuscript to printing, resulting in a wider transmission of knowledge, distribution and availability of texts, and (ii) a change in thought, attitude and practice, bringing about the scientific revolution in Europe. "Developments in science during the scientific Revolution of the sixteenth and, especially, the seventeenth centuries defined modern science in many ways" (Puente-Castelo, 2016: 167). This period was one of gradual development and change in attitude, policy, opinion and new tendencies, with *reason* considered the most important and basic foundation for everything else. Modern science is essentially born in this period. Scientists of the late Romantic period could be labelled as amateurs in comparison to present-day ones, in that their methods could be considered not so systematic and rigorous as is currently expected. However, this period also represents a shift, since scientists were for the first time somehow rationally trained, having free time for research and reflection on the shortcomings of scholastic philosophy,[2] and this had a considerable quantitative and qualitative impact on the Age of Enlightenment.

The ideological crisis that takes place between 1680 and 1720 in European thought leads to the Enlightenment and hence to modern science. The works of the English philosophers John Locke and Francis Bacon, the French René Descartes and the Dutch Baruch Spinoza, would all be influential in establishing Enlightenment ideas and the scientific method. For instance, Bacon would emphasise the importance of observation and experimentation and the need to avoid opinions. This is also the time when academic journals emerged.

Thus it was not until the eighteenth century, in the Enlightenment Age,[3] that History begins to find both a wider audience and acceptance as a branch of science. To the Enlightenment we also owe the discovery of the great intellectual traditions of China, as well as its historians, historical traditions and uses, with its historiography in particular being unequalled in quality. The new intellectual atmosphere was also remarkable because of the special value it gave to the scientific method and to the philosophical reductionist approach, continuing what the Renaissance had already brought into existence. In the mid-eighteenth century the particular methods and forms of Philosophical History became fashionable in England. In fact, it had been “Voltaire himself who gave the title of *philosophy of History* to the first part of his *Essai sur les Moeurs et L’Esprit des Nations*” (Rosenthal, 1955: 151), using the term not only to refer to the past life of mankind but also to its culture, including in this second reinterpretation of the term the traditional activities, conventional habits, ethic values, customs and practices of human societies. This latter use of the term, then, represents Voltaire’s *Philosophy of Culture* and leads to a focus on cultural aspects of History and knowledge. On the other hand, the French Montesquieu (1689–1755) would focus on the social and institutional aspects of the discipline. Likewise, the French philosophers Jean-Jacques Rousseau and Denis Diderot, the Italian philosopher and jurist Cesare Beccaria, the Spanish Benito Feijoo, the German philosopher Immanuel Kant and the Scottish writers David Hume (author included in CEPhiT), Henry Home, Adam Smith, Adam Ferguson, Thomas Reid (both Ferguson and Reid are included in CHET) and John Millar, all represent the practice of Philosophy in the Enlightenment Age.

---

[1] See Karen O'Brian (2012) for a further account on this.

[2] A system of critical Philosophy that dominated European thought from 1100 to 1700.

[3] "Enlightenment" is more a retrospective label than the one used in this period by its representative authors. It refers to most of the eighteenth century although some of its historians belong to the late seventeenth century and others to the early nineteenth century.

The eighteenth century is a turning point for History in that there arises a new conception of the field. The Member of Parliament and English historian Edward Gibbon (1737–1794) now writes on History and his works are widely read; his historical writings serve to disseminate History as a source of knowledge, and indeed his works are still read today. He channelled the aims and scope of historical writing into a new philosophical understanding, as was also the case with the Scottish writer David Hume (1711–1776), as we can see from his six-volume *History of England*. As O’Brian (2012: 518) notes:

> Gibbon absorbed and greatly accelerated the general transformation of English historical writing in this period towards a new, ‘philosophical’ conception of its purpose and scope. This transformation owed an enormous amount to the Scottish Enlightenment in general, and to David Hume’s History of England (1754–62) in particular.

During the eighteenth century, the Scottish William Robertson (1721–1793) “showed the way to a somewhat different kind of philosophical History that set political events in a richer, broader context of legal and social developments, arts, sciences, and ‘manners’” (O’Brien, 2012: 521). There is, then, a fresher and more dignified view of History as a means of understanding people’s nature and actions. “This more contemplative view of the nature and value of historical knowledge certainly provided an entrée for female readers” (O’Brien, 2012: 522). Catharine Macaulay (1731–1791), one of the authors whose work is sampled in CEPhiT (the Philosophy subcorpus of the Coruña Corpus; see Moskowich *et al.*, 2016), is the first recorded British historian adopting a philosophical approach, and many others followed. Her *Histories of England* in fact outsold David Hume's work, and in addition to this, her periodical publications on political, war and other matters were also extremely popular. As such she can be regarded as a prominent, solid and intelligent historian.

Scottish historians – or philosophers – intended to disavow old models of writing History, being instead somewhat innovative in their approach, as reflected in the works of authors such as William Robertson and the already mentioned David Hume, Adam Ferguson, Adam Smith and John Millar. They were known to each other socially and maintained an intellectual dialogue which would have an impact on future generations of historians during the nineteenth century, not only in Britain but also on the continent.

The movement of Romanticism followed on from the Enlightenment at the end of the eighteenth century and continued through the following century. Romanticists were interested in all the aspects and likely consequences of an excessive dependence on reason. They argued that it was an error of Enlightenment to have ignored the close connection between History and tradition. In the late eighteenth century a more philosophical way for writing History was sought. In the late 1760s Joseph Priestly (also included in the *Coruña Corpus of English Scientific Writing*) was lecturing on History at the Warrington Academy[4] and described a framework and pattern to be followed for the description of historical knowledge that would lead humanity to share common values and principles, independently of their place of origin.

The nineteenth century was a time in which historical writing in Britain reached its peak, history books matching the sales of novels by the end of the century. The German Georg Wilhelm Friedrich Hegel (1770–1831) had initiated a non-religious perspective for historical-philosophical writing, extending the idea, a widely welcomed one, that philosophical History is the right way to write *History*. One of the most acclaimed historians of the nineteenth century was the German Leopold von Ranke whose methodology would be a path to follow for many historians.

It is in the nineteenth century that History became an acceptable and approved academic discipline, and also when its standards and methods were established in a way which was intended to be permanent. Historiography became part of History and was historicised. As Tucker (2009: 39) observes, "historiography is more productive than ever and is becoming ever more and more specialized and diverse. Simultaneously, the world is becoming globalized and world History is becoming increasingly relevant". In this way, and "as part of this process, questions regarding the nature of historical knowledge were posed more and more frequently. This can be understood as the starting-point of the philosophy of historiography" (Tucker, 2009: 37). In the nineteenth century female historians are trying to find a place and to make a living alongside their male counterparts, but finding insuperable barriers here, as the prolific and well-respected historian Mary Hays (1759–1843) describes in *Female Biography* (1803). Unlike their male counterparts, female historians, however serious and prolific, would tend to disappear, simply erased from literary History and absent from contemporary works. Thus it was that Hays published *Historical Dialogues for Young Persons* in 1806, as well as *The History of England* with Charlotte Smith, *Memoirs of Queens* in 1821, and penned many other historical works, children books, biographies and works in other genres. In 1818 the prolific historian and translator Lucy Aikin (1781–1864) (included in CHET) published *Memoirs of the Court of Queen Elizabeth*. Many historical works and other genres would follow, in a career which had begun in 1801. The historian, biographer, translator, novelist, compiler and poet Elizabeth Benger (c.1775–1827) wrote the biographies of *Anne Boleyn* in 1821 and *Mary, Queen of Scots* in 1823. In 1809 she had written the innovative and socially relevant poem *The Abolition of Slave Trade*, although this was considered ridiculous by the press at the time. In 1828 *A Comparative View of the Social Life of England and France* was published by the social historian, biographer and editor Mary Berry. Many of these female historians have been erased from major History books, yet their individual efforts brought them a place in other *Histories*, ones that that have perpetuated their names and works. As Crespo (2012: 31) states, "it is possible to regard science as a social construct in which particular epistemic communities developed under the protection of individual effort".

---

[4] In Cheshire. Active from 1756 to 1782.

Many institutions and societies concerned with historical and literary knowledge had been firmly established by the end of the nineteenth century. As already stated above, for the first time, and from the early part of this century, works on History were widely read, the modern profession had evolved and a diverse array of institutions were devoted to this recently accepted academic discipline.

It is not our aim to list all of these here but to give an account of some of the most influential ones in relation to History and other related branches of knowledge during the period 1700–1900. Many of these societies have had an enduring presence in the field, while others, more ephemeral and short-lived, were important in terms of what they produced. Some focused on a very specific area of interest. It is also important to recall the historical role that such institutions played in forming a male-only grid for the professionalisation of the discipline, its organisation, and the confinement of female practitioners to the amateur ranks. Most of these learned institutions were intended for the use and fellowship of men only, with membership barred to women. The only exceptions in the nineteenth century were the Zoological Society of London, founded in 1829, and the Royal Entomological Society, founded in 1833. Indeed, in Britain it was not until the first equal opportunities legislation in 1919 that women gradually came to be admitted into learned societies and gained employment as scholars, albeit to a limited extent. Previous to that, female historians were rejected as members or fellows of these Societies on the grounds of their sex. This legislation was not without problems, since many obstacles for females arose and further laws would be needed (Pugh, 1992). We might recall that it was only one year prior to this, in 1918, that women over 30 had been granted the right to vote.

The Royal Society of London had been founded in 1663 as a way of gathering scholars in London for meetings at which they would discuss science and natural knowledge. Its reputation increased over the years, and indeed was presided over by Isaac Newton during the first quarter of the eighteenth century. At this time, it became usual practice for the government to base its decisions related to science on the knowledge and advice of this exclusively male society. The rigidly male-only profile of the society persisted through the nineteenth century, and indeed female historians would not be admitted until the mid-twentieth century.

The Society of Antiquaries was founded in 1707 in London, with the main goal of studying British antiquities. It gathered many old historic manuscripts and collections of paintings, leading to the birth of the National Portrait Gallery in 1856, as a means of storing and exhibiting portraits of historically relevant British people. The Society of Antiquaries of Scotland was founded in 1780 by David Steuart Erskine (1742–1829), with the purpose of promoting, cultivating and preserving the study of natural and other sciences related to History, antiquities and Scottish records. In 1824 The Royal Asiatic Society of Great Britain and Ireland was founded in London. It aimed to research and promote science and literature relating to Asia. All its presidents and notable fellows, past and present, have been men, and female historians continue to be not considered for certain posts. In 1826 the Zoological Society of London was founded. Its goal was to collect animals free of their natural habitat and environment, to kill some of them for study and enslave others for human entertainment and leisure. It was instituted expressly "for the promotion of natural knowledge", as Scherren (1905: 1) puts it.

In 1833 The Royal Entomological Society was founded in London. The study of insects and the distribution and dissemination of entomological knowledge were its main goals. With this in mind, the Society established a library, a collection of dead insects, and a scholarly journal, the *Transactions of Entomological Society of London*. The Public Record Office was born in 1838 with the intention of collecting public court and governmental records and preserving them in good archival conditions. When founded, it was headed by the so-called Master of the Rolls, an experienced judge. The first of these was Lord Henry Bickersteth Langdale (1783–1851), who in turn employed Sir Francis Palgrave (1787–1861), archivist and historian, with overall responsibility as Deputy Keeper of the Office.

In 1838 the Camden Society was founded in London, and the Historical Society in 1868. Later, in 1897, these would merge under the new name of the Royal Historical Society. Again, the society was for men only as women were not allowed to enter or to become fellows. The Royal Historical Society would prove to be a major turning point in relation to the professionalisation of History:

> By the 1890s it was taking on more public, professional responsibilities – organising national events (such as the Domesday Commemoration of 1886), overseeing the teaching of History in schools and universities, and working closely with the British Museum and the Public Record Office – a tradition that continues today, with the British Library and the National Archives. (*The Royal Historical Society,* 2014)

The Ælfric Society was established in London in 1842 with the aim of publishing the works of Anglo-Saxon scholars and old ecclesiastical writers of note.

In 1864 the Early English Text Society was founded by Frederick James Furnivall (1825–1910). It was dedicated to the printing and editing of Old, Early and Middle English historical records, with all types of historical manuscripts intended to be thus published. In a similar vein, the Royal Commission on Historical Manuscripts was founded in 1869 with the aim of locating private historical records, as well as describing their contents. It subsequently merged with the Public Record Office, the guardian of British records since 1838, leading to The National Archives.

In 1886 the Scottish History Society was founded. It focused its efforts on the publishing of previously unpublished original, historical records from Scotland, including previously unpublished scholarly editions of Scottish History.

The learned British Record Society was founded in 1889. Its aim was to publish historical records and its indexes, as well as to conserve these. In 1898 its Scottish part became independent under the name Scottish Record Society. In 1899 a proposal to create a British Academy to promote historical, philosophical and philological studies is made. However, this would not be realised until 1902 when the British Academy came into being. Many other societies and institutions were founded during the nineteenth century that played a major role in the professionalisation of History. Since it is not our goal here to list and describe them all, we will simply name some of the most influential of those not mentioned above, together with the year in which they were founded: the Ray Society (1844) focusing on natural History; the Hakluyt Society (1846) concerning records of voyages and geography; the Harleian Society (1869) focussing on heraldry and genealogy records; the London Topographic Society (1880); the Pipe Roll Society, which dealt with old archival rolls; the Oxford Historical Society (1884) related to the publishing of the historic archival records of Oxfordshire; The British Association for the Advancement of Science (1831).

It was common for papers presented at these societies and institutions to be published in the periodical press or in volumes issued at regular intervals by the society itself. Many current academic journals have their origins in the proceedings of these societies. It was in 1886 that papers on History began to be published in Britain under the auspices of *The English Historical Review*,[5] this being the oldest scholarly journal of History in the country which remains active today. It was founded by the historian James Bryce (1838–1922), following the German journal *Historische Zeitschrift* (1859), the French *Revue Historique* (1876) and the *Rivista Storica Italiana* (1884). Yet it was back in 1665 that the first academic journal in Britain was published periodically, the *Philosophical Transactions of the Royal Society*. Other journals would follow over the course of the eighteenth, nineteenth and twentieth centuries with an exponential rate of growth.

All these journals accompanied the professionalisation of the discipline in Europe. They failed to attain a general readership but did reach scholars and those with a more specialised interest, establishing high standards of quality in the discipline.[6] At a time prior to periodicals, scholars would write their discoveries in letters, often in Latin and sent to other scholars and friends, perhaps finally being published in the transactions of an academy or society. In the seventeenth century "the journal probably was not accepted as a definitive form of publication and […] scholars frequently gathered together their scattered journal contributions into the form of a book" (Kronick, 1962: 61).

During the twentieth century many new universities and colleges were set up, by which time History had already earned its place as an academic discipline in Europe. As a measure of this, in the 1990s British historians published some 20,000 historical works about the eighteenth century (Corfield, 2011: 1). The discipline, then, had been firmly established.

4. Conclusion

Our aim here has been to present an overview of the History of *History* and its related jargon in the period covered by the *Coruña Corpus of English Scientific Writing*, 1700–1900. We note in particular that a differentiation between the terms *History, Philosophy* and *Historiography* was not clear in the period 1700–1900, and that to some extent this continues to be the case, with the terms sharing certain elements and indeed merging at times.

---

[5] For a further account on the History of the journal see Doris S. Goldstein (1986).

[6] See Philippa Levine (1986) for a further account on this.

Writers of History in this period were eager to know the past better and to write about it. They turned History into a discipline and professionalised it, with women historians left out of this process to a great extent. Men followed a professional path, whereas women had no way in, and instead were forced to follow an individual route of the so-called *amateur*, with the result that the impact they had on the writing of History remains opaque and difficult to perceive. In the late Modern period historians achieved such prominence in society that many scholarly societies were born, and a number of the current academic journals were established. Again, most of these excluded female intellectuals, this despite the evident quality of women's writing and also the substantial sales that many female writers enjoyed. Apart from the evident misogyny of the period, it is clear that in the period 1700–1900 historians in general came increasingly to be a good example of the historiographical achievements of the discipline as one of the high points in the History of History.

Projects such as CHET, the History sub-corpus of the *Coruña Corpus*, are of great interest in that they constitute a "resource for the study of scientific writing in English long before it became the *lingua franca* of science" (Moskowich, 2012: 35). As Joan C. Beal (2012: 12) has observed, "historians have long recognized the long eighteenth and nineteenth centuries as worthy of investigation". The various papers in the present collection are testament to the considerable amount of research still to be done on academic disciplines during this period.

References

Allan, David. (2012). Scottish Historical Writing of the Enlightenment. In Masayuki Sato; Jose Rabasa; Edoardo Tortarolo & Daniel Woolf (Eds.), *The Oxford History of Historical Writing*. Vol. 3. (497–517). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199219179.003.0025

Alvarado, Alice. (2012). Left out: Women's Role in Historiography and the Contribution of Mary Ritter Beard. *Saber and Scroll*, 1, 2, 89–103.

Banner, James M. (2011). *Being a Historian. An Introduction to the Professional World of History*. Cambridge: Cambridge University Press.

Beal, Joan C. (2012). What is language? In Charles Barber, Joan C. Beal & Philip A. Shaw (Eds.), *The English Language. A Historical Introduction* (1–30). Cambridge: Cambridge University Press.

Corfield, Penelope J. (2011). The exploding galaxy: Historical studies of Eighteenth-century Britain. *Journal of Eighteenth-century Studies*, 34, 4, 517–526. https://doi.org/10.1111/j.1754-0208.2011.00449.x

Crespo, Begoña. (2012). Astronomy as Scientific Knowledge in Modern England. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy "playne and simple". The writing of Science between 1700 and 1900* (15–34). Amsterdam/Philadelphia: John Benjamins.

Garret, Aaron. (2014). *The Routledge Companion to Eighteenth Century Philosophy*. New York: Routledge. https://doi.org/10.4324/9781315815558

Goldstein, Doris S. (1986). The Origins and Early Years of the English Historical Review. *English Historical Review*, 101, 398, 6–19. https://doi.org/10.1093/ehr/CI.CCCXCVIII.6

Hesketh, Ian. (2011). *The Science of History in Victorian Britain: Making the Past Speak*. London: Pickering & Chatto.

Kronick, David A. (1962). *A History of Scientific and Technical Periodicals. The Origins and Development of the Scientific and Technological Press 1665–1790*. New York: The Scarecrow Press.

Levine, Philippa. (1986). *The Amateur and the Professional: Antiquarians, Historians and Archaeologists in Victorian England, 1838–1886*. Cambridge: Cambridge University Press.

Marwick, Arthur. (2001). *The Nature of History: Knowledge, Evidence, Language*. London: Palgrave, Macmillan.

Moskowich, Isabel. (2012). Patterns of English Scientific Writing: adjectives and other building-blocks. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy 'playne and simple': The Writing of Science between 1700 and 1900* (79–92). Amsterdam: John Benjamins.

Moskowich, Isabel. (2016). Philosophers and Scientists from the Modern Age: compiling the Corpus of English Philosophy Texts (CEPhiT). In Isabel Moskowich, Gonzalo Camiña-Rioboo, Inés Lareo, & Begoña Crespo (Eds.), *'The Conditioned and the Unconditioned: Late Modern English Texts on Philosophy* (1–23). Amsterdam: John Benjamins.

Moskowich, Isabel; Gonzalo Camiña, Inés Lareo, & Begoña Crespo (Eds.) (2016). *The Conditioned and the Unconditioned*. Amsterdam/Philadelphia: John Benjamins.

O'Brien, Karen. (2001). The History Market. In Isabel Rivers (Ed.), *Books and their Readers in Eighteenth-Century England: New Essays* (105–134). Leicester: Leicester University Press.

O'Brien, Karen. (2012). English Enlightenment Histories, 1750–c.1815. In Masayuki Sato, Jose Rabasa, Edoardo Tortarolo, & Daniel Woolf (Eds.), *The Oxford History of Historical Writing. Vol. 3* (518–535). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199219179.003.0026

Puente-Castelo, Luis. (2016). Explaining the use of *if…then…*structures in CEPhiT. In Isabel Moskowich, Gonzalo Camiña-Rioboo, Inés Lareo, & Begoña Crespo (Eds.), *'The Conditioned and the Unconditioned: Late Modern English Texts on Philosophy* (167–181). Amsterdam/Philadelphia: John Benjamins.

Pugh, Martin. (1992). *Women and the Women's Movement in Britain 1914–1959*. London: Hampshire.

Rosenthal, Jerome. (1955). Voltaire's Philosophy of History. *Journal of the History of Ideas*. Vol. 16, 2, 151–178. https://doi.org/10.2307/2707660

The Royal Historical Society (2014). History of the Royal Society. *Working for History and Historians*. London: University College London. Accessed [19/10/2017] from http://royalhistsoc.org/about/history/

Scherren, Henry. (1905). *The Zoological Society of London*. London: Cassell and company.

Smith, Bonnie G. (2000). *The Gender of History: Men, Women, and History Practice*. Harvard: Harvard University Press.

Sreedharan, E. (2000). *A Textbook of Historiography. 500BC to AD 2000*. Bangalore: Orient Longman.

Tucker, Aviezer (Ed.) (2009). *A Companion to the Philosophy of History and Historiography*. Oxford: Wiley-Blackwell.

Vann, Richard T. (2018). Historiography (entry). *Encyclopaedia Britannica*. Accessed [19/10/2017] from https://www.britannica.com/topic/historiography

Woolf, Daniel. (2012). Historical Writing in Britain from the Late Middle Ages to the Eve of Enlightment. In Masayuki Sato, Jose Rabasa, Edoardo Tortarolo, & Daniel Woolf (Eds.), *The Oxford History of Historical Writing*. Vol. 3 (473–496). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199219179.003.0024

Zammito, John. (2009). Historians and Philosophy of Historiography. In Aviezer Tucker (Ed.) *A Companion to the Philosophy of History and Historiography* (63–84). Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781444304916.ch5

CHAPTER 2

“There were always Indians passing to and fro”

Notes on the representation of Native Americans in CHET documents

Marina Dossena
Università degli Studi di Bergamo

1. Introduction

Ideology can be shown to have always played a very important part in the way in which people, countries, and even languages have been represented: stereotypical images normally derive from oversimplified and ideologically biased views, and even language policies can be based on such views, which often derive from political considerations. An early example of this is to be found in the Statutes of Iona, which in 1609 prescribed that the education of eldest children born in the Hebrides should be carried out in the Lowlands, so that they would learn English and thus help overcome the “ignorance and incivility” perceived to exist in the Scottish islands:

> it being undirstand that the ignorance and incivilitie of the saidis Iles hes daylie incressit be the negligence of guid educatioun and instructioun of the youth in the knowledge of God and good letters […] it is inactit that every gentilman or yeaman within the said Ilandis, or any of thame, haveing childerine maill or femell, and being in goodis worth thriescore ky, sall put at the leist their eldest sone, or haveing no children maill thair eldest dochter, to the scullis in the Lowland, and interneny and bring thame up thair quhill that may be found able sufficientlie to speik, reid and wryte Inglische.
>
> (*Register of the Privy Council*, 1609, Vol. IX, 28–29, *apud* Innes (1993))

Gaelic had been associated with “incivilitie” for at least a century, as shown for example in the *Flyting of Dunbar and Kennedy*, of 1508. However, this attitude was rooted in political considerations: soon after the Statutes of Iona, an Act of the Privy Council of 10th December 1616 also established parish schools with the following premises:

https://doi.org/10.1075/z.225.02dos

© 2019 John Benjamins Publishing Company

> Forsameikle as the Kingis Majestie having a speciall care and regaird that the trew religioun be advancit and establisheit in all the pairtis of this kingdome and that all his Majesties subjectis especiallie the youth, be exercised and trayned up in civilitie, godliness, knawledge, and learning, that the vulgar Inglishe toung be universallie plantit, and the Irische language, whilk is one of the cheif and principall causes of the continewance of barbarite and incivilitie amongis the inhabitantis of the Ilis and Heylandis, may be abolishit and removeit; […].
>
> (Act of the Privy Council of 10th December 1616, *apud* Innes (1993))

Closer to our times, the twentieth century has provided even more tragic examples of how ideological bias can lead to the restriction of both languages and cultural traditions, and indeed the attempt to eradicate them through genocide. Nor is it anachronistic to use the word ‘genocide’ with reference to the attacks on the life and culture of aboriginal peoples in America and Australia throughout late Modern times. A long line of events based on political views, cultural attitudes, and ultimately ideological stance can be shown to have characterized the ways in which the relationship between Europeans and non-Europeans, and indeed among Europeans, has unfolded over the centuries.

In what follows ideology will be investigated in a sample of documents concerning the representation of Native Americans in eighteenth- and (mostly) nineteenth-century treatises, travelogues, lectures and letters; special attention will be paid to materials currently available in the *Coruña Corpus of History English Texts* (henceforth CHET, see Moskowich *et al.* (2019)), although other materials can also be considered, such as those in the *Evans Collection of Early American Imprints* and elsewhere. After an overview of the materials at hand, I will focus on those in which Native Americans are mentioned, in order to assess the ideological quality of such representations. My methodology will combine the analysis of corpus data with qualitative observations on lexical choices and how such defining strategies present contiguities with those observed in previous studies (see Moskowich *et al.* (2019) and Dossena, 2013, 2015, and 2016); these findings will be interpreted following the discourse-historical approach outlined by Reisigl & Wodak (2009) and Wodak (2014).

Within this framework, the importance of looking at textual occurrences in ‘context’ is of paramount significance: the social, political, and historical circumstances in which the texts were produced provide extralinguistic variables that can account for the attitudes conveyed by the linguistic choices observed in the documents. These choices also help maintain (or – on the other hand – might attempt to change) the very same circumstances in which they occurred; it is therefore crucial to consider the texts in relation to their interdiscursive potential, in order to see how they position themselves in the development of persistent cultural discourses of discovery, conquest, settlement and identity.

In this chapter, first of all I will present an outline of the texts under investigation, then I will focus on those in which Native American culture is presented; in the central section of my study I will concentrate on the key topics discussed in these texts and what ideological representations they convey; finally, some concluding remarks will be offered.

2. The documents under investigation

CHET comprises different text types: in the eighteenth-century section there are mostly treatises (14), but also three essays, a biography, a travelogue and a 'narrative';[1] similarly, treatises account for the majority of texts in the nineteenth-century section (14 again); in addition, there are two lectures and two textbooks, a biographical catalogue and a journal article. This greater attention to genres addressing students or even lay audiences, rather than other scholars, is relevant in a study of ideology in historical texts, as it may shed light on how views could be construed and then propagated. In the nineteenth-century section there is also greater variety both in terms of authors (six women as opposed to two in the eighteenth-century section) and in terms of place of publication: in the eighteenth-century section, out of twenty texts making up the sample, only two were published in America, two in Ireland, and two in Scotland, while most were published in England; in the nineteenth-century section, instead, three texts were published in the US, two in Canada, four in Ireland, and one in Scotland, leaving 50% of the sample for texts published in England.

Of all these texts, however, relatively few (8 altogether) deal with Native Americans, which makes quantitative analyses hardly relevant; even so, they may be of interest from a range of other perspectives. First of all, there is only one lecture (Adams, 1770), while the other seven texts are treatises; of these, two are authored by women (Warren, 1805 and Callcott, 1828). As for place of publication, two texts were published in London (Adolphus, 1802 and Callcott, 1828), one was published in Chicago (Breese, 1884), three were published in Boston (Penhallow, 1726, Adams, 1770 and Warren, 1805), and two were published in Toronto (Gray, 1872 and Kingsford, 1887).

Such remarks are very relevant for a qualitative analysis of the materials at hand: for instance, in both treatises and lectures, i.e. in text types whose argumentative quality can be remarkable, it is important to note the place of publication, as in texts in which conflict is narrated, attitudes may depend on which side of the conflict the author supported. Similarly, attitudes may and often do vary over time – when thirteen British colonies become the United States of America, the narration of the events that led to their independence may be expressed in more favourable terms in American accounts than in English ones.

---

[1] Classified by compilers as a "lecture". Similarly, compilers have classified the biographical catalogue in the nineteenth-century samples as a "dictionary".

Restricting our area of interest again to the documents in CHET in which Native American culture is represented, the following overview can be offered, starting from Table 1 below, in which authors, titles, dates and places of publication are listed.

Table 1. Documents discussing Native Americans in CHET

Publication			
Date	**Place**	**Author**	**Title**
1726	Boston	Penhallow, Samuel	*The history of the wars of New-England, with the Eastern Indians. Or, A narrative of their continued perfidy and cruelty, from the 10th of August, 1703. to the peace renewed 13th of July, 1713. And from the 25th of July, 1722. to their submission 15th December, 1725. Which was ratified August 5th 1726.*
1770	Boston	Adams, Amos	*A concise, historical view of the perils, hardships, difficulties and discouragements which have attended the planting and progressive improvements of New-England.*
1802	London	Adolphus, John	*The History of England, from the Accession of King George the Third, to the Conclusion of Peace in the Year one thousand seven hundred and eighty-three.*
1805	Boston	Warren, Mercy Otis	*History of the Rise, Progress, and Termination of the American Revolution interspersed with Biographical, Political and Moral Observations.*
1828	London	Callcott, Mary	*A Short History of Spain.*
1872	Toronto	Gray, John H.	*Confederation; or, The political and parliamentary history of Canada, from the conference at Quebec, in October, 1864, to the admission of British Columbia, in July, 1871.*
1884	Chicago	Breese, Sidney	*Early History of Illinois.*
1887	Toronto	Kingsford, William	*The History of Canada.*

While most titles appear to be merely descriptive, Penhallow (1726) and Adams (1770) do provide their readers with an indication of what attitudes their texts will convey: their titles include lexical items in which danger, hardships, belligerence and cruelty are evoked, thus eliciting the readers' emotional participation and indirectly invoking solidarity with their viewpoints. In this sense, such documents can be fruitfully investigated following the tenets of Appraisal theory (see Hommerberg, 2011; Martin & White, 2005; White, 2007, 2015), which provides a comprehensive and accurate framework for the study of phenomena in which the interaction of semantics and pragmatics is particularly important.

Appraisal theory studies the linguistic construction of authorial voices and textual personae starting from the presupposition that all texts interact with one another, no matter how implicitly or explicitly, and respond to one another with the expression of *Attitude* (e.g., emotional or affectual responses), *Engagement* (i.e., acknowledging, ignoring or rejecting different view-points, for instance employing evidentiality, concessive forms and presumptions), or *Graduation* (i.e., strengthening or downtoning statements or their semantic focus). Expressions of Attitude comprise three sub-systems: *Affect* (relating to emotion), *Judgement* (relating to the implicit or explicit evaluation of behaviour with respect to social norms), and *Appreciation* (relating to the evaluation of objects). In the (often argumentative) texts under discussion here both Affect and Judgement can be identified in varying degrees, as they enhance the persuasive quality of the discourse.

In the next section the main topics addressed in these documents will be outlined, in order to see what persuasive strategies are employed. Special attention will be paid to lexical choices and the kind of affective response they may elicit; however, other aspects will also be considered, such as references to the role of interpreters and scouts, whose role as intermediaries could often be crucial, and what cultural traits are highlighted. For reasons of space, only some salient examples will be provided; however, many of them will be quoted at length, as in several cases a broader context is useful to assess how lexical choices convey stance by means of their interaction also with other textual features, such as greater attention paid to specific themes, e.g. warfare and its rituals. In this sense, the contents take shape on the basis of linguistic features which, in turn, guide the readers' interpretation through the connotations they provide. Among such themes we find warfare, which is the main object of interest in most texts, but also language and interactions with subjects whose role was often of intermediation between Native and Euro-American subjects – interpreters, of course, but also scouts, guides, missionaries and even fur trappers.

3. Native Americans in CHET

3.1 Cruelty, warfare, and other clichés

European readers had been extremely interested in the discoveries of the New World right from the start. In England, for instance, the accounts collected by Richard Hakluyt in *Divers Voyages Touching the Discoverie of America* (1582) and in *The Principall Navigations, Voyages and Traffiques of the English Nation* (1589) provided both entertainment and useful pieces of information on what resources could be accessed in those exotic territories. Indeed, in 1584 Hakluyt had published "*A Particuler Discourse Concerning the Greate Necessitie and Manifolde Commodyties That Are Like to Growe to This Realme of Englande by the Westerne Discoueries Lately Attempted*", in which Elizabeth I was invited to support the colonization schemes of Walter Raleigh. In the text, economic, political and religious aims mingled and supported each other, as English and Spanish antagonism concerned all three fields. Among the arguments that were put forward, some (1, 2 and 11 below) address religious issues pertaining to the strife between Catholic and Protestant countries; others – 3, 4 and 5 – make economic considerations concerning both employment and trade; finally, others – 6 and 15 – refer to the importance of taking this new and possibly unique opportunity to prevent damage from opposing nations on account of the greater wealth that the new colonies will afford:

1. That this westerne discoverie will be greately for the inlargement of the gospell of Christe whereunto the Princes of the refourmed relligion are chefely bounde amongest whome her Majestie is principall.
2. That all other englishe Trades are growen beggerly or daungerous, especially in all the kinge of Spaine his Domynions, where our men are dryven to flinge their Bibles and prayer Bokes into the sea, and to forsweare and renownce their relligion and conscience and consequently theyr obedience to her Majestie.
3. That this westerne voyadge will yelde unto us all the commodities of Europe, Affrica, and Asia, as far as wee were wonte to travell, and supply the wantes of all our decayed trades.
4. That this enterprise will be for the manifolde imploymente of nombers of idle men, and for bredinge of many sufficient, and for utterance of the greate quantitie of the commodities of our Realme.
5. That this voyage will be a great bridle to the Indies of the kinge of Spaine and a means that wee may arreste at our pleasure for the space of teime weekes or three monethes every yere, one or twoo hundred saile of his subjectes shippes at the fysshinge in Newfounde Iande.
6. That the rischesse that the Indian Threasure wrought in time of Charles the late Emperor father to the Spanishe kinge, is to be had in consideracion of the Q.

> moste excellent Majestie, leaste the contynuall commynge of the like threasure from thence to his sonne, worke the unrecoverable annoye of this Realme, whereof already wee have had very dangerous experience.
>
> […]
>
> 11. That the Spaniardes have executed most outragious and more then Turkishe cruelties in all the west Indies, whereby they are every where there, become moste odious unto them, whoe woulde joyne with us or any other moste willingly to shake of their moste intollerable yoke, and have begonne to doo it already in dyvers places where they were Lordes heretofore.
>
> […]
>
> 15. That spedie plantinge in divers fitt places is moste necessarie upon these luckye westerne discoveries for feare of the daunger of being prevented by other nations which have the like intentions, with the order thereof and other reasons therewithall alleaged. (Hakluyt, 1584)

In this list of points Native populations are only mentioned indirectly in 11, when Spanish cruelty is evoked, and a claim is made that those same populations would welcome English rule, in order to shake off that "most intolerable yoke".

In actual fact, all European settlers attempted to subdue Native populations, although in times of war different nations would side with different armies. In the CHET documents, for instance, great attention is paid to the French and Indian wars, and all accounts include anecdotes about ambushes, battles, the captivity of hostages, the enemies' cruelty and the heroic resistance of the side whose story is being recounted. Some passages exploit what would become clichés of captivity narratives and present almost graphic details of the tortures that were inflicted by one side on the other. William Kingsford's text[2] is a case in point: the negotiations he describes while mixing them with comments on Indian torture strategies is reminiscent of representations that were stereotypical in novels, dime novels, and even paintings (see Rosso, 2016 and Cartosio, 2016). In the excerpt below, for instance, not only does he describe what tortures could the prisoner expect, but he also offers a snapshot in which enemy scalps are waved as a non-verbal act of defiance:

---

[2] William Kingsford (London 1819 – Toronto 1898), historian, was born in London. In 1837 he went to Canada with his regiment and obtained his discharge in 1840; in 1841 he entered the office of the city surveyor of Montreal and later he was employed on the Hudson River Railway, on the Panama Isthmus Railway, and on the Grand Trunk Railway. In 1873 he was appointed by the Canadian government engineer in charge of the harbours of the Great Lakes and the St. Lawrence; but in 1879 he was dismissed and he devoted his later life to historical research. The first volume of his *History of Canada* appeared in 1887, and the tenth and last volume in 1898. (Wallace, 1948: III, 337–338).

> Marguerie was en parole, bound to return, and the life of Godefroy depended on his presence, if war were persevered in. He himself ran the chance of having his fingers cut off one by one with oyster shells, to be scalded, to receive red hot ashes on the scalped skull, and to be burned by a slow fire. He cast aside all personal considerations and recommended a refusal of the demand. […]. In the interval communication was held with the Indians through Father Ragueneau and the interpreter Nicolet, both speaking Huron, which the Iroquois understood; the latter knowing perfectly the language. […]. The Indians saw that they had nothing to hope for, as their force was outnumbered; so waving some Algonquin scalps in token of defiance, they prepared to return. (CHET, Kingsford, 1887: 174–175)

Kingsford describes the Iroquois' aim "to prevent all settlement in the territory adjoining their homes; to make it desolate, so that neither the white men, nor rival tribes should live in their neighbourhood to be a threat to them" (1887: 185). This is in fact a legitimate aspiration, and the author admits that their motive is "one of personal security" (Kingsford, 1887: 185). Nonetheless, the Iroquois are described as formidable enemies and the torture of their prisoners is almost routine in its elliptical summary; hostages expect death daily, but make "hair-breadth escapes", and the qualifier draws attention to the (almost) fictional quality of the narration:

> On arriving at the Mohawk village there were the usual tortures, nails torn away, fingers cut off, the captives forced to run the gauntlet, with all the known systematic Indian cruelty. […]. Jogues lived a year in captivity, daily expecting death. He attended some Indians to Albany, where, by the help of the Dutch, he managed to secrete himself in a vessel, and after several hair-breadth escapes reached New York. (CHET, Kingsford 1887: 185)

On the other hand, the same fighting and torture techniques could be adopted by all sides; in the following excerpt, the Natives are the victims and some get scalped by Europeans; cruelty is also justified on religious grounds, as a Jesuit is seen as both an accomplice of the Natives who sided with the French and as an opponent in terms of doctrine. Finally, comments are also offered on the attitude of a Quaker man, who refused to join the garrison and whose family suffered as a result. Clearly the author's views are not conveyed as explicitly in this case as in the case of the Jesuit, but the text clearly implies that ridiculing military power can have dire consequences:

> The Indians were under amazing Terror; […]. Our Men purſued them ſo warmly, that ſeveral were ſlain on the ſpot; more got into their Canoos, & others ran into the River; which was ſo rapid and the falls in ſome places ſo great, that many of them were drowned. […]. The number of the dead which we ſcalpt, were twenty ſix, beſides Mounſieur Ralle the Jesuit, who was a Bloody Incendiary, and Inſtrumental to moſt of the Miſchiefs that were done us, by preaching up the Doctrine of meriting Salvation by the deſtruction of Hereticks. […]. About ſeventy French Mohawks

> were now making a deſcent on our Frontiers, who divided into ſeveral Parties and kill'd a great number of Cattel. Some of them fell on the Houſe of John Hanſon of Dover, who being a ſtiff Quaker, full of Enthuſiaſm, and ridiculing the Military Power, would on no account be influenced to come into Garriſon; by which means his whole Family (then at home,) being eight in number, were all kill'd and taken. But ſome time after his Wife and two or three of his Children were redeemed with conſiderable Pains and Expence. (CHET, Penhallow, 1726: 106)[3]

As for contrasts between different nations, Kingsford provides some comments on the territories that were affected and some stress is placed on the fact that the Iroquois generally attacked "when the odds were in their favour"; however, it may be difficult to ascertain whether this was actually the case, or if it was the explanation offered by those who had been defeated by them:

> By 1649 there were no more Hurons to destroy. Accordingly, the whole strength of the Iroquois was turned against the Saint Lawrence. They had even, as has been remarked, attacked the Abenakis. Throughout Canada, the inhabitants lived in continual danger. The Indians even ravaged the Island of Orleans. Individually, great gallantry was shewn. The men who were brought face to face with the Iroquois never failed to give a good account of themselves. As a rule, they were overwhelmed by numbers, for the Iroquois seldom attacked, but when the odds were in their favour. (CHET, Kingsford, 1887: 195)

Ideological stance is thus seen to characterize representations of warfare; while this may be predictable on account of the argumentative quality of the discussion, it may be less obvious to find ideology also in the linguistic representations that will be the object of the next section.

3.2 Ideology in language representation: Wigwam words and eloquent speeches

Like in art (Cartosio, 2010, 2016), also in linguistic representations Native Americans could have two diametrically opposed images: savage warriors who spoke harsh and guttural languages, or idealised, 'vanishing Indians' (Dixon, 1913), whose eloquence could be (and often was) compared with that of Classical orators (Dossena, 2015). Of course both images were equally based on myths construed by Euro-American commentators; however, they proved highly influential. Although Bellin (2000: 6) is careful to stress that "Indian 'voices' in American texts exist in translation – […] [in the sense that] they have been transmitted in an alien modality and languages, [and] they have been shaped by the visions and ideologies of Euro-America,"[4] they have been employed to convey specific attitudes to readers across time. For instance, when the right of Euro-Americans to acquire territories had to be defended, the argument was often made that Native populations were declining anyway: Gilbert Malcolm Sproat[5] claimed that "the experience of the Jesuits in California, and of the earliest settlers in the American and British territories on the North Pacific, affords proof of the tendency of the savages to extinction, even before white people went amongst them" (1868: 274–275). However, Sproat himself found that contact with colonists was nefarious, as change was "not to civilization, but to that abased civilization which is, in reality, worse than barbarism itself. He [the Indian] is a vain, idle offensive creature, from whom one turns away with a preference for the thorough savage in his isolated condition" (1868: 285).

---

[3] According to the entry in *Encyclopedia Britannica*, Samuel Penhallow (St Mabon, Cornwall, 1665 – Portsmouth, New Hampshire, 1726) emigrated to Massachusetts in 1686. At first he was commissioned by the Society for the Propagation of the Gospel in New England to study the Indian languages and to preach to the Natives, but he later took on various administrative offices. In 1714 he was appointed a justice of the superior court of judicature and from 1717 until his death he was chief justice of that court. His *History of the War of New England with the Eastern Indians, or a Narrative of their Continued Perfidy and Cruelty* (1726) covers the years from 1703 to 1726.

Sproat's contradictory attitude is mirrored in an equally ambivalent attitude to language that is found in many books, in which the Natives' 'natural eloquence' was exalted, although neither the reading public nor perhaps many compilers could access it in its original form. In fact, commentators tried to express their evaluations of Native languages on the basis of what they could expect their readers to recognize and understand. John McIntosh called Algonquin "the Italian of the western continent" on account of its vowel system, and emphasized positive traits which were often found in elocution manuals, such as elegance, harmony, "richness of expression, […] variety of turns, […] propriety of terms, and […] regularity" (1853: 93–94). Similarly, Henry Rowe Schoolcraft[6] labelled Algonquin "the court language of the Indian" (1857: 673), thus applying Euro-American models of 'politeness' and sociolinguistic evaluation.

---

4. A study of the linguistic reliability of transcriptions and translations is beyond the scope of this paper; however, interesting comments are provided by Merrell (2006, 2013: 60–63), Guthrie (2007), and McNenly (2014).

5. Sproat (1834–1913) was a Canadian businessman, government agent, Indian reserve commissioner, magistrate, and author of Scottish origin (Foster, 2003).

6. Schoolcraft (1793–1864) was an American explorer and ethnologist; best known for his discovery of the source of the Mississippi River, as superintendent of Indian affairs for Michigan in 1836 he concluded a treaty with the Ojibwa in which they ceded much of northern Michigan to the United States (*Encyclopedia Britannica*).

Interest in Native languages was also found on the part of missionaries, who had drawn up grammars and dictionaries of those languages for decades (Eliot, 1666; Stevens, 1956; Swiggers, 2009); however, their aim was to learn to communicate effectively, not necessarily to set down linguistic notes so that languages could be preserved. Still, it is on account of that very aim that a certain degree of reliability can be expected; in other text types, authenticity could simply be evoked by means of recognisable lexis, which Lepore (2001) has called "wigwam words" – i.e., Native items that function as overt markers of cultural identity, and which are not normally translated, though they could be, as they contribute to stylistic effect.

Like other texts, also those in CHET include such vocabulary; indeed, the word *wigwam* itself occurs in Kingsford's text in two instances in which the word *village* or *settlement* would not convey the same sense of estrangement: in both cases it refers to the place where the protagonists were held captive, and therefore a familiar lexical item would not be consistent with the misery that is meant to be conveyed. In the second example, however, religious associations and a greater sense of understanding emerge from the protagonist's decision to return after the embassy is over:

> Shortly afterwards, Père Bressani was made a prisoner, equally to be tormented. He was given to a family to replace a grandfather; but he was of little use in an Indian wigwam, so the Dutch managed to ransom him at the cost of three hundred livres, […]. The embassy reached the Mohawk. Jogues […] was impressed by his reception at the scene of his former sufferings. He was promised a constant welcome in the wigwam of the family in which he had passed so many months of misery, and finding some Christian Huron Indians among the Iroquois, he administered to them religious consolation. The consideration which he received determined him to return to the Mohawks, after communicating the result of his mission.
>
> (CHET, Kingsford, 1887: 185, 188)

As for the evocation of gruesome authenticity, a notorious lexical item, *tomahawk*, features in the narration of how the same missionary is killed. In this paragraph the intention to fight the French is summarized in the metaphorical phrase "to take up the hatchet against them": this time, the name of the weapon is given in English – a lexical choice that may derive from the wish to avoid repetitions, but which may reserve the Native item for a more brutal context, so as to stress the savage quality of the act:

> Jogues proceeded to carry out his purpose of returning to the Hudson. As he ascended the Richelieu he met canoes on the river, from whom the reports of the state of feeling of the Mohawks was unsatisfactory. There had been a poor crop: there was more than the usual sickness, and the cause was assigned to the bewitching influences created by the box left behind by Jogues. It had been stated that the intention of the French was to destroy the Mohawks, and that they took this form

> of doing so, and a strong appeal had been made to take up the hatchet against them. […]; Jogues […] was seized and carried prisoner to the old spot. There was not unanimity with regard to him; some of the tribe wished to save his life. On the evening of the 18th October, 1646, Jogues was invited to a supper; and as he was entering the tent, his skull was cleft in two by a tomahawk.
> (CHET, Kingsford, 1887: 189)

On the other hand, exotic forms of communication are not restricted to vocabulary, as gestures may function as powerful pragmatic moves. Sincerity, trustworthiness and gratitude, for instance, may be encoded by the donation of skins. Even more significant from the pragmatic point of view is the case of wampum belts, which are used to validate predications (Merrell, 2013: 63–65), as in the following example:

> as a token of the sincerity of the Six Nations, we do here, in the name of all, present our great queen, with the belts of wampum. […] "To confirm to you that we will see your request executed, we lay down this string of wampum in return for yours." Then turning to the Delawares, holding a belt of wampum in his hand, he spoke to them as follows: […]. This string of wampum serves to forbid you, your children, and grandchildren to the latest posterity, for ever, meddling in land affairs; neither you, nor any who shall descend from you, are ever hereafter to presume to sell any land; for which purpose you are to preserve this string, in memory of what your uncles have this day given you in charge. (McIntosh, 1853: 270, 296, 297)

Another very significant gesture is smoking a calumet; in the following instance, smoking together bears witness to the sincerity of the participants and is interpreted as a sign of good will on both sides; the ritual is expected to dispel animosity like smoke is expected to vanish into air:

> I smoke this pipe in evidence of my sincerity. If you are sincere, you will receive it from me. My only desire is, that we should smoke it together […]. When this pipe touches your lip, may it operate as a blessing upon all my tribe. May the smoke rise like a cloud, and carry away with it all the animosities which have arisen between us. (McIntosh, 1853: 284–285)

Breese, another author in CHET, describes a scene in which the European protagonist refuses to be the first to offer the pipe, lest it may give an impression of weakness. Indeed, this choice results in admiration on the part of his opponents, who manifest this with other instances of non-verbal communication, such as rubbing the friars' feet with bear's oil and buffalo fat, and feeding them three morsels of meat as a politeness move.[7]

---

[7] Sidney Breese (Whitesboro, NY, 1800 – Pinkneyville, Perry County, Ill., 1878) was a senator for Illinois, then served as Speaker of the Illinois House of Representatives and finally became Chief Justice of the Illinois Supreme Court (see the *Biographical Directory of the US Congress*).

> La Salle formed his boats in a line across the river, and assumed an appearance as formidable as his little fleet would allow. Some of the Indians fled, some seized their arms, but La Salle, alone and unattended, was in the midst, but he did not present the calumet, lest it might be regarded as an evidence of conscious weakness. Struck by his courage and bold bearing, the savages, though thousands in number, presented that mystic symbol to him, and soon a friendly intercourse was effected. They rubbed the uncovered feet of the friars with bear's oil and the fat of the buffalo, and then fed them with meat, putting, with great ceremony, the first three morsels into their mouths, as a mark of great civility. (CHET, Breese, 1884: 109)

Given that cultures were so distant, there was always the possibility that gestures could be misunderstood, with unpredictable but often disastrous outcomes. In this respect, then, the role of interpreters proved crucial in many circumstances, though of course the level of their linguistic competence cannot be assumed to have been homogeneous. Some examples of their activity are discussed in the next section.

3.3 Interpreters and scouts, missionaries and trappers

In many cases the role of interpreter was played by Natives belonging to different nations; in other cases, scouts, trappers and indeed missionaries could conduct negotiations in languages they had acquired through contact and self-teaching. However, their aims were typically utilitarian, and it cannot be excluded that translation could sometimes be less than *bona fide*, although of course deliberate misrepresentations could have serious consequences too. In the excerpts below, different instances of the various situations that might occur are provided by the authors quoted in CHET; in the first one, the scouts are from the Miami nation and inform Peorias of the arrival of Europeans who are allied with their Iroquois enemies; however, the Europeans' communicative skills manage to overcome distrust, thus indirectly drawing attention to their supposedly superior pragmatic competence:

> Although some scouts from the Miamis, […], informed them that La Salle's designs were hostile, and that he was in league with their ancient enemy, the Iroquois, […], yet his sagacity, self-possession and consummate address overcame their distrust, whilst the meek bearing and unaffected simplicity of the ill-clad, bare-footed friars excited their sympathy, won their confidence, and aided, essentially, in placing their relations upon the most amicable footing. (CHET, Breese, 1884: 110)

His *Early History of Illinois*, published posthumously, is an account of the ninety years of French presence in Illinois, from 1673 to 1763.

The author goes on to remark that the Peorias "seem to have been a humane and an inoffensive people, by no means warlike or treacherous, and living in no dread but of the Iroquois" (1884: 110). Indeed, comments on the varying levels of hostility that existed among the different nations towards one another (not least on account of the alliances that could be formed with European traders and settlers) is often the object also of Kingsford's remarks – see the example below:

> The hostility of the Iroquois had even been intensified by the presence of the French. The Hurons indeed came to be regarded as the allies and supporters of the hated white man on the Saint Lawrence, against whom constant war was to be waged. Every settlement above Three Rivers was an additional threat to the Indian. The River Richelieu, discharging into the Saint Lawrence, was looked upon by the Iroquois as the boundary of French power. Such limit might be tolerated for the purpose of barter and trade. The establishment of the French at Montreal furnished an additional cause for exterminating the Hurons. […]. The missions lasted fifteen years. Their operations had gradually widened, […]. Nevertheless, they were asking for more priests; when it was physical support which was needed to stay the destruction of the people who, to a greater or less degree, they were striving to influence to accept their teaching. (CHET, Kingsford 1887: 201)

Kingsford observes that the Hurons were about to disappear, and yet the missionaries who were supposed to look after their souls did not ask for sufficient support also to look after their physical preservation. Looking back on the controversies of the previous century, the author comments quite extensively on how settlements were perceived as threatening, and the kind of antagonism existing between Hurons and Iroquois is distinctly reminiscent of the representations found in James Fenimore Cooper's novels. This is seen also in the note on an interpreter whose heroic behaviour, strength and endurance evoke those of Natty Bumppo, the protagonist of *The Leatherstocking Tales* (1827–1841):

> François Marguerie and Thomas Godefroy fell into an Iroquois ambush. The former appears as one of the most chivalrous characters of these early days: in a critical period recommending a course of action in order to save the community, at the risk of being subjected to a slow death amid continuous torture. He was an interpreter, having gained a knowledge of the Huron language, and is mentioned by de Brébœuf, when but twenty-two, as possessed of strength and endurance remarkable even in the savage community in which he lived.
> (CHET, Kingsford, 1887: 174–175)

As a matter of fact, contact in Quebec was both frequent and indeed sought after, as the fur trade required interaction between European settlers and Native nations; as Breese (1884) remarks, this was closely linked to christianisation, and in this case the stereotypical metaphor of "forest children" is employed to designate the latter, thus recurring to a paternalistic label which was also mirrored in references to 'Father' when a Euro-American leader was meant (Dossena, 2015):

> The Hurons frequently visited Quebec, and as they supplied the costly furs and the valuable peltries, they were much caressed, particularly by the Directors of the Company of New France, who made large profits by their friendship. […]. These [Jesuit priests] are the first missionaries of whom we have any record, to light the fires of Christianity in the wilds of the upper lakes, and the mission they established near Lake Iroquois, and denominated Saint Joseph, was the first dwelling place of the members of the Company of Jesus – of those who worshiped the true God in those barbarous regions. The forest children flocked to it in crowds, and in the bosoms of many, sentiments of devotion were enkindled, which soon led to their conversion. (CHET, Breese, 1884: 74)

This draws attention to the fact that the fur trade was not devoid of political and even religious interferences; Breese (1884) also remarks on the competition between Protestant and Catholic adventurers for an obviously indispensable support of Native nations – see the quotation below:

> At the time of their [the Jesuits'] advent to Quebec in 1622, dissensions existed there to an alarming extent, caused by some French Protestants engaging in the fur trade, and who, if successful, might secure the favor of the natives to such a degree as to deprive the Catholic portion of the adventurers from any participation in it. (CHET, Breese, 1884: 98)

Kingsford (1887), instead, provides an outline of what arrangements were made by a French company established in the first half of the seventeenth century for the settlement of the Montreal area. Fur trading was ceded to local inhabitants except in specific areas; in addition, a fixed quantity of beaver skins had to be given to an earlier company. What is even more interesting is the number of emigrants that was expected to be sent from France to Canada every year – however, how voluntary such emigration could be, how it was encouraged, and how those strategies compared with mechanisms established in Britain for the promotion of emigration, would deserve more in-depth investigation in other archival documents:

> It was in 1645, that a change was made in the organization of the Company, […]. The fur trade was ceded in community to the inhabitants, excluding that of Acadia, Miscou, and Cape Breton. […]. Twenty emigrants were to be sent annually to Canada from France, and one thousand lbs. of beaver skins were to be paid each year to the old Company. (CHET, Kingsford, 1887: 188)

4. Concluding remarks

A preliminary quantitative approach to CHET showed that only a relatively small number of texts could be considered, in order to answer research questions on how Native Americans were represented in the corpus. Instead, what was much more fruitful was a qualitative investigation of the texts at hand, as they could be read on the basis of both a Discourse-Historical approach and within the framework of Appraisal theory.

These methodological choices were found to be appropriate for a discussion of texts which are broadly categorised as ‘historical’, but from many of which objectivity can hardly be expected, not least because the treatises themselves proclaim their argumentative quality already on their title pages. Readers are thus invited to share their point of view and sympathise with the authors’ opinions, whether it is to express judgements on the actions of both enemies and allies or to respond affectively to narrations of cruelty, heroic deeds, and other events eliciting emotional participation.

Given the mutable, but always asymmetrical, power relationship existing between Europeans and Native nations, the texts are seldom exempt from generalisations concerning the ways in which the new (and very exotic) world is represented. Indeed, lexical choices often reinforce stereotypical images even in the situations where a greater ethnographic interest might appear between the lines. All Native nations are represented in the same ambivalent terms that are found throughout late Modern times – as fierce warriors and savage enemies, but also as cultures doomed to extinction: in this sense the texts in CHET follow patterns found in other collections. In addition, European antagonisms are also found to contribute to the ways in which both North and South American Natives are discussed. As both France and Spain were enemies for British and American readers, their involvement results in triangular relationships in which the representation of Native nations varies depending on their being allies of one side or the other and/or on their being victims of French and Spanish settlers.

Although the texts discussed here were published over two centuries, were authored by both men and women on both sides of the Atlantic, and belonged to apparently different genres, the ideology underpinning them seems to be consistent despite the differences that could be suggested by heterogeneity in time and place of publication, authorship, and textual typology. While implicitly or explicitly reinforcing tenets of Euro-American superiority, it was a long-lasting ideology that could also convey a certain degree of ‘imperialist nostalgia’ – and it would not be till the late twentieth century that such views would begin to be addressed critically.

Acknowledgements

This study is based on analyses conducted within the 2016–18 project 'Etiquetado electrónico de textos científico-técnicos en lengua inglesa entre los siglos XVII y XX (4): Coruña Corpus', financed by the Spanish Ministerio de Economía y Competitividad; programa estatal de fomento de la investigación científica y técnica de excelencia – subprograma estatal de generación de conoscimento (FFI2016-75599-P) and coordinated by prof. Isabel Moskowich, Departamento de Filoloxía Inglesa, Universidade da Coruña. In addition, it has benefited from the investigation of materials collected for the 2015–2018 project 'Knowledge Dissemination across Media in English: Continuity and Change in Discourse Strategies, Ideologies, and Epistemologies' (PRIN 2015TJ8ZAS_004), financed by the Italian Ministry of Education and led by prof. Marina Bondi, Università di Modena e Reggio Emilia.

References

Primary sources

Dixon, Joseph K. (1913). *The vanishing race*. Garden City, NY: Doubleday, Page, & Co.

Eliot, John. (1666). *The Indian grammar begun: Or, an essay to bring the Indian language into rules, for help of such as desire to learn the same, for the furtherance of the Gospel among them*. Cambridge: Printed by Marmaduke Johnson.

Evans Collection of Early American Imprints, Accessed [30/11/2017] from http://quod.lib.umich.edu/e/evans

Hakluyt, Richard. (1584). *A particuler discourse concerning the greate necessitie and manifolde commodyties that are like to growe to this realme of Englande by the westerne discoueries lately attempted*, Accessed [15/03/2019] from www.let.rug.nl/usa/documents/before-1600/richard-hakluyt-discourse-of-western-planting-1584.php

McIntosh, John. (1853 [1843]). *The origins of the North American Indians: With a faithful description of their manners and customs, both civil and military, their religions, languages, dress and ornaments. New edition*. New York: Nafis and Cornish.

Moskowich, Isabel; Inés Lareo, Paula Lojo Sandino, & Estefanía Sánchez-Barreiro (Comps.) (2019). *Corpus of History English Texts*. A Coruña: Universidade da Coruña. Accessed [15/03/2019] from https://ruc.udc.es/dspace/handle/2183/21849

Schoolcraft, Henry Rowe. (1857). *Historical and statistical information respecting the history, condition and prospects of the Indian tribes of the United States*. Philadelphia: Lippincott, Grambo & Co.

Sproat, Gilbert Malcolm. (1868). *Scenes and studies of savage life*. London: Smith, Elder & Co.

Secondary sources

Bellin, Joshua D. (2000). *The demon of the continent. Indians and the shaping of American literature*. Philadelphia: The University of Pennsylvania Press.

Biographical Directory of the US Congress, Accessed [30/11/2017] from http://bioguide.congress.gov/

Cartosio, Bruno. (2010). Ways of the West: History and literature / Fiction and truth. In Angela Locatelli (Ed.), *The Knowledge of Literature* IX (39–56). Bergamo: Bergamo University Press.

Cartosio, Bruno. (2016). American artists look West. In Marina Dossena & Stefano Rosso (Eds.), *Knowledge dissemination in the long nineteenth century: European and transatlantic perspectives* (9–25). Newcastle: Cambridge Scholars.

Dossena, Marina. (2013). "John is a good Indian": Reflections on Native American culture in Scottish popular writing of the nineteenth century. In Carla Sassi & Theo van Heijnsbergen (Eds.), *Within and without empire: Scotland across the (post)colonial borderline* (185–199). Newcastle: Cambridge Scholars.

Dossena, Marina. (2015). (Re)constructed eloquence. Rhetorical and pragmatic strategies in the speeches of Native Americans as reported by nineteenth-century commentators. *Brno Studies in English*, 41/1, 5–28. https://doi.org/10.5817/BSE2015-1-1

Dossena, Marina. (2016). America through the eyes of nineteenth-century Scots: The case of ego documents and popular culture. In Marina Dossena & Stefano Rosso (Eds.), *Knowledge dissemination in the long nineteenth century. European and transatlantic perspectives* (45–64). Newcastle u. T.: Cambridge Scholars.

Encyclopedia Britannica, Encyclopædia Britannica, inc. Accessed [15/03/2019] from https://www.britannica.com/

Foster, Hamar. (2003–) "Sproat, Gilbert Malcolm," in *Dictionary of Canadian Biography*, vol. 14. University of Toronto/Université Laval. Accessed [15/03/2019] from http://www.biographi.ca/en/bio/sproat_gilbert_malcolm_14E.html

Guthrie, Thomas H. (2007). Good Words: Chief Joseph and the Production of Indian Speech(es), Texts, and Subjects. *Ethnohistory*, 54/3, 509–546. https://doi.org/10.1215/00141801-2007-005

Hommerberg, Charlotte. (2011). Persuasiveness in the discourse of wine: The rhetoric of Robert Parker. Linnaeus University Dissertations 71/2011. Gothenburg: Linnaeus University Press.

Innes, Ewan. (1993). The Social, Economic & Political Reasons for the Decline of Gaelic in Scotland. In *ScottishHistory.com*. Accessed [15/03/2019] from www.scottishhistory.com/articles/highlands/gaelic/gaelic_page1.html

Lepore, Jill. (2001). Wigwam Words. *American Scholar*, 70/1, 97–108.

Martin, James R. & Peter R. R. White. (2005). *The language of evaluation. Appraisal in English*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230511910

McNenly Scarangella, Linda. (2014). Foe, friend, or critic. Native performers with Buffalo Bill's Wild West Show and discourses of conquest and friendship in newspaper reports. *American Indian Quarterly*, 38/2, 143–176.

Merrell, James H. (2006). "I desire all that I have said … may be taken down aright": Revisiting Teedyuscung's 1756 treaty council speeches. *The William and Mary Quarterly*, 63/4, 777–826.

Merrell, James H. (2013). Conversations in the woods: Western Pennsylvania, 1722–1762. In Brian DeLay (Ed.), *North American borderlands* (59–87). New York: Routledge.

Reisigl, Martin & Ruth Wodak. (2009). The discourse-historical approach. In Ruth Wodak & Michael Meyer (Eds.), *Methods of CDA* (87–121). London: Sage.

Rosso, Stefano. (2016). The winning of the western: Early dissemination of a literary genre. In Marina Dossena & Stefano Rosso (Eds.), *Knowledge dissemination in the long nineteenth century: European and transatlantic perspectives* (27–44). Newcastle: Cambridge Scholars.

Stevens, C. J. (1956). Soul savers, grammarians, and the red man. *American Speech*, 31/1, 40–51. https://doi.org/10.2307/454015

Swiggers, Pierre. (2009). David Zeisberger's description of Delaware morphology. *Historiographia Linguistica*, 36/2–3, 325–344. https://doi.org/10.1075/hl.36.2-3.08swi

Wallace, W. Stewart (Ed.) (1948). *The Encyclopedia of Canada*. Toronto: University Associates of Canada.

White, Peter R. R. (2007). Appraisal. In Jan-Ola Östman & Jef Verschueren (Eds.), *Handbook of pragmatics online*, s.v. Amsterdam & Philadelphia: John Benjamins. Accessed [15/03/2019] from www.benjamins.com/online/hop

White, Peter R. R. (2015). *The Appraisal website*. Accessed [15/03/2019] from www.grammatics.com/appraisal/

Wodak, Ruth. (2014). Discourse-historical approach (DHA). In Karen Tracy (Ed.), *International encyclopedia of language and social interaction* (275–287). Oxford: Elsevier.

CHAPTER 3

An introduction to CHET, the *Corpus of History English Texts*

Isabel Moskowich
Universidade da Coruña

> So I hope that the significance of history will not be thought to lie in its general propositions, as if these were the flower and fruit of the whole endeavour, but that its value will be seen to consist in its taking a familiar, perhaps commonplace theme, an everyday melody, and composing inspired variations on it … and thus disclosing in the original theme a whole world of profundity, power, and beauty. For this, however, there is required above all great artistic facility, creative vision, loving absorption in the empirical data. (Nietzsche, *Human, All Too Human,* 93)

1. Historical knowledge

The sudden growth and development of the scientific method, itself giving rise to the scientific revolution, made clear that the various disciplines of scholarly investigation were not all the same. In particular, the observational and natural sciences, and indeed any field that used mathematics as a universal language, came to occupy a distinct position in the realm of knowledge, one not so readily available to other areas. An intense interest grew in discoveries, in the exploration of the world (and beyond), and in the classification of the phenomena within it. As a result of this, disciplines not grounded in the scientific method, those with no methodological tenets of replicability or the reliance on mathematical certainties, seemed to evolve more slowly in the Modern period, and indeed are still not regarded as "pure" science today. Such is the case with History.

From the Old French *estoire*, the term *history* originally denoted the relation of incidents, either true or false, and was first introduced into English in the 14th century. The fact that no distinction between *history* and *story* existed in the English language until the middle of the eighteenth century (1756, according to the *OED*) may serve to illustrate the limits of the discipline, and of the object of study itself, prior to this time. However, the etymology of the term can ultimately be traced to ancient Greek, where ἱστορία, *istoria*, referred to "a learning or knowing by inquiry;

https://doi.org/10.1075/z.225.03mos

© 2019 John Benjamins Publishing Company

an account of one's inquiries, history, record, narrative" (Harper, 2010–2017), and was related to *historein* ("inquire") and *histor* ("wise man, judge"). This in turn suggests that History relied more heavily on facts and records than on the word of established, foundational authorities, as was the case with most fields of knowledge before and indeed after Scholasticism. Perhaps it is this characteristic of History that makes it difficult to classify as a science, a point which has been widely recognised, in that "no other discipline has its portals so wide open to the general public as history. In no other discipline is the transition from the dilettante to the professional so gradual" (Huizinga *et al.*, 1984: 39).

Also, according to the definitions of science that emerged from the scientific method, History could not be considered a science in that historical facts cannot be verified by the repetition of experiments. It may thus be safer to cling to the idea that History deals with the interpretation of the past and that, hence, it is a subjective endeavour, its conclusions never directly verifiable. Yet the discipline's object of study, the limits therein, are also curiously fuzzy and hard to pin down. The traditional definition of History as the study of events affecting Humankind since the emergence of writing (thus distinguishing it from pre-history) is itself unsatisfying, in that historical studies, both in the academic realm and more broadly, cover periods for which no written record exists. Technological advances allow present-day researchers to use alternative methods for the reconstruction and analysis of past events (carbon 14 dating, forensic techniques, etc…). Moreover, recent years have seen an interesting branching of the field, with the emergence of Cultural History, Political History, Social History, History of Science, etc., and in these new disciplines we can easily perceive a variety of slightly different objects of study and approaches. In the Modern Period no such branches existed. In contrast, History as we understand it today was not very different from medieval physics (a branch of philosophy) and eighteenth-century Natural Philosophy, concerned with the study of mankind and the natural world from a general point of view. Besides, closely related disciplines, such as what we today know as Historiography (or the approach to the methodologies used in historical studies) did not exist in their own right in the Modern Period. The word *historiography* was originally a synonym for *history* itself, in that a historiographer was simply an alternative term denoting a historian (or even 'chronicler') in the early Modern period, as shown by the creation of the title of Historiographer Royal in England in 1660 (Saslow, 1978).[1] It was not until the eighteenth century and the Enlightenment that Historiography/History began to resort to rigorous methods, under the influence of Voltaire in France and Hume in England in their (failed) fight for a rationalistic approach to historical studies

---

[1] Dryden himself occupied this post from 1670 (Ham, 1935).

(Jaki, 1993; De Syon, 1999; see also Alfaya this volume). At that time, and perhaps with the help of Edward Gibbon's *The Decline of the Roman Empire* (sampled in the *Corpus of History English Texts*), History became an independent discipline. During the nineteenth century the prime focus for scholars was that of institutional change and how constitutional governments had developed. The nineteenth-century idea of academic History indeed only came to be clearly defined in the last quarter of the twentieth century (Salevouris & Furay, 2015). However, the academic shift towards professionalisation had its beginnings in Germany in the nineteenth century (Hoefferle, 2011) with new teaching methods in universities such as Göttingen, and was accompanied by the birth of the first scholarly journals on the subject. Indeed, throughout the second half of the nineteenth century an abundance of explanatory models of History appeared. Let us bear in mind that this was a time in which the biological world was increasingly understood in terms of evolution and scientific progress, rather than being guided by eternal and immutable laws, and progressive historians thus began to see the present in the past and to reject the view that the living world was static and law-driven.

2. Compilation principles in CHET

Considering the ideas discussed above, it does not seem unreasonable to include the field of History in the *Coruña Corpus of English Scientific Writing* (CC). However, we have done this after considering at length the issue of whether History and Historiography are in fact sciences at all. The scientific method would certainly exclude the area of historical study, in that historical events cannot be directly verified. Yet such a strict application of any such boundaries would also exclude a field such as Theoretical Physics from the sciences. The *Corpus of History English Texts* (CHET) is a specialised corpus in the sense proposed by Connor & Upton (2004: 2):

> Instead of being compiled for the representativeness of language across a large number of communicative purposes, specialized corpora often focus on one particular genre e.g. research papers, letters or business requests or specific situation (e.g. academic lectures, office communications in business).

The texts we have included in CHET vary a great deal in nature, although they all have in common the fact that they narrate events in the past that are related to human communities. In the spirit of maximal inclusivity, we have compiled samples of texts that present-day historians would probably consider to be records, rather than texts on History in a strict sense. Such is the case with the 1739 travelogue by Elizabeth Justice, *A Voyage to Russia: describing the Laws, Manners, and Cuſtoms, of that great Empire, as govern'd, at this preſent, by that excellent Princeſs, the*

Czarina. Shewing the Beauty of her Palace, the Grandeur of her Courtiers, the Forms of Building at Petersburgh, and other Places: with ſeveral entertaining Adventures, that happened in the Paſſage by Sea, and Land. Another text included, *The history and antiquities of the cathedral church of Salisbury; illustrated with a series of engravings, of views, elevations, plans, and details of that edifice: also etchings of the ancient* (Bigland, 1810),[2] is a dictionary or catalogue of terms about engravings and their evolution in relation to historical facts. However awkwardly such texts sit within our current notion of academic History, their inclusion is justified on the grounds that they reflect the field as it was understood in the eighteenth and nineteenth centuries, and thus they constitute a solid grounding for the corpus.

Not surprisingly, the texts in the earlier part of CHET are more likely to seem awkward in this sense. Hence, John Oldmixon's *Memoirs of Ireland from the Restoration, to the Present Times*, published in 1716, is classified as a (historical) treatise, just like Montagu Burrows' *The History of the Foreign Policy of Great Britain* (1895). The corpus thus reflects the dynamic nature of the discipline and, above all, the kinds of communicative formats or genres within it as they adapt their forms to different functions and audiences at a time when there are no clear methodological tendencies within the discipline, in contrast to Astronomy, Philosophy or other fields of knowledge.

As the third discipline-specific corpus within the larger *Coruña Corpus* project, CHET was planned and compiled following the same guidelines as the two previously compiled subcorpora, CETA (*Corpus of English Texts on Astronomy*) and CEPhiT (*Corpus of English Philosophy Texts*). All are intended to be mutually complementary, in that they all contain samples of English texts published between 1700 and 1900 (Moskowich and Parapar, 2008). As with the previous two subcorpora, CHET is a specialised corpus as it is also "delimited by a specific register, discourse domain or subject-matter" (De Beaugrande, 2001: 11). Although specialised, CHET is not small (Aston, 1997) as "the term *specialized corpora* does not necessarily imply that the corpora in question are small" (Flowerdew, 2004: 10), and all the sub-corpora in the CC contain around 400,000 words.

The whole CC, with its different sub-corpora, has been designed to complement other corpora which are also specialised and diachronic (Moskowich, 2011), among these *A Representative Corpus of Historical English Registers* (ARCHER, 1650–1999), *The Lampeter Corpus of Early Modern English Tracts (1640–1740), Middle English*

---

2. Works such as diaries and travelogues were no less important than entirely methodological works. As Nietzsche noted, method is useless if it does not provide some creativity. "Having thus become critical of speculative philosophy, he sought to develop an analysis that could explain historical change by attending to the opposite of universals: local realities" (Johnson, 2017). An impeccable method does not necessarily lead to an equally impeccable analysis.

Medical Texts (MEMT, 1350–1500) and *The Helsinki Corpus of English Texts* (up to 1710). In terms of domain, the *Lampeter Corpus* represents Science in general, and thus all the parts of the CC are more specific. The CC is also more specific than the *Helsinki Corpus of English Texts*, originally compiled for the study of the evolution of English in general rather than for the study of the evolution of particular registers. Although both the CC and ARCHER focus on scientific writing, they differ in that ARCHER includes material extracted from the *Philosophical Transactions* whereas CC offers a representation of longer formats and different genres.

The intention of the compilers of CHET was to offer linguistic material that allowed scholars to investigate not only the language of late Modern English writers on History, but also to provide the tools to go one step further and explore how these writers interacted not only with their epistemic communities but with the reading public in general, in the sense that History at the time was perhaps open to a broader readership than other disciplines.

During the Modern English period, knowledge in all its branches, and with all its limits, was viewed from an entirely different perspective from our own twenty-first century one. As noted elsewhere (Crespo & Moskowich, 2010) boundaries between the fields of knowledge were fuzzy in the past, and the methods that are now characteristic of science and academic research were only just beginning to be used in (a few) experimental works. As was the case with the previous two subcorpora of the CC, on Astronomy and Philosophy, such a notion presents a considerable challenge in the compilation of a corpus representative of scientific and academic English. In order to be as inclusive as possible in identifying and compiling the samples, we sought a balance between Modern and present-day perspectives on History. In doing so, and as with the previous subcorpora, UNESCO's (1988) classification of scientific fields has been combined with our knowledge of the texts and their time of composition prior to making any decisions.

As the third of the subcorpora in the overall project, CHET has many characteristics common to its predecessors, these principally derived from the general principles of compilation. One of these, for instance, is that each text sample is accompanied by a metadata file that can be directly accessed from the corpus software or used as a filter of information (see *About the Corpus of History English Texts* in the INFO tab of the Coruña Corpus Tool https://ruc.udc.es/dspace/handle/2183/21849). This metadata includes information on both the author and on the text sampled.

CHET follows the general structure and principles of the CC, with two 10,000-word text extracts collected for each decade, in the belief that this ensures an optimal representativeness. Therefore, each century is represented by a total of some 200,000 words. As seen in Table 1 below, CHET in fact includes 201,938 words for the eighteenth century and 202,486 for the nineteenth, totalling 404,424.

Table 1. Words in CHET

Century	N of words
Eighteenth century	201938
Nineteenth century	202486
TOTAL	**404424**

We have also followed CC principles in including only edited and printed prose texts. We have used first editions whenever possible and, when not, we have chosen versions dating from less than thirty years after the original publication of the work, a stipulation which reflects Kytö, Rudanko & Smittenberg's claim that language change can be perceived within 30-year periods (2000: 92). We have also selected fragments from different parts of the works sampled, so as to avoid the skewing of data due to the stylistic repetition of rhetorical patterns within specific sections of a work.

As part of the CC, the time-span of CHET is delimited by extra-linguistic factors relating to three aspects of life: political changes, scientific changes and communicative changes. In the case of social and political changes, these limits are established by the Restoration of the English monarchy in the late seventeenth-century and the end of the Victorian era. Regarding scientific changes, the emergence of the scientific method is of paramount importance. Accompanied by the foundation of the Royal Society of London, the so-called Scientific Revolution brought with it ideas of clarity and simplicity, such as the basic linguistic or stylistic lines in the presentation of scientific work, and had an impact on all disciplines. However, such an impact was not simultaneous across disciplines, and indeed the texts gathered in CHET are a good example of the late adaptation to the new trends in historical writing.

The upper time-limit is set at 1900, this date reflecting a moment in time when significant events in the history of science marked dramatic changes (Moskowich, 2012). These include the discovery of the electron in 1896, the crisis in mechanical physics, Planck's proposal of quantum mechanics, and Einstein's paper on the Special Theory of Relativity (1905); of course, we might expect such changes to have been accompanied by changes in the prevailing scientific discourse.

CHET contains samples by twenty authors from the eighteenth century and twenty from the nineteenth century, as detailed out in Table 2 below:

Table 2. Authors in CHET

Date	Author	Title
1704	Tyrrell, James	*The General History of England both Ecclesiastical and Civil: containing the reign of Richard II, taken from the most ancient records, manuscripts and printed historians, with an appendix…* vol III, second part. Printed for W. Rogers. London.
1705	Anderson, James	*An historical essay, shewing that the crown and Kingdom of Scotland, is imperial and independent. Wherein the gross mistakes of a late book, intituled, The superiority and direct dominion of the imperial crown and kingdom of England, over the crown and kingdom of Scotland, and of some other books to that purpose are exposed. With an appendix.* Edinburgh: printed by the heirs and successors of Andrew Anderson. To be sold by the booksellers of Edinburgh.
1710	Crawfurd, George	*A Genealogical Hiſtory of the Royal and Illuſtrious Family of the Stewarts, from the Year 1034 to the Year 1710. Giving an Account of the Lives, Marriages and Issue of the moſt Remarkable Perſons and Families of that Name. To which are prefixed, First, a General Description of the Shire of Renfrew, the Peculiar Reſidence and ancient Patrimony of the Stewarts: and, secondly, a Deduction of the Noble and Ancient Families, Proprietors there for upwards of 400 Years, down to the preſent Times: Containing the Deſcent, Original Creations, and moſt Remarkable Actions of their reſpective Anceſtors; alſo the Chief Titles of Honour they now enjoy; with their Marriage and Iſſue, continued down to this preſent Year, and the Coat of Arms of each Family in Blazon*. Edinburgh: printed by James Watson.
1716	Oldmixon, John	*Memoirs of Ireland from the Restoration, to the Present Times. Containing I. An Account of the Designs of the Tories in England and Ireland, to ruin the Protestant Interest there, by breaking the Act of Settlement, and other Acts made for its Security, in 1660, &amp; seq.* *II. A Conspiracy to massacre the Protestants, in 1674.* *III. A Plot for a French Invasion, and to betray the strong Cities and Ports to the Invaders.* *IV. The Debates, concerning that Plot in the Parliament of England, and the Proceedings against the Earl of Tyrone, and others thereupon.* *V. Tyrconnel's Cruel and Arbitrary Government.* *VI. The Tyrannical Reign of the late King James, and his Treaty with Lewis the XIVth, to deliver up that Kingdom to him.* *VII. Some Facts of the Wars in Ireland in 1689, 1690, 1691.* *Never before Printed. With Lifts of King James's Officers Civil and Military; of his Popish Parliament in 1689, and of King William's Parliament in 1692.* London: printed for J. Roberts.

(*continued*)

Table 2. (*continued*)

Date	Author	Title
1721	Strype, John	*Ecclesiastical Memorials; Relating chiefly to Religion, and the Reformation of it, and the Emergencies of the Church of England, under King Henry VIII. King Edward VI. and Queen Mary the First. In three volumes. Volume I. All which Being New, and Such as hitherto Escaped our Writers and Historians, will Communicate much more Light to those great Transactions in this Kingdom: And moreover Discover further the Inclinations and Influences of the respective Princes; The Embassies and Correspondencies with Foreign Potentates and States, chiefly with respect to Religion: The Oppositions made to it; The Troubles and Persecutions of the Professors of it: The Tempers, Practices and Events of the Two Cardinals, Wolset and Pole, and other Prelates and Great Men of Both Parties, in the respective Reigns: Besides, Accounts of Convocations, Royal and Episcopal Visitations, Ecclesiastical Constitutions, Books from time to time set forth; with various other Matters worthy of Note and Observation. In three volumes. With a Large Appendix to each Volume, containing Original Papers, Records, &c.* London: printed for John Wyatt.
1726	Penhallow, Samuel	*History of the Wars of New England with the Eastern Indians. Or, a Narrative of their continued Perfidy and Cruelty, from the 10th of August, 1703. To the Peace renewed 13th of July, 1713. And from the 25th of July, 1722. To their Submission 15th December, 1725. Which was ratified August 5th 1726.* Boston: printed by T. Fleet, for S. Gerriſh.
1732	Horsley, John	*Britannia romana: or, The Roman antiquities of Britain: in three books. Britannia Romana: or the Roman Antiquities of Britain: In three books.* *The I. Contains the History of all the Roman Transactions in Britain, with an account of their legionary and auxiliary forces employed here, and a determination of the stations per lineam valli; also a large description of the Roman walls, with maps of the same laid down from a geometricalsurvey.* *II. Contains a compleat Collection of the Roman Inscriptions and Sculptures which have hitherto been discovered in Britain, with the letters engraved in their proper shape and proportionate size, and the reading placed under each; as also an historical account of them, with explanatory and critical observations.* *III. Contains the Roman Geography of Britain, in which are given the originals of Ptolemy, Antonini Itinerarium, the Notitia, the anonymous Ravennas, and Peutinger's table, so far as they relate to this island, with particular essays on each of these ancient authors, and the several places in Britain mentioned by them.* *To which are added, a Chronological Table, and Indexes to the Inscriptions and Sculptures after the manner of Gruter and Reinesius.* Also Geographical Indexes both of the Latin and English names of the Roman places in Britain, and a General Index to the work. *The whole illustrated with above an hundred Copper Plates.* London: printed for John Osborn and Thomas Longman.

Table 2. (*continued*)

Date	Author	Title
1739	Justice, Elizabeth	*A Voyage to Russia: describing the Laws, Manners, and Cuſtoms, of that great Empire, as govern'd, at this preſent, by that excellent Princeſs, the Czarina. Shewing the Beauty of her Palace, the Grandeur of her Courtiers, the Forms of Building at Petersburgh, and other Places: with ſeveral entertaining Adventures, that happened in the Paſſage by Sea, and Land.* York: printed by Thomas Gent
1740	Bancks, John	*The history of the life and reign of the Czar Peter the Great Emperor Of All Russia And Father Of His Country in India.* London: printed for J. Hodges
1745	Hooke, Nathaniel	*The Roman History, from the building of Rome to the Ruin of the Commonwealth. Illustrated with Maps and other Plates.* Vol. II. Book IV. London: printed by James Bettenham
1750	Chapman, Thomas	*An essay on the Roman Senate*. Cambridge: printed by J. Bentham
1760	Birch, Thomas	*The Life of Henry Prince of Wales, Eldest Son of King James I. Compiled chiefly from his own Papers, and other Manuscripts, never before published.* London: printed for A. Millar.
1762	Scott, Sarah	*The History of Mecklenburgh, from the Firſt Settlement of the Vandals in that Country, to the Present Time; including a Period of about Three Thouſand Years*. London: printed for J. Newbery.
1770	Adams, Amos	*A concise, historical view of the perils, hardships, difficulties and discouragements which have attended the planting and progressive improvements of New-England; with a particular account of its long and destructive wars, expensive expeditions, &c.: With reflections, principally, moral and religious.: In two discourses, preached at Roxbury on the general fast, April 6. 1769: And published at the general desire of the hearers.* Boston: printed and sold by Kneeland and Adams.
1775	Anderson, Walter	*The History of France. From the Commencement of the Reign of Henry III. and the Riſe Of the Catholic League; to the Peace of Vervins, and the Establishment Of the famous Edict of Nantes, In the Reign of Henry IV. Together with The moſt intereſting Events in the Hiſtory of Europe, during that Period*. London: printed for the autor.
1780	Cornish, Joseph	*The life of Mr. Thomas Firmin, citizen of London*. London: printed for J. Johnson.
1788	Gibbon, Edward	*The history of the decline and fall of the Roman Empire*. Volume the tenth. Dublin: Luke White.
1790	Gifford, John	*The History of England from the earliest Times to the Peace of 1783.* Vol. I. Book III. London: printed for Harrison and Co.

(*continued*)

Table 2. (*continued*)

Date	Author	Title
1795	Adams, John	*A view of universal history, from the creation to the present time. Including an account of the celebrated revolutions in France, Poland, Sweden, Geneva, &c. &c. Together with an accurate and impartial narrative of the late military operations and other important events.* Vol. II. London: printed for G. Nearsley.
1800	Stock, Joseph	*A narrative of what passed at Killalla, in the County of Mayo, and the parts adjacent, during the French invasion in the summer of 1798. By an eye witness.* Dublin: printed by T. Baylis.
1802	Adolphus, John	*The history of England from the accession of King George the Third, to the conclusion of peace in the year one thousand seven hundred and eighty-three.* Vol. III. London: printed for T. Cadell, Jun. and W. Davies.
1805	Warren, Mercy Otis	*History of the rise, progress and termination of the American revolution. Interspersed with Biographical, Political and Moral Obſervations. In three volumes*. Vol. I. Boston: printed by Manning and Loring, for E. Larkin
1810	Bigland, John	*The history of Spain, from the earliest period to the year 1809. In two volumes*. Vol.1. London: printed for Longman, Hurst, Rees, and Ormd.
1814	Britton, John	*The history and antiquities of the cathedral church of Salisbury; illustrated with a series of engravings, of views, elevations, plans, and details of that edifice: also etchings of the ancient monuments and sculpture: including biographical anecdotes of the bishops, and other eminent persons connected with the church*. London: printed for Longman, Hurst, Rees, Orme, and Brown.
1820	Hardiman, James	*The history of the town and county of the town of Galway, from the earliest period to the present time. Embellished with several engravings. To which is added a copious appendix, containing the principal charters and other original documents*. Dublin: printed by W. Folds and sons.
1828	Callcott, Maria	*A Short history of Spain. In two volumes*. Vol. II. London: John Murray.
1833	Aikin, Lucy	*Memoirs of the Court of King Charles the First. In two volumes*. Vol. I. London: printed for Longman, Rees, Orme, Brown, Green, and Longman.
1839	Petrie, George	*On the History and Antiquities of Tara Hill*. Dublin: Royal Irish Academy.
1840	Smyth, William	*Lectures on Modern History, from the Irruption of the Northern Nations to the Close of the American Revolution. In two volumes*. Vol II. Cambridge: J. & J. J. Deighton.
1844	D'Alton, John	*The history of Drogheda, with its environs; and an introductory memoir of the Dublin and Drogheda railway. In two volumes*. Vol. I. Dublin: printed by M. H. Gill.

Table 2. (*continued*)

Date	Author	Title
1855	Masson, David	*Medieval history*. London: printed by W. and R. Chambers.
1857	Sewell, Elizabeth Missing	*A first history of Greece*. New York: printed by D. Appleton and Company.
1860	Freer, Martha Walker	*History of the reign of Henry IV. King of France and Navarre. In two volumes*. London: printed by Hurst and Blackett, publishers.
1862	Bennett, George	*The History of Bandon*. Cork: Henry and Coghlan.
1872	Gray, John Hamilton	*Confederation; or, The Political and Parliamentary History of Canada, from the Conference at Quebec, in October, 1864, to the Admission of British Columbia, in July, 1871. In two volumes*. First volume. Toronto: Copp, Clark & Co., printers.
1875	Killen, William Dool	*The ecclesiastical history of Ireland. From the earliest period to the present times*. Vol. II. London: printed by Macmillan and co. / R. Clay, sons, and Taylor, printers.
1884	Breese, Sidney	*The Early History of Illinois, from its Discovery by the French, in 1673, until its Cession to Great Britain in 1763. Including the Narrative of Marquette's Discovery of the Mississippi*. Chicago: E. B. Myers & Company.
1887	Kingsford, William	*The history of Canada*. Vol. I. [1608–1682.]. Toronto: Rowsell & Hutchinson.
1893	Cooke, Alice M.	"The Settlement of the Cistercians in England". *The English Historical Review*, Vol. 8, No. 32. (625–648)
1895	Burrows, Montagu	*The History of the Foreign Policy of Great Britain*. Edinburgh and London: W. Blackwood and sons.

The authors and their works cited in the list above are intended to be representative of the field of History in a broad sense. As already noted, we have compiled two *ca*. 10,000-word samples of printed material per decade. The list of authors includes eight women, these occurring throughout the corpus, from 1739 to 1893, a considerable number compared to the two in CETA or the three in CEPhiT.

The preference for first editions seeks to reflect the language as actually used by the author. Also, no translations have been included, as these are likely to carry interference from the source language, which would certainly alter any possible analyses and findings for which the data were used. Excluding translations implies a particular obstacle in terms of collecting eighteenth-century works, in that many authors still wrote in Latin and had their works later translated into English or made the translations themselves. It is also a general practice for the CC to include just one sample per author, to avoid individual idiosyncrasies appearing as general linguistic trends.

In what follows, we will look at a number of variables and their distribution in CHET. In relation to the texts themselves, we will see how samples are distributed according to genre or communicative format. As regards authors and authorship, we will consider their sex and the place where they acquired their writing habits.

Our classification of communicative formats or genres in the CC is based on Görlach's description of the evolution of what he calls text-types (2004). In CHET in particular we have identified the formats set out in Table 3 below:

Table 3. Genres in History texts

Genres in CHET	Samples	Words
Treatise	28	283002
Essay	3	30312
Travelogue	1	10005
Lecture	2	30120
Textbook	2	20203
Article	1	10730
Dictionary	1	10017
Biography	1	10035
TOTAL		**404424**

Looking at the texts which we have been able to compile here, we see that works dealing with historical matters during the late Modern English period were basically written within eight formats, and that these were not equally popular. The formal and functional restrictions imposed by discipline are clearly observed here, in that the genres Travelogue and Biography are not to be found in either of the previously compiled subcorpora (CETA and CEPhiT), and on-going work suggests that it is unlikely that they will be found in those to come. Perhaps because the status of History as a discipline is yet to become well defined, some of the genres here occur only rarely, as is the case with the 1780 biography by Cornish, the 1814 work by Britton, a half-catalogue, half-dictionary of technical terms, or the 1893 article by Alice Cooke, Article being a genre which was still highly restricted to the experimental and observational sciences at the time. For similar reasons of the diachronic dynamics of the genre we included just one sample of Travelogue, an account of a voyage by Elizabeth Justice (1739), a format that was dying out at the time, to be replaced by travel guides and documentaries. Perhaps due to the non-experimental character of History, we did not find many essays, and textbooks were similarly scarce, perhaps due to History being a field open to the general public. However, there were a huge number of treatises available, although here we ought to note that the Treatise is not necessarily a highly formal or technical format; on the contrary, the term is often used by authors themselves to refer to almost any piece of writing of considerable length. The two centuries show a markedly different distribution in terms of formats, this perhaps a direct consequence of the evolution of the field itself. In fact, although we find five genres represented in each section of CHET, they are rather different. This way, the five genres to be found in the eighteenth-century samples are Essay, Treatise, Lecture, Travelogue and Biography. The five genres found in the following period are not the same five as there are extracts of Treatise, Dictionary, Lecture, Textbook and Article. Contrary to our previous *Corpus of English Philosophy Texts*, where a wider variety of genres appeared in the nineteenth century, History texts seem to use a similar range of formats over time, albeit with some disappearing or emerging over the course of the two centuries. Thus, the formats Treatise and Lecture occur across the whole period, whereas Essay, Biography and Travelogue are represented only in the eighteenth century, and Article, Dictionary and Textbook only in the nineteenth.

As for variables directly relating to authors, we will deal first with geographical distribution, that is, the places where the authors were educated and acquired their writing habits. Each metadata file accompanying the document of a text sample provides information about the places of an author's education, from general to specific, so that writers can be classified as North-American or European, and then relating to particular territories and cities. Figure 1 sets out information regarding the latter, what we call Place 2 in metadata files, that is, the country or territory of a writer's education, as this is perhaps the most pertinent classification. More than half the authors (24) in CHET were educated in England, six in Ireland, and five in Scotland. More broadly, Europe produced far more works than North America. As with CEPhiT, historical reasons might go a long way towards explaining this: North America, after a convulsive eighteenth century, was perhaps more concerned with the practical application of scientific advances and economic recovery, and consequently was less motivated to explore other branches of knowledge.

Sex is the second parameter here. As might have been expected, not many women wrote and published works on History in the period, although it was not in fact one of those fields restricted to men. As already mentioned, there are eight female authors (representing 23% of the total) as seen in Figure 2 below, which is certainly far more than in other disciplines.

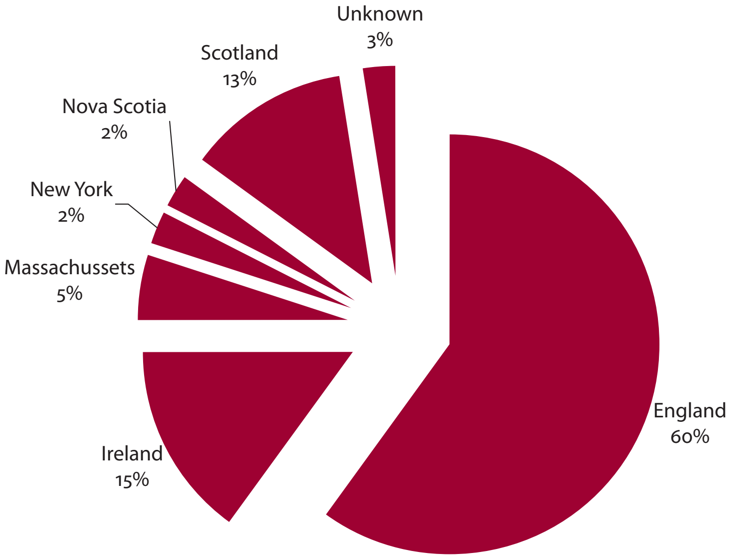


Figure 1. Geographical distribution in CHET

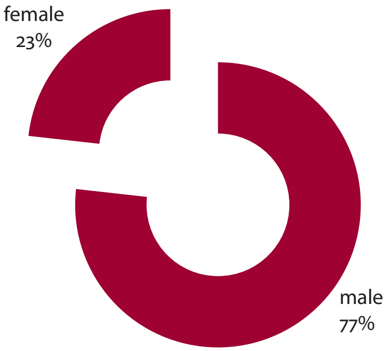


Figure 2. Proportion of words by female/male authors

There seems to be evidence here of changes in society, in that the eighteenth century saw few women publishing works under their own name compared to the following century. In fact, there are only two female authors in the eighteenth-century part of the corpus (Elizabeth Justice with her Travelogue and Sarah Scott with a Treatise on the history of Mecklenburgh). In the following century, by contrast, we managed to collect six representative samples of History texts by women, of which only one seems to be a less formal format (the Textbook by Sewell), the rest being treatises, plus one article. This evolution in time can be seen in Figure 3:

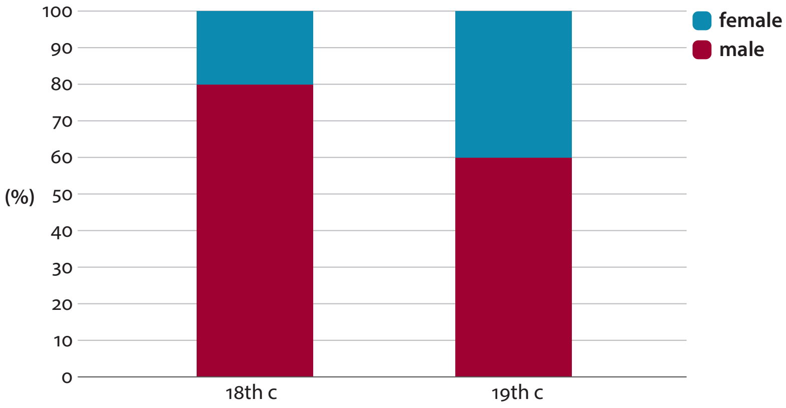


Figure 3. Sex of authors in CHET by century

Women were excluded from institutional science, and publishing works on History or any sort of learned theme was still considered a masculine activity, with the work of women often not taken seriously (Herrero, 2007: 75). Perhaps because of the type of discipline History was at the time, CHET does not reflect the scarcity of overt female activity as greatly as other fields do.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Aston, Guy. (1997). Small and large corpora in language learning. In Barbara Lewandowska-Tomaszczyk, & Patrick J. Melia (Eds.), *Practical applications in language corpora* (51–62). Lodz: Lodz University Press.

Connor, Ulla & Thomas A. Upton. (2004). Introduction. In Ulla Connor & Thomas A. Upton (Eds.), *Discourse in the Professions. Perspectives from corpus linguistics* (1–8). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/scl.16.01con

Crespo, Begoña & Isabel Moskowich. (2010). *CETA* in the Context of the Coruña Corpus. *Literary and Linguistic Computing*, 25 (2), 153–164. https://doi.org/10.1093/llc/fqp038

De Beaugrande, Robert. (2001). Large corpora, small corpora, and the learning of "language". In Moshen Ghadessy, Alex Henry, & Robert L. Roseberry (Eds.), *Small Corpus Studies and ELT. Theory and practice* (3–28). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.5.04bea

De Syon, Guillaume. (1999). Voltaire. In Kelly Boyd (Ed.), *Encyclopedia of Historians and Historical Writing*, vol 2. (1270–72). Oxford: Taylor & Francis.

Flowerdew, Lynne. (2004). The argument for using English specialized corpora to understand academic and professional language. In Ulla Connor & Thomas A. Upton (Eds.), *Discourse in the Professions. Perspectives from corpus linguistics* (10–33). Amsterdam: John Benjamins. https://doi.org/10.1075/scl.16.02flo

Görlach, Manfred. (2004). *Text-Types and the History of English*. Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110197167

Ham, R. G. (1935). Dryden as Historiographer-Royal: The Authorship of His Majesties Declaration Defended, 1681. *Review of English Studies*, 11 (43), 284–298. https://doi.org/10.1093/res/os-XI.43.284

Harper, Douglas. (2010–2017). *Online Etymology Dictionary*. Accessed [07/12/2016] from http://www.etymonline.com/

Herrero López, Concepción. (2007). Las mujeres en la investigación científica. *Criterios*, 8, 73–96.

Hoefferle, Caroline. (2011). *Essential Historiography Reader*. Boston: Pearson.

Huizinga, Johan, James Holmes, & Hans van Marle. (1984). The task of Cultural History. *Men and Ideas: History, the Middle Ages, the Renaissance* (17–75). Princeton (NJ): Princeton University Press.

Jaki, Stanley L. (1993). History as Science and Science in History. *The Intercolegiate Review, Fall*, 20–27.

Johnson, Marcus. (2017). *Nietzsche, Philology and Genealogy*. Accessed [24/02/2017] from http://www.academia.edu/30482423/Nietzsche_Philology_and_Genealogy

Kytö, Merja, Juhani Rudanko, & Erik Smitterberg. (2000). Building a Bridge between the Present and the Past: A Corpus of 19-century English. *ICAME*, 24, 85–97.

Moskowich, Isabel. (2011). 'The Golden Rule of Divine Philosophy' exemplified in the Coruña Corpus of English Scientific Writing. *Lenguas para fines específicos*, 17, 167–198.

Moskowich, Isabel. (2012). "A smooth homogeneous globe" in *CETA*: Compiling Late Modern Astronomy Texts in English. In Nila Vázquez (Ed.), *Creation and Use of Historical English Corpora in Spain* (21–37). Cambridge: Cambridge Scholar Publishing.

Moskowich, Isabel & Javier Parapar. (2008). Writing Science, Compiling Science. The Coruña Corpus of English Scientific Writing. In María Jesús Lorenzo Modia (Ed.), *Proceedings from the 31st AEDEAN Conference* (531–544). A Coruña: Universidade da Coruña.

Nietzsche, Friedrich. (1996). *Human, All Too Human. A Book for Free Spirits*. Cambridge: Cambridge University Press.

OED Online (2009). Oxford: Oxford University Press. Accessed [07/12/2016] from http://dictionary.oed.com

Salevouris, Michael J. & Conal Furay. (2015). *The Methods and Skills of History: A Practical Guide*. Oxford: John Wiley and Sons.

Saslow, Edward L. (1978). Dryden as Historiographer Royal, and the Authorship of 'His Majesties Declaration Defended'. *Modern Philology*, 75 (3), 261–272. https://doi.org/10.1086/390790

UNESCO (1988). *Proposed International Standard Nomenclature for Fields of Science and Technology*. UNESCO/ROU257 rev. 1. Paris.

CHAPTER 4

Typical linguistic patterns of English History texts from the eighteenth to the nineteenth century

An information-theoretic approach

Stefania Degaetano-Ortlieb, Katrin Menzel and Elke Teich
Universität des Saarlandes

1. Introduction

In this paper, we investigate the diachronic linguistic development of history texts from the late Modern English (lModE) period as depicted by the CHET corpus. These history texts are scientific texts belonging to the humanities. The question we approach is whether they show unique linguistic patterns distinguishing them from other scientific fields and from 'general' language of the lModE period.

In history texts, we can expect numerous noun phrases, in particular terms and lexical items that specify places, time periods, power relations or identity. These texts are subject-oriented texts and will probably contain a high number of proper nouns for people, collective structures, and communities, institutions, geographical locations or unique objects. According to Coffin (1997) one particular mode of representation of history texts is the narrative one, which as Werlich (1975, 1983) and later Biber (1988), among others, show is marked by particular linguistic features (e.g. verbs in the past tense, activity and process verbs, temporal expressions, temporal conjunctions as well as third person pronouns, deictic expressions, and coreference-chains). Also, history texts will typically draw parallels or analogies between the present and the past and emphasise the relevance of historical events to present circumstances.

In the analyses, our focus will be on the phrase/clause level, considering that scientific writing is prominently marked by nominal rather than verbal style as well as embedded structures within nominal phrases (e.g. Halliday, 1988; Biber *et al.*, 1999; Degaetano-Ortlieb & Teich, 2016). As these history texts are classified as scientific texts, they should also show prevalence for nominal style when compared to 'general' language. A comparison to other scientific texts will show

https://doi.org/10.1075/z.225.04deg

© 2019 John Benjamins Publishing Company

whether history texts make use of particular phrase/clause structures distinguishing them from other scientific disciplines. For this, we use *relative entropy*, an information-theoretic measure that captures differences between two probability distributions, comparing the CHET corpus to a general language corpus and a scientific text corpus that cover the same time span.

In a second part, we look in more detail at possible diachronic linguistic changes focusing on lexical productivity of typical patterns for history texts as opposed to other scientific and general language texts. Here, we follow the assumption that typical patterns can serve two purposes. They can either realise conventionalised language use (e.g. formulaic expressions) and are thus typical of that genre/register or they function as a particular device to communicate informational content. A conventionalised purpose would be reflected in a less productive use more confined to particular lexical items, while a content-oriented purpose would result in a more productive use reflected in a variety of lexical items. For this, we use *surprisal*, an information-theoretic measure that allows us to calculate the probability of a lexical item based on its previous context. As already shown in Degaetano-Ortlieb & Teich (2016), surprisal is a good indicator of lexical productivity. Thus, in our analyses, we pose the following questions:

a. Which linguistic patterns are typical of history texts of the lModE period as depicted by the CHET corpus in comparison to other scientific texts and general language texts?
b. Do these typical patterns become more or less productive over time in history texts as opposed to scientific and general language texts?

To answer these questions, we make use of three diachronic corpora with British texts from the lModE period (1700 and 1900) representing different specialised and non-specialised text genres which will be described in more detail in Section 2. Primarily, we compare the CHET corpus depicting specialised language within the field of history (Crespo & Moskowich, 2015) to specialised scientific language in a corpus of research articles from journals dedicated to the natural sciences (RSC; Kermes *et al.*, 2016) and with a non-specialised corpus that was compiled with the aim to represent general lModE usage (CLMET; De Smet *et al.*, 2015).

Diachronic change in lModE is still a relatively underresearched area compared to other stages of the English language. There seem to be far more studies, for instance, on early Modern English. For lModE, there are a few introductory textbooks (e.g. Görlach, 1999, 2001; Tieken-Boon van Ostade, 2009) and overviews on grammar-writing in the eighteenth and nineteenth century (Görlach, 1998; Tieken-Boon van Ostade, 2008). Only a few authors have explored specific linguistic phenomena for the whole period of lModE in a systematic way. Although empirical, corpus-based studies on lModE are still rare, there is considerable recent interest in this area. There are some noteworthy recent research results on the characteristics of lModE that are based on the analysis of corpus data. Certain authors have obtained valuable insights from a qualitative and quantitative perspective on certain linguistic phenomena although they typically analysed only relatively small, specific corpora (e.g. Arnaud (1998) on the development of the progressive in letters by around 20 authors from the nineteenth century, a sufficiently frequent and unambiguous grammatical construction in this type of texts). Some article-length research reports and a few monographs and edited volumes also cover various aspects of lModE (e.g. Kytö *et al.*, 2006; Hundt, 2014). Since 2001, there have been six International Conferences on Late Modern English (LMEC) that resulted in several volumes of proceedings and thematic publications, focusing especially on selected grammatical and syntactic phenomena and, to a certain extent, on sociolinguistic or genre-specific variation (e.g. Dossena & Jones, 2003; Dossena & Tieken-Boon van Ostade, 2008; Pérez-Guerra *et al.*, 2007; Tieken-Boon van Ostade & van der Wurff, 2009).

The richly annotated corpora that we use in our study have been collected only recently and provide significant research potential as they permit detailed empirical and exploratory analyses to be conducted. They have inspired genre-specific studies on lModE such as Monaco's comparison between philosophical and life sciences texts (Monaco, 2016) or Moskowich's analysis of stance markers in academic texts from the domains of astronomy, philosophy and history (Moskowich, 2017). Our analysis will complement the findings from publications such as the above-mentioned works and possibly lead to new insights into the linguistic changes of the lModE period uncovering general, across-the-board shifts as well as genre-specific developments.

The paper is structured as follows. In Section 2, we present our data and an overview of our methodology. Section 3 shows selected analyses of typicality and lexical productivity to illustrate our corpus-based approach informed by information-theoretic measures. Section 4 concludes the paper with a summary and envoy.

2. Data and methodology

2.1 Corpora

For our analysis, we chose three diachronic corpora of lModE texts – the *Corpus of History English Texts* (CHET), the *Corpus of Late Modern English Texts* (CLMET), and the *Royal Society Corpus* (RSC). From these corpora we select only overlapping periods of 1700–1850 for comparison.

The *Corpus of History English Texts* (CHET), beta version, is a subcorpus of the *Coruña Corpus of English Scientific Writing* (Crespo & Moskowich, 2015). It consists of around 400,000 words, which were extracted from 40 works from the field of history published in the eighteenth and nineteenth centuries (two samples of approximately 10,000 words per decade). The corpus design includes six genres: Treatise, Essay, Lecture, Textbook, Article and Travelogue (see Moskowich, this volume). Treatises make up the largest proportion of this specialised corpus in terms of number of texts and number of words. The other text types account for up to 10% each of the total words in the corpus. The texts, for instance, include extracts from works by British historians such as George Crawfurd, Montagu Burrows or Martha Walker Freer. Around 20% of the corpus consists of texts by female authors. A few texts by North American historians are also included in the data, but most works on history were produced in Europe during this period. Therefore, most texts of this corpus are mostly written by British authors. The corpus was sampled as a collection of extracts from introductions, central chapters and conclusions from texts by historians from different fields to be as representative and balanced as possible. Metadata for each text include sociolinguistic information on the author, such as age, sex and place of education.

The *Corpus of Late Modern English Texts* (CLMET; De Smet *et al*., 2015) is a larger collection of texts by British authors drawn from the Oxford Text Archive and the Project Gutenberg. It was compiled with the aim to complement smaller historical corpora and to mirror general lModE usage. This corpus covers the period from 1710 to 1920 and is subdivided into three main subperiods of 70 years each. The corpus version that we used for our analysis (V3.1) consists of *ca*. 40 million words from different genres of formal written British English. A considerable number of texts in this resource had been drawn from prose fiction so that the corpus is slightly biased towards literary texts. It is still a relatively genre-balanced corpus, including narrative fiction and non-fiction, drama, letters, treatises as a subtype of scientific texts as well as some unclassified texts. There is some variation in terms of authorial social background. The corpus includes texts by men and women from diverse social classes. Nevertheless, most authors from that period represented in the collection are men from the socially and economically dominant classes. The maximum amount of text per author is 200,000 words. Where possible, different texts of the respective authors were sampled. The corpus contains works by authors such as Winston Churchill, Emily Brontë and Charles Darwin.

The *Royal Society Corpus* (RSC; Kermes *et al*., 2016) is a large specialised corpus of complete scientific journal articles. This corpus includes the digitised texts of the *Philosophical Transactions* and the *Proceedings of the Royal Society of London* published between the middle of the seventeenth century and the second half of the nineteenth century. The *Philosophical Transactions* as the first and longest-running

English scientific journal, represent various scientific disciplines that became more specialised over time. The authors were typically British fellows of the Royal Society, sometimes also foreign members or other outstanding scientists of the time. Most articles in the period covered were written by men. Text examples from the corpus are research articles by natural philosophers and scientists such as Thomas Henry Huxley, Caroline Herschel or Benjamin Franklin. The RSC version used in our study (V3.6.0) has a size of about 32.5 million tokens. The RSC corpus is distributed through CLARIND-D. For query it is also available via the Saarbrücken CQPweb interface (http://corpora.clarin-d.uni-saarland.de/cqpweb/).

All corpora used in our analysis have been released in XML format, are annotated at different linguistic levels (tokens, parts of speech, lemma) with TreeTagger (Schmid, 1995), segmented into sentences, and encoded in the Corpus Query Processor (CQP; Evert & Hardie, 2011).

2.2 Detecting and inspecting variation with information-theoretic measures

2.2.1 *Feature detection*

Our focus is on detecting phrase/clause patterns that are typical of history texts compared to general language texts and other scientific texts. Instead of selecting specific predefined patterns to be analysed, our approach aims at finding patterns that are differently used in a data-driven fashion. For this, we use relative entropy, more precisely Kullback-Leibler Divergence (KLD; Kullback & Leibler, 1951).

In comparison to relative frequency, which measures the probability of a linguistic unit in a corpus, relative entropy measures the probability of a unit in corpus *A* compared to the probability of the same unit in corpus *B*. This is formalized as follows:

$$D(A||B) = \sum_i p(unit_i|A) log_2 \frac{p(unit_i|A)}{p(unit_i|B)}$$

where $p(unit_i|A)$ is the probability of a unit in corpus *A*, and $p(unit_i|B)$ the probability of that unit in corpus *B*. The $log_2 \frac{p(unit_i|A)}{p(unit_i|B)}$ denotes the difference between both $p(unit_i|B)$ probability distributions (i.e. the ration between the two distributions). This ratio is then weighted with the probability of $p(unit_i|A)$ so that the sum over all occurrences of $unit_i$ gives the *relative* entropy.

The main advantage of using relative entropy over frequency is that it directly gives us a comparative measure, which we do not obtain by considering the relative frequency of a unit in a corpus. In information-theoretic parlance, relative entropy measures bits of transmitted information, more precisely the number of additional bits of information needed to process a linguistic unit in corpus *A* based on its probability distribution in corpus *B*. The more two probability distributions differ, the higher the number of bits (higher relative entropy value). The more similar the two probability distributions, the lower the number of bits (lower relative entropy value). Based on the value obtained by relative entropy for each linguistic unit, we can determine which units are the most *typical* ones of a particular distribution. Thus, in a comparison between the use of a linguistic unit in two corpora, the higher the relative entropy value of a unit for one corpus (e.g. CHET) in comparison to another corpus (e.g. RSC), the more typical that unit is for the given corpus.

The linguistic units in our case are phrase/clause patterns based on sequences of three part-of-speech tags (i.e. part-of-speech trigrams, e.g. noun followed by preposition followed by noun), but any linguistic unit could be used (phoneme, morpheme, word, etc.). We chose part-of-speech trigrams as bigrams would have been too short to depict phrase/clause structure and four-grams would have been too long and thus would have produced too sparse data (i.e. they would have been too infrequent). Tri-grams proved to work well also in other diachronic studies (Culpeper & Kytö, 2010; Kopaczyk, 2013; Degaetano-Ortlieb & Teich, 2016; Degaetano-Ortlieb *et al.*, 2018) and we use them as indicators to identify phrase/clause patterns.

2.2.2 *Feature inspection*

After obtaining typical phrase/clause patterns for history texts, we inspect the diachronic development of these patterns in terms of their lexical productivity, i.e. whether the lexical items used in the pattern show high variation or are more confined to particular lexical items pointing to a more conventionalised usage. For this, we use surprisal.

In comparison to the frequency of a lexical item in a corpus, surprisal measures how probable a lexical item is given a particular preceding context in a corpus. Take for example the word *tea*, which is a relatively low frequency word (50.35 per million tokens) in comparison to *man* (1917.99 per million tokens) in the CLMET corpus. However, in the context of *cup of* the probability of having *man* is lower (0 occurrences in CLMET) than of *tea* (135 occurrences in CLMET). Surprisal captures exactly this as it measures the probability of a unit in a particular context, formalised as:

$$S(unit) = -log_2 p(unit|context)$$

where $p(unit|context)$ is the probability of a unit (e.g. *tea*) given a context (e.g. *cup of*). This probability is considered in a logarithmic space (here logarithm to the base 2 to simplify calculation).

In terms of information theory, surprisal calculates the number of bits transmitted by a linguistic unit in a context. The more bits are transmitted the more surprising and thus informative a linguistic unit is in the given context. Thus, in our example *man* following *cup of* would be quite surprising while *tea* would be quite predictable. In our case, the linguistic units are the words (i.e. unigrams) in a corpus. As context we define three preceding words, which allows us to account for enough context for capturing possible phrase/clause patterns and to avoid a rare context (which would be the case if we considered a longer span of preceding words).[1] Surprisal is calculated for each corpus and time period separately.

Considering that we focus on phrase/clause patterns, we also calculate the surprisal of a pattern. For this, we first average the surprisal of each lexical item of a lexical realisation (lexR) for a given pattern. For example, lexical surprisal of the expression *cup of tea* (which is one lexical realisation of the pattern *noun preposition noun*) would consist of the sum of *S(cup), S(of)* and *S(tea)* divided by three (i.e. the average surprisal of the expression):

$$S(lexR) = \frac{S(cup) + S(of) + S(tea)}{3}$$

We then take the mean of each expression realising the *noun preposition noun* pattern, which gives us $S(pattern)$:

$$S(pattern) = \frac{\sum_{i=1}^{n} S(lexR)_i}{n}$$

where n is the number of lexical realisations. For each pattern, surprisal is measured for each corpus and time period separately. The range of surprisal values obtained for a pattern (i.e. all values of *S(lexR)* of a pattern) is divided into three groups based on quartiles. Thus, we obtain high, middle and low surprisal ranges for a pattern. High *S(pattern)* indicates high lexical variation of a pattern and thus a productive use. Low *S(pattern)* indicates low lexical variation of a pattern and thus a more confined use to specific lexical items. We then compare the percentage of high, middle and low *S(pattern)* across corpora and time periods. Over time, percentage of high, middle and low *S(pattern)* may change indicating how the pattern changes in terms of its lexical productivity.

---

[1] Some studies on very formulaic genres focus on longer lexical bundles and identical repetitive strings of words in order to look for textual standardisation and "evidence of units of routine language use" (e.g. Stubbs & Barth (2003: 62ff) analysing the number of identical 7-grams in religious texts – such patterns would be too long for our purposes and clearly lead to data sparsity).

3. Analyses

In this section, we present two analyses. First, we explore linguistic patterns that are typical of history texts in the CHET corpus by comparison with general language as represented by the CLMET corpus and with scientific language as represented by the RSC corpus. Second, we show linguistic diachronic developments in terms of lexical productivity of typical patterns in the CHET corpus and compare developments of the same patterns across the other two corpora.

3.1 Typical linguistic patterns of CHET

To analyse whether history texts have unique linguistic patterns distinguishing them from other scientific fields and from general language of the late Modern English period, we use relative entropy as described in Section 2.2.1.

Comparison of history texts with general language is shown in Table 1, presenting the top five typical patterns of CHET and CLMET, respectively. In history texts (CHET), nominal structures are clearly typical showing also a preference for embedded structures by the use of prepositional phrases within noun phrases. General language texts (CLMET) show a clear preference for involved style by the use of verb phrases combined with the personal pronoun *I* and modality or mental state verbs (e.g. *I should like, I am afraid, I think*). Thus, on a scale from informational to involved style, history texts are positioned more towards the informational end when compared to general language.

Comparison between history (CHET) and other scientific texts (RSC) is presented in Table 2, showing the top five typical patterns. Two patterns typical of CHET are identical to the comparison to CLMET: nominal phrase structures with prepositional embedding and with proper nouns (NN.IN.NP and NP.IN.NP). Showing up in both comparisons indicates that these structures are particularly typical to CHET texts. In fact, in history texts entities such as persons, locations and the like are a main part of the informational content presented.[2]

---

[2] Note that TreeTagger, the part-of-speech tagger used for tagging our corpora, tends to assign the NP (proper noun) tag to nouns beginning with a capital letter, which is also sometimes problematic in Present-day English corpora. Therefore, we will query NN (common noun) and NP (proper noun) together as one group in our following analysis in Section 3.2. This is particularly important for the earlier data before 1750 when nouns that authors or printers wanted to make more prominent or that belong to certain semantic categories, e.g. animate nouns, general abstract nouns, concrete objects, were often written with an initial capital letter (cf. Osselton, 1985).

Table 1. Top five patterns for history (CHET) and general language texts (CLMET)

History texts (CHET)		**General English (CLMET)**	
Pattern	**Example**	**Pattern**	**Example**
NN.IN.NP	*reign of King; king of England*	PP.VVP.RB	*I do not; I know not*
DT.NN.IN	*the death/name of*	PP.VVP.PP	*I think I; I think it*
IN.DT.NN	*in the year; of the king*	PP.MD.VV	*I should like/think*
NP.IN.NP	*King of Denmark; Duke of Gloucester*	PP.VBP.JJ	*I am sure/afraid/glad*
NN.IN.DT	*part/account of the*	PP.VBP.RB	*I am not; they are not*

DT: determiner, IN: preposition or subord. conj., JJ: adjective, MD: modal verb, NN: common noun sg., NP: proper noun sg., PP: personal pronoun, RB: adverb, VBP: verb *be* present non-3rd p., VV: verb base form, VVP: verb present non-3rd p.

Table 2. Top five patterns for history (CHET) and scientific texts (RSC)

History texts (CHET)		**General English (CLMET)**	
Pattern	**Example**	**Pattern**	**Example**
NN.IN.NP	*reign of King; king of England*	IN.DT.NN	*of an inch; of the body*
NP.IN.NP	*King of Denmark; Duke of Gloucester*	NN.IN.DT	*part of the; Duke of Gloucester; surface of the*
IN.PP.VBD	*that it was; that he was*	DT.NN.IN	*the action/surface of*
VVD.TO.VV	*refused/began to take*	DT.JJ.NN	*the same time/manner*
TO.VV.PP$	*to make/use/take/save/support his*	IN.DT.JJ	*on/in/at the same*

DT: determiner, IN: preposition or subord. conj. / that complementizer, JJ: adjective, NN: common noun sg., NP: proper noun sg., PP: personal pronoun, PP$: possessive pronoun, TO: to, VBD: verb *be* past, VV: verb base form, VVD: verb past tense

In addition, *that*-clauses and *to*-infinitive clauses (see Examples (1) and (8)) are typical of history texts in comparison to natural science texts. Both types of clauses follow mental state, speech act and reporting verbs, typically in the past tense (see Examples (1) to (4)), indicating a narrative style of writing as described by Biber (1988). *That*-clauses also follow denotations of (un)certainty and possibility (Examples (5) and (6)). *To*-infinitives may also signal narration expressing the purpose of an action (Examples (7) and (8)).

(1) In fine, they **perceived**, *that it was* only the Fear of the Spaniards Victory in Italy that let the Cause. (CHET. Strype, 1721: 94 [126 (10169)])

(2) They possessed themselves of the Arsenals and Ammunition, and **seemed** *to threaten* a general Destruction. (CHET. Bancks, 1740: 10 [130 (1723)])

(3) He **said**, they were of his Sect and Opinions; and *that they were learned*, before he companied with them. (CHET. Strype, 1721: 77 [126 (1551)])

(4) He also **promised** *to reduce* the royal forests to their ancient boundaries, to correct all tyrannical encroachments […].
(CHET. Gifford, 1790: 180 [139 (1155)])

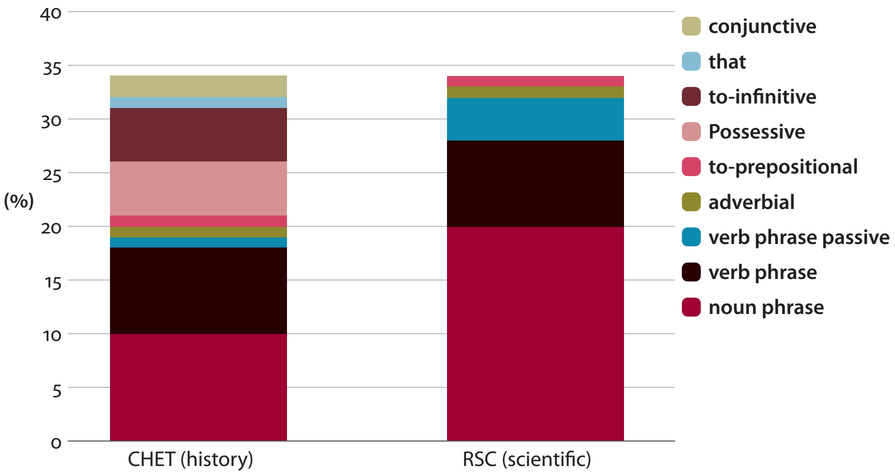


Figure 1. Phrase and clause types for history (CHET) and other scientific texts (RSC) within the top 30 typical patterns

(5) […] **who had it only from some common Reports then current**, *that he was* Strangled one day presently after Dinner with a Towel […].
(CHET. Tyrrell, 1704: 964 [122 (7906)])

(6) As to the precise time of Barscube's Descent, I cannot determine; **but this much I certainly know**, *that they were* a younger Son of that Noble Family.
(CHET. Crawfurd, 1710: 76 [124 (711)])

(7) Some of them **ran** *to extinguish* the fire, and others to repel the Scots […].
(CHET. Adams, 1795: 93 [140 (7055)])

(8) When the prospect was finally closed, men of spirit and energy **refused** *to* remain in a country where they were held unfit to receive the right of citizens.
(CHET. Killen, 1875: 241 [157 (DEL)])

Looking at the top 30 typical patterns and categorising them into phrase/clause types (see Figure 1), besides noun, verb, and adverbial phrases (maroon, brown and blue), we can see that CHET when compared to other scientific texts is particularly characterized by to-infinitives and that-clauses (as shown above) as well as by possessive noun phrases and subordinated or coordinated structures with conjunctions (beige). Phrases including possessives are a device used for back reference to persons/group of persons and the like, who are described in the texts (see

Example (9) and (10)). Conjunctions are used to elaborate on the content presented (see Example (11)).

(9) To consummate the adversity of the states, the assassination of their protector, the Prince of Orange, was perpetrated at Delf, by Balthasar Gerard, a Burgundian; worked up to the horrid villainy by a fanatical rage; and who, to testify *his* obduracy, amidst the tortures he suffered, held up *his* hand, burnt to a cinder, to make the sign of the cross.
(CHET. Anderson, 1775: 117 [136 (3061)])

(10) One of their spies had been taken and imprisoned within the place; and after he had set fire to *his* prison, he broke out of it, and running to the walls, called aloud to *his* countrymen to pursue *their* advantage.
(CHET. Adams, 1795: 93 [140 (7024)])

(11) […]; he resigned all the Temple-Lands in Scotland in Queen Mary's Hands, *and* obtained an Erection thereof into a Temporal Lordship, with the Dignity of Lord Torphichen, an. 1563. (CHET. Crawfurd, 1710: 76 [124 (576)])

In summary, while history texts are clearly different from general language by the use of a nominal style of writing, when compared to other scientific texts, they show typical phrase and clause types that indicate a stylistic type of writing serving the purpose of narration of historical events, persons involved, and the actions taken.

3.2 Productivity of typical patterns

In this section, we inspect the diachronic development of typical patterns of history texts as represented by the CHET corpus. In particular, we focus on whether these patterns become more or less productive over time, i.e. if they show more or less lexical variation. For this, we use surprisal as described in Section 2.2.2.

3.2.1 *Nominal patterns with prepositional phrases*

As shown above, particularly typical of history texts are prepositional phrases embedded in nominal phrases (proper/common noun (NP/NN) followed by a preposition (IN) followed by a proper/common noun (NP/NN): NounPrepositionNoun). We calculate surprisal (i.e. how predictable a word is based on its previous context) of each lexical item, *S*(*word*).[3] In addition, we consider the surprisal of each lexical realization of a pattern, *S*(*lexR*), i.e. surprisal of the three lexical items divided by three (as we consider sequences of three parts of speech). We then inspect the surprisal range of the whole pattern (i.e. S(NounPrepositionNoun)). For a better comparison of surprisal across time periods, we use quartiles dividing the range into three parts: high, middle, and low surprisal values (cf. Section 2.2.2). We inspect the range distribution over time to see whether the lexical items of the pattern belong to the high or low surprisal range. High surprisal points to high variation indicating productive use, low surprisal points to low variation and a more confined use to specific lexical items.

---

[3] I.e. surprisal of lexical realizations of NN/NP, surprisal of lexical realisations of IN, and surprisal of lexical realisations of NN/NP.

Figure 2 shows the percentage of surprisal ranges (high, middle, low) for the whole pattern across time periods for the CHET corpus. In 1700, around 46% of the lexical items belong to low, around 38% to high, and around 16% to middle surprisal. Towards 1850, the percentage of high surprisal increases, while the percentage of low surprisal decreases. Thus, while the pattern had a more restricted (conventionalized) use especially around 1750, over time the pattern becomes more productive, i.e. its lexical realisations show greater variation.

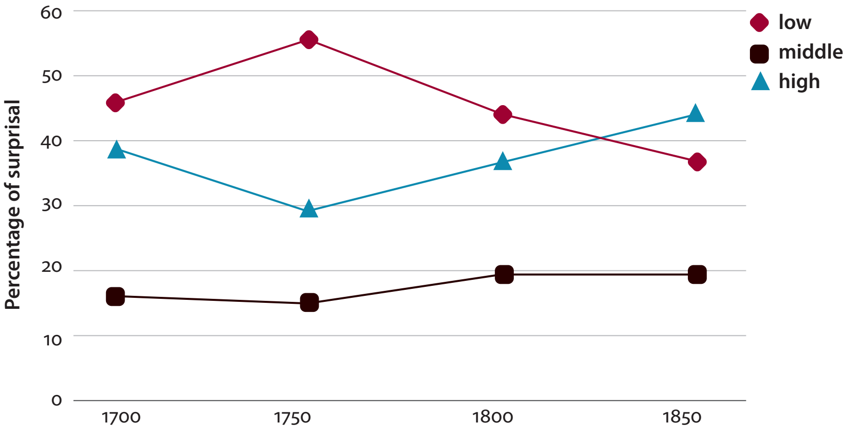


Figure 2. Percentage of surprisal range of the NounPrepositionNoun pattern for CHET

Comparing this to the same pattern in the RSC corpus (see Figure 3), we see a different tendency. While in the period of 1700 the percentage of high and low surprisal is similar (both around 40%), towards the 1850 period the percentage of high surprisal decreases (around 30%) and low surprisal increases (up to 50%). Thus, the use of this pattern within natural science texts becomes more conventionalised over time.

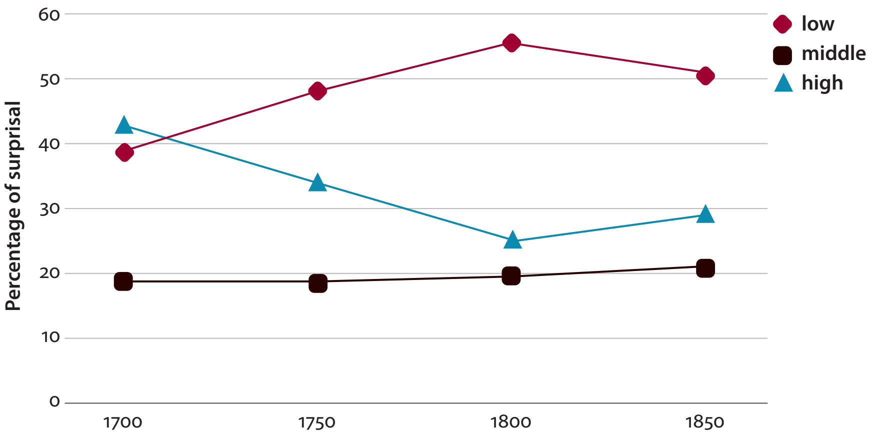


Figure 3. Percentage of surprisal range of the NounPrepositionNoun pattern for RSC

In the CLMET corpus (see Figure 4), surprisal of the pattern remains relatively stable over time, showing around 40% to 50% of surprisal for the low and high ranges. Thus, in general language the pattern does not show any remarkable diachronic changes in terms of lexical productivity.

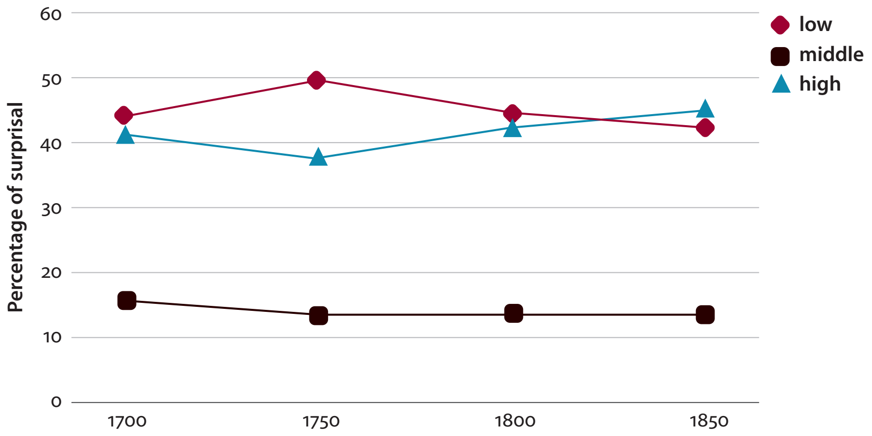


Figure 4. Percentage of surprisal range of the NounPrepositionNoun pattern for CLMET

Let us now compare the lexical realizations of the pattern for low surprisal values across corpora and time periods. Table 3 shows the lexical realisations for CHET, RSC and CLMET for 1700 and 1850. Focusing on 1700, we can see how the lexical realisations of the nominal pattern in CHET texts are dominated by *of*-genitive constructions in offices, titles and positions of persons, typically connecting a high rank title and a prominent toponym in multi-word concepts that function as semantic units (e.g. *King of England, Bishop of London*), while in the RSC and the CLMET a higher diversity of lexical items is used in the noun slots of this pattern. To a certain extent, CHET texts are characterized by stylistic devices such as synecdoche or metonymy (e.g. *Crown of England*) and resemble the CLMET texts in this aspect as they are also characterised by the use of stylistic devices and figurative language in contrast to natural science texts composed in the plain-style tradition of the Royal Society. The lexical items in these patterns in both specialised corpora reflect the main topics of the respective domains. On the one hand, these topics are related to the political history of the time and its recent past in the CHET corpus, i.e. power relations in England and Europe, monarchy and the struggle for supremacy as well as privileges of nobility and the church. On the other hand, in the RSC texts these patterns reveal common terms that are related to every-day materials and their properties, *of*-constructions as historical multi-word terms for chemical substances and compounds of major importance, e.g. *spirit of Wine* (alcohol) or *oil of vitriol* (sulphuric acid) and quantities and units of measurements (e.g. *quantity of Water/Air/Matter* or *grain of Sand, drop of Water*). In both the RSC and the CLMET texts, we find some rather general concepts implying human analysis and reflection. However, the scientific texts slightly seem to favour formulaic expressions relating to the sphere of facts (e.g. *matter of fact*), while in the general language texts the opinions of individuals seem to play a more important role in such phrasal patterns functioning as semantic units (e.g. *way of thinking).*

Comparing this to the 1850 period, all corpora share the same most general pattern *point of view* which almost functions as one single word. Here, a closer look into the data reveals that this term is embedded in different types of prepositional phrases or noun phrases in each corpus. In the RSC texts, it is often directly preceded by an adjective typical of scientific texts (*analytical, geological, physiological, chemical* etc.). It can also have a more literal meaning (e.g. *the eye being placed in a proper point of view*). In general language texts, it is sometimes preceded by evaluating adjectives or adjectives with social relevance such as *moral, religious, social, economical, national*. In many cases, the phrases it is embedded in seem to be shorter as in the RSC and are only preceded by a pronoun (e.g. *this, my, his, her*) referring to an individual (e.g. *Mr. Spencer's/my husband's/Erica's point of view*). In the CHET corpus, *point of view* sometimes refers to the perspective of different social actors or occurs in argumentative contexts (e.g. *The argument, looked at from a slightly different point of view, would seem convincing.*)

Considering diachronic tendencies (compare upper and lower part of Table 3), these seem different across corpora. CLMET does not show extremely remarkable changes. There are mainly rather general *of*-constructions in both sections. *sum of money* is still in the list in 1850, *glass of wine* has been replaced by *cup of tea* due to the increasing popularity of tea consumption. Interestingly, *man* and *life* have disappeared from the list that were prominent in 1700 (e.g. *man of honour/sense/war, time/course of life*) and specific names of institutions have occurred in the list (e.g. *Bank of England, House of Commons*) as well as general nouns as part of certain phrases (e.g. *sort/kind of thing*).

CHET is less characterized by official titles of people and events (e.g. *feast, death*) than before and shows a more diverse use of the nominal pattern with personal names in *of*-constructions that are part of phrases related to family relationships and the life of specific people (e.g. *reign of George* [*I.* or *II.*], *son of Richard* [*Wight, a Protestant clergyman, and grandson of Thomas Wight, who was also a clergyman*]). Concrete and abstract nouns in this pattern are still related to political concepts and power (e.g. *king, duke* and *throne, reign, act, parliament*). The third element in this pattern in CHET is still frequently a toponym (*Spain, Normandy, France* etc., but not *England* or *London* as in 1700). Also, there are more formulaic expressions and terms unrelated to the topic of the texts than before (*point of view, matter of course*). The RSC has shifted towards a stronger terminological usage and more specific nouns in concepts describing the physical world, but the *of*-construction is also dominant (e.g. *iodide of ethyl, axis of rotation, centre of gravity*). Scientific disciplines and factual knowledge have evolved, but some concepts remain important over time (common environmental substances such as *water* or *air*, quantities, measurements or changes of e.g. *temperature* or *time*). For some of these phrasal constructions there is a preference for shorter nominal compounds in Present-Day English (*silver nitrate*), but scientific texts from this period of lModE are already characterised by a certain preference for short terms (*nitrate of silver* being actually a nitrate of oxide of silver).

Considering the lexical realisations of the nominal pattern with high surprisal (see Table 4), we can see how these expressions differ quite strongly from those with low surprisal (compare with Table 3) for all three corpora. There is much more lexical variation for each part of speech in Table 3. Considering the surprisal profiles (see Figures 2 to 4 above), in CHET this more varied usage increases, while in RSC it decreases and in CLMET it remains relatively stable. Moreover, while in the RSC the pattern with high surprisal seems to be confined to the function of expressing locations in both time periods (e.g. *Scawby near Brigg, Walling near Aldermarston*), in CHET the use is much more varied (locations, time, occupation, etc.).

Table 3. Most frequent lexical realizations of the NounPrepositionNoun pattern from the low surprisal range

CHET			RSC			CLMET		
Realization	**Raw freq.**	**Mean $S(lexR)$**	**Realization**	**Raw freq.**	**Mean $S(lexR)$**	**Realization**	**Raw freq.**	**Mean $S(lexR)$**
				1700				
King of England	7	3.90	*spirit of Wine*	49	3.95	*way of thinking*	23	4.20
Bishop of London	5	4.88	*Degree of Heat*	42	3.54	*sum of money*	22	3.73
Crown of England	4	3.35	*space of Time*	35	2.99	*man of honour*	21	3.29
King of France	4	3.60	*quantity of Water*	28	2.17	*course of life*	20	4.57
Duke of Lancaster	4	3.73	*grain of Sand*	24	1.48	*time of life*	19	4.96
Duke of York	4	3.75	*quantity of Air oil*	21	2.14	*man of sense*	18	3.54
reign of King	4	4.68	*of vitriol quantity*	21	2.63	*Glass of Wine*	16	3.38
Duke of Ormond	4	5.98	*of Matter matter*	20	1.86	*peace of mind*	16	4.52
feast of St. death	3	3.43	*of fact drop of*	19	2.61	*man of War*	15	4.51
of William	3	4.46	*Water*	17	1.44	*part of mankind*	15	3.64
				1850				
point of view reign	5	4.40	*point of view quantity*	47	4.01	*point of view state*	66	3.42
of George duke of	4	4.20	*of air iodide of ethyl*	39	4.56	*of mind House of*	41	4.29
Normandy order	4	4.08	*quantity of water*	38	5.05	*Commons sort of*	37	4.16
of St.	4	4.40	*increase of temperature*	38	4.19	*thing Bank of*	33	3.45
King of Spain	3	3.86	*centre of gravity length*	35	5.35	*England sum of*	29	4.43
history of Ireland	3	4.20	*of time*	35	5.02	*money kind of thing*	28	4.29
matter of course	2	2.65	*axis of rotation*	34	4.22	*side by side*	26	3.90
son of Richard	2	4.20	*inch in diameter nitrate*	34	3.54	*cup of tea*	22	4.44
throne of France	2	4.00	*of silver*	33	3.66	*rate of interest*	21	3.71
Act of Parliament	2	3.65		32	4.20		21	5.00

Table 4. Top ten lexical realizations of the NounPrepositionNoun pattern from the high surprisal range

CHET		RSC		CLMET	
Realization	**S(*lexR*)**	**Realization**	**S(*lexR*)**	**Realization**	**S(*lexR*)**
		1700			
Shillings per Pound	14.00	*Scawby near Brigg*	17.33	*disproportion betwixt ourself*	18.33
knowledge between Arthur	13.67	*Walling near Aldermarston*	17.33	*taverns near temple-bar*	17.67
Edgar under Seal	13.67	*Caerleon upon Usk*	17.00	*bagnio near Bloomsbury*	17.00
drawing off kill	13.67	*Medway near Sheerness*	17.00	*coke upon lyttleton*	17.00
sterling into Fife	13.33	*Stanford near Lutterworth*	17.00	*Saxony against Sweden*	17.00
Business than Rice	13.00	*Edinburgh over-against Leith*	17.00	*ballance betwixt profit*	16.67
damage from Lightning	13.00	*Burgh under Stanemoor*	16.67	*bees-wax at Philadelphia*	16.67
drop by drop	13.00	*pulpit at Narborough*	16.67	*deference betwixt incredulity*	16.67
homage by Edgar	13.00	*Barton near Kettering*	16.67	*Stanislaus into Poland*	16.67
boundary unto Gaul	13.00	*Alessandro for Alexandro*	16.67	*Tunbridge until Tuesday*	16.67
		1850			
journey below Lake	14.33	*anthracite near Bideford*	18.67	*cephalopods besides orthoceras*	19.00
journey below Peoria	14.00	*Anthropoidea except Hapale*	18.67	*bastion after bastion*	18.00
distance below Peoria	13.67	*fortress near Malaga*	18.33	*ill-behaviour during service-time*	18.00
merchant from Bristol	13.33	*Eastbourne around Beachy*	18.33	*korosko via Murat*	18.00
gentry around Bandon	13.33	*Glenelg along Glen*	18.00	*macrurous before brachyurous*	18.00
press throughout British	13.33	*Crofton near Whitby*	17.67	*derry down derry*	18.00
absorption into Prussia	13.00	*escarpment above Rocca*	17.67	*ceyx among kingfishers*	17.67
December until March	13.00	*Spey near Laggan*	17.67	*convexity into concavity*	17.67
intercourse between Australia	13.00	*Chellaston near Derby*	17.67	*legend respecting robespierre*	17.67
precaution against danger	13.00	*tusk near Charmouth*	17.67	*Rotterdam through schiedam*	17.67

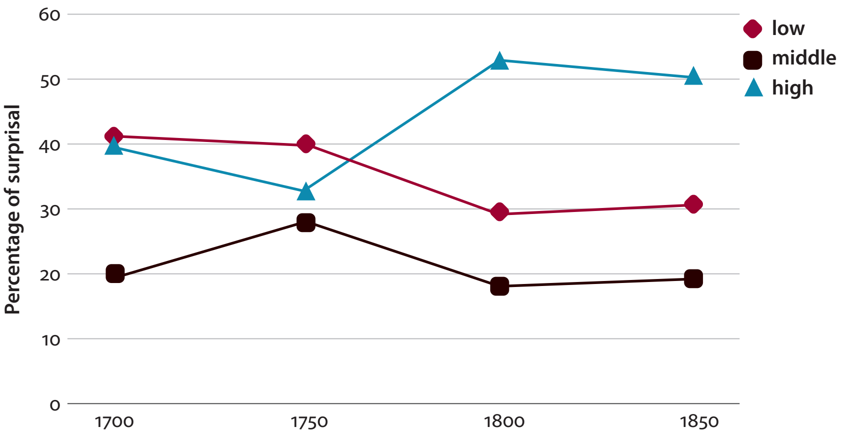


Figure 5. Percentage of surprisal range of the VVD TO VV pattern for CHET

In summary, this typical nominal pattern shows a different tendency for each corpus in terms of lexical productivity. While lexical productivity in CHET increases widening its usage to more variate lexical items, in RSC it decreases pointing to a more conventionalised usage of the pattern, and in CLMET it remains relatively stable over time.

3.2.2 To-*infinitive*

A second type of pattern prominently typical of history texts as depicted by the CHET corpus is the *to*-infinitive (cf. Section 3.1). The ranking obtained from the relative entropy value shows us that a very particular *to*-infinitive pattern is the most typical one: a *to*-infinitive preceded by a past tense verb (VVD.TO.VV pattern in Table 2). Based on this, we select this pattern and inspect the diachronic percentage distribution of surprisal ranges across time periods and corpora to inspect the pattern's productivity.

Considering the CHET corpus (see Figure 5), while the percentage of surprisal for high, middle and low ranges seems relatively stable in the first two time periods, towards 1850 the percentage of high surprisal increases up to 50%. This clearly indicates a more varied lexical usage of the pattern over time. If we compare this to the RSC and CLMET corpora, surprisal ranges are much more stable over time for this pattern (compare Figure 5 with Figures 6 and 7), the percentage of high and low surprisal ranges being relatively similar (around 40% to 45%).

Let us now consider the distribution of surprisal ranges for each single part of speech of the VVD.TO.VV pattern in the CHET corpus to see which part of speech is affected by higher variation contributing to a higher productivity of the pattern. Figure 8 shows the percentage of high, middle and low surprisal for each part of speech. For both verb forms (VVD and VV), from 1800 onwards, high surprisal increases, pointing to higher lexical variation over time. For comparison consider surprisal ranges of the *to*-particle were the distribution remains relatively stable over time. Clearly, the *to*-particle does not have any lexical variation. The slight changes are due to the fact that the *to*-particle might be more or less predictable given its previous context. This example shows nicely how surprisal captures differences between function and content words in terms of lexical productivity.

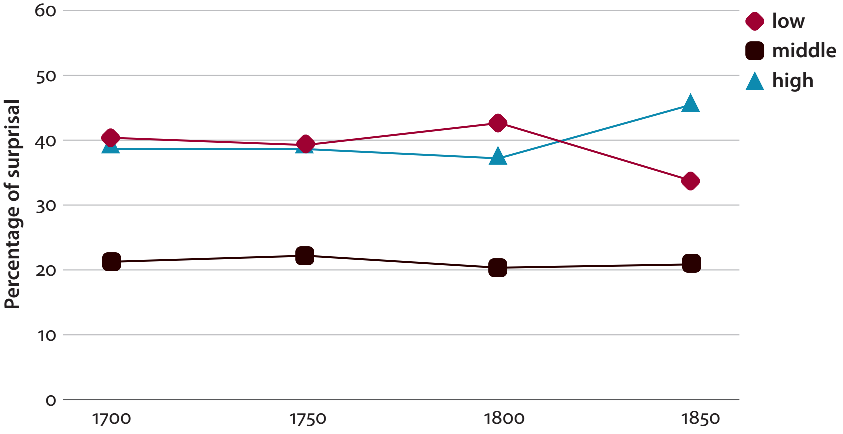


Figure 6. Percentage of surprisal range of the VVD TO VV pattern for RSC

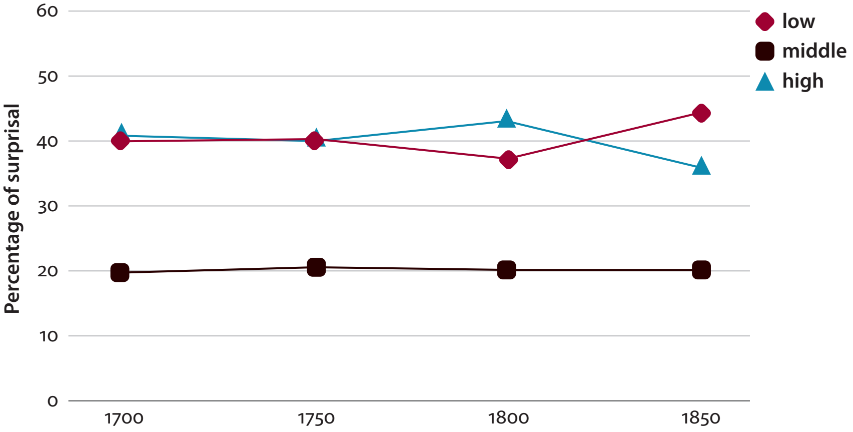


Figure 7. Percentage of surprisal range of the VVD TO VV pattern for CLMET

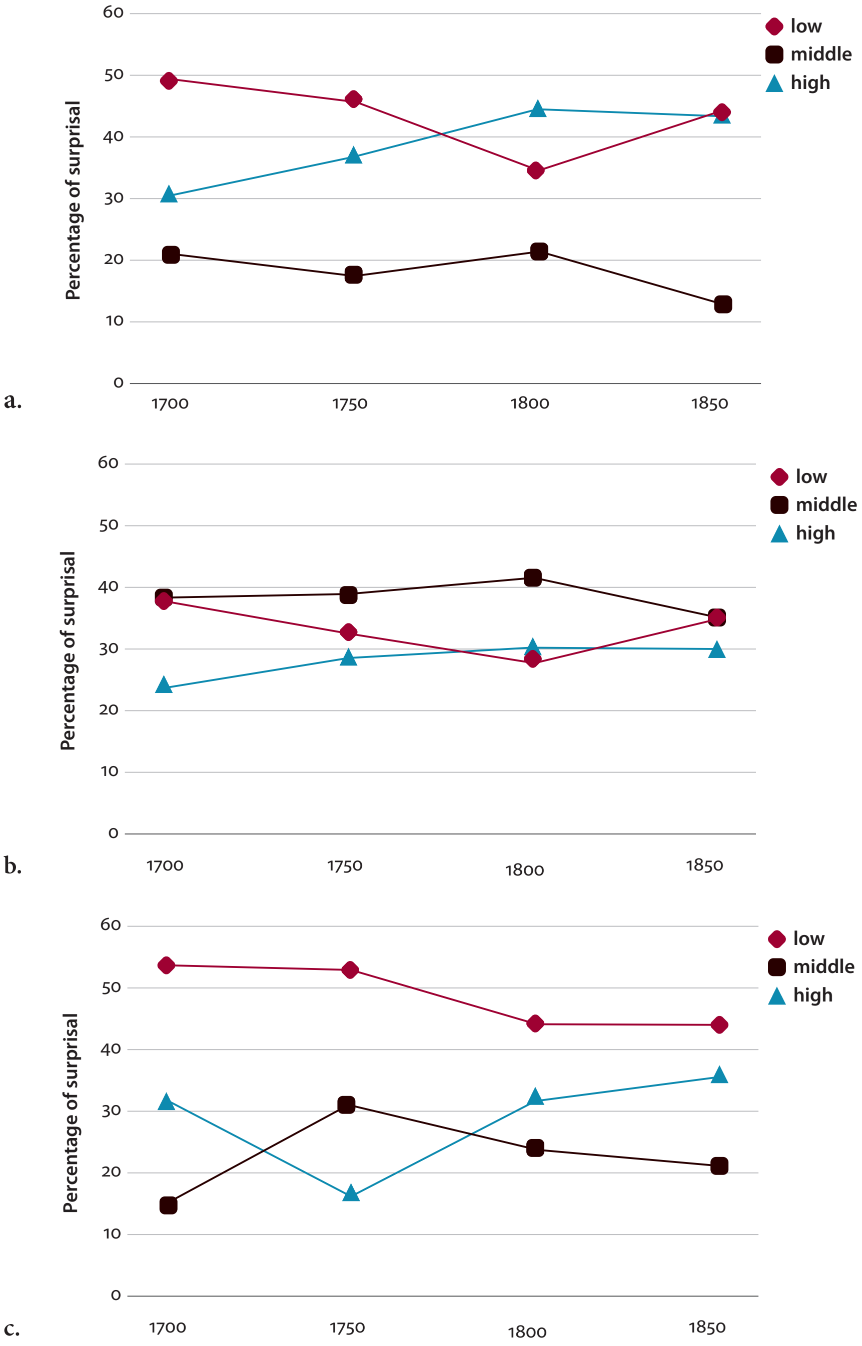


Figure 8. Percentage of surprisal range for each single part of speech of the VVD TO VV *to*-infinitive pattern in CHET

To inspect possible reasons for an increase in high surprisal for the past tense verb position (VVD), we look at its preceding context as it affects its predictability (i.e. its surprisal). Comparing the preceding context of the past tense verb (VVD) in the periods of 1800 and 1850, for low surprisal (i.e. high predictability of the verb) personal pronouns are dominant (65.57% mostly realized as *he*). For high surprisal of the past tense verb (i.e. low predictability of the verb), nominal phrases prevail (32.0% mostly realized as proper nouns such as *Socrates*), followed by personal pronouns (27.8% mostly realized as *they*) and various adverbs (19.6%, e.g. *painfully, invariably, gallantly*). Thus, while this *to*-infinitive pattern was mostly used in contexts preceded by a *he* subject, over time its productivity rises as it is used also in other more specific contexts (nominal phrases depicting specific persons or groups of persons referred to by plural personal pronouns). Most frequent lexical realizations of past tense verbs with low surprisal in the pattern used with *he* subjects are *refused, hoped, determined*, while with high surprisal and more specific subjects *menaced, authorized, concurred* are used.

Considering the development of the infinitive verb (VV), the top three verbs with low surprisal are general verbs (*make, take* and *give*), which remain the same over time. Infinitive verbs with high surprisal in this *to*-infinitive pattern instead are quite varied with realizations such as *to harangue* in *the archbishop tried to harangue* and *to discourse* in *Clement then commenced to discourse*. Interestingly, if the infinitive verb falls into the high surprisal range, in most cases the preceding past tense verb has also high surprisal.

Thus, the rise in productivity of this particular *to*-infinitive pattern seems to be motivated by the fact that the range of topic/subjects discussed is more varied over time. Also its frequency rises over time from 686.86 per million to 1305.73 per million doubling its frequency. This implies that as the *to*-infinitive, which serves the purpose of elaboration, is increasingly used over time, its lexical productivity rises.

4. Summary and discussion

We have presented a data-driven approach for the investigation of history texts as depicted by the CHET corpus for the lModE period. We pursued mainly two research goals: (1) identify typical phrase/clause patterns of CHET and (2) analyse their lexical productivity over time.

Instead of preselecting specific linguistic patterns to be analysed, we aimed at finding out from the data itself which phrase/clause patterns are uniquely used in the CHET texts when compared to other scientific texts (RSC corpus) and general language texts (CLMET corpus). To approximate phrase/clause patterns, we used part-of-speech trigrams, which have proved appropriate already in previous studies (Culpeper & Kytö, 2010; Kopaczyk, 2013; Degaetano-Ortlieb & Teich, 2016; Degaetano-Ortlieb *et al.*, 2018, Degaetano-Ortlieb & Teich, 2018; Degaetano-Ortlieb & Teich, 2019). For detection of typical patterns, we have used relative entropy, which enabled us to directly measure typicality of part-of-speech trigrams for one corpus compared to another (cf. Section 2.2.1). In comparison to general language, history texts prominently use nominal patterns in combination with prepositional phrases relating to persons, locations and the like. This confirmed our assumption that history texts would resemble a more informational writing style when compared to general language texts, which reflect a more verbal and involved style. The same nominal patterns are also typical of CHET when compared to other scientific texts (mainly from the natural sciences). This confirms a relatively unique use of these nominal patterns by CHET. In addition, comparison with other scientific texts has also shown particular clause patterns to be typical of history texts: in particular *to*-infinitives and *that*-clauses. Both clause types function as a narrative device for elaboration of the informational content presented. In addition, also possessive and conjunctive patterns where shown to be typical of CHET (cf. Section 3.1).

Lexical productivity was analysed for the most typical nominal and *to*-infinitive patterns. To measure lexical productivity, we used surprisal inspecting how predictable a lexical item is based on its previous context (cf. Section 2.2.2). Low surprisal indicates low productivity (i.e. a more conventionalised writing style and high predictability of a lexical item), while high surprisal indicates high productivity (i.e. more lexical variation and low predictability of a lexical item).

Considering the nominal pattern (NounPrepositionNoun), diachronically for CHET lexical productivity goes up as the percentage of high surprisal increases for this pattern, while the percentage of low surprisal decreases. While around 1700 *of*-genitive constructions prevailed denoting e.g. titles and positions of persons, over time the variation of lexical items used in the pattern increases allowing also a more varied usage denoting locations, time, occupation, family relationship, etc. (cf. Section 3.2.1). This tendency was not observed for either scientific or general language texts. For scientific texts (RSC) a more conventionalised usage is observed diachronically, especially due to consolidation of terminology in the investigated period. For general language (CLMET), lexical productivity of this nominal pattern remained relatively stable over time.

For the *to*-infinitive pattern (verb-past-tense *to* verb-infinitive), a similar tendency to the nominal pattern can be observed: lexical productivity rises considerably over time. A deeper analysis of the productivity of each part of speech of this particular *to*-infinitive pattern showed that the rise in productivity of both verbal forms seems to be related to a more varied usage of subjects preceding the pattern (cf.

Section 3.2.2). Again, this tendency was not visible in scientific or general language texts, as both remained relatively stable in terms of lexical productivity over time.

Overall, we have shown which patterns are typical of History texts using the CHET corpus as well as a rise in lexical productivity due to a more varied usage of the typical patterns investigated.

While here we have focused on history texts, our methodology can be applied to any type of comparative analysis envisaging an exploratory approach (see e.g. Degaetano-Ortlieb & Teich, 2016, 2017; Degaetano-Ortlieb *et al.*, 2018). For example, regarding diachronic linguistic change, latent linguistic patterns of change can be detected from the data itself and linguistic patterns involved in genre/register/text type/intra-textual variation could be inspected (cf. Degaetano-Ortlieb & Teich 2018, 2019).

References

Arnaud, Ren. (1998). The Development of the Progressive in 19th Century English: A Quantitative Survey. *Language Variation and Change*, 10, 123–52. https://doi.org/10.1017/S0954394500001265

Biber, Douglas (Ed.) (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan. (1999). *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Coffin, Caroline. (1997). Constructing and Giving Value to the Past: An Investigation into Secondary School History. In Coffin, C. *Genre and Institutions: Social Processes in the Workplace and School* (196–229). London, New York: Continuum Studies in Language and Education. Continuum.

Crespo, Begoña & Isabel Moskowich. (2015). A Corpus of History Texts (CHET) as Part of the Coruña Corpus Project. *Proceedings of Corpus Linguistics* (14–23). St. Petersburg, Russia.

Culpeper, Jonathan & Merja Kytö (Eds.) (2010). *Early Modern English Dialogues: Spoken Interaction as Writing*. Studies in English Language. Cambridge: Cambridge University Press.

De Smet, Hendrik, Susanne Flach, Jukka Tyrkkö, & Hans-Jürgen Diller. (2015). Corpus of Late Modern English Texts (version 3.1). In *Functional and Cognitive Linguistics: Grammar and Typology (FunC)*. KU Leuven. Accessed [16/11/2017] from https://perswww.kuleuven.be/u0044428/clmet30.htm

Degaetano-Ortlieb, Stefania & Elke Teich. (2016). Information-based Modeling of Diachronic Linguistic Change: From Typicality to Productivity. In *Proceedings of the 10th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities at ACL* (165–173). Berlin, Germany. Accessed [16/11/2017] from http://aclanthology.coli.uni-saarland.de/pdf/W/W16/W16-2121.pdf. https://doi.org/10.18653/v1/W16-2121

Degaetano-Ortlieb, Stefania & Elke Teich. (2017). Modeling Intra-textual Variation with Entropy and Surprisal: Topical vs. Stylistic Patterns. *Proceedings of the Joint LaTeCH and CLfL Workshop at ACL* (68–77). Vancouver, Canada. Accessed [16/11/2017] from http://www.aclweb.org/anthology/W17-2209

Degaetano-Ortlieb, Stefania & Elke Teich. (2018). Using Relative Entropy for Detection and Analysis of Periods of Diachronic Linguistic Change. In *Proceedings of the 2nd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences*, Humanities and Literature at COLING2018. Santa Fe, NM, USA. ACL. Accessed [08/07/2019] from https://aclweb.org/anthology/papers/W/W18/W18-4503/

Degaetano-Ortlieb, Stefania & Elke Teich. (2019). Towards an Optimal Code for Communication: The Case of Scientific English. *Corpus Linguistics and Linguistic Theory*, open access, pp. 1–33.

Degaetano-Ortlieb, Stefania, Hannah Kermes, Ashraf Khamis, & Elke Teich. (2018). An Information-Theoretic Approach to Modeling Diachronic Change in Scientific English. In Carla Suhr, Terttu Nevalainen, & Irma Taavitsainen (Eds.), *From Data to Evidence in English Language Research* (258–281). Leiden: Brill.

Dossena, Marina & Charles Jones (Eds.) (2003). *Insights into Late Modern English*. Bern: Peter Lang.

Dossena, Marina & Ingrid Tieken-Boon van Ostade (Eds.) (2008). *Studies in Late Modern English Correspondence: Methodology and Data*. Bern: Peter Lang.

Evert, Stefan & Andrew Hardie. (2011). *Twenty-first Century Corpus Workbench: Updating a Query Architecture for the New Millennium*. Accessed [16/11/2017] from http://eprints.lancs.ac.uk/62721/

Görlach, Manfred. (1998). *An Annotated Bibliography of Nineteenth-Century Grammars of English*. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/lisl.26

Görlach, Manfred. (1999). *English in Nineteenth-Century England: An Introduction*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511627828

Görlach, Manfred. (2001). *Eighteenth-Century English*. Heidelberg: Universita¨tsverlag C. Winter.

Halliday, M. A. K. (1988). On the Language of Physical Science. In Mohsen Ghadessy (Ed.), *Registers of Written English: Situational Factors and Linguistic Features* (162–177). London: Pinter.

Hundt, Marianne (Ed.) (2014). *Late Modern English Syntax*. Studies in English Language. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139507226

Kermes, Hannah, Stefania Degaetano-Ortlieb, Ashraf Khamis, Jörg Knappen, & Elke Teich. (2016). The Royal Society Corpus: From Uncharted Data to Corpus. In *Proceedings of the Language Resources and Evaluation Conference (LREC)* (1928–1931). Portorož, Slovenia.

Kopaczyk, Joanna (Ed.) (2013). *The Legal Language of Scottish Burghs: Standardization and Lexical Bundles*. Oxford Studies in Language and Law. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199945153.001.0001

Kullback, Solomon & Richard A. Leibler. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86. https://doi.org/10.1214/aoms/1177729694

Kytö, Merja, Mats Rydén & Erik Smitterberg (Eds.) (2006). *Nineteenth-Century English: Stability and Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486944

Monaco, Leida Maria. (2016). Was Late Modern English Scientific Writing Impersonal? Comparing Philosophy and Life Sciences Texts from the Coruña Corpus. *International Journal of Corpus Linguistics*, 21(4), 499–526. https://doi.org/10.1075/ijcl.21.4.03mon

Moskowich, Isabel. (2017). Pronouns as Stance Markers in the Coruña Corpus: An Analysis of the *CETA*, *CEPhiT* and *CHET*. In Francisco Alonso-Almeida (Ed.), *Stancetaking in Late Modern English Scientific Writing. Evidence from the Coruña Corpus* (73–91). Colección Scientia [Applied Linguistics]. Valencia: Servicio de Publicaciones de la Universidad Politécnica de Valencia.

Osselton, Noel Edward. (1985). Spelling-Book Rules and the Capitalization of Nouns in the Seventeenth and Eighteenth Centuries. *Historical & editorial studies in Medieval & Early Modern English for Johan Gerritsen* (49–61).

Pérez-Guerra, Javier, Dolores González-Álvarez, Jorge L. Bueno Alonso, & Esperanza Rama-Martínez (Eds.) (2007). *'Of Varying Language and Opposing Creed' New Insights into Late Modern English*. Bern Peter Lang.

Schmid, Helmut. (1995). Improvements in Part-of-Speech Tagging with an Application to German. In *Proceedings of the ACL SIGDAT-Workshop*, Kyoto, Japan.

Stubbs, Michael & Isabel Barth. (2003). Using Recurrent Phrases as Text-type Discriminators: A Quantitative Method and some Findings. *Functions in Language*, 10(1), 61–104. https://doi.org/10.1075/fol.10.1.04stu

Tieken-Boon van Ostade, Ingrid (Ed.) (2008). *Grammars, Grammarians and Grammar-writing in Eighteenth-century England*. Berlin: De Gruyter. https://doi.org/10.1515/9783110199185

Tieken-Boon van Ostade, Ingrid. (2009). *An Introduction to Late Modern English*. Edinburgh: Edinburgh University Press.

Tieken-Boon van Ostade, Ingrid & Wim van der Wurff (Eds.) (2009). *Current Issues in Late Modern English*. Number 77 in Linguistic Insights. Studies in Language and Communication. Bern: Peter Lang.

Werlich, Egon (Ed.) (1975). *Typologie der Texte*. Wiebelsheim: Quelle and Meyer.

Werlich, Egon (Ed.) (1983). *A Text Grammar of English*. Wiebelsheim: Quelle and Meyer.

CHAPTER 5

Exploring the narrative dimension in late Modern English History texts

Leida Maria Monaco
Universidade da Coruña

> Ask a historian to explain something and they will usually tell you a story.
> (Geoffrey Roberts, 2001. *The History and Narrative Reader*)

1. Introduction

Whereas the academic register in general is often described as highly structured and rather impersonal (Hyland, 1998; Gotti, 2001), the historical register is usually characterised as having a narrative character. History tends to be identified, first and foremost, with the narration of past events (Van Seters, 1997; White, 1984, 1987; Munslow, 2007; Alfaya, this volume), and in many ways the work of a historian can be compared to that of a storyteller (Mandelbaum, 1967: 414), especially in terms of the linguistic characterisation of the historical register. However, whereas narration "is a manner of speaking as universal as language itself, and narrative is a mode of representation so natural to human consciousness that to suggest it is a problem might well appear pedantic" (White, 1984: 1), the place of narration in historical writing has been an intensely debated topic since the late nineteenth century.

Benedetto Croce, for instance, considered that narrative was essential to history, and that a good historian had first of all to be a good narrator (Croce, 1893 [2017]), this being a commonly shared idea from the time of Herodotus to the first half of the twentieth century (Marenco, 2009). However, Croce argued that narratives are only "alive" if supported by the corresponding "living documents", otherwise they are "nothing but a complex of empty words or formulas asserted by an act of the will" (Croce, 1921: 17–18). By the second half of the 1900s, theorists of historiography such as Gallie (1964), Dray (1971), Mink (1970) and White (1973) defended the importance of narrative in historical writing and were sometimes criticised for giving "too literary" a view of "a discipline that sought to be objective and scientific" (Carr, 1991: 7). For Louch (1969: 55–56), narrative "stands proxy for

https://doi.org/10.1075/z.225.05mon

© 2019 John Benjamins Publishing Company

experience", filling the gap between "discontinuous and radically different states: bud then flower, seed then sprout, cocoon then moth". Thus, when extended to the grand sweep of history, narration is the means of explaining how a local political order becomes nationalist, how a feudal economy becomes capitalist, how an agricultural society shifts to become an industrial one or how a sacred way of life turns secular (Louch, 1969: 55). Strictly speaking, therefore, narrative can be defined as "the organization of material in a chronologically sequential order and the focusing of the content into a single coherent story, albeit with subplots" (Stone, 1979: 3); in the case of historical writing, then, "transforming the past into history" (Rüsen, 2005: 2).

On the other hand, if viewed as a form of discourse, narration may serve a number of other functions. For instance, its purpose in a historical account is not necessarily restricted to telling a story, but may also be used to describe a situation or to analyse a historical process (White, 1984, 1987). Biber (1988) identified narration as one of the basic dimensions of the English language, with some registers, like fiction and history, being essentially narrative, yet that certain other registers also present narrative characteristics, as seen in their pervasive linguistic features. Looking specifically at scholarly prose, Conrad (1996, 2001) studied the narrative component in biology and history research articles and textbooks. She showed that whereas history textbooks and articles were clearly more narrative than those from biology, both disciplines employed narration, albeit with different purposes: the former to tell a story, the latter to describe evidence and procedures. Gray's (2011) more recent study shows that disciplines in the humanities (other than history), such as applied linguistics, political science and theoretical philosophy, use narration more frequently than the so-called hard sciences, such as biology or physics. Again, this appears to be linked to the way in which evidence is presented in different disciplines, in that writing in the humanities tends to resort to narration to describe what has happened during a research process or a particular situation, whereas in the hard sciences the passive voice tends to be used to describe both procedures and scientific phenomena (Gray, 2011: 147–152).

However, another recent diachronic study of register variation in English scientific discourse (Monaco, 2017) has shown that narrativity was not a common feature of philosophical writing in the eighteenth century, which at that time tended to deal with "timeless" questions such as morality, marriage or the existence of God, all of these being discussed mainly in the present tense. In fact, in that study both philosophy and the two other scientific disciplines analysed – astronomy and life sciences – appeared to become more narrative in the nineteenth century. This might be related to the fact that several of the genres used in those disciplines at the time, treatises and letters, were widely used for the popularisation of knowledge for a mass readership, and several genres were also intended specifically for a female audience. Thus, telling a story may well have been seen as a means of making science more understandable to a lay audience. This clearly contrasts to earlier findings in Atkinson (1999), which show how scientific prose gradually lost narrative features throughout the eighteenth and nineteenth centuries; the reason for this apparent contradiction must surely lie in the question of which sub-registers are used in an analysis (see Biber & Gray, 2013); in the case of Atkinson (1999), these were research articles published in the *Philosophical Transactions* of the Royal Society of London, and therefore were not aimed at a non-specialised public.

All of the studies mentioned above (Biber, 1988; Conrad, 1996, 2001; Atkinson, 1999; Gray, 2011; Monaco, 2017) use Multidimensional Analysis, developed by Biber (1988) in an attempt to provide a linguistic description for the different registers of the English language, and which would be subsequently applied to other languages and/or registers, resulting in models of variation specific to those languages or registers. In this chapter, our aim is to analyse narrative features of texts from the eighteenth-century part of the beta version of *Corpus of History English Texts*, using the same model of dimensional variation used in a previous analysis (Monaco, 2017) of the three other linguistic disciplines currently available in the *Coruña Corpus* (astronomy, philosophy and life sciences). Section 2 gives a brief account of the theoretical-methodological framework of Multidimensional Analysis, and sets out the linguistic features under investigation, while Section 3 offers a description of the corpus and methodology used. The results are presented and discussed in Section 4, and some concluding remarks are offered in Section 5.

2. The multidimensional analysis and the "narrative" dimension

In an attempt to describe the differences between speech and writing, Biber (1988) undertook a comprehensive corpus study of both written and spoken Present-Day English, known as Multidimensional (hereafter MD) Analysis. His methodology entailed using factor analysis, a multivariate statistical technique of data reduction, on a large set of lexical and semantic variables (that is, linguistic features) retrieved from a substantial corpus of texts. Factor analysis groups several linguistic features into factors, showing how such linguistic features tend to appear together in some texts because they share an underlying communicative function. Also, each factor usually has two sets of features in binary opposition, called positive and negative, reflecting opposite ends of a communicative function (e.g. involved vs. informational; abstract vs. non-abstract). Each factor is thus interpreted as a dimension of variation with regard to the communicative function it reveals, and each text in turn has a factor loading, positive or negative, according to its individual co-occurrence patterns, and can be characterised with respect to each dimension of variation as more or less informational, more or less abstract, etc. In this pioneering MD study, five clear dimensions of variation were identified: (1) Informational vs. involved production; (2) Narrative vs. non-narrative concerns; (3) Explicit vs. situation-dependent reference; (4) Overt expression of persuasion, and (5) Abstract vs. non-abstract style.

Over the past thirty years, many studies have replicated Biber's (1988) original MD analysis on both general corpora (Lee, 1999) and more specialised ones (Atkinson, 1999; Biber, 2001; Xiao, 2009; Gray, 2011; Grieve, 2016), in order to trace variation across and also within registers. Despite the fact that each register usually reveals different variation patterns, many of these studies have shown that, while some dimensions are characteristic of a particular language or a particular discourse domain, "two linguistically similar parameters of variation have emerged repeatedly across MD studies: a basic oral/literate parameter of variation, and a narrative/non-narrative dimension" (Biber, 2016: 16). Monaco (2017) analysed variation across three late Modern English scientific registers in three subcorpora of the *Coruña Corpus* – those of Astronomy in CETA (Moskowich & Crespo, 2012), Philosophy in CEPhiT (Moskowich *et al.*, 2016) and Life Sciences in CELiST (in beta version at the time; see Lareo, 2011). Four dimensions of variation were found: (1) "Involved/persuasive vs. impersonal style", (2) "Argumentative vs. descriptive focus", (3) "Elaborate vs. non-elaborate discourse", and (4) "Narrative vs. non-narrative discourse", revealing once more the presence of the two constant parameters (oral/literate and narrative/non-narrative) in dimensions 1 and 4, respectively. In the current study, we will ask to what extent the narrative dimension – Dimension 4 – is present in the beta version of CHET (the eighteenth-century section) as a means of clarifying whether the history subcorpus of the *Coruña Corpus* is actually more narrative (as we might expect it to be) than the other three subcorpora here.

Monaco's (2017) Dimension 4 "Narrative vs. non-narrative discourse" contains two opposing sets of linguistic features.[1] Those with strong positive loadings are time adverbs, perfect aspect and past tense (these three being clear markers of narration) and, with slightly weaker positive loadings, general adverbs, the *seem*/*appear* copula, place adverbs, downtoners and hedges, all of which also tend to appear in narrative texts. The opposite end of the dimension – non-narrative – contains features with negative loadings: *be* as main verb and present tense, as well as phrasal coordination. Illustration of each linguistic feature is given in Examples (1)–(13), all of which have been retrieved from CHET:

---

[1] The linguistic features searched for in Monaco (2017), and hence also in the present study, are those originally analysed in Biber (1988). Closed lists of time, place and general adverbs, hedges and downtoners can be found in Biber (1988: 235–239).

Positive linguistic features

- Time adverbs

(1) and for the better Procurement thereof, the King *now* sent the Duke of Gloucester over (CHET, Tyrrell, 1704: 952 [122 (1135)])

- Perfect aspect

(2) The first hints that William and Mary *had ascended* the throne, aroused them to spirited measures (CHET, Adams, 1770: 27 [135 (5667)])

- Past tense

(3) We know not what effect this prelate's boldness *produced*, farther than that Edward *grew* very cautious in his proceedings (CHET, Adams, 1795: 80 [140 (1047)])

- General adverbs

(4) It was recommended under the name of the Grand Scheme; and was *certainly* the favourite one of Philip II (CHET, Anderson, 1775: 123 [136 (4804)])

- Seem/appear

(5) He always kept his heart open, and never *appears* to have formed a design of amassing any particular sum (CHET, Cornish, 1780: 24 [137 (3926)])

(6) This treaty *seems* to have been the foundation of the many ruinous connections into which the Scots afterwards entered with France (CHET, Adams, 1795: 91 [140 (6086)])

- Place adverbs

(7) he was sent *away* about seven in the evening, on a poor jade ill accoutred, to travel all night with the French (CHET, Stock, 1800: 50 [141 (4550)])

- Downtoners

(8) This was *partly* owing to the hereditary hatred which had long subsisted between the Normans and Anjouvins (CHET, Gifford, 1790: 191 [139 (2153)])

- Hedges

(9) It was an high pyramidal cap of cloth or silk, *almost* concealed by a profusion of pearls and jewels (CHET, Gibbon, 1788: 105 [138 (7293)])

Negative linguistic features

- *be* as main verb

(10) The constitution of a people, in this and the like cases, *is* capable of entire freedom, nay, not capable of any other settlement (CHET, Chapman, 1750: 8 [132 (DEL)])

- Present tense

(11) In this Truce *there is not* any mention of any Homage, Liege-Sovereignty or Ressort reserv'd to the French King (CHET, Tyrrell, 1704 [122 (2551)])

- Phrasal coordination

(12) From this the River of Greif hath its Course some short way Eastward, till meeting at the Church of Inchennan with White-Cart, they *mix and empty* themselves into the River of Clyde, a little below that Church (CHET, Crawfurd, 1710: 955 [124 (236)])

(13) for, as their greatest strength lay in their *horse and elephants*, to abandon the plain country, and post themselves in *high and steep* places (CHET, Hooke, 1745: 35 [131 (3782)])

As can be appreciated in the examples above, some of these linguistic features tend to co-occur, which shows – at least, up to a point – their mutually dependent character, although we will see in Section 4 that those with weaker loadings tend to appear with both positive and negative features, more or less indiscriminately. The retrieval of each of these linguistic features from the corpus is described in the next section.

3. Corpus and methodology

Only the eighteenth-century part of CHET (which contains a total of twenty texts) has been dealt with in the present study. The linguistic features discussed in the previous section were searched for with the Coruña Corpus Tool (CCT; see Moskowich, this volume; also Camiña & Lareo, 2016) and the Saarbrücken CQPweb interface (http://corpora.clarin-d.uni-saarland.de/cqpweb/coruna_chet/; see Degaetano *et al.*, this volume; Hardie, 2012). Raw frequencies of each linguistic feature in each text were normalised per 1,000 words, then standardised to a mean score of 0.0 and a standard deviation of 1.0, the resulting scores thus being comparable with those in the reference corpus (Monaco, 2017), using the following formula[2] (see Biber & Finegan, 2001: 111–112):

$$z = (x - \mu)/\sigma$$

---

[2] A relative standard score for a feature is obtained by subtracting the mean frequency for that feature in the reference corpus from the normalised frequency of that feature in the target corpus, divided by its standard deviation, again, in the reference corpus.

The resulting standard scores (also called z-scores) were then combined to form dimension scores for each text in the following way: scores for all the positive features and those for the negative features were added together separately, after which the negative features were deducted from the positive ones. This resulted in individual dimension scores for each text, which were then plotted along the dimension in question (here Dimension 4, as shown in Section 4). A positive dimension score for a text indicates that it contains more narrative features than the average for the reference corpus, whereas a negative score characterises it as relatively non-narrative. However, in this study we want to look at variation across genres, rather than within individual texts, and hence a dimension score was calculated for each genre by summing the dimension scores for all the texts in a genre, then dividing by the number of texts. Likewise, an overall dimension score for eighteenth-century CHET has been calculated for each period of 50 and 25 years between 1700 and 1800 so that any diachronic progression could be appreciated.

The results are discussed in what follows.

4. Analysis of data

4.1 Variation across time and disciplines

Figure 1 plots Dimension 4 scores for CHET and three other subcorpora from the *Coruña Corpus* (CETA, CEPhiT and CELiST) in the first and second halves of the eighteenth century, with the aim of describing the register of History in contrast to those of Astronomy, Philosophy and Life Sciences in the same period. Positive scores (i.e. above zero) indicate the presence of narrative features in a given register with respect to the mean (i.e. zero, which represents the average content of narrativity in the total reference corpus). Negative scores, on the contrary, indicate a relatively low frequency of narrative features – and subsequently a higher presence of non-narrative ones – relative to the mean.

First, as can be seen in Figure 1, CHET presents a far higher narrative score than the other subcorpora for both halves of the century, clear evidence that the History register was already notably more narrative than the other three registers included here by the eighteenth century. Diachronically, CHET varies very little, although moving in a more ‘narrative’ direction, from a positive score of 4.5 to 5.0. Philosophy, conversely, begins at the relatively high negative score of −3.1 and goes down to −3.3, thus appearing to become even less narrative over the course of the century. As noted in the Introduction, eighteenth-century philosophers resorted to narration very rarely, this due largely to the fact that their subject matter tended to be of a universal nature and not restricted by time (topics such as morality, causality,

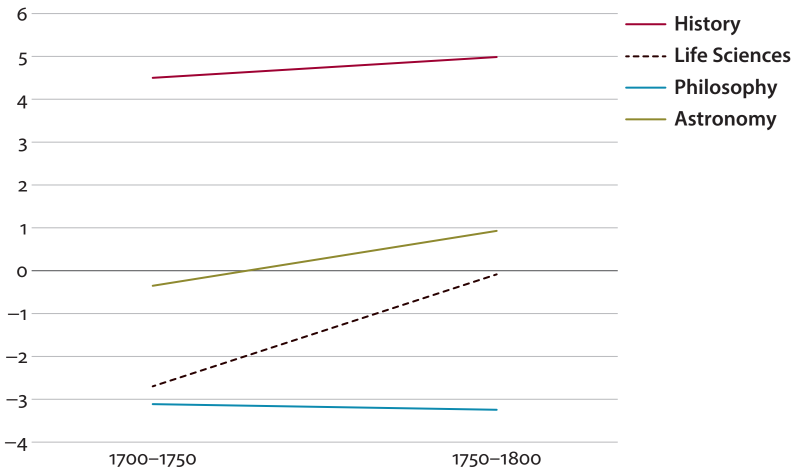


Figure 1. Dimension 4 scores for the first and second halves of eighteenth-century CHET with respect to three other scientific registers from the *Coruña Corpus*

God, etc.), and therefore tended to be dealt with in the present tense. Conversely, an abundance of negative features (all of them *italicised*) in eighteenth-century CEPhiT can be appreciated in Example (14):

> (14) Virtue, of all Objects, *is* the most *valuable and lovely*; and accordingly this Species of Philosophers *paint* her in the most amiable Colours, borrowing all Helps from *Poetry and Eloquence*, and treating their Subject in an *easy and obvious* Manner, such as *is* best fitted to please the Imagination, and engage the Affections (CEPhiT, Hume, 1748: 1 [91 (90)])

The Life Sciences, in turn, gain narrative features in the second half of the eighteenth century, attaining an overall score close to the reference corpus mean. Astronomy, which also appears to present a more narrative discourse in the second half of the century, does so less sharply, starting with an overall score only slightly below the mean (–0.3) and moving close to a positive score of 1.0. Although both eighteenth-century Astronomy and Life Sciences are seen here to be less narrative than History writing, it has been suggested (Monaco, 2017: 215) that the relatively higher occurrence of narrative features in the discourse of these two disciplines appears to reflect the lasting importance of experimental reports in the natural sciences, at least in some of the genres used here, although it has also been shown that the experimental article was gradually losing its narrative component during this period (Gotti, 2001).

In order to complement Figure 1 and show a more gradual diachronic evolution, Figure 2 below presents a progression of Dimension 4 scores for the same four subcorpora of the *Coruña Corpus* – CHET, CETA, CEPhiT and CELiST – through timespans of twenty-five years. Whereas thirty-year timespans are considered to be more reliable in tracking language change (Kytö *et al.*, 2000), we have resorted to periods of twenty-five years in order to capture maximum variation over the century.

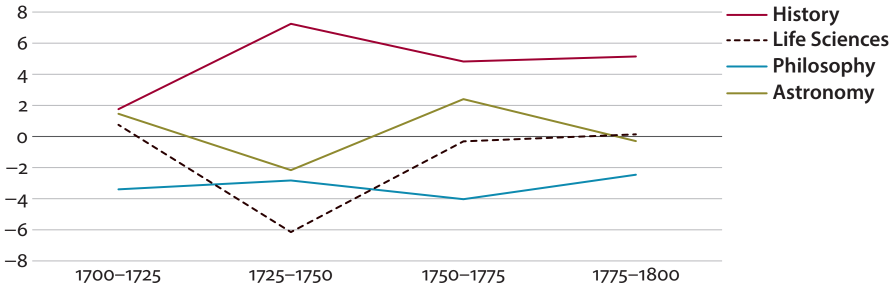


Figure 2. Diachronic evolution of eighteenth-century CHET on Dimension 4 "Narrative vs. non-narrative discourse" over four 25-year periods, compared to three other disciplines from the *Coruña Corpus*

Contrary to Figure 1, Figure 2 shows a far less pronounced difference between CHET and the rest of the subcorpora, with the exception of the period 1725–1750. In fact, at the beginning of the century, CHET has a narrative score only 1 unit higher than Life Sciences and less than 0.5 higher than Astronomy, these three disciplines remaining close to the mean. However, unlike Philosophy, which maintains a relatively low narrative score throughout the century, the other disciplines vary in this respect over time. Astronomy and Life Sciences, especially the latter, present a sharp descent in the first half of the century, whereas History moves in the opposite direction, gaining five units here. Such differences are illustrated in the following text excerpts (henceforth, positive features will be marked in **bold**, and negative ones in *italics*):

(15) Hence it *is* that all the Constellations **have changed** their Places; thus, the Constellation of Aries, that in Hipparchus's Time, **was** near the Vernal Equinox, *is* **now** removed a whole Sign towards the East, and *is* got into the Sign or Portion of the Ecliptic called Taurus, and so every Constellation, **has**, since the first Observation **changed** Place with the following: But the Portions of the Ecliptic *retain* still the same Names which they **had** in the Time of Hipparchus (CETA, Gordon, 1726: 69 [4 (1074)])

(16) ANOTHER dark-dealer the publisher of the late Plagiarie Treatise, entitled, an impartial account of the Affairs of Scotland, &c. to make it more current, **has thrown in** into the Title page, Some remarkable Instances that *may* give light into the Dependency of Scotland, on the Crown of England; which *is* mere juggle: For in this Book, *there's* nothing has the least Tendency to Homage; Only *there's* a passing hint of a Rumor spread abroad by the Partisans…
(CHET, Anderson, 1705: 101 [123 (9803)])

(17) Abinthium vulgar. It *grows* to be three Foot high; the Stalks *are* hoary full of a white Pith, ye Leaves are a Willow green **above**, & a light hoary **underneath**; the Flowers *are* yellow. It *grows* in Lanes and waste Places, and *flowers* in *July and August*. The *Leaves & Tops are* used; they *purge* Melancholy Humours, *provoke* Urine, *restore* an Appetite that is lost by Drinking
(CELiST, Blackwell, 1737: 5 [48 (1707)])

(18) December 25th about sixty **laid** siege to St. Georges Garrison, where they **continued** thirty Days, and **were** not a little flusht with the expectation of Success; for at their first coming they **took** two Soldiers, who **gave** an account of the state of Matters: But Mr. Canady the commanding Officer being one of uncommon Courage and Resolution, **stood** his Ground till Col. Westbrook arrived, who **soon put** them to a rout (CHET, Penhallow, 1726: 99 [127 (3886)])

Excerpts (15) and (16) are from early eighteenth-century Astronomy and History texts, respectively. While the former appears to contain a more or less balanced proportion of narrative and non-narrative features, alternating present simple and present perfect tenses, the latter is an example of mainly present-tense historical discourse with very little narration, which reflects the relatively low narrative score for CHET at the beginning of the century. Extracts (17) and (18), in turn, represent the discourse of Life Sciences and History within the period 1725–1750. Here, the text on botany (17) appears to contain anything but narration, consisting instead of a description of the characteristics and behaviour of a plant. On the contrary, the abundance of narrative features in (18) is characteristic of a historical text, one which in this case is indeed telling a story, about one of the Wars of New England.

As can be seen in Figure 2, in the second half of the century the gap between CHET and the rest of the subcorpora diminishes once more, even though the former only falls by two units. Even though at this point there are clear differences among the four disciplines, the distribution of positive and negative features seems to be more balanced, with Philosophy maintaining a very low score, Life Sciences staying close to the average of the overall distribution of narrative features in the corpus, and Astronomy becoming more narrative than Life Sciences, with a positive score of 2.5, although not reaching the level of History. By 1800, Philosophy, Life Sciences and History appear to be moving slowly in the narrative direction, this in contrast to the sudden decrease in Astronomy to the average level, this notwithstanding the fact that all the disciplines except Philosophy have an overall higher score in the second half of the century, as seen in Figure 1, above. These differences between Figures 1 and 2 illustrate the familiar phenomenon that large-scale variation trends sometimes hide rather sharp period-internal variation on a smaller scale. Nevertheless, the five texts per twenty-five-year period used here are perhaps insufficient for such findings to be considered representative, and hence conclusions must be drawn with caution.

In order to explore these variation patterns in greater depth, in the following section we will consider internal variation in eighteenth-century CHET from a different perspective, that of genre.

4.2 Variation across genres within eighteenth-century CHET

As introduced in Chapter 3 (see also Moskowich, 2017), the twenty texts contained in the eighteenth-century section of CHET are from five different genres (Görlach, 2004): fourteen treatises, three essays, one lecture, one biography and one travelogue.

Figure 3 plots individual scores for each of these genres on Dimension 4, ranging from narrative to non-narrative, with zero representing the average Dimension 4 score for the total distribution of the features in the reference corpus (CETA, CEPhiT, and CELiST).

As can be seen, four of the five genres in eighteenth-century CHET present positive scores (except for Travelogue, with a negative score of −0.3), which indicates that all these genres have a larger proportion of narrative features than the average in the other three subcorpora. The only travelogue in our corpus is a text written by a woman, Elizabeth Justice, in 1739, hence it belongs to the second quarter of the century – the period where CHET appears to present the highest proportion of narrative features, as we have seen in Figure 2 – and yet, curiously, this text contains passages which are far more descriptive than narrative, as Example (19) illustrates:

(19) *There's* two Rows of Pews down the Middle; and, at the lower End, fronting the Altar, *there's* a large Pew for the Common People. On the Left Hand, *there's* a Place railed in; and in it a very fine Brass Font, which *has* on it all the Coats of Arms belonging to the Gentlemen who **gave** it; and there *is* near this Place a Pew, where the Bishop sits. There *is* over these Places an Organ; the Pipes of which *are* Brass, and it *is* finely painted (CHET, Justice, 1739: 5 [129 (887)])

The abundance of description in this travelogue is to be expected if we consider that an account of the author's travels must necessarily contain detailed descriptions of

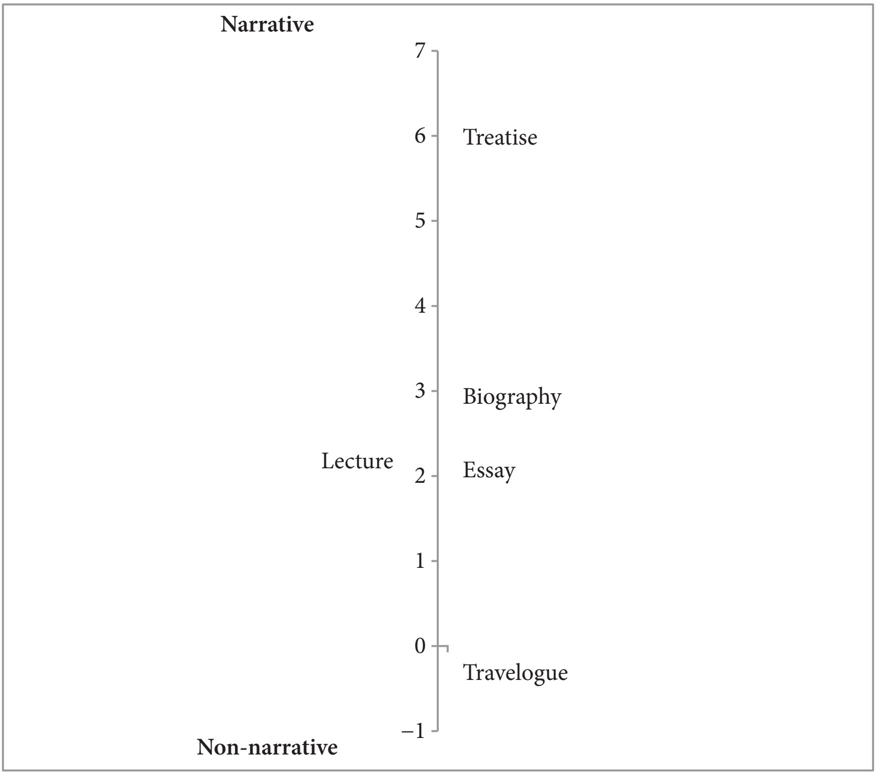


Figure 3. Dimension 4 scores for five genres in eighteenth-century CHET

the places visited. However, the fact that this second timespan (1725–1750) contains a text with a negative score (–0.3) implies that the other four texts here must have very high positive scores, so as to yield a mean period score of 7.2, and indeed they do (see Appendix). The genre with the highest narrative score is treatise (mean 5.9), but this does not in itself suggest that all the treatises in eighteenth-century CHET are narrative. Rather, the fourteen texts in the corpus which are treatises show a wide range of internal variability, as can be seen in Examples (20)–(22), all taken from treatises:

(20) (…) for as he was made consul on the first of July, so on the other hand it must **have been** much about the same time in the year following, or **rather later, before** he **arrived** in Britain, as the historian expressly *assures* us. If therefore Agricola **left** Britain **at first** in the year 74, the year **before** Cerialis **quitted** the government; he must according to Tacitus's account **have come** over **again** to Britain in the year 78, when the summer ***was*** **far** advanced

(CHET, Horseley, 1732: 47 [128 (5589)])

(21) But the vanity of the Queen-mother of France **disposed** her, without waiting the conclusion of it, to enter into a particular agreement with Don Antonio, and to furnish him with some soldiers for an expedition to Tercera; the chief island of the Azores in the Atlantic sea, where the Spaniards **had been** repulsed. This ***was*** **only** the prelude to a grand armament, which she **determined** to send into that part of the world, and which **was** accordingly equipped the following year. It **consisted** of near sixty *ships of war and transports*, with five thousand regular troops aboard (CHET, Anderson, 1775: 119 [136 (3544)])

(22) Yet our enjoyments *are* confined by a narrow circle; and, whatsoever *may be* its value, the luxury of life *is* possessed with more innocence and safety by the master of his own, than by the steward, of the public fortune. In an absolute government, which levels the distinctions of *noble and plebeian* birth, the sovereign *is* the sole fountain of honour; and the rank, both in *the palace and the empire, depends* on the *titles and offices* which are *bestowed and resumed* by his arbitrary will (CHET, Gibbon, 1788: 104 [138 (6946)])

The first of these extracts, (20), contains some 11 instances of narrative features in 83 words of text (past tense: *arrived, left, quitted*; perfect aspect: *have been, have come*; time adverbs: *later*, *at first, before, again*; other adverbs: *rather, far*). Rather than a story, though, this passage seems to offer speculation on how and when certain events took place within a specific time frame in the past (*If therefore Agricola left Britain in… he must according to Tacitus have come over again…*). Extract (21), in turn, contains a slightly less dense proportion of 7 narrative features in a total of 97 words (past tense: *disposed, was* (2), *determined, consisted*); perfect aspect: *had been repulsed*; adverbs: *only*), offering a description of the actions of the Queen-mother of France, preceded by a short analysis of her character (*the vanity… disposed her*), and followed by various descriptions, either of a place plus an allusion to something that had previously happened there (*where the Spaniards had been repulsed*), or of the ammunition used in the expedition (*It consisted of near sixty ships…*).

It should be noted, though, that both (20) and (21) also contain non-narrative features, such as the verb *to be* (albeit in the past), which was marked in both bold and italics, and an instance of present tense in (20) and of phrasal coordination in (21). This shows that a relatively strong presence of positive features in a passage does not automatically exclude every single negative feature from that context, but, rather, reduces their proportion. In contrast, Excerpt (22), also taken from a treatise, contains solely non-narrative features (present tense: *depends*, and present tense and *be* as main verb, in combination; and instances of phrasal coordination, such as *noble and plebeian, titles and offices*, or *bestowed and resumed*; see also Núñez-Pertejo, this volume), and exemplifies a passage in which the state of the art of a particular period of time is analysed and reflected on in the present tense. Also in this case, the socio-political situation described, plus the year – and indeed the century – in which this treatise was written appears to suggest that the historian may well have been describing the very times in which he was living. However, most treatises are narrative, (22) being an example of the few relatively non-narrative samples of the genre in the corpus.

If we look once more at Figure 2, we note that essay, lecture and biography show only moderately narrative scores, compared to the overall score for treatises, with the first two remaining close to 2.0 units above the reference corpus mean. Example (23) below is taken from one of the essays in the corpus, while (24) and (25) are from the only lecture and biography samples, respectively:

(23) The gods therefore were consulted in form; and sacrifices being offered, prayers were addressed to Jupiter, and the other tutelary Deities, that, if they **approved** of the new king, they would testify their approbation by some auspicious sign; which being done by a flash of lightning on the left, the people **immediately proceeded** to elect in form, and create him King, for whom the Gods **had** so openly **declared** (CHET, Chapman, 1750: 4 [132 (593)])

(24) We *are* come to a period memorable in the annals of Europe, for the peace of Utrecht. This country **had now**, with very small intermissions, **been** in a state of war for **near** forty years. In those long wars the country **lost** five or six thousand of their bravest youth; and it was found, that in the space of fifty years, the number of inhabitants in the province **had** not **doubled** (CHET, Adams, 1770: 37 [135 (8309)])

(25) In the year 1660, Mr. Firmin **married** a citizen's daughter with five hundred pounds as a portion, which though not a large sum, **was** to him who **knew** so well how to improve it, a valuable acquisition. The great expense of supporting a family in this age of *dissipation and luxury renders* many young traders, and indeed persons of all professions, very averse from matrimony (CHET, Cornish, 1780: 6 [137 (938)])

On an initial view of these extracts, it may seem difficult to distinguish between the genres, except perhaps in the case of (25), where we are given details of the marriage of Mr. Firmin, the man portrayed in Joseph Cornish's biography *The Life of Mr. Thomas Firmin, Citizen of London*, published in the last quarter of the eighteenth century. Yet although narration is pervasive throughout the text, this passage also contains a reflection on the economic implications of marriage at the time, written in the present tense. Excerpt (23), on the other hand, is taken from an essay about the Roman senate by Thomas Chapman, published in 1750. This passage narrates an episode from the history of Rome, the coronation of Romulus as king. The historian is here describing a tradition of the period through a detailed narration of Romulus's artful manipulation of the people by strengthening their belief in the king's divine nature, the text enlivened with a generous use of irony.

Finally, Example (24) belongs to a lecture by Amos Adams (1770) on the colonisation of New England, and this passage appears to be a moment of contextualisation, where the author appeals to the audience's shared knowledge of the history of Europe by means of an inclusive *we* and the use of the present tense (*We are come to a period memorable in the annals of Europe*), after which he recounts, through narration in the past perfect, all the plagues that had occurred in America and how long they had lasted by the time the Treaty of Utrecht was signed in Europe. Bearing in mind that lecture used to be a genre written to be delivered orally in public, the appeal to the audience seems to be justified. Despite their differences, all three extracts here contain a rather moderate number of narrative features, in contrast with some of the treatises, where past tense and place and time adverbs are much more frequent.

We have seen, then, that despite having an overall Dimension 4 score notably higher than the rest of the subcorpora analysed in Monaco (2017), eighteenth-century CHET presents quite a wide range of internal variation, not only among genres but also within them. These include treatise, which at that time appears to have been a rather eclectic format "often used by authors themselves to refer to almost to any piece of writing of considerable length" (Moskowich, this volume). However, the text examples discussed here have also shown that narrative features tend to be present in nearly all the genres, although some contain a more balanced proportion of both narrative and non-narrative features, often alternating stories with reflections on the state of affairs of the period in question. A summary of our findings is given in the final section, below.

5. Concluding remarks

In this chapter we have looked at narrative and non-narrative features in twenty texts drawn from the eighteenth-century section of CHET, following Biber's (1988) MD Analysis and using Monaco's (2017) study of CETA, CEPhiT and CELiST as a reference. Relative standard scores for Dimension 4 "Narrative vs. non-narrative discourse" were compared, firstly, across the four subcorpora – which represent the disciplines of Astronomy, Philosophy, Life Sciences, and History – and, subsequently, within history writing through an analysis of variation across genres.

The first analysis included the variable time, results being plotted across two periods of fifty years and across four periods of twenty-five years. The former revealed that eighteenth-century CHET is notably more narrative than the other three subcorpora, corroborating the hypothesis that historical discourse is expected to contain a larger proportion of narrative features than other types of scholarly discourse (in this case, Astronomy, Philosophy and Life Sciences). Excerpts from the texts have shown that narrative features are not necessarily restricted to offering relations of events, but may be also used for speculation and deduction (see, also, Puente-Castelo, this volume). The latter analysis, on the other hand, showed diachronic variation over the course of the century through twenty-five-year timespans. Although we accept, following Kytö *et al.* (2000), that periods of twenty-five years may be too short to account for language change, and hence that results should be treated with caution, it is nevertheless interesting to note that both History and two other disciplines – Astronomy and Life Sciences – appear to gradually gain narrative features through the century. In the case of CHET, the movement in the narrative direction appears to be quite pronounced, especially considering that in the first twenty-five years its narrative score was only marginally higher than that of CETA and CELiST (CEPhiT being generally non-narrative across the whole century).

The results of this analysis also show a sharp increase in narrative features for CHET during the second quarter of the century. The explanation for such a phenomenon appears to lie in the variable of genre, which was used in a subsequent analysis. Genre scores revealed that treatises are, overall, much more narrative than essays, and also more so than the three other genres – lecture, biography and travelogue – although this latter genre was represented by just one sample in our corpus. Also, four of the five texts included in the period 1725–1750 in CHET are, unsurprisingly, treatises, all with scores above 8.5, as can be seen in the Appendix. The travelogue, by contrast, presents a negative score (albeit a very low one), this seemingly due to its abundant descriptions of places, people and local customs. On the other hand, the essays, the lecture and the biography present more moderate positive scores, showing a more balanced alternation of narration of actions and events with analyses and reflections on the different situations described, the latter often being expressed in the present tense.

Finally, having looked at individual excerpts from texts when analysing variation across genres, we have also noted that there is, in fact, a great deal of variation within some genres. A good example of this genre-internal variation is treatise, with some texts having negative scores (this is also the case with one of the essays; see Appendix), and some having very moderate narrative scores, although in fact most treatises do have very high scores, which explains their relatively high average score. However, as already seen in Gray (2011) and Monaco (2017), in the present study we have found that even in some of those texts with very high narrative scores it is still possible to find present tense and/or other features from the opposite subset (in this case, non-narrative), although usually in a much more reduced proportion. This lack of rigidity with regard to the narrative vs. non-narrative duality appears, once more, to reflect the non-rigid character of late Modern English scholarly discourse, at a time when rhetorical boundaries were not yet fully defined and register characteristics were only beginning to emerge.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Atkinson, Dwight. (1999). *Scientific discourse in sociohistorical context: The Philosophical Transactions of the Royal Society of London, 1675–1965*. Mahwah, NJ: Erlbaum.

Biber, Douglas. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, Douglas. (2001). Dimensions of variation among eighteenth-century speech-based and written registers. In Susan Conrad & Douglas Biber (Eds.), *Variation in English: Multi-Dimensional studies* (200–214). Harlow: Longman.

Biber, Douglas. (2016). Using multi-dimensional analysis to explore cross-linguistic universals of register variation. In Marie-Aude Lefer & Svetlana Vogeleer (Eds.), *Genre- and register-related discourse features in contrast* (7–34). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/bct.87.02bib

Biber, Douglas & Edward Finegan. (2001). Intra-textual variation wihin medical research articles. In Susan Conrad & Douglas Biber (Eds.), *Variation in English: Multidimensional studies* (108–123). Harlow: Longman.

Biber, Douglas & Bethany Gray. (2013). Being specific about historical change: the influence of sub-register. *Journal of English Linguistics*, 41, 104–134. https://doi.org/10.1177/0075424212472509

Camiña, Gonzalo & Inés Lareo. (2016). Editorial policy in the Corpus of English Philosophy Texts: Criteria, conventions, encoding and other marks. In Isabel Moskowich, Gonzalo Camiña, Inés Lareo, & Begoña Crespo (Eds.), *'The Conditioned and the Unconditioned': Late Modern English texts on philosophy* (45–60). Amsterdam/Philadelphia: John Benjamins.

Carr, David. (1991). *Time, narrative, and history*. Bloomington: Indiana University Press.

Conrad, Susan. (1996). Academic discourse in two disciplines: Professional writing and student development in biology and history. Unpublished doctoral dissertation. Flagstaff: Northern Arizona University.

Conrad, Susan. (2001). Variation among disciplinary texts: A comparison of textbooks and journal articles in biology and history. In Susan Conrad & Douglas Biber (Eds.), *Variation in English: Multi-Dimensional Studies* (94–107). Harlow: Longman.

Croce, Benedetto. (1921). *Theory and history of historiography*. Translated by Douglas Ainslie. London: George G. Harrap & Company.

Croce, Benedetto. (2017 [1893]). *La storia ridotta sotto il concetto generale dell'arte*. Milano: Adelphi Edizioni.

Dray, William Herbert. (1971). On the nature and role of narrative in historiography. *History and Theory*, 10(2), 153–171. https://doi.org/10.2307/2504290

Gallie, Walter Bryce. (1964). Narrative and historical understanding. *History and Theory*, 6(2), 149–172. https://doi.org/10.2307/2504278

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

Görlach, Manfred. (2004). *Text types and the history of English*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197167

Gray, Bethany. (2011). Exploring academic writing through corpus linguistics: When discipline tells only part of the story. Unpublished PhD dissertation. Northern Arizona University, Flagstaff, AZ.

Grieve, Jack. (2016). *Regional variation in written American English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139506137

Hardie, Andrew. (2012). CQPweb – combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3), 380–409. https://doi.org/10.1075/ijcl.17.3.04har

Hyland, Ken. (1998). *Hedging in scientific research articles*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.54

Kytö, Merja; Juhani Rudanko & Erik Smitterberg. (2000). Building a bridge between the present and the past: A corpus of 19th-century English. *ICAME Journal*, 24, 85–97.

Lareo, Inés. (2011). Colocacións con *make*, *take* e *do* + nome nun corpus do século XIX de textos ingleses científicos e literarios escritos por mulleres. *Cadernos de Fraseoloxía Galega*, 13, 191–214.

Louch, A. R. (1969). History as narrative. *History and Theory*, 8/1, 54–70. https://doi.org/10.2307/2504189

Mandelbaum, Maurice. (1967). A note on history as narrative. *History and Theory*, 6/3, 413–419. https://doi.org/10.2307/2504424

Marenco, Franco (2009). Narrative in History / History in Narrative. In Block de Behar, Lisa; Paola Mildonian; Jean-Michel Djian; Djelal Kadir; Alfons Knauth; Dolores Romero López & Marcio Seligmann Silva (Eds.), *Comparative literature: Sharing knowledges for preserving cultural diversity. Encyclopedia of Life Support Systems. Vol. I*. Oxford: Eolss Publishers.

Mink, Louis. (1970). History and fiction as modes of comprehension. *New Literary History*, 1, 541–558. https://doi.org/10.2307/468271

Monaco, Leida Maria. (2017). A multidimensional analysis of late modern English scientific texts from the Coruña Corpus. Unpublished PhD dissertation. University of A Coruña.

Moskowich, Isabel. (2017). Genre and change in the Corpus of History English Texts. *Nordic Journal of English Studies*, 16(3), 84–106.

Moskowich, Isabel & Begoña Crespo (Eds.) (2012). *Astronomy 'playne and simple': The writing of science between 1700 and 1900*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.173

Moskowich, Isabel; Gonzalo Camiña, Inés Lareo, & Begoña Crespo (Eds.) (2016). *'The Conditioned and the Unconditioned': Late Modern English texts on philosophy*. Amsterdam/Philadelphia: John Benjamins.

Munslow, Alun. (2007). *Narrative and history*. New York: Palgrave McMillan. https://doi.org/10.1057/978-1-137-01943-1

Roberts, Geoffrey. (2001). *The history and narrative reader*. London/New York: Routledge.
Rüsen, Jörn. (2005). *History: Narration, interpretation, orientation*. New York/Oxford: Berghahn Books. https://doi.org/10.2307/j.ctt1x76fc2
Stone, Lawrence. (1979). The revival of narrative: Reflections on a new old history. *Past & Present*, 85, 3–24. https://doi.org/10.1093/past/85.1.3
Van Seters, John. (1997). *In search of history: Historiography in the ancient world and the origins of biblical history*. Winona Lake, Indiana: Eisenbrauns.
White, Hayden. (1973). *Metahistory: The historical imagination in nineteenth-century Europe*. Baltimore & London: The John Hopkins University Press.
White, Hayden. (1984). The question of narrative in contemporary historical theory. *History and Theory*, 23(1), 1–33. https://doi.org/10.2307/2504969
White, Hayden. (1987). *The content of the form: Narrative discourse and historical representation*. Baltimore/London: Johns Hopkins University Press.
Xiao, Richard. (2009). Multidimensional Analysis and the study of Word Englishes. *World Englishes*, 28(4), 421–450. https://doi.org/10.1111/j.1467-971X.2009.01606.x

Appendix. Individual dimension 4 scores per text

Text	Genre	Score
1704 Tyrrell	Treatise	7.67
1705 Anderson	Essay	−3.53
1710 Crawfurd	Treatise	−3.42
1716 Oldmixon	Treatise	2.32
1721 Strype	Treatise	5.74
1726 Penhallow	Treatise	8.56
1732 Horsley	Treatise	10.40
1739 Justice	Travelogue	−0.31
1740 Bancks	Treatise	8.78
1745 Hooke	Treatise	8.78
1750 Chapman	Essay	5.36
1760 Birch	Essay	4.40
1762 Scott	Treatise	7.61
1770 Adams	Lecture	2.18
1775 Anderson	Treatise	4.56
1780 Cornish	Biography	2.94
1788 Gibbon	Treatise	−5.20
1790 Gifford	Treatise	9.18
1795 Adams	Treatise	8.49
1800 Stock	Treatise	10.33

CHAPTER 6

Time and history

A preliminary approach to binomials in late Modern English Astronomy and History texts

Paloma Núñez-Pertejo
Universidade de Santiago de Compostela

1. Introduction

Despite being a frequent linguistic and stylistic phenomenon throughout the history of English (Sauer, 2017: 278), and mentioned in the literature as early as the nineteenth century, phraseological units usually referred to as *binomials* constitute "a relatively under-researched field" (Sauer & Schwan, 2017a: 83),[1] and seem never to have been a central focus of interest (cf. Mollin, 2014: 10), perhaps on the grounds that the stylistic ideal for scholars throughout the twentieth century was "simplicity and not richness" (Sauer, 2018a: 86). Research into this area of language, however, seems to be increasing, attested by the growing number of specialized publications here, although according to Sauer (2014: 27), "a comprehensive history of the use of twin-formulae in English literature remains to be written". The aim of this paper is to contribute to the history of English binomials by providing a preliminary account of their use in two different disciplines, history and astronomy, during the late Modern English period (1700–1900). Recent work on binomials has concentrated on Old English (Chapman, 2017; Fulk, 2017; Kotake, 2017; Ogura, 2017; Zagórska, 2017), Middle English (Krygier, 2017; Kubaschewski, 2017; Schenk, 2017),[2] early Modern English (Bach, 2017; Doty & Wicklund, 2017; Lehto, 2017; Rutkowska, 2017; Sprau, 2017), and contemporary English (cf. Mollin, 2014, 2017; Schaefer,

---

[1] For extensive reviews of the literature on binomials, cf. Mollin (2014: 10–13); Sauer (2014: 26–27); Sauer (2017: 279); Sauer & Schwan (2017a: 86–89); Sauer (2018a: 86), and most notably Kopaczyk & Sauer (2017: 351–356), among others.

[2] "Binomials have certainly been a feature of English from Old English to the present" (Sauer & Schwan, 2017a: 89), although they were especially popular in the 15th century, both in prose and in poetry (cf. Sauer, 2018b: 62).

https://doi.org/10.1075/z.225.06nun

© 2019 John Benjamins Publishing Company

2017). The period between 1700 and 1900, though, has been somewhat neglected (cf., however, Tyrkkö, 2017). It is my concern here to fill this gap and to offer a descriptive account of the use of binomials in this period in the history of English. Specifically, the data are drawn from two historical sub-corpora comprising a selection of history and astronomy texts (cf. Section 3 below); the comparison between the binomials in these two disciplines, it is hoped, will shed light on potential register-dependent features of this frequently occurring linguistic phenomenon.

This chapter is organized as follows: Section 2 offers an overview of binomials by providing a definition of the term (2.1) and by considering a number of different parameters, such as the formal structure and word class of members (2.2), function (2.3), and semantic structure (2.4). Section 3 sets out the methodology used, while Section 4 provides a picture of how some of these pairs were used in astronomy and history texts in the late Modern English period. Finally, Section 5 summarizes the main conclusions of the study.

2. Binomials

2.1 Definition

"[T]hinking in pairs seems to be a common human trait" (Sauer & Schwan, 2017b: 192), and this is linguistically expressed by means of *binomials* or phrasal units "whose shape, motivation and function bridge the *spoken* and *written* medium, bringing together *phonology, semantics, syntax, style* and *etymology*" (Kopaczyk & Sauer, 2017: 1; my emphasis). Hence binomials are "a natural phenomenon of all language use and can be considered a universal" (Tyrkkö, 2017: 296), this illustrated by the fact that the study of these combinations has been approached from different perspectives in different languages (cf., among others, Koch, 1983 on Arabic lexical couplets; Danet, 1984 on binomials in Hebrew; Lambrecht, 1984 on German binomial expressions; Klégr, 1991 on Czech binomials; Okamoto, 1993 on nominal repetitive constructions in Japanese; Rodríguez-Sánchez, 2013 on Spanish binomials; Mmadike, 2014 on irreversible binomials in Igbo; Gorgis & Al-Tamimi, 2015 on binomials in Iraqui and Jordanian Arabic).

A classical definition of binomials, also known as *doublets, conjoined phrases, word pairs, tautological pairs, repetitive (word) pairs, twin-formulae, fixed coordinates* and *freezes* (cf. Sauer, 2014; Kopaczyk & Sauer, 2017: 7; Sauer & Schwan, 2017a, b), is that of a "sequence of two words pertaining to the same form-class, placed on an identical level of syntactic hierarchy, and ordinarily connected by some kind of lexical link" (Malkiel, 1959: 113; cf. also Bhatia, 1993: 108), and presenting some sort of semantic relation, e.g. *heaven and earth, rise and fall, high and* *low, by and by*, etc. In other words, they are "coordinated word pairs whose lexical elements belong to the same word class, and which do not transcend syntactic boundaries" (Mollin, 2014: 7). In this paper, the term *binomial* has been adopted, since, as Sauer (2014: 26) notes, it can be used in a wider sense than some of the alternatives available (*twin-formulae, fixed coordinates, freezes*), and also because it is a neutral term which "leaves it open whether they are formulaic and have a long history, or whether they have been newly created, and whether they are tautologic or not" (Sauer & Schwan, 2017a: 84).

Likewise, Mollin agrees that the umbrella term *binomial* should be used to refer to coordinated word pairs in general, whether fixed (idiomatic) or not, while the more restrictive *twin-formula* or *freeze* "should be reserved for the special class of irreversible binomials" (2014: 8), whose constituent order is fixed (e.g. *bread and butter*). On these lines, some authors mention a "cline of fixedness", along which binomials would range from more rigid, *frozen*, irreversible structures to others which are "created on the spur of the moment" (Sauer, 2017: 280). This latter type, also referred to as *current, ad hoc* binomials (cf. Klégr & Čermák, 2008: 42), seems to be more frequent in English than the *idiomatic* type, while the reverse seems to be true of other languages, e.g. Czech (cf. Klégr & Čermák, 2008: 45–46), in which the use of established or fixed binomials appears to be more widespread.[3]

It should be noted, however, that not all sequences of two words of the same class and coordinated by a conjunction are binomials; the two coordinated words need to be linked by some semantic relation (cf. Kopaczyk & Sauer, 2017: 3; cf. also Section 2.4 below).

Apart from binomials or doublets (*black and white, up and down*), multinomials have also been discussed in the literature, including triplets/trinomials as in (1), (2) below, quadrinomials/quatruplets (3), quintuplets (4), sextuplets, septinomials and indeed combinations consisting of up to ten elements (cf. Sauer, 2017: 281). The relation between binomials and multinomials is more than simply an apparent one, since the latter "can sometimes be seen as concatenations of two or more binomials" (Sauer, 2018b: 63), or mere extensions of binomials (Sauer, 2017: 280); trinomials and longer sequences, in turn, are often regarded as "instances of enumeration" (Sauer, 2014: 28), or even lists (Kopaczyk & Sauer, 2017: 3). This, of course, is a question that merits closer scrutiny in itself (for more information on multinomials, cf. Malkiel, 1959: 120–121; Klégr & Čermák, 2008: 46–47; Sauer, 2017; Sauer, 2018b: 62–63).

---

[3] While *doublets, conjoined phrases* and *word pairs* seem to be rather vague labels, *tautological pairs, repetitive pairs, twin-formulae, fixed coordinates* or *freezes* are, in turn, more restrictive and "stress the formulaic and fixed nature of binomials" (Kopaczyk & Sauer, 2017: 7). Thus, tautological word pairs "would exclude binomials that contain antonyms; formulae or freezes would exclude binomials that are not formulaic but were created on the spur of the moment" (Sauer, 2017: 278).

(1) hatred, love and envy (Strype, 1721: 78 [126 (2254)])

(2) right, holy and just (Strype, 1721: 88 [126 (DEL)])

(3) astronomy, geography, navigation, and other mathematical sciences (Morden, 1702: Title)

(4) hunger, scarcity, ill-favoured lodging, ill air, and many other incommodities (Strype, 1721: 91 [126 (8395)])

In the present study, the focus will be on combinations of just two elements, thus leaving aside more complex structures, despite the evident interest of these.

2.2 Formal structure of binomials and word class of constituents

Although binomials are typically made up of two words that belong to the same class (e.g. nouns, as in *heaven and earth*; adjectives, as in *good and bad*, etc.) and are linked by a conjunction, variations of this basic or prototypical word A + conjunction + word B pattern exist:

a. ***the*** *king and* ***the*** *queen*, or ***any forged*** *deed or* ***false*** *oath* (Oldmixon, 1716: 64 [125 (5520)]) (*extended* binomials), in which the nouns are preceded by determiners and adjectives; the latter, as well as nouns in the genitive, can also be found in postmodifying position, i.e. following nouns;
b. *she was* ***bold*** *in battle* ***and prudent*** (*split* binomial), in which the two adjectives are separated or *split* by a prepositional phrase (cf. Sauer, 2018a: 87);
c. *good, honest* (*reduced* binomial, or binomial without a linking conjunction);
d. *profound* ***in wisdom****, firm* ***in purpose*** (*shortened plus extended* binomial), in which the two adjectives are followed or postmodified by two prepositional phrases "but the conjunction has been omitted" (Sauer, 2018a: 88).[4]

Although "these coordinated pairs have the same function as the basic binomials" (Sauer, 2018b: 66), they are not all included in my analysis. In fact, the wide diversity of morphosyntactic patterns, more complex than the basic pattern word A + conjunction + word B, has led Rutkowska to conclude that "there is no clear-cut division between binomials and semantically related coordinated phrases" (2017: 181), and indeed, the distinction is often blurred. In the present study, apart from basic binomials, I have also included combinations with some additional modification, or some degree of complexity, such as *of the day and of the night* and *meer fable and monkish invention* (Anderson, 1705: 82 [123 (6373)]). Generally, these patterns are not excessively complex and can be seen as *prototypical* or *genuine* binomials (cf. Sauer, 2017: 283; Sauer, 2018a: 88); in fact, *of the day and of the night* is considered to be a variation of the simpler structure *day and night* (cf. Kopaczyk & Sauer, 2017: 371). In turn, longer and more complex sequences, such as (5) below, were excluded from the count, as were certain other patterns that did not comply with the word class requirement, such as *by chance or fortuitously*, that is, two clearly synonymous adverbials realized by a prepositional phrase and an adverb respectively (cf. Sauer, 2018a: 88–89 on binomials with a mixed word class). These are all undoubtedly very interesting sequences, and deserve further research. In addition, I have excluded cases of juxtaposition, that is, *reduced* binomials (e.g. *come, go*), as well as some prepositional phrases (e.g. *day after day*).

---

[4] For a detailed account on the formal structure and properties of binomials, cf. Sauer & Schwan (2017a: 91); Sauer (2017: 282–283); Sauer (2018a: 87–88); Sauer (2018b: 65).

(5) **three shillings** upon every tun of wine, **and twelve pence** upon every pound of merchandise. (Tyrrell, 1704: 950 [122 (30)])

Regarding word class, combinations of nouns, adjectives and verbs are among the most frequent. In adjective + adjective pairs, these are usually attributes and are commonly assigned to things and people (cf. Sauer, 2018b: 65). By contrast, other combinations, such as adverbs and prepositions, are far more infrequent.

In terms of connectors, *and* is usually the preferred option, followed by *or*, while other connecting elements, such as *neither…nor, but*, and even *&* (ampersand), as in *civil & religious*, are far more rarely attested (cf. Sauer & Schwan, 2017a: 93; Sauer, 2018a: 89).

2.3 Function

A defining feature of binomials is that of repetition, attested at different levels, (structural, semantic, phonological, etc.; cf. Kopaczyk & Sauer, 2017: 15–17). Those pairs showing the highest degree of repetition are classified as *core* binomials (*wear and tear*), whereas those not showing all types of repetition are classified as *peripheral* (*staff and equipment*; cf. Kopaczyk & Sauer, 2017: 16–17).

Binomials have been attested since Old English times as serving "to create a rich, ornate and elevated style",[5] especially in literary texts, while they were also used "for emphasis" (Sauer, 2014: 28; cf. also Sauer & Schwan, 2017a: 86; Sauer, 2018a: 86; Sauer, 2018b: 63), especially in terms of alliteration, which, according to

---

[5] *Copia verborum* (cf. Sauer, 2017: 279).

Klégr & Čermák (2008: 56), seems to be “the dominant means of formal cohesion” in Old English binomials.[6] Their aesthetic and prosodic qualities have also been noted by Malkiel (1959), given their high frequency in poetic texts (cf. also Klégr & Čermák, 2008: 55). Stylistically, they are used to show both explanatory and creative variation (cf. Klégr & Čermák, 2008: 57). In this respect, Sauer & Schwan (2017b: 192; cf. also Sauer, 2018b: 63) distinguish between *stylistic* and *factual* binomials, the former used for stylistic purposes, e.g. *encouraged and fomented*, and the latter to reflect reality, in that they are assumed to be based on factual knowledge, e.g. *longitude and latitude*.

Apart from those mentioned above, binomials also served other purposes, such as etymological ones, in facilitating comprehension of Latin and French terms or loan-words through the inclusion of native words in glosses and hence providing some sort of explanation (or even translation) of the foreign terms in the vernacular language, this seen in both Old and Middle English. Another plausible function of binomials was to clarify difficult arguments and to increase precision, that is, to make them “as comprehensive and unambiguous as possible” (Sauer, 2017: 279–280) and therefore to avoid “possible misunderstandings” (Sauer, 2018b: 63), especially in legal and legislative texts.[7] Thus, the use of binomials constitutes “an extremely effective linguistic device to make the legal document precise as well as all-inclusive” (Bhatia, 1993: 108). This also applies to theological and scientific (or pseudo-scientific) texts, all of which exhibit interesting combinations of binomials in different periods of English, including Modern and Present-Day English (cf. Sauer, 2017: 278; Sauer & Schwan, 2017a: 89). The relationship between binomials and conventionalized and formulaic language, then, is evident, but their use is not limited to this kind of usage.

2.4 Semantic structure

According to Sauer & Schwan (2017b: 187), “three basic semantic categories of binomials are fairly clear […], namely synonymy, antonymy and complementarity” (cf. also Sauer, 2018a: 98–101; Sauer, 2018b: 69–71). The former includes combinations of words with a similar (denotational) meaning, and these words are usually, but not always, interchangeable. In other cases, the two words have exactly the same

---

[6] In oral societies, however, binomials “probably served as an aid to memory” (Sauer, 2017: 279). Moreover, alliteration seems to have fulfilled the same function (as an aid to memory) in legal binomials (cf. Sauer, 2018b: 66).

[7] In fact, the use of binomials in legal English has been explored extensively (cf., among others, Gustafsson, 1984; Bhatia, 1993; Frade, 2005; Iglesias-Rábade, 2007; Kopaczyk, 2009, 2012).

(denotational and connotational) meaning, i.e. they make up so-called tautological units, such as *aches and pains* (cf. Kopaczyk & Sauer, 2017: 12), these being relatively rare, since there are usually stylistic and frequency-of-use differences between the members of the pair (cf. Sauer, 2017: 287). According to Sauer, it is "often difficult to distinguish between synonymy and tautology" (2018a: 98), the latter of course being easily identified in pairs such as *more and more*, that is, whenever a word is repeated.

Antonymy, in turn, refers to words which have opposite meanings, and can be further characterized as being "absolute", "strict", or "mutually exclusive" (e.g. *life ~ death*), "gradable" (*hot ~ cold*), or "converse" (*buy ~ sell*); more examples can be found in Sauer (2017: 288) and Sauer & Schwan (2017b: 189). Antonyms in binomials are usually more concrete than the higher unit to which they refer, which is typically more abstract (cf. Sauer, 2017: 288; Sauer, 2018a: 99), e.g. *day and night* or *men and women*, which refer to the whole day and to all mankind respectively; they are said to refer to parts of a whole, of a higher unit (cf. Sauer, 2017: 288; Sauer, 2018a: 99; Sauer, 2018b: 70).

Finally, those related pairs "that are neither clearly synonymous nor clearly antonymous" (Sauer & Schwan, 2017b: 190; cf. also Sauer, 2018b: 69) are generally classified as complementary. Thus, from a semantic standpoint, complementarity is clearly a very heterogeneous group with many subgroups (cf. Sauer, 2017: 287), whose members may express meanings such as cause and effect or result, more general plus more specific, more specific plus more general, sequence of actions, etc.[8] However, complementarity is considered within antonymy by Kopaczyk & Sauer (2017), who add contiguity to the semantic structure of binomials, together with synonymy and antonymy. Contiguity would thus be used "as an umbrella term for other semantic relations which cover various degrees of hyponymy/hyperonymy, sequential and causal relationships, metonymic and metaphoric extensions of meaning, etc." (Kopaczyk & Sauer, 2017: 12). Malkiel (1959), in turn, distinguishes between synonymy, opposition (antonymy), complementation, as in *food and drink*, and succession, as in *rise and fall*.

Mollin's account of the semantic structure of binomials, however, differs from the previous ones in that she establishes a connection between semantics and register on the one hand, and semantics and frequency, on the other. Thus, antonymy and complementarity/complementation have to do with register, that is, with the contents of a given text, while synonymy could contribute to "an increasing or decreasing frequency of binomials" (2017: 294), whose use would be stylistically motivated, this coinciding with Sauer's claim that "pairs consisting of synonyms occur mainly among the 'stylistic' binomials" (2018b: 69).

---

[8] For a full description and classification of complementary binomials, cf. Sauer (2017: 289–291); Sauer (2018a: 99–101); Sauer (2018b: 70–71).

3. Methodology

This study of binomials in the late Modern English period (1700–1900) is based on data drawn from the *Corpus of English Texts on Astronomy* (CETA; cf. Moskowich & Crespo, 2012; cf. also Moskowich, Lareo, Camiña-Rioboo & Crespo, 2012) and the beta version of the *Corpus of History English Texts* (CHET; cf. Moskowich, 2016, 2017; this volume), twin sub-corpora of the *Coruña Corpus of English Scientific Writing*, compiled under the same principles. The texts in CETA (astronomy) belong to what is traditionally known as the "hard" sciences, while the texts in CHET (history) belong to the so-called "soft" sciences (cf. Moskowich, 2017: 98).

Each sub-corpus comprises samples of ca. 10,000 words, roughly 20,000 words per decade and discipline, i.e. 200,000 words per century and discipline, which amounts to a total of ca. 800,000 words for both CETA and CHET. The exact number of words in the two sub-corpora are indicated in Tables 1 and 2 below:

Table 1. Words per century in CETA

Century	Number of words
18th c.	208,079
19th c.	201,830
Total	**409,909**

Table 2. Words per century in CHET

Century	Number of words
18th c.	201,938
19th c.	202,486
Total	**404,424**

For the present study, I selected the earliest five works in CETA and the earliest five in CHET. These were all published during the first quarter of the eighteenth century, between 1702 and 1726, and comprise a total of 101,115 words. The corpora include specific software (CCT) which makes it possible to retrieve words and sequences of words easily, although for our purposes here it was necessary to read the texts themselves. Therefore, all the binomials attested in these works were retrieved manually and are all listed in their Present-Day English spelling. Thus, *cœlestial and terrestrial* appears as *celestial and terrestrial*, and *reign'd and dyed* as *reigned and died*.

The samples chosen for the current analysis are from texts by the following authors in Tables 3 and 4, listed as they appear in the corpora metadata, and with the (abbreviated) titles of their works (cf. Crespo, 2012: 30):

Table 3. Selection of authors and texts from CETA

ASTRONOMY (CETA)			
Author	**Title**	**Year**	**Words**
Henry Curson	*The theory of sciences illustrated*	1702	10,247
Robert Morden	*An introduction to astronomy, geography, navigation, and other mathematical sciences*	1702	10,154
William Whiston	*Astronomical lectures*	1715	9,939
John Harris	*Astronomical dialogues between a gentleman and a lady*	1719	9,907
George Gordon	*An introduction to geography, astronomy and dialling*	1726	10,437
			50,684

Table 4. Selection of authors and texts from CHET

HISTORY (CHET)			
Author	**Title**	**Year**	**Words**
James Tyrrell	*The general history of England, both ecclesiastical and civil*	1704	10,089
James Anderson	*An historical essay*	1705	10,066
George Crawfurd	*A genealogical history of the royal and illustrious family of the Stewarts*	1715	10,111
John Oldmixon	*Memoirs of Ireland from the Restoration, to the present times*	1716	10,077
John Strype	*Ecclesiastical memorials*	1721	10,088
			50,431

4. Results

4.1 Binomials in the *Corpus of English Texts on Astronomy* (CETA)

Astronomy, "or the observation of stars, the heavens, and planetary motions" (Crespo & Moskowich, 2010: 157), has traditionally been classified as one of the oldest sciences, as well as a part of the seven liberal arts (cf. Sauer, 2018b: 62). In fact, the *grounds and principles* of the seven liberal arts are detailed in Curson's *The theory of sciences illustrated* (1702). Among the main concerns of astronomy we find the study of a range of phenomena relating to stars, planets, comets, galaxies, such as their movement, and evolution, as well as other phenomena outside the Earth's atmosphere; that is, anything that has to do with the formation and development of the universe (cf. Cantos & Vázquez, 2012: 123–124). According to Crespo, from the seventeenth to the nineteenth centuries, with the establishment of modern science, there was a shift in the concerns of astronomy; not only were the position and motion of the sun, the moon and the planets of interest, or the effects of the Sun on the Earth, but also issues such as the physical, chemical and meteorological description of celestial objects, in addition to studies of light, including brightness and temperature (2012: 18). Other, somewhat less scientific topics that also featured in astronomy texts of the time included the influence of the stars on people's fate, the signs of the zodiac and their connection with the twelve months of the year, etc. (Sauer, 2018b: 61).[9]

Another main concern in the late seventeenth and early eighteenth centuries was to make scientific knowledge more accessible, to "popularize" science and "to include not only the rich or professional groups but also the middle class who demanded instruction" (Crespo & Moskowich, 2010: 157). Hence the vernacular (English) as a means of the transmission of scientific knowledge replaced Latin, until now the language of learning and culture.

Many authors' writings on scientific topics were characterized by a complex and dense style, although others preferred simplicity and clarity, as recommended by the Royal Society. As already noted (cf. Section 2.3 above), binomials were commonly found in formal registers, illustrative of a literary style, yet they also occurred in more informal, less elevated styles such as popular science (cf. Sauer, 2018b: 63), astronomy being no exception. In fact, during the Middle and Early Modern English periods, binomials were frequently attested in texts such as almanacs, these "addressed to laymen and used on a daily basis" (Rutkowska, 2017: 175).

In the following sections some the features that characterize binomials in the *Corpus of English Texts on Astronomy* (CETA) are discussed. It should be noted that "with a phenomenon as fluid and mercurial as binomials the figures are necessarily approximative and show tendencies rather than hard data" (Klégr & Čermák, 2008: 42). Thus, a total of 669 different pairs or types were retrieved from the five works selected for analysis (cf. Section 3 above), although some of these combinations occur more than once, and the number of tokens is 776 (see Table 5 below). Hence, *ad hoc* binomials, that is, binomials created on the spur of the moment and used only once by an individual writer, are more frequent than fixed, idiomatic binomials, which we might expect to appear more than once (cf. Section 2.1 above).

4.1.1 *Word class of constituents and connectors*

As has been noted above, noun + noun combinations illustrate the most commonly recurring type of binomial. In CETA, 422 different types of pairs of nouns were recorded (500 tokens), these including *heavens and earth, planets and stars, day and*

---

[9] According to Rutkowska, "it has been assumed that texts dealing with the influence of the planets on people's character traits and destinies belong to astrology, whereas the description of the position of the planets in the sky are counted as texts on astronomy" (2017: 193).

night, line or axis, latitude or distance, parts or regions. Coordinated names such as those of planets, as in *Saturn and Mars*, or stars, such as *Sirius and Procyon*, are a special subgroup within nominal binomials, and will be considered here, despite studies such as Rutkowska's (2017), in which they are excluded on the grounds that they are too specific and context-bound (cf. also Sauer, 2018a: 88). Although most combinations are attested just once, some are repeated twice or even more; *sun or star*, for example, occurs 19 times, followed by *day and night* (7), *body or aspect* (7), *earth and planet* (6), *sun and moon* (6), *meridian and horizon* (4), *degree and minute* (4), etc. Another recurring pattern, this time combining two *-ing* forms, is *rising and setting* (6 tokens), although in one case the coordinated *-ing* forms show verbal, not nominal, properties, and the pair was accordingly classified as a verb + verb binomial. Something similar applies to *rising or setting* (3 tokens), since in one of the examples the *-ing* forms are not nominal but adjectival:

(6) amplitude is the distance of the ***rising or setting*** *sun or star*
(Gordon, 1726: 71 [4 (1370)])

Here, *rising* and *setting* premodify the following (and very popular) noun + noun binomial *sun or star*; the two *-ing* forms in (6) were therefore grouped under the adjective + adjective pairs.

Apart from those more formulaic nominal pairs, a few nouns are attested in different binomials, as either the first or second element. This, according to Sauer (2018b: 72), shows the flexibility of such combinations as well as the tendency of some nouns to be more often used in binomials than others. Some examples follow:

a. magnitude: *number and magnitude, magnitude and brightness, magnitude or splendor*;
b. distance: *distances and diameters, periods and mean distances, deviation or distance*;
c. motion: *motion and gravity, phases and motions*;
d. heaven: *heavens and earth, earth and heavens, heavens or spheres*;
e. star: *stars and constellations, star or constellation, constellations or sets of stars, circles and stars, signs or stars*;
f. earth: *earth and water, earth or water, sun and earth, earth and sun*;
g. day: *month and day, day or night*;
h. planet: *sun and planet, planet and sun, stars or planets*;
i. place: *order and place, place and annual motion, place or line*;
j. sphere: *sphere and globe, orb or sphere*;
k. globe: *hole and globe, globes or stars*;
l. sun: *sun and stars, sun or moon*;
m. moon: *moons or satellites, satellites or moons*;

n. latitude: *longitude and latitude, latitude or declination*;
o. degree: *signs and degrees, parts or degrees*.

All the examples listed so far, and most of the pairs analyzed, are prototypical binomials, in that they illustrate the basic noun + noun structure. However, there are other less prototypical combinations that show variations in this basic pattern and whose structure is more complex. Examples include *the arctic* ***circle and*** *the antarctic* ***circle*** (the noun *circle* is preceded by the adjectives *arctic* and *antarctic*); *longest* ***day and*** *shortest* ***night***; also *shortest* ***day, and*** *longest* ***night*** (here two superlative adjectives precede *day* and *night*); in ***mother*** *(of Andromeda)* ***and wife*** *(of Cepheus)*, two PPs follow the nouns *mother* and *wife* respectively, while in *new stars, or stars new appearing*, the first noun (or rather NP) seems to be modified by the second.[10]

Nouns are followed by adjectives less frequently in the corpus (159 types, 172 tokens). Examples include *east and west, terrestrial and celestial, solid and liquid, blank or plain, spiral or elliptical, parabolic or elliptical, ptolemean or aristotelian, ptolemean or copernican*. Among the binomials that occur more than once, we find *north or south* (6 occurrences), *north and south* (3), and *right and oblique* (3). Some of the adjectives that are present in different pairs, either as the first or the second element, are:

a. true: *true and proper, true and apparent*;
b. bright: *greatest and most bright, great and bright, large and bright*;
c. plain: *brief and plain, plain and concise, blank or plain*;
d. long: *shortest and longest, long and winding*;
e. strong: *more strong and lively, strong and vivid, strong or furious*;
f. fixed: *fixed and inmoveable, fixed and erratic, fixed and unmoved*;
g. certain: *definite and certain, certain and definite, most certain and infallible*.

Forms in *-ing*/*-ed* were classified either as adjectival (Example (7)) or verbal (8) according to the context:

> (7) To conclude, in this they are particularly distinguished from the planets, which have no such *twinkling or glimmering* light. (Curson, 1702: 349 [0 (2070)])
>
> (8) These catalogues have been much *increased and corrected* by our modern astronomers. (Gordon, 1726: 67 [4 (676)])

Apart from basic adjective + adjective pairs, there are examples of more marginal combinations, such as two different PPs following the adjectival *-ed* form and the adjective *erratic*, respectively, as in (9), or pairs of adjectives preceded by the adverb *partly*, as in (10) and (11):

---

[10] My most sincere gratitude to Prof. Hans Sauer (personal communication) for his useful advice on how to deal with more complex patterns, and also for his invaluable help to clarify and classify some marginal or borderline cases (cf. Section 2.2).

(9) ***fixed** as the stars, **or erratic** as the planets* (Morden, 1702: 19 [1 (4496)])

(10) *partly **earthy and** partly **fiery*** (Morden, 1702: 19 [1 (4572)])

(11) *partly **liquid and** partly **solid*** (Morden, 1702: 20 [1 (4609)])

The third most frequent group, after nouns and adjectives, is that of verb + verb binomials (50 types, 61 tokens). Almost all verb + verb pairs occur only once, e.g. *stands and struts, shift and change, burn and consume, twinkle or glimmer, work or perform, bissected or dichotomized*, with the exception of *rise and set*, which occurs on 9 occasions (also *rises and sets* and *rises or sets*), and *change and enlarge*, which is attested twice. There are no cases of binomials consisting of auxiliary verbs in the corpus, but pairs consisting of verb + phrasal verb/phrasal verb + verb (e.g. *to enumerate and set down, carried off or destroyed*) are occasionally attested.

Other word classes, such as prepositions, adverbs and numerals,[11] are found far more rarely (a total of 38 types and 43 tokens here), including *before and after, before or after* (preposition + preposition), *certainly and infallibly, really or actually* (adverb + adverb), *thirds and fourths, 5 or 6* (numeral + numeral). No instances of binomials combining two pronouns were found.

Table 5 below shows the word class of the constituents in binomials for the astronomy texts:

Table 5. Word classes of constituents distributed per author in CETA (types and tokens)

Word class	**Curson**	**Morden**	**Whiston**	**Harris**	**Gordon**	**Total**
	Types/ Tokens	**Types/ Tokens**	**Types/ Tokens**	**Types/ Tokens**	**Types/ Tokens**	**Types/ Tokens**
Noun + Noun	118/142	108/123	52/56	77/84	67/95	422/500
Adj. + Adj.	37/38	54/56	28/30	19/22	21/26	159/172
Verb + Verb	7/9	9/9	9/10	15/19	10/14	50/61
Other	4/4	3/3	7/7	15/20	9/9	38/43
TOTAL	**166/193**	**174/191**	**96/103**	**126/145**	**107/144**	**669/776**

In terms of connection or coordination, *and* is the most frequent conjunction, present in almost 400 out of the 669 types retrieved (460 tokens). *Or*, although less commonly recorded, is also very frequent, occurring in 265 types (311 tokens). Other connectors, including *nor* (*heat nor cold*; *East nor West*; *no other **use nor end***) are rare (only 5 tokens overall). It could be argued that the high frequency of *or* as a connector is somehow register-dependent, in that an astronomical/astrological term often needs further clarification and authors provide an alternative term or an explanation, as in *Cazimi or the heart of the sun* (Curson, 1702: 385 [0 (7923)]). Cazimi is a Medieval astrological term referring to planets that are so close to a conjunction with the sun that they are, almost literally, "in the heart" of the sun. It is a Latin transliteration of the Arabic term *kaṣmīmī*, "as if in the heart"; the Arabic term itself appears to be a translation of the Greek *egkardios* (ἐγκάρδιος), "in the heart". Other, similar examples include *Via Lactea, or Milky Way* (Morden, 1702: 38 [1 (9061)]), and *Aries, or the Ram* (Harris, 1719: 30 [3 (5681)]), in which Latin loan-words (*Via Lactea, Aries*) are explained or translated by means of native ones (*Milky Way* and *Ram*, respectively; cf. Section 2.3 above).

---

[11] I have classified numerals apart, as with adverbs, prepositions, etc., although some numerals could clearly have been classified along with adjective + adjective binomials (cf. Sauer, 2018a: 89).

Finally, it is worth noting that binomials (nominal and adjectival) are used quite consistently in the (long) titles of the works under scrutiny, as in Curson's (1702) *The theory of sciences illustrated; or, the* ***grounds and principles*** *of the seven liberal arts*; also in Morden's (1702) *An Introduction to astronomy, geography, navigation, and other mathematical sciences made easie by the* ***description and uses*** *of the* ***cœlestial and terrestrial*** *Globes*, and in Gordon's *An introduction to geography, astronomy, and dialling. Containing the most useful elements of the said sciences, adapted to the meanest capacity, by the* ***description and uses*** *of the* ***terrestrial and celestial*** *globes.*[12]

4.1.2 *Semantic structure*

As noted in Section 2.4, three broad categories can be distinguished to account for the semantic relation between the elements of a binomial, synonymy, antonymy and complementarity, the latter comprising pairs that are neither synonymous nor antonymous and having many subgroups. As already pointed out (cf. Section 2.4 above), many of the antonymous pairs can also be classified as complementary in that they refer to parts of a whole, that is, to a higher unity. Thus, the binomial *men and women* refers to people in general, to Mankind, and the combination *day and night* refers to the whole day, a period of twenty-four hours (cf. Section 2.4 above).

A relatively large number of binomials cannot be included here, these mainly pairs of names, prepositions and conjunctions, "whose semantic structure could not be analysed in terms of the semantic categories" (Mollin, 2014: 36), and also because the relation between the elements is difficult to define. Thus, in *England and France*, it is clear that the two countries have had close relations throughout their histories, either as allies or as enemies, but trying to assign a specific semantic relation to this pair (and to many others) is a complex issue.

---

[12] Some of these titles also include multinomials, not discussed in this paper. Moreover, the binomial *description and uses* appears in the titles of both Morden's and Gordon's works.

Words in synonymous pairs convey *very* similar meanings, while words in tautologous pairs have *exactly* the same meaning; the distinction between synonymy and tautology is, however, far from transparent, and the differences between the two remain quite elusive. Some examples of synonymous binomials (whether tautologous or not) are:

a. nouns: *vapours and exhalations* (also *exhalations and vapours*), *way or course, dish or plate, shape and form, girdles or belts, greatness or bigness, fairs or markets, compound and not simple* (synonymy is here indicated by negating the second adjective), *60 minutes, or one hour*; also *a year, or 365 days/twelve months, or a year's time* (the second element explains or makes the first one more precise), *magnet, or loadstone* (a loan-word is explained or translated by means of a native word; cf. Section 2.3),[13] a ***scarcity*** *of olives* ***and dearth*** *of oil* (a split synonymous binomial);
b. adjectives: *twinkling or glimmering, improper or offensive, favourable and auspicious, vain and false, secret and hidden, fixed and inmoveable, quick and swift*;
c. verbs: *twinkle and sparkle, twinkle or glimmer, shift and change, fix and set, elevate or raise, vanish and become invisible*.

In turn, examples of antonymous binomials, i.e. pairs of words with opposite meanings, include:

a. nouns: *day and night* (=the whole day), *heaven and earth* (=the entire creation), *ascension or descension, heat nor cold, light and darkness, friends or enemies, appearing and disappearing* (antonymy is here conveyed by means of the negative prefix *dis-* in the second element);
b. adjectives: *solid and liquid, shortest and longest, good or bad, lucky and unlucky* (antonymy is again expressed by means of a negative prefix, *un-*, in the second element; cf. Sauer 2018a: 99), *occult and latent, great or small, artificial and natural, north and south, increasing and decreasing*;
c. verbs: *rise and set* (antonymous verb + verb pairs are very infrequent in the CETA corpus).

Finally, complementary relations between the members of the pairs include, among others, the following (Sauer, 2018a: 99–101; Sauer, 2018b: 70–71):

a. parts of a whole or co-hyponyms in a semantic field, e.g. *astronomy and geography, philosophers and astronomers, astronomers and astrologers* (referring to branches of science), *morning or evening, days and weeks, months and years, girdle and sword, hills and valleys, Greeks and Romans*;

---

[13] Cf. also Section 4.1.1 on the high frequency of *or* as a connector of elements in binomials in astronomy texts.

b. generally positive binomials: *love and friendship, master and teacher, harmony, and proportion, care and diligence, diligent and faithful, right and orderly*;
c. generally negative binomials: *theft and cruelty, rapture and violence; bold or furious, terrible and threatening, dull and tedious, silly and crampt, vain and foolish*;
d. cause-effect or effect-cause: *burn and consume, revived and improved*;
e. more general to more specific: *ceremonies and duckings, miles or furlongs*;
f. more specific to more general: *tempest, and terrible weather*;
g. sequence of actions: *enumerate and set down, discover and show, made and finished up, stands and struts, heard and read.*

4.2 Binomials in the *Corpus of History English Texts* (CHET)

The period running from 1700 to 1900 shows "a general change in society" (Moskowich, 2017: 85) and coincides with the beginning of the Restoration period and the founding of the Royal Society. All the texts in the beta version of the *Corpus of History English Texts* (CHET) belong to the realm of history writing, a discipline which has been defined in different ways and from different perspectives over the course of time. Agreement on when historicism and history writing emerged seems not to have been reached, and while some accounts claim that its beginnings can be found "in the Age of Reason" (Moskowich, 2017: 88), others allude to the importance of Gibbon's *The history of the decline and fall of the Roman Empire* (1788), on the basis of the author's use of historical sources rather than relying on previous accounts of the same historical facts. The tenet that "History should in fact be exact, impartial and critical" (Moskowich, 2017: 89) was maintained in the nineteenth century, a century in which Historiography, that is, "the methods and techniques used to describe those recorded past events" (Moskowich, 2017: 89) took its modern form and came to be distinguished from History itself, "the facts occurring in the past and somehow recorded" (Moskowich, 2017: 89). However, neither historiography nor history were regarded "real" sciences in the eighteenth and nineteenth centuries (cf. Moskowich, 2016: 113).

In what follows, the use of binomials in CHET is described according to the same parameters used for their description in CETA (cf. Section 4.1 above). A total of 569 different pairs or types were considered, although the number of tokens here is almost 700 (see Table 6 below), which again seems to be a clear indication that *ad hoc* binomials, i.e. those occurring just once, are more frequently used than fixed, idiomatic ones (cf. Section 2.1 above).

4.2.1 *Word class of constituents and connectors*

As was the case in CETA, the majority of binomials in CHET are made up of nouns. A total of 369 different noun + noun combinations (477 tokens) were recorded in the corpus, such as *law and justice, towns and castle, accusation and impeachment, princes or kings, offence or displeasure, land or house*, etc. These examples clearly illustrate how the nouns forming nominal pairs in CHET differ from those in CETA, in that they do not refer to the same topics (stars, constellations, planets), but to issues such as justice, right, order, government, kingdom, crown, etc. Moreover, combinations of names (*Lancaster and York, Essex and London*), once again retrieved as a special subgroup within noun + noun binomials (cf. Section 4.1.1 above), are far less frequent than in astronomy texts. Most pairs of nouns are attested only once, but others are repeated twice or more, among these *day and place* (8 tokens), *king and cardinal* (5), *sect and learning* (5), *commission and dispensation* (5), *reader and teacher* (4), *crown and kingdom* (3), *men and women* (3), *daughter and heiress* (3). A very special case is that of *son and heir*, which occurs on 50 different occasions in Crawfurd's *A genealogical history of the royal and illustrious family of the Stewarts* (1710). In fact, out of the 97 tokens of noun + noun binomials recorded from this work, 50 are instances of *son and heir*, a pair that shows a high degree of fixedness and formulaicity, at least at the time when the text was written.

In addition, as was the case in CETA, a few nouns seem to be recurrent in different binomials, either as the first or second element:

a. justice: *right and justice, legality and justice, law and justice*;
b. king: *king and government, king and nobles*;
c. power: *power and glory, heart, or power, privileges and powers*;
d. (e)state: *estate and power, royal state and liberty, royal estate or liberty*;
e. will: *will and command, will and assent, will and favour*;
f. honour: *honour and safety, honour and respect, honour, or revenue*;
g. friend: *friends and servants, friends and followers, friends and relations, relations and friends*;
h. charter: *charters and deeds, charters or deeds, charters and writings, charters and grants*;
i. successor: *son and successor, brother and successor, monks and successors*;
j. error: *errors and superstitions, articles and errors, errors and opinions*.

Likewise, examples illustrative of less well-defined and more marginal binomials, with intervening elements before and after the nouns, also occur in CHET: *the same* ***demands, and*** *much the same* ***offers****, his royal* ***benignity, and*** *gracious* ***seigniory****, court-****intrigues or*** *popular* ***factions*** (Tyrrell, 1704: 951 [122 (192)]), ***grants*** *of the kings of England,* ***and bulls*** *of Popes* (Anderson, 1705: 56 [123 (1480)]), *a forged*

deed, or** a false **witness (Oldmixon, 1716: 53 [125 (3195)]), *public **notary, and** scribe of the acts* (Strype, 1721: 77 [126 (1882)]).

Following nouns, adjectives constitute the second most frequent word class (103 types, 116 tokens). Examples here include *generous and communicative, great and noble, valuable and acceptable, true or false, dogmatic or daring, guilty or blameworthy*. Among the pairs that occur more than once, we find *spiritual and temporal* (5 tokens), *learned and worthy* (3; also *worthy and learned*), *diverse and many* (3), *ancient and modern* (2), *noble and ancient* (2), *genuine and true* (2; also *true and genuine*), *reverend and learned* (2). Additional examples of adjectives present in different binomials are:

a. great: *great and mighty, great and eminent, greater and smaller*;
b. full: *entire and full, full and complete*;
c. arbitrary: *arbitrary and illegal, illegal and arbitrary, violent and arbitrary*;
d. false: *false and unfaithful, counterfit and false*.

Examples of *-ing* and *-ed* forms classified as adjectives include *honoured and knowing* (Example (12)), while the two coordinated *-ed* forms in Example (13) fulfil a verbal function, and hence the pair was classified as a verb + verb binomial:

(12) by the *honoured and knowing* Sir James Dalrymple.
(Anderson, 1705: 57 [123 (1661)])

(13) In this treasury is to be found, a great number of charters, *granted or witnessed* by the predecessors or branches, of many of our ancient families.
(Anderson, 1705: 54 [123 (1271)])

As in CETA, the third most frequently attested word class, after nouns and adjectives, is that of verbs (75 types, 80 tokens), as in *place and rank, go and come, supply and reform, tarry or abide, molest or trouble, invest or inventory*. Very few pairs are repeated more than once, among them *taken and reputed*, which occurs on three occasions (the reverse pair *reputed and taken* occurs once), *adjudged and declared* (2), *admonished and exhorted* (2; *exhorted and admonished* also occurs once), and *said and affirmed* (2). Pairs including auxiliary verbs were not found in CHET.

Other word classes, such as prepositions (*in and unto*), adverbs (*here and there*) and numerals (*second and third*), are far less frequent (22 types, 25 tokens), while no instances of binomials consisting of two pronouns were recorded.

The distribution of the different word classes of the elements found in binomials in history texts is shown in Table 6:

Table 6. Word classes of constituents distributed per author in CHET (types and tokens)

Word class	Tyrrell	Anderson	Crawfurd	Oldmixon	Strype	Total
	Types/ Tokens	Types/ Tokens	Types/ Tokens	Types/ Tokens	Types/ Tokens	Types/ Tokens
Noun + Noun	95/103	96/104	36/97	61/64	81/109	369/477
Adj. + Adj.	27/31	39/46	13/13	6/6	18/20	103/116
Verb + Verb	22/23	16/16	3/3	6/6	28/32	75/80
Others	3/3	4/4		7/8	8/10	22/25
TOTAL	**147/160**	**155/170**	**52/113**	**80/84**	**135/171**	**569/698**

Regarding connectors, *and* is again the most frequent conjunction, occurring in 489 types (613 tokens). While in astronomy texts *or* was very common, present in 265 types and 311 tokens, it occurs far less frequently in history texts, with only 75 types (80 tokens). Other connectors (e.g. *nor* in *no honour, nor worship*) are very rare (only 5 instances overall), another example here being *donors & witnesses* (Anderson, 1705: 55 [123 (1403)]), the two nouns connected by means of an ampersand.

As was the case in CETA, noun + noun and adjective + adjective binomials (and multinomials) also have a strong presence in the (long) titles of the works selected from CHET. Thus, Tyrrell's *The general history of England, both **ecclesiastical and civil**: Containing the reign of Richard II*; Anderson's *An historical essay, shewing that the **crown and kingdom** of Scotland, is **imperial and independent***; and Crawfurd's very long *A genealogical history of the **royal and illustrious** family of the Stewarts, from the year 1034 to the year 1710. Giving an account of the lives, marriages and issue of the most remarkable **persons and families** of that name. To which are prefixed, first, a general description of the shire of Renfrew, the **peculiar residence and ancient patrimony** of the Stewarts: and, secondly, a deduction of the **noble and ancient** families, proprietors there for upwards of 400 Years, down to the present times: Containing the descent, original creations, and most remarkable actions of their respective ancestors; also the chief titles of honour they now enjoy; with their **marriage and issue**, continued down to this present year, and the coat of arms of each family in blazon.*

4.2.2 *Semantic structure*

In what follows, some examples of synonymous (tautologous or not) binomials, and pairs consisting of antonyms, are listed. As in CETA, some of the most illustrative binomials from CHET are included here:

a. synonymous noun + noun pairs: *nobility and gentry, treaty, or agreement, knights, or gentlemen, plots and conspiracies, subsidies and aids, repertory or inventory, notes or transcripts, proofs and evidences, fable and forgery, hurry and haste, tyranny and oppression, asses and mules, quietness and tranquility, weaver, or maker of cloth* (in this case, the second element provides an explanation of the first noun);
b. synonymous adjective + adjective pairs: *false and unfaithful, entire and full, full and complete, true and genuine, genuine and true, true and undoubted, counterfeit and false, diverse and sundry, guilty, or blameworthy*;
c. synonymous verb + verb pairs: *saved and preserved, mistaken, or erred, exalt and aggrandize, ratifies and confirms, encouraged and fomented, molest or trouble*;
d. antonymous noun + noun pairs: *lords and ladies, king and queen, temporality and spirituality, sea and land, land and sea, ancestors and descendents, day and night, men and women*;[14]
e. antonymous adjective + adjective pairs: *ecclesiastical and civil, civil and ecclesiastical, true or false, ancient and modern, greater and smaller*;
f. antonymous verb + verb pairs: *go and come* (again, antonymous verb + verb pairs are hardly attested in the CHET corpus).

Complementary relations, in turn, are illustrated below:

a. parts of a whole or co-hyponyms in a semantic field: *philosophy and mathematics, prior and convent, judge and jury, camp and court*;
b. generally positive binomials: *assistance and protection, respect and amity, prosperous and happy, dutiful and loyal, wise and moderate*;
c. generally negative binomials: *melancholy and discontent, violence, and treachery, destroyed or pillaged, feared and hated*;
d. cause-effect or effect-cause: *sickness and death, cost and charges, marriage, and coronation* (also a sequence of actions), *recovered and lived*;
e. more specific to more general: *orchards and gardens; jewels, and other moveables*; *briberies, and other sinister practices*;
f. sequence of actions: *met and saluted, place and rank, reigned and died, embalm and aromatize, condemned and executed, killed and eaten, pursued and found*.

---

[14] Note that most pairs listed here could also be classified as complementary (cf. Sections 2.4 and 4.1.2).

5. Final remarks

The analysis of binomials in eighteenth-century astronomy and history texts retrieved from CETA and CHET, two components of the *Coruña Corpus of English Scientific Writing*, has shown that this is a frequent linguistic phenomenon, used by different authors writing on a range of scientific topics. Thus, in the five works selected from CETA, a total of 669 types (776 tokens) were recorded, while a smaller number of types (569) and tokens (698) were recorded in the five works selected from CHET.

The majority of binomials retrieved from the two subcorpora can be regarded as instances of *hapax legomena*, that is, *ad hoc* combinations created by the authors on the spur of the moment and used only once. Of course, there are also many examples of pairs that occur two or more times. In astronomy texts, for example, *sun or* star is attested on 19 occasions, followed by *day and night* (7), *body or aspect* (7), *earth and planet* (6), and *sun and moon* (6). In history texts we also find recurring patterns, whose integrating elements clearly differ from those attested in astronomy texts, especially in the case of noun + noun combinations, as can be seen in pairs such as *day and place* (8 tokens), *king and cardinal* (5), *sect and learning* (5), *commission and dispensation* (5), *reader and teacher* (4). A very special case in CHET is that of *son and heir*, used by the same writer on 50 different occasions; this could serve as an illustration of a rather fixed, formulaic, even idiomatic kind of binomial. Moreover, the fact that a few nouns may occur in different binomials, either as the first element or the second, shows the flexibility of some of these combinations, as well as the tendency of certain nouns to be more often used in binomials than others (cf. Sauer, 2018b: 72).

The majority of the examples retrieved from both subcorpora are cases of nominal binomials (422 types, 500 tokens in CETA; 369 types, 477 tokens in CHET), followed by adjectives (159 types, 172 tokens in CETA; 103 types, 116 tokens in CHET) and verbs (50 types, 61 tokens in CETA; 75 types, 80 tokens in CHET). Other combinations, including pairs of adverbs, prepositions or numerals are, in turn, much rarer, or non-existent (pairs of auxiliary verbs or pronouns). In addition, binomials consisting of names (e.g. *England and France*) are far more frequently attested in CETA than in CHET.

And is by far the most frequent conjunction used to connect elements in binomials, followed by *or*; others, such as *but, neither…nor* and ampersand, appear far less frequently. An interesting finding here is that connecting *or* occurs much more often in astronomy than in history texts. This is due to the fact that in many cases the writer needs to explain or clarify a previous term by using an alternative form, or simply by translating it into his native language, which would justify the choice of *or* to connect the two members of the pair.

Semantically speaking, three broad categories were distinguished to account for the semantic relation established between the elements in the binomials, namely synonymy (whether exact or not), antonymy and complementarity, the latter comprising pairs that are neither synonymous nor antonymous but somehow related, and whose elements may show special relations, such as cause and effect, effect and cause, part-whole, co-hyponymy, etc. It is often the case that antonymous pairs are also complementary, in that they refer to parts of a whole, to a higher unity (e.g. *heaven and earth*).

The fact that binomials constitute a frequent linguistic phenomenon in both astronomy and history texts makes them ideal candidates for further research. Questions such as alliteration and repetition, function and motivation, etymologic structure (i.e. whether these pairs are made up of loan-words, native words, or a combination of both), and (ir)reversibility, among others, might usefully be explored in order to obtain a more complete picture and better understanding of their use, not only in the late Modern English period but also in other periods of the history of English.

Acknowledgements

For generous financial support I am grateful to the Spanish Ministry of Science, Innovation and Universities (grant FFI2017-86884-P) and the Regional Government of Galicia (grants ED431B 2017/12 and ED431D 2017/09). I am most indebted to Professor Hans Sauer for his generosity and helpful advice on binomials. Thanks are also due to Elena Seoane for her valuable comments on an earlier version of this paper.

References

Bach, Ulrich. (2017). ‘I do make and ordayne this my last wyll and testament in manner and forme Folowing’: Functions of binomials in Early Modern English protestant wills”. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (222–240). Cambridge: Cambridge University Press.

Bhatia, Vijay K. (1993). *Analysing genre: Language use in professional settings*. London: Longman.

Cantos, Pascual & Nila Vázquez. (2012). Subject specific vocabulary in astronomy texts: A diachronic survey of the *Corpus of English Texts on Astronomy*. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy ‘playne and simple’: The writing of science between 1700 and 1900* (123–153). Amsterdam/Philadelphia: John Benjamins.

Chapman, Don. (2017). Fixity and flexibility in Wulfstan’s binomials. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (41–62). Cambridge: Cambridge University Press.

Crespo, Begoña. (2012). Astronomical discourse in 18th and 19th century texts: A new-born model in the transmission of science. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy 'playne and simple': The writing of science between 1700 and 1900* (57–77). Amsterdam/Philadelphia: John Benjamins.

Crespo, Begoña & Isabel Moskowich. (2010). CETA in the context of the *Coruña Corpus*. *Literary and Linguistic Computing*, 25(2), 153–164. https://doi.org/10.1093/llc/fqp038

Danet, Brenda. (1984). The magic flute: A prosodic analysis of binomial expressions in legal Hebrew. *Text*, 4, 143–172. https://doi.org/10.1515/text.1.1984.4.1-3.143

Doty, Kathleen L. & Mark Wicklund. (2017). 'Shee gave Selfe both Soule and body to the Devill': The use of binomials in the Salem Witchcraft Trials. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (241–260). Cambridge: Cambridge University Press.

Frade, Celina. (2005). Legal multinomials: Recovering possible meanings from vague tags. In Vijay K. Bhatia, Jan Engberg, Maurizio Gotti, & Dorothee Heller (Eds.), *Vagueness in normative texts* (133–156). Bern: Peter Lang.

Fulk, Robert D. (2017). Pragmatic and stylistic functions of binomials in Old English. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (27–40). Cambridge: Cambridge University Press.

Gorgis, Dinha T. & Yaser Al-Tamimi. (2015). Binomials in Iraqi and Jordanian Arabic. *Journal of Language and Linguistics*, 4(2), 135–151.

Gustafsson, Marita. (1984). The syntactic features of binomial expression in legal English. *Text*, 4/1–3, 123–141.

Iglesias-Rábade, Luis. (2007). Twin lexical collocations in legal Late Middle English. *Studia Anglia Posnaniensia*, 43, 17–47.

Klégr, Aleš (1991). A note on binomials in English and Czech. *Prague Studies in English*, 19, 83–88.

Klégr, Aleš & Jan Čermák. (2008). Binomials in an historical English literary perspective: Shakespeare, Chaucer, Beowulf. In Martin Procházka & Jan Čermák (Eds.), *Shakespeare between the Middle Ages and modernity: From translator's art to academic discourse. A tribute to Professor Martin Hilský*. MBE (40–62). Prague: Charles University, Faculty of Arts.

Koch, Barbara Johnstone. (1983). Arabic lexical couplets and the evolution of synonymy. *General Linguistics*, 23, 51–61.

Kopaczyk, Joanna. (2009). (Multi-word) units of meaning in 16th-century legal Scots. In Rod W. McConchie, Alpo Honkapohja & Jukka Tyrkkö (Eds.), *Selected Proceedings of the 2008 Symposium on New Approaches in English Historical Lexis* (88–95). Somerville, MA: Cascadilla Proceedings Project.

Kopaczyk, Joanna. (2012). Repetitive and therefore fixed? Lemmatic bundles and text-type standardization in 15th-century administrative Scots. In Hans Sauer & Gaby Waxenberger (Eds.), *English Historical Linguistics 2008. Vol. 2: Words, texts and genres* (189–209). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.324.11kop

Kopaczyk, Joanna & Hans Sauer (Eds.) (2017). *Binomials in the history of English: Fixed and flexible*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316339770

Kotake, Tadashi. (2017). Binomials or not? Double glosses in Farman's gloss to the *Rushworth Gospels*. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (82–97). Cambridge: Cambridge University Press.

Krygier, Marcin. (2017). Binomial glosses in translation: The case of the *Wycliffite Bible*. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (159–172). Cambridge: Cambridge University Press.

Kubaschewski, Elisabeth. (2017). Binomials in Caxton's *Ovid* (Book I). In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (141–158). Cambridge: Cambridge University Press.

Lambrecht, Knud. (1984). Formulaicity, frame semantics, and pragmatics in German binomial expressions. *Language*, 60(4), 753–796. https://doi.org/10.1353/lan.1984.0004

Lehto, Anu. (2017). Binomials and multinomials in Early Modern English parliamentary acts. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (261–278). Cambridge: Cambridge University Press.

Malkiel, Yakov. (1959). Studies in irreversible binomials. *Lingua*, 8, 113–160. https://doi.org/10.1016/0024-3841(59)90018-X

Mmadike, Benjamin I. (2014). *A study of Igbo gender-based irreversible binomials*. Nnamdi Azikiwe University, Open Science Repository.

Mollin, Sandra. (2014). *The (Ir)rerversibility of English Binomials. Corpus, Constraints, Developments*. Amsterdam/Philadelphia: John Benjamins.

Mollin, Sandra. (2017). Developments in the frequency of English binomials 1600–2000. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the History of English: Fixed and flexible* (281–295). Cambridge: Cambridge University Press.

Moskowich, Isabel. (2016). Lexical richness in modern women writers: Evidence from the *Corpus of History English Texts*. *Revista Canaria de Estudios Ingleses*, 72, 111–128.

Moskowich, Isabel. (2017). Genre and change in the *Corpus of History English Texts*. *Nordic Journal of English Studies*, 16(3), 84–106.

Moskowich, Isabel & Begoña Crespo (Eds.) (2012). *Astronomy 'playne and simple': The writing of science between 1700 and 1900*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.173

Moskowich, Isabel, Inés Lareo, Gonzalo Camiña-Rioboo & Begoña Crespo (Comp.) (2012). *A Corpus of English Texts on Astronomy (CETA)*. Amsterdam: John Benjamins.

Ogura, Michiko. (2017). Binomials, word pairs and variation as a feature of style in Old English poetry. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (63–81). Cambridge: Cambridge University Press.

Okamoto, Shigeko. (1993). Nominal repetitive constructions in Japanese: The 'tautology' controversy revisited. *Journal of Pragmatics*, 20, 433–466. https://doi.org/10.1016/0378-2166(93)90039-R

Rodríguez-Sánchez, Ignacio. (2013). Frequency and specialization in Spanish binomials *N y N*. *Procedia-Social and Behavioral Sciences*, 95, 284–292. https://doi.org/10.1016/j.sbspro.2013.10.649

Rutkowska, Hanna. (2017). Binomials in several editions of the *Kalender of Shepherdes*, an Early Modern English almanac. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (175–200). Cambridge: Cambridge University Press.

Sauer, Hans. (2014). Twin-formulae and more in late Middle English: *The Historye of the Patriarks*, Caxton's *Ovid*, Pecock's *Donet*. In Michael Bilynsky (Ed.), *Studies in Middle English: Words, Forms, Senses and Texts* (25–46). Frankfurt am Main: Peter Lang.

Sauer, Hans. (2017). Milton's binomials and multinomials in *Samson Agonistes*. In Artur Kijak, Andrzej M. Łęcki & Jerzy Nykiel (Eds.), *Current Developments in English Historical Linguistics. Studies in Honour of Rafal Molencki* (275–303). Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Sauer, Hans. (2018a). Binomials in the Middle English and Early Modern English versions of Bocaccio's *De mulieribus claris*. In Margaret Connolly & Raluca Radulescu (Eds.), *Editing and Interpretation of Middle English Texts. Essays in Honour of William Marx* (83–105). Turnhout, Belgium: Brepols Publishers. https://doi.org/10.1484/M.TT-EB.5.114035

Sauer, Hans. (2018b). Flexible and formulaic: Binomials and multinomials in the Late Middle English *The Wise Book of Philosophy and Astronomy*. *Acta Philologica*, 50, 61–78.

Sauer, Hans & Birgit Schwan. (2017a). Heaven and earth, good and bad, answered and said: A survey of English binomials and multinomials (part I). *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 134 (1), 93–96.

Sauer, Hans & Birgit Schwan. (2017b). Heaven and earth, good and bad, answered and said: A survey of English binomials and multinomials (part II). *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 134 (2), 185–204.

Schaefer, Ursula. (2017). On the linguistic and social development of a binomial: The example of *to have and to hold*. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (322–343). Cambridge: Cambridge University Press.

Schenk, Ulrike. (2017). Binomials in Middle English poetry: *Havelok, Ywain and Gawain, The Canterbury Tales*. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (125–140). Cambridge: Cambridge University Press.

Sprau, Melanie. (2017). Binomials and multinomials in Sir Thomas Elyot's *The Boke Named the Gouernour*. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (201–221). Cambridge: Cambridge University Press.

Tyrkkö, Jukka. (2017). Binomials in English novels of the Late Modern Period: Fixedness, formulaicity and style. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (296–231). Cambridge: Cambridge University Press.

Zagórska, Paulina. (2017). Lexical pairs and their function in the *Eadwine Psalter* manuscript. In Joanna Kopaczyk & Hans Sauer (Eds.), *Binomials in the history of English: Fixed and flexible* (98–121). Cambridge: Cambridge University Press.

CHAPTER 7

“Were this eſtimation, however, to be depended on”

Inversion conditionals as evidence of paradigmatic change in CHET

Luis Puente-Castelo
Universidade da Coruña

1. Introduction

The second half of the seventeenth century witnessed the emergence of alternatives to Scholasticism, the main paradigm of knowledge during the Middle Ages. The rise of Empiricism, which focused on the real world (instead of classical texts) as the object of science, constituted the first step towards science as it is understood at present, and led to an important increase in the production of knowledge. These changes in science were reflected in scientific discourse, as writers abandoned the rigid, logic-based, and introverted scholastic forms of argumentation in favour of a new model, the so-called “rhetoric of immediate experience” (Atkinson, 1996: 335), characterised by extensive narrations of phenomena from the physical world.

However, perhaps the most important change to science and scientific discourse was that it became a social endeavour: reference to authoritative texts was no longer a sufficient claim to veracity; rather, authors felt the need to persuade their colleagues of the validity of their claims (Bazerman, 1988; Allen, Qin & Lancaster, 1994; Atkinson, 1996, 1999). From that moment, scientific discourse has been characterised by being interpersonal in nature, moving beyond the mere statement of factual content and introducing a series of discourse strategies that seek to achieve the best possible reception for one’s ideas among the potential audience (Myers, 1989; Swales, 1990), strategies such as avoiding confrontation and categorical statements, trying to reach consensus, and being humble and polite (Hyland, 1996, 1998, 2000). These in turn correspond to a series of linguistic features, such as the use of modal verbs, *verba dicendi*, and probability adverbs, among others.

Conditionals are one of these strategies. They have been defined as “a rhetorical device for gaining acceptance for one’s claims”, by means of which “scientists try to

https://doi.org/10.1075/z.225.07pue

© 2019 John Benjamins Publishing Company

reach a consensus with their readers" (Warchal, 2010: 141). Conditionals are considered to be a particularly valuable resource in scientific discourse (Carter-Thomas & Rowley-Jolivet, 2008: 191), in that they are enormously versatile, in both formal and functional terms. Inversion conditionals are a prime example of this versatility, as they represent a wholly different strategy from that of the most usual conditional subordinators, and a strategy which has persisted in use until the present day. Moreover, previous studies (Puente-Castelo, 2017) have found that the evolution in the use of inversion conditionals over time in several disciplines (astronomy, philosophy, life sciences) shows interesting parallelisms with the evolution in the scientific practices in each of these disciplines, and, more particularly, in the paradigmatic change from Scholasticism towards contemporary science.

The aim of this paper, then, is to try to ascertain whether the evolution of the use of inversion conditionals in the beta version of CHET follows the same pattern as in CETA, CEPhiT and CELiST,[1] the subcorpora on astronomy, philosophy, and life sciences, respectively, of the *Coruña Corpus of English Scientific Writing*. In order to do so, selected conditional particles will be searched in CHET, and the results of the search will then be classified according to different linguistic and extra-linguistic parameters. The results will be analysed looking in particular at the diachronic evolution in the use of inversion conditionals over time.

In what follows, Section 2 presents an analysis of the functional versatility of conditionals, while Section 3 focuses on their formal variability. Section 4 discusses previous work (Puente-Castelo, 2017) that shows the parallelisms between inversion conditionals and paradigmatic evolution. Section 5 then presents the corpus and the methodology used, Section 6 examines the results, and Section 7 provides some tentative conclusions.

2. Conditionals and their versatility as a resource for scientific writing

Despite having frequently been pigeonholed into four stereotypical types, the so-called "first, second and third" and "zero" conditionals (Hwang, 1979; Maule, 1988; Fulcher, 1991; Ferguson, 2001; Jones & Waller, 2010), the conditional structures used in real instances of language can in fact be characterised by their very notable variability, both formally and functionally.

The core function of conditionals is to indicate "the relationship between different segments of text and to make the readers recognise this relation" (Warchal, 2010: 146). They establish a link, most frequently causal, between two statements, and they also commonly introduce information about the perceived probability of the link, this by means of the use of particular verb forms or other evaluative linguistic elements, such as frequency adverbs. This has several applications in scientific writing, with conditionals being used to express well-known causal relationships such as mathematical equalities, formulae, or so-called "universal-truths", as shown in (1) below. Moreover, conditionals can also be used to express the relationship between phenomena and their consequences (2), not only in terms of content but also at the text level, this leading to their use to establish links between different sections of the text (3). They can also be used to state pre-requisites or instructions (4).

---

[1] While CETA and CEPhiT have already been published, CELiST is still in a beta version.

(1) Given that $x = y$, then $n(x + a) = n(y + a)$ must also be true. (Quirk *et al.*, 1985: 1090)

(2) Had he not seen the car coming, he would have been killed. (Biezma, 2011: 555)

(3) If perceptions of change had been measured, then the findings may have been different. (Warchal, 2010: 144)

(4) Patients entered the study if they satisfied the WHO criteria for stroke. (Ferguson, 2001: 71)

On the other hand, the inherent non-assertiveness of conditionals (Carter-Thomas & Rowley-Jolivet, 2008: 191) makes them particularly appropriate for formulating hypotheses and theories and, as seen in (5) below, for introducing tentative claims or conclusions (Ferguson, 2001: 61; Gabrielatos, 2010: 1) as well as for evaluating different options or examining the consequences of different courses of action.

(5) If a patient has an early failure from a low anterior resection, they may be able to be retrieved by resection. (Carter-Thomas & Rowley-Jolivet, 2008: 200)

However, once Scholasticism and its model of using authoritative classical sources had been abandoned, authors instead had to persuade each other of the veracity of their claims, as noted above. Thus, conditionals also helped writers to show solidarity and respect towards their peers, both directly by means of the expression of respect, humility and politeness, as in (6) below, and indirectly, in the recognition of the work of others and in the consideration of competing points of view or alternatives to one's own claims (Declerck & Reed, 2001; Carter-Thomas & Rowley-Jolivet, 2008: 191), as in Example (7). In this sense, they can also be used to manage interactions with one's audience as signposting devices, anticipating "the author's intentions and the development of the text" (Carter-Thomas & Rowley-Jolivet, 2008: 194) for the reader, as can be seen in Example (8) below.

(6) If I may be quite frank with you, I don't approve of any concessions to ignorance. (Quirk *et al.*, 1985: 1095)

(7) Even if health care providers are diligent in keeping current with genetic medicine, the interpretation of the results of genetic testing is often complex. (Carter-Thomas & Rowley-Jolivet, 2008: 202)

(8) Now if we go to patients who experienced mucositis toxicity. (Carter-Thomas & Rowley-Jolivet, 2008: 194)

Besides these uses, the inherent non-assertiveness of conditionals is also useful to recognise uncertainty, as well as to avoid categorical expressions by toning down claims, making their validity conditional on the factors expressed in the protasis. An example of this is shown in (9) below. Conditionals are also used to create argumentative spaces in which one's claims hold, acting as space-builders and delimiting the scope of such claims (Fauconnier, 1994; Dancygier, 1998; Carter-Thomas & Rowley-Jolivet, 2008), and thus also avoiding categorical uses, as can be seen in Example (10).

(9) Finally (if this is important), the S1 meaning can be converted into an S meaning to recover a more intuitive object to represent the meaning of the original sentence. (Warchal, 2010: 148)

(10) As such, it can be said to belong to modality if the category is defined as the expression of the speaker's attitude or stance. (Warchal, 2010: 148)

3. Beyond *if*: The formal variability of conditionals

The functional versatility of conditionals is mirrored by their formal variability. Conditionals prototypically consist of two constituents: the protasis, whose presence is obligatory and which contains a conditional marker; and the apodosis, which may be elided. These constituents are prototypically clauses, but they can also be phrases or single words, and they may appear in different positions in relation to one another, with the protasis before, after, or in the middle of the apodosis, or by itself. They can be inserted at different grammatical levels, and they contain a very important number of different combinations of verb forms.

However, despite the fact that many studies on conditionals focus mainly or exclusively on *if* (Werth, 1997; Ford, 1997; Facchinetti, 2001; Ferguson, 2001; Carter-Thomas & Rowley-Jolivet, 2008; Warchal, 2010; Hesabi, Dehaghi & Shahnazari, 2013), and sometimes also *unless*, conditional constituents can be introduced by several other conditional subordinators. Quirk *et al.* (1985: 1089), for instance, include many other subordinators, such as *as long as, so long as, assuming* *(that), given (that), in case, in the event that, just so (that), on condition (that), provided (that), providing (that)*, and *supposing (that)*. Declerck & Reed (2001: 21) also consider *on the understanding that* and *lest* as conditional elements. A recent examination of the literature (Puente-Castelo, 2017: 52) also found that *so (that)* presented conditional uses until at least 1859.

3.1 Inversion conditionals

There are also conditional structures without a subordinator, however. Conditionals may be formed by means of the inversion of the operator in the protasis, as shown in Example (2) above, repeated here:

(2) *Had he* not seen the car coming, he would have been killed.
(Biezma, 2011: 555)

The main operators allowing this inversion with a conditional interpretation in Present Day English are *were, should,* and *had* as in (2) above, although Declerck & Reed (2001) also include *be* (11) and Quirk *et al.* (1985: 1093–1094) *might* or *could* with an adverb such as *but* or *just*, as in (12) below.

(11) […] be I right or wrong about that part of the story […]
(Jackson 1998: 40 in Declerck & Reed, 2001: 28)

(12) Might/Could I but see my native land, I would die a happy man.
(Quirk *et al.,* 1985: 1094)

However, as the array of operators which allowed inversion with a conditional interpretation during the period 1700–1900 was "more varied" than it currently is (Rissanen, 1999: 308–309), it is necessary to add some other operators to the ones above. For instance, several probability modals such as *may* or *would*, as well as *did* or *is*, were used in the past to express conditionality by inversion and have since lost this conditional use. Among these, *did* was in use until at least the first part of the twentieth century, as shown in Example (13) below, found in Iatridou & Embick (1994: 138)

(13) There are other articles, to which, did time permit, we might draw attention.
(Curme, 1931: 327)

3.2 Some grammatical characteristics of inversion conditionals

Inversion conditionals are different from the other types of conditionals in that they lack a conditional particle, and thus the different nuances of meaning introduced by these particles are absent. Instead, inversion conditionals show some *grammatical* restrictions.

First and most obviously, since the particle experimenting inversion is the operator in the verb form, which is finite by nature, non-finite and verbless clauses are not possible with inversion conditionals (Quirk *et al.,* 1985: 1090), unlike the options with subordinator. Also impossible (Quirk *et al.,* 1985: 1094: Iatridou and Embick, 1994: 137) is using the contracted negation *n't* with the inverted operator, although negation with *not* seems indeed possible, as shown in (14)–(15) below:

(14) *Hadn't he seen the car coming, he would have been killed.
(Iatridou and Embick, 1994: 137)

(15) Had he not seen the car coming, he would have been killed.
(Biezma, 2011: 555)

Several authors have also noted that there are no apodosis-less conditionals introduced by inversion (Biezma, 2011: 566). This has been related to the fact that inversion conditionals cannot be clefted or used as free-standing answers to questions (Iatridou & Embick, 1994: 141), as shown in (16) below:

(16) *It is had John come that Mary would have left.
(Iatridou and Embick, 1994: 141)

Likewise, there is also some controversy about the possibility of using inversion conditionals with *only* and *even* (17), with Iatridou & Embick (1994) arguing that this is impossible whilst Jong-Bok (2011: 7) claims that, at least with *even*, the use of inversion is possible and has indeed been found in corpus data (18).

(17) (*Only) had I thought that would I have called him. (Jong Bok, 2011: 20)

(18) Even had she known his birthday, she wouldn't have scoured Alburquerque thrifts. (COCA 2006 FIC in Jong Bok, 2011: 20)

Despite inversion conditionals lacking the conditional nuances of meaning introduced by each different conditional subordinator, some authors have still tried to identify particular meanings for inversion conditionals which are different from their hypothetical *if* equivalents. Iatridou & Embick (1994) argued that inversion conditionals can only appear in a protasis whose information is old and known to be false when it is uttered, while Biezma (2011: 59) claims that the use of inversion marks "the antecedent proposition as previously considered/entertained", thus relating it to previous discourse. According to Biezma, this would explain why sometimes inversion conditionals are interpreted as a reproach, as in (19) below, since they are used to remind the recipient that the protasis should have been fulfilled and was not, as the information in the protasis is marked as previously entertained.

(19) Mom: Had you done your chores, you would be able to go with your friends.
(Biezma, 2011: 568)

Finally, some instances of double conditional marking (with both *if* and inversion) have also been found, as shown in (20) below:

(20) The old man would never have left her here if had he seen it.
(COCA 2008 FIC in Jong Bok, 2011: 19)

4. Inversion conditionals in use: Previous studies

A previous study (Puente-Castelo, 2017) on the use of conditionals in texts on astronomy, philosophy and life sciences from the eighteenth and nineteenth centuries found that the CETA, CEPhiT, and CELiST subcorpora of the *Coruña Corpus of English Scientific Writing* contain 213.35 cases of inversion conditionals per million words, accounting for 6.91% of all conditionals therein. As shown in Figure 1 below, *were* is the most frequent among the conditional operators in these corpora, accounting for half the uses of inversion conditionals, at 106.17 uses per million words. *Had* accounts for 54.32 uses per million (25% of uses of inversion conditionals), whilst *should* presents 25.51 cases per million (12% of uses). *Did* and *could* are less frequently used (13.99 cases per million, 7%, and 9.88 per million,

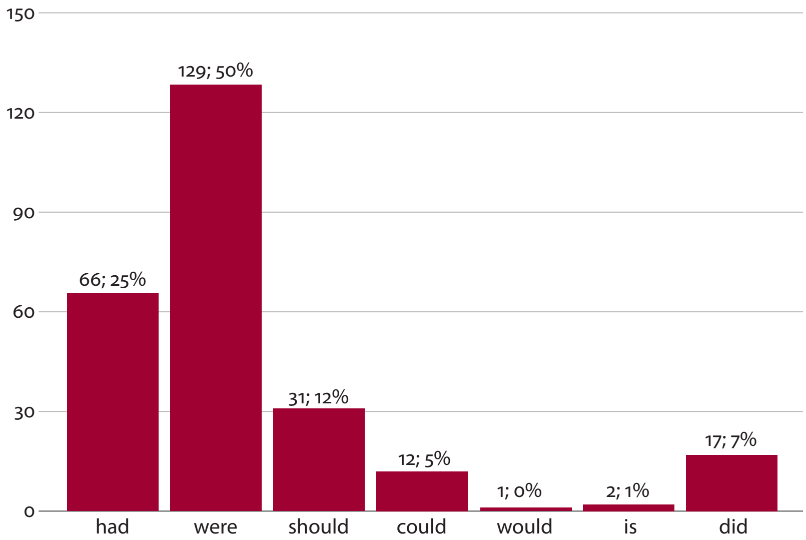


Figure 1. Use of different types of inversion conditionals in CETA, CEPhiT and CELiST (Puente-Castelo, 2017: 160)

5%, respectively), while *is* presents 2 cases and *would* a single instance, all three of these from samples in the first decade of the period under analysis (1700–1710). No cases with *might, may* or *be* were found.

Regarding their diachronic evolution, both *if* and inversion conditionals decrease in use over time, while *unless* and peripheral conditionals increase, as shown in Figure 2 below. Inversion conditionals fall from 234.95 uses per million in the eighteenth century to 189.66 in the nineteenth. Linear regression tests with results from all samples show that these differences are not statistically significant.

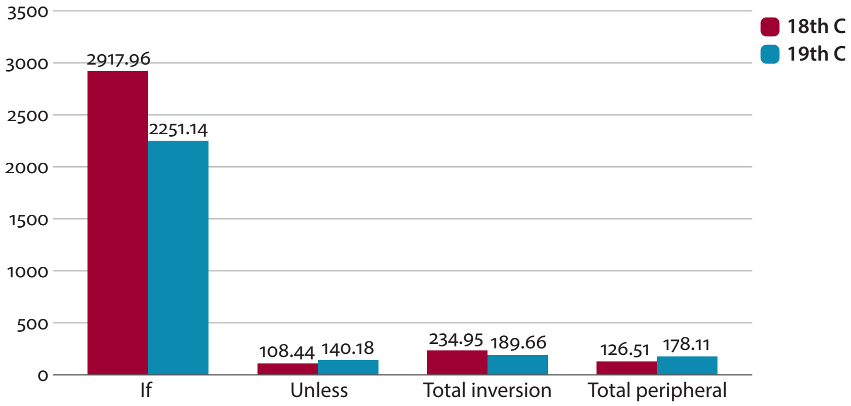


Figure 2. Use of types of conditionals per century in CETA, CEPhiT and CELiST, normalised figures [$N = 1{,}000{,}000$] (Puente-Castelo, 2017: 161)

The results in Puente-Castelo (2017) show the different evolution of the strategies with which conditionals can be introduced, in that one of these, inversion, diminishes in use over time, while the other, using a subordinator, increases. However, the results also show that there are significant differences in the variety of individual elements which can be used with either of these strategies, as the array of subordinators increases over time, with new particles such as *in case* (first use in the samples analysed in 1774), *lest* (1790), *assuming* (1845), and *on condition* (1867); while, on the other hand, several operators stopped being used to express conditionality by inversion, with the final use of *is* in 1700 and *would* in 1702, hinting at the first steps towards the present-day scenario in which only *had, were*, and *should* are commonly used to express conditional inversion.

Regarding discipline, inversion conditionals are more common in philosophy texts, and particularly in female-authored philosophy texts, than in any other discipline, a tendency which extends to all operators other than *should*, which is more common in texts on life sciences. This may perhaps be linked to the conservative nature of philosophy compared to the other disciplines.

In terms of their form, inversion conditionals present a majority of uses with a protasis-apodosis order (74.80%), and a notably low frequency of middle-protasis uses (3.87%). There are also some particular uses regarding verb form choice, as there is no single use with present simple or present subjunctive in the protasis, a feature related to the constraint imposed by the operators capable of expressing inversion. A summary of these verb forms is shown in Table 1.

Table 1. Use of verb forms combinations in inversion conditionals in CETA, CEPhiT and CELiST (Puente Castelo, 2017: 205)

Verb form combination	N Uses	Verb form combination	N Uses
Past simple, would	86	Past perfect, could present perfect	7
Past simple, could	27	Should, would	7
Past perfect, would present perfect	26	Past perfect, past perfect	6
Past simple, should	12	Past simple, present simple	6
Should, will	10	Other 27 combinations	64
Past simple, might	7		

Finally, turning to their function in discourse, most inversion conditionals (86%) have been found to encode hypothesizing features, as shown in Figure 3 below. This contrasts with the almost non-existent known-fact conditionals, the most categorical ones, which represent only three of the total of 258 inversion conditionals in CETA, CEPhiT and CELiST.

Apart from hypothesizing conditionals, the next most frequent functions are relevance conditionals (5.42%), concessive conditionals (3.87%) and scope-restricting conditionals (3.48%). There is also a single use of a non-committal conditional.

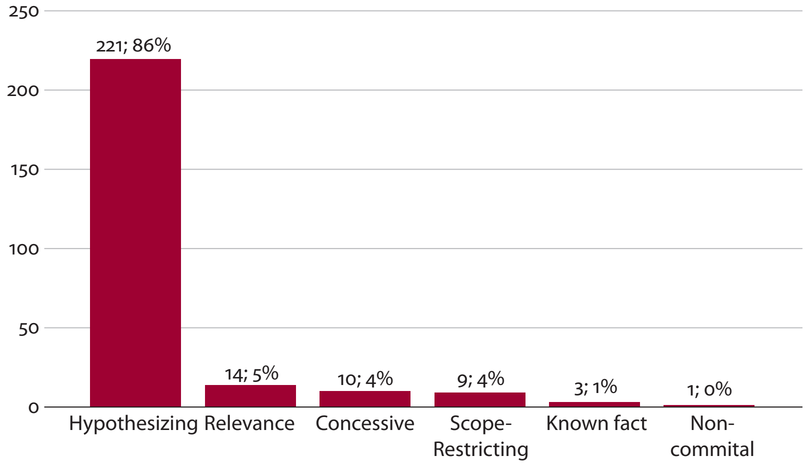


Figure 3. Functions of inversion conditionals in CETA, CEPhiT, and CELiST

It is also noticeable that despite there being no difference in meaning between *if* and inversion conditionals, several conditional functions (metalinguistic, politeness, directive) only occur with *if*.

4.1 Inversion conditionals as evidence of paradigmatic change

These results in Puente Castelo (2017) show how one of the strategies to encode conditionality, inverting the operator, sees diminished productivity over time, with a less frequent use overall, a reduced number of possible operators, and its use appearing in more specialised contexts. Meanwhile, the other strategy, subordination, takes its place, increasing in both use and variability, with some brand new conditional subordinators appearing over time.

This evolution may be related to the process of the substitution of Scholasticism as a paradigm of knowledge, which is in here modelled as a two-step process in which scholastic features were first generally abandoned, before new features appear at a later stage.

Following this model, the abandonment of scholastic features characteristic of the first stage can be noted in the perceptible decrease in the use of conditionals in general and the contraction in their variability, with a decrease in the use of categorical conditionals, in the proportion of use of inversion conditionals, and in a lower number of possible verb form combinations. It would also explain the shift from letters to articles as the preferred genre for scientific expression. This process manifests itself at different rates in different disciplines, reflecting the different rhythms at which the various disciplines evolved from the old scholastic model towards contemporary scientific discourse. For instance, philosophy texts, which are probably the most influenced by scholastic practices, present a higher use of conditionals in general, whilst astronomy texts, one of the first disciplines to adopt new scientific practices, are characterised by a lower use here.

Later, during the new process of expansion, a series of new peripheral conditional particles are introduced, as well as new conditional functions such as scope-restricting conditionals, which allow authors to mitigate claims in a more covert way by means of the invocation of shared-knowledge, as shown in Example (10) above, repeated here:

(10) As such, it can be said to belong to modality if the category is defined as the expression of the speaker's attitude or stance. (Warchal, 2010: 148)

These changes show how the different formal types of conditionals in general, and inversion conditionals in particular, are then especially illustrative of the paradigmatic evolution of scientific writing over time, and raise a new question: does this evolution in the use of inversion conditionals mirroring the paradigmatic evolution also occur in other disciplines?

5. Corpus and methodology

The corpus used for the present analysis is the full two-century period of the beta version of CHET, containing 40 text samples and some 404,424 words. In order to extract the data, selected conditional particles, shown in Table 2 below, were searched in CHET using the *Coruña Corpus Tool* (CCT) (Parapar & Moskowich, 2007). All conditional particles, and not just inversion operators, were searched, in order to make the examination of the proportion of use of each type of conditional possible.

Table 2. Elements introducing clauses with conditional interpretations during the period 1700–1900 (taken from Puente-Castelo, 2017: 53)

Central conditional subordinators	*If, unless*
Peripheral conditional subordinators	*As long as, so long as, assuming (that), given (that), in case, in the event that, just so (that), lest, on condition (that), on the understanding that, provided (that), providing (that), supposing (that), so (that).*
Operators allowing inversion with conditional interpretation.	*Had, were, should, might, could, may, would, is, be, did*

The resulting data then underwent a process of manual disambiguation, given that the available POS-tagged version of CHET (De Gaetano *et al*., this volume) does not distinguish between conditional and non-conditional uses in search results. Consequently, interrogative uses of *if*, comparative uses of *as long as* and *so long as*, uses of *assuming, given, provided, providing*, and *supposing* as full verbs, and all the uses of *as if*, among others, were eliminated.

This process of disambiguation also affects inversion conditionals. Using any of the operators as query words produces a substantial number of results, among which only a minority are conditional uses, thus making the disambiguation process far longer and painstaking.

After the process of disambiguation, the total number of conditionals found in the corpus was 558. The inversion conditionals were then classified according to a series of linguistic and non-linguistic parameters. The latter include the genre of the works, their date of publication and the sex and origin of their authors, whilst the former comprise the different combinations of verb forms in both constituents, the order of these constituents (initial, final, or middle protasis, or apodosis-less conditionals), and the function that the conditional plays in discourse.

In order to classify conditionals according to the function they play in discourse, the typology presented in Puente-Castelo (2017) was used. This typology, shown in Table 3 below, distinguishes eleven different conditional functions in discourse, looking particularly at the mitigating or non-mitigating role of the conditional, as well as the interpersonal features of conditionals, as explained in Section 2 above.

Table 3. Typology of conditionals proposed in this work (Puente-Castelo, 2017: 107)

Type	**Function**	**Example**
Known fact	To state widely accepted facts and mathematical truths.	Given that $x = y$, then $n(x + a) = n(y + a)$ must also be true.
Hypothesizing	To state the likelihood of an apodosis given a protasis.	If a patient has an early failure from a low anterior resection, they may be able to be retrieved by resection.
Scope-Restricting	To describe the scenario or build the argumentative space in which the claims made can be held, either by defining a concept or specifying the universe to which the claim relates.	As such, it can be said to belong to modality if the category is defined as the expression of the speaker's attitude or stance.
Method	To narrate completed methodological procedures or to introduce instructions.	If 10% or more of the malignant nuclei were stained, the slide was scored as negative.
Rhetorical	Strong assertions which take the form of conditional structures.	If they are Irish, I'm the Pope. He's ninety if he's a day.
Concessive	To introduce an impediment for the fulfilment of the apodosis, under which, nevertheless, it held.	Our point still goes through if the minimal phrase containing both parts of this idiom is always headed by a verb. …the use of change predicates is possible precisely because they apply to the virtual entities, if not to the actual entities that ultimately ground them.
Directive	To present an obligatory desirable course of events as if it were optional and not compulsory.	Now if we go to patients who experienced mucositis toxicity…
Politeness (Speech act)	To introduce a conventional expression of politeness.	If I may be quite frank with you, I don't approve of any concessions to ignorance.

Table 3. (*continued*)

Type	Function	Example
Relevance (Speech act)	To explain the circumstances under which the statement of the apodosis is relevant.	Finally (if this is important), the S1 meaning can be converted into an S meaning to recover a more intuitive object to represent the meaning of the original sentence.
Meta-linguistic (Speech act)	To make a comment on the wording of the discourse.	His style is florid, if that's the right word.
Non-committal (Speech act)	By authors, to distance themselves from the claims of others.	Chomsky's views cannot be reconciled with Piaget's, if I understand both correctly.

6. Results and discussion

As noted above, the total number of conditionals in CHET after disambiguation is 558, a proportion of 1379.74 per million words. From these, 62 (11.11%) are inversion conditionals (153.30 per million words, a proportion notably lower than in the other three corpora). As shown in Figure 4 below, 75.08% of conditional uses in CHET contain *if*, 5.01% *unless*, and 8.78% are peripheral conditionals.

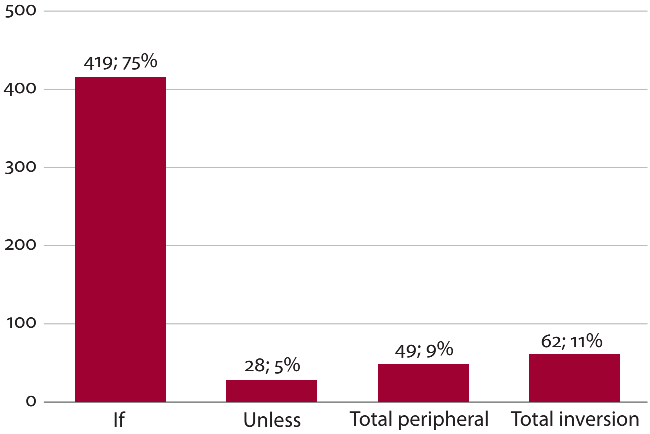


Figure 4. Use of types of conditionals in CHET

Among inversion conditionals, the most common operator is *had*, with 44 uses (70.96%). As shown in Figure 5 below, the next more common particles are *should* (12 uses, 19.35%), and *were* (3 uses, 4.83%). *Did* presents two uses and *could* a single one. No inversion uses of *might, may, would, is*, or *be* were found. These results contrast sharply with those in Puente-Castelo (2017), in which *were* is the most frequent operator for inversion conditionals in CETA, CEPhiT and CELiST, (50% of all the uses of inversion conditionals), while in CHET the most frequent is *had*, at almost 71% of uses.

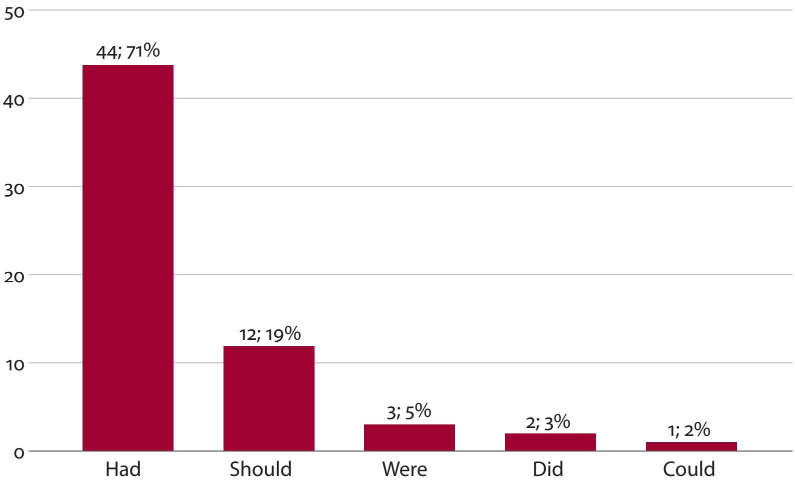


Figure 5. Use of different types of inversion conditionals in CHET

This striking difference is best seen in Figure 6 below, in which three groups can be distinguished. Texts on history present a higher than usual use of *had* (71%), with *should* in second position (19%) and *were* third (5%). Texts on life sciences present a moderate use of both *were* and *should* (34%) with *had* in third position (21%), whilst texts on both astronomy and philosophy present a higher use of *were* (56 and 54%, respectively), followed by *had* (23 and 29%) and *should* (10 and 3%), although it should be noted that philosophy texts contain around twice as many conditionals as any of the other disciplines.

This is not the only difference in the use of inversion conditionals between history texts and those from other disciplines. Whereas the use of inversion conditionals in CHET remains relatively stable over time, with the other three subcorpora there is a noticeable decrease. Although the average use of inversion conditionals in each century show some differences (163.41 uses per million in the eighteenth century and 143.21 uses per million in the nineteenth century), the evolution is not so clear when the results are analysed by decade, as shown in Figure 7 below. In fact, statistical tests here, which are non-significant, show that diachronic evolution accounts for only 0.3% of the total variation, thus pointing to a model in which diachronic evolution has little importance in the variability of inversion conditionals.

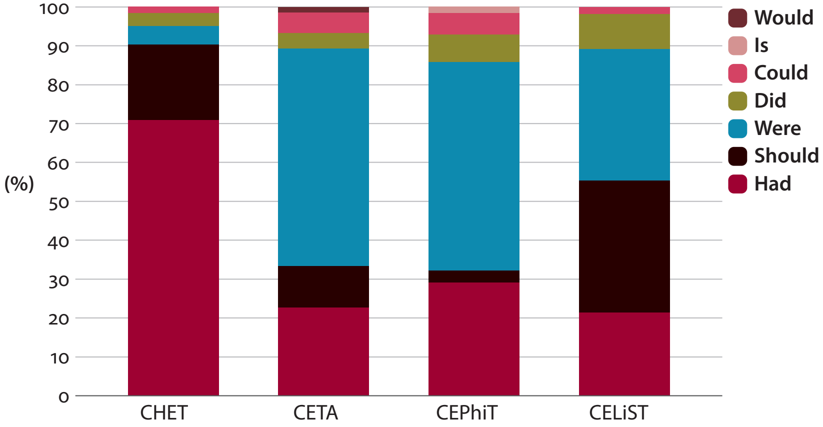


Figure 6. Proportional use of different types of inversion conditionals in CHET, CETA, CEPhiT and CELiST

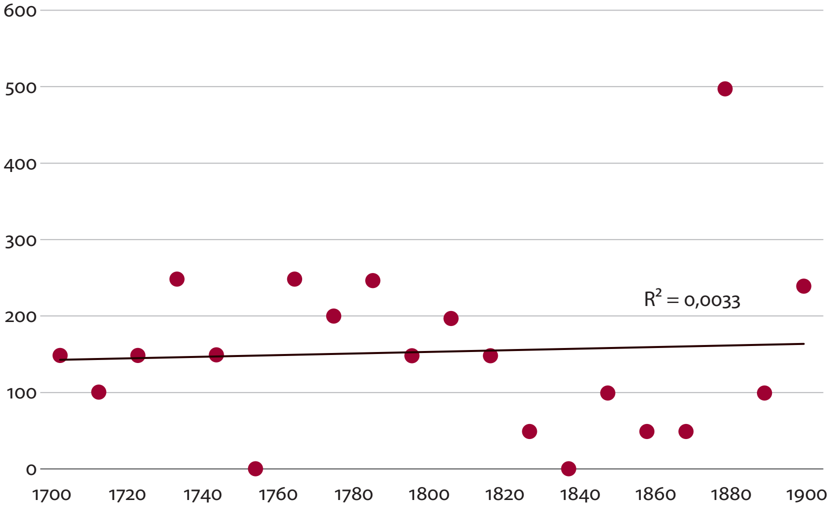


Figure 7. Diachronic evolution in the use of inversion conditionals in CHET

The analysis of the results according to the other extra-linguistic parameters included in this study shows few significant differences. Men use inversion conditionals slightly more often than women (161.02 vs. 122.70 uses per million), but the small number of women (eight) in the corpus makes it difficult to extract definite conclusions from the data, since a particular author's idiosyncrasies might easily skew the results. This scarcity in the number of elements in any given category also affects the analyses according to the genre of the samples (in which the category

"Others" is the one with the most uses, 332.70 per million), and to the nationality of the author (in which Irish authors use inversion conditionals much more frequently, 214.35 uses per million).

However, there are more differences between the uses of inversion conditionals in CHET and in the other three corpora (CETA, CEPhiT, CELiST) regarding the different linguistic parameters addressed.

In terms of the functions of the conditionals in discourse, hypothesizing conditionals are once again the most common function, with 85.48% of all uses, and the second most used conditionals are once again relevance conditionals, at 4.83%. These percentages are very similar to those in the other three corpora. However, as shown in Figure 8 below, non-committal conditionals, which presented a single use in the other three corpora, are the joint-second most used conditional function in CHET, at 4.83%. There are also two conditionals (3.22%) with a concessive function, and a single scope-restricting conditional. Examples (21) and (22) show two cases of relevance and non-committal conditionals, respectively, found in the corpus.

(21) Should these relations of our domestic writers […] excite any doubt in the mind of the reader, it should be remembered that Tacitus […] relates, […]
(Hardiman, 1820: 34 [146 (1649)])

(22) yet the reaſons that are offer'd by ſome moderns againſt the truth of it may perhaps excuſe our incredulity, ſhould we look upon it as a meer fiction.
(Hooke, 1745: 47 [131 (9495)])

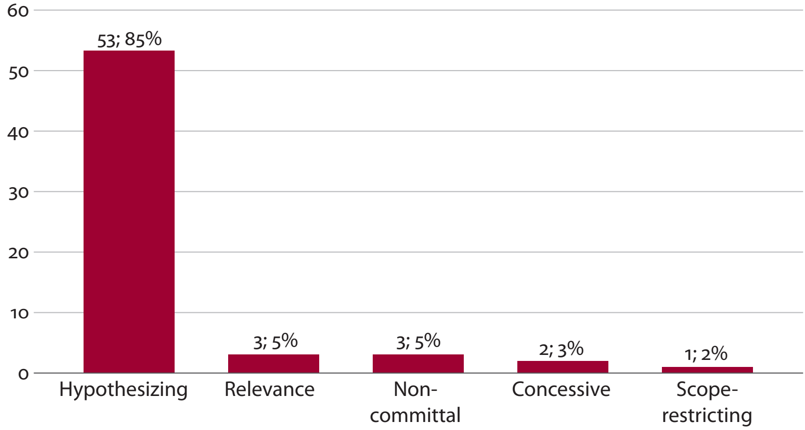


Figure 8. Functions of inversion conditionals in CHET

Regarding the order of constituents, initial conditionals are once again the most common, with 56.45% of uses, with final conditionals at 37.09% and middle conditionals at 6.45%. These results show a lower use of initial conditionals and a higher use of final conditionals than in the other three corpora (74.80% and 21.31%, respectively).

Finally, turning to verb form combinations, the differences here are very striking, as set out in Table 4 below. Conditionals with a *past perfect* in the protasis and *would* + *present perfect* in the apodosis represent 38.70% of all uses, with a further 20.96% of variants of this type with different modal verbs.

Table 4. Use of verb forms combinations in inversion conditionals in CHET

Verb form combination	N Uses	Verb form combination	N Uses
Past perfect, would present perfect	24	Should, would	3
Past perfect, might present perfect	7	Should, past simple	3
Past perfect, must present perfect	4	Other 14 combinations	21

These results, of course, are contingent on the majority of uses with *had*, as mentioned above.

7. Conclusions

The analysis of the results has shown notable differences in the use of inversion conditionals between history texts and other disciplines. Particularly, the higher use of *had* and the diachronic stability of the results are clearly divergent from the results in other subcorpora.

These results, and specifically the lack of diachronic evolution in the use of inversion conditionals in CHET, point to a scenario in which inversion conditionals cannot be as readily used as evidence of paradigmatic change in history texts as in astronomy, philosophy, and life sciences texts. There may be several reasons for this, including the possibly more conservative nature of historical texts or the longer-lasting influence of Scholasticism in historical writing, extending beyond the end of the period under analysis and hence outside the scope of the present study.

However, the important differences between history texts and those from other disciplines regarding the use of inversion conditionals, and particularly the proportion of use of different operators, may reflect interesting disciplinary particularities. The higher use of *had* in history texts may be evidence of the possibly more narrative nature of history texts (see also Monaco, this volume), requiring authors to speculate on past events and thus necessarily to use the past perfect tense, which in turn would lead to a higher use of hypothesizing conditionals, as indeed we have seen in the data here. This could be proved if, in an analysis of other types of conditionals in CHET, these are found to show a similarly high proportion of past perfect uses in the protasis. But this must be left for future research.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Allen, Bryce, Jian Qin & Frederik Wilfrid Lancaster. (1994). Persuasive Communities: A Longitudinal Analysis of References in the Philosophical Transactions of the Royal Society, 1665–1990. *Social Studies of Science*, 24/2, 279–310. https://doi.org/10.1177/030631279402400204

Atkinson, Dwight. (1996). The Philosophical Transactions of the Royal Society of London, 1675–1975: A sociohistorical discourse analysis. *Language in Society*, 25, 333–371. https://doi.org/10.1017/S0047404500019205

Atkinson, Dwight. (1999). *Scientific discourse in sociohistorical context: The Philosophical Transactions of the Royal Society of London, 1675–1975*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

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

Biezma, María. (2011). Conditional inversion and givenness. *Proceedings of SALT*, 21, 552–571. https://doi.org/10.3765/salt.v21i0.2591

Carter-Thomas, Shirley & Elizabeth Rowley-Jolivet. (2008). If-conditionals in medical discourse: from theory to disciplinary practice. *Journal of English for Academic Purposes*, 7, 191–2057. https://doi.org/10.1016/j.jeap.2008.03.004

Curme, George O. (1931). *Syntax*. Boston: D. C. Heath.

Dancygier, Barbara. (1998). *Conditionals and Prediction: Time, knowledge and causation in conditional constructions*. Cambridge: Cambridge University Press.

Declerck, Renaat & Susan Reed. (2001). *Conditionals: A comprehensive empirical analysis*. Berlin and New York: Mouton de Gruyter. https://doi.org/10.1515/9783110851748

Facchinetti, Roberta. (2001). Conditional Constructions in Modern English Legal Texts. In Maurizo Gotti & Marina Dossena (Eds.), *Modality in Specialised Texts* (133–150). Bern: Peter Lang.

Fauconnier, Gilles. (1994). *Mental Spaces*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511624582

Ferguson, Gibson. (2001). If you pop over there: a corpus-based study of conditionals in medical discourse. *English for Specific Purposes*, 20, 61–82. https://doi.org/10.1016/S0889-4906(99)00027-7

Ford, Cecilia E. (1997). Speaking conditionally: some contexts for *if*-clauses in conversation. In Angeliki Athanasiadou & René Dirven (Eds.), *On Conditionals Again* (386–413). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.143.21for

Fulcher, Glenn. (1991). Conditionals revisited. *ELT Journal*, 45, 164–168. https://doi.org/10.1093/elt/45.2.164

Gabrielatos, Costas. (2010). A corpus-based examination of English if-conditionals through the lens of modality: Nature and types. Unpublished PhD dissertation. Lancaster: Lancaster University.

Hesabi, Akbar; Morteza Dehaghi & Mohammadtaghi Shahnazari. (2013). A comparative Analysis of the Frequency and Function of If-Clauses in Applied Linguistics and Chemistry Articles. *Journal of Language Teaching and Research*, 4, 186–191. https://doi.org/10.4304/jltr.4.1.186-191

Hwang, Myong Ok. (1979). A semantic and syntactic analysis of if-conditionals. Unpublished MA thesis. Los Angeles: University of California Los Angeles.

Hyland, Ken. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, 17/4, 433–454. https://doi.org/10.1093/applin/17.4.433

Hyland, Ken. (1998). *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.54

Hyland, Ken. (2000). *Disciplinary Discourses: Social interactions in academic writing*. Harlow, Essex: Pearson Education.

Iatridou, Sabine & David Embick. (1994). Conditional Inversion. *Proceedings of the North Eastern Linguistic Society*, 24, 133–147. Cambridge, MA: MIT Press.

Jones, Christian & Daniel Waller. (2010). If only it were true: the problem with the four conditionals. *ELT Journal*, 65, 24–32. https://doi.org/10.1093/elt/ccp101

Jong Bok, Kim. (2011). English Conditional Inversion: A Construction-Based Approach. *Language and Information*, 15, 13–29.

Maule, David. (1988). “Sorry, but if he comes, I go”: Teaching conditionals. *ELT Journal*, 42, 117–123. https://doi.org/10.1093/elt/42.2.117

Myers, Greg. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10/1, 1–35. https://doi.org/10.1093/applin/10.1.1

Parapar, Javier & Isabel Moskowich. (2007). The Coruña Corpus Tool. *Revista del Procesamiento de Lenguaje Natural*, 39, 289–290.

Puente-Castelo, Luis. (2017). On conditionality: A corpus-based study of conditional structures in Late Modern English scientific texts. Unpublished PhD Dissertation. A Coruña: Universidade da Coruña.

Quirk, Randolph; Sydney Greenbaum, Geoffrey Leech, & Jan Svartvik. (1985). *A comprehensive grammar of the English language*. London: Longman.

Rissanen, Matti. (1999). Syntax. In Roger Lass (Ed.), *Cambridge History of English vol. 3* (187–331). Cambridge: Cambridge University Press.

Swales, John. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Warchal, Krystyna. (2010). Moulding interpersonal relations through conditional clauses: Consensus-building strategies in written academic discourse. *Journal of English for Academic Purposes*, 9, 140–150. https://doi.org/10.1016/j.jeap.2010.02.002

Werth, Paul. (1997). Conditionality as cognitive distance. In Angeliki Athanasiadou & René Dirven (Eds.), *On Conditionals Again* (97–114). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.143.14wer

CHAPTER 8

Modal verb categories in CHET

Francisco Alonso-Almeida and Francisco J. Álvarez-Gil
Universidad de Las Palmas de Gran Canaria

1. Introduction

Scientific writing is, and so is academic writing in general, often regarded as depersonalised discourse, as depersonalisation seems to be strongly associated to degrees of objectivity in relation to findings and conclusions put forward in scientific prose. The sense of being objective is somehow interpreted in terms of how defocalised the person responsible for making given claims is. A look at earlier stages of the scientific literature demonstrates that there is not such a thing as depersonalised scientific prose, as it is almost unreasonable to think that no one may be held responsible for their texts, especially when texts are not anonymous. In this paper, our objective is to analyse the use and function of modal verbs in scientific texts from the late Modern English period with examples excerpted from the beta version of the *Corpus of History English Texts* (CHET), a subcorpus of *the Coruña Corpus of English Scientific Writing*.

Even if history may be considered as a less technical discipline if compared with other disciplines, say, astronomy, mathematics, or chemistry, authors of history texts recourse to argumentation and reasoning in order to formulate new knowledge from evidence and critical thinking. Modal verbs constitute a rich resource to develop meaning, and they can certainly not avoid reflecting different degrees of perspectivisation. As we shall see, given the extraordinary number of modal verbs that may be registered in one piece of scientific writing, objectivity understood in the sense of detachment from claim is certainly contested, and it is more logical to expect linguistic devices reflecting the authors' stance framing their subjective critical thinking.

The organisation of this paper is as follows. Section 2 describes our notion of modality and modal verbs. After this, we offer a description of the method and the data used in this study. The discussion of findings is given in Section 4. Finally, the conclusions are presented.

https://doi.org/10.1075/z.225.08alo

© 2019 John Benjamins Publishing Company

2. Modality and modal categories

Modality can be broadly defined as the linguistic encoding of the speaker's or writer's beliefs and attitudes towards the proposition manifested. In other words, modality refers to the status of the proposition in context. Palmer (1986: 2) refers to modality as a *vague* notion but admits that "something along the lines of Lyon's (1977: 452) 'opinion or attitude' of the speaker is promising". There are several ways in which modality can be manifested in the language other than modal verbs. These can be lexical or grammatical. Palmer (1986: 33ff) describes modal verbs, mood, and particles and clitics as examples of grammatical marking of modality. The lexical marking includes adverbs and other related expressions that may show the speaker's attitude towards the proposition.

There seems to be sufficient critical agreement that modality is divided into epistemic and deontic modality in the fashion of Lyon's (1972), as put forward in Palmer (1986: 19). Epistemic is "concerned with matters of knowledge or belief, on which basis speakers express their judgements about state of affairs, events or actions" (Hoye, 2009: 100). Deontic modals refer to the "necessity of acts in terms of which the speaker gives permission or lays and obligation for the performance of actions at some time in the future" (Hoye, 1997: 43).

Another way of looking at modality is the distinction between (a) propositional modality and (b) event modality put forward in Palmer (2001). The former concerns the speaker's or writer's judgment regarding the proposition manifested. The latter is related to the speaker's or writer's attitude towards a likely event in the future. Propositional modality is divided into *epistemic* (judgement about the factual status), and *evidential* (evidence for factual status is given). In our study of modals, we shall refer to modals used as evidentials in our description of examples, as this category is frequently subsumed under the domain of epistemic modality. Our view is substantially different from the one described in Palmer (1986) and Chafe (1986) *inter alia*, as explained in Alonso-Almeida (2015a). We feel, however, this chapter is not the right place to raise and develop this issue. It will suffice here to indicate the classification of evidentials according to the nature of the source of evidence.

Evidentials can be reported and sensory, and Willet (1988) has classified them according to (a) direct evidence (visual, auditory, and sensory), (b) indirect evidence (reported or inferred from reason or results). This categorisation of evidentials seeks the purpose of relating the nature of the evidence to assert a particular propositional content (*P*) or the mode of knowing with different degrees of truth. As pointed out in Alonso-Almeida (2015b) and Alonso-Almeida & Carrió-Pastor (2015), it seems to us that the expression of information source reports primarily on aspects of the author's elaboration of *P*, i.e. involvement and perspectivisation, rather than the evaluation of *P* in terms of truth. This evaluation is, in our view, a pragmatic side effect, and it is not necessarily in the mind of the writer. Another option is to refer to the debate that modals can be monosemous or polysemous (Collins, 2009: 23, cf. references therein), and so there is the need to establish which modal meaning predominates in a given form.

The second big category of modality in Palmer (2001) is event modality, and includes a further twofold distinction, namely deontic and dynamic. In deontic modality, conditions are internal, i.e. moral senses, or external. Senses of obligation and permission tend to depend on external ones. Collins (2009: 22) states that deontic modality "occurs when the factors impinging on the actualisation of the situation referred to in the utterance involve some type of authority – as when a person or a set of rules or a social convention is responsible for the imposition of an obligation or a granting of permission". In dynamic modality, conditions can be both internal or external, and it involves senses of willingness and ability on the part of the speaker or writer. Our notion of dynamic modality exemplifies the sense of potentiality and disposition.

Cases of dynamic modality are often misinterpreted in favour of epistemic or purely deontic readings, as the boundary between one and the others are certainly fuzzy and hence these analyses, even if debatable, rely on the interpreters' expectations rather than on an accurate evaluation of the facts. One master proof to elucidate cases of, say, epistemic possibility and dynamic possibility is the fact that the second does not undergo actualisation. The instance in (1) taken from *the Bok of Ypocras of lyf and deyth* (MEMT, in Taavitsainen & Pahta (2011), my emphasis) is a good example of how the categorisation of *may* will largely depend on the writers'/hearers' view of the facts and their analyses of the contextual cues. In this sense, both epistemic and dynamic nuances are acceptable as long as the interpreter understands that a man has the chances of curing himself providing he takes the adequate drugs, and this sense is epistemic. It could be also the case that this modal particle is here reporting on the potentiality of medicines to cure a person, or the disposition of a person to cure himself by using medicines. In these two cases, there is no actualisation of the events described. In the case of the epistemic interpretation, may clearly refers to the chances of *P*, i.e. "cure him", to occur. Dynamic possibility, in our view, happens to be a more plausible disambiguation here, as professional medicine was not an available solution for everyone during the medieval period. For this reason, recipes were copied here and there, gathered either in the form of beautifully-decorated compendia, in loose pieces of parchment, torn pieces of parchment, or as notes in margins, flyleaves and intentionally blank spaces of books.

(1) A man *may* cur hyme with medysignes 'A man may cure him(self) with medicines' attack.
(*Bok of Ypocras of lyf and deyth*, MEMT, Taavitsainen & Pahta (2011))

Dynamic modality, as earlier stated in Alonso-Almeida (2015b), is a frequent guest in technical and scientific texts, as it represents a powerful means to talk about the potentialities and disposition of people, gadgets, events, experiments and scientific phenomena in general to behave. It is predictable then that dynamic modals prevail in texts of a scientific nature, as the ones studied in this chapter.

There are other classifications of modality, and these include more modal categories than the two macro-types explained above. These taxonomies follow different established pragmatic, philosophical/logical or cognitive criteria. We will not refer to any of these in our present paper, and we will broadly categorise the modals excerpted from CHET according to (a) epistemic, (b) deontic, and (c) dynamic modality. Epistemic necessity modals to express inferentiality is described in the section devoted to epistemic modality.

2.1 Modal verbs

There are various ways in which modality can exhibit, as said at the beginning of this section, one of which is by the use of modal verbs. The criteria for identifying modal verbs involve certain morphological, syntactic, and semantic principles, which Denison (1993: 292ff) summarises, as follows:

1. modal verbs do not present non-finite forms,
2. tense-distinction takes place in the majority of these verbs, although traditional past forms may contextually carry a different meaning,
3. modal verbs do not inflect for third person singular present indicative,
4. most modals present contracted forms for the negative (can't, won't, mustn't), and some of them also show a phonological reduction in form of a clitic form ('ll, 'd, for shall/will and would, respectively),
5. they show no imperative forms,
6. they are followed by a bare infinitive,
7. modal verbs affect the complete proposition in which they are embedded,
8. dialectally, more than one modal verb co-occur syntagmatically, and
9. as operators, they share a same set of NICE properties: 'They can be negated by a following n't/not, take part in the subject-verb inversion, survive post-verbal ellipsis, and be stressed for emphatic polarity'.

The study of modal verbs in the history of English is quite controversial (Campbell, 1959; Allen, 1975; Kemenade, 1992). Scholars do not seem to agree as for the exact dating of this innovation in the language, and whether it is an innovation, or it already existed in Old English in the form of preterite-present verbs. The way in which the issue is methodologically approached may change our view of the facts. Whereas Lightfoot (1979) and Roberts (1985) offer a syntactic explanation as to the development of this closed class of verbs, Plank (1984) and Haspelmath (1994) consider grammaticalisation as the correct path to describe the emergence of modals. Plank (1984) seems not to be exactly a follower of the generative 'Principle of Transparency' posed by Lightfoot (1979), but Haspelmath (1994) seeks to reconcile the methods on the basis of a combination of a formal syntactic account, and a functional perspective that includes grammaticalisation, as the latter may give evidence of functional selection concerning variation. This debate is obviously beyond the scope of the present paper, but we thought convenient to note this controversy here, even if the report presented in this study may give the impression of certain established use of modal verbs.

3. The texts, the method, and results

3.1 The texts and the method

The *Corpus of History English Texts* (CHET) in its beta version is the only corpus we have analysed for the present research. We are not involved in a disciplinary contrast at this stage, as modal verbs seem to appear copiously in scientific and technical texts. Besides computerised analyses, manual inspection and some degree of introspection are certainly needed to evaluate each modal in its context in order to categorise results effectively. For this reason, we focus on one corpus leaving the rest in the *Coruña Corpus* for future research. The Editors have already provided readers with an extensive description of CHET, as a subcorpus of the *Coruña Corpus of English Scientific Writing* project, and this description applies in the present work. So, the reader is referred to Moskowich (this volume) for more information on aspects related to size of texts, dates, authorship, textual genres included in CHET, representativeness, sociological details, among some other particulars.

As indicated above, we have used corpus tools for retrieving the instances of what Biber *et al.* (1999: 483) call central modal verbs; these are *can, could, may, might, shall, should, will, would*, and *must*. Periphrastic semi-modals have not been considered at present. For the computerised analysis and for excerpting samples, we have used the suite The *Coruña Corpus Tool* (CCT), as it provides information for gaining statistics to weigh tendencies of usage and a larger contextual selection to evaluate the exact meanings of the modals for categorisation, respectively. Registering exact modal meanings has been carried out manually, as each case deserved individual attention and contextual interpretation.

The merits of this paper are above all to demonstrate the (ab)use of one modal meaning category versus the others, and to determine which specific modal verbs are in the lead for each semantic category. This information should be useful to appreciate modal changes in progress as well as to understand certain modal verb active uses in scientific and technical writing in the eighteenth and nineteenth centuries, which might be still now characteristic of the register or, conversely, they have died out.

3.2 Results

The number of modal verbs found from the corpus interrogation are given in Table 1 below:

Table 1. Modal verbs in CHET

Modal verb	Occurrences
can	153
could	596
may	328
might	443
must	278
shall	153
should	602
will	153
would	795
Total	**3501**

These forms were identified to show any of the three broad modal categories described above in Section 2, thus:

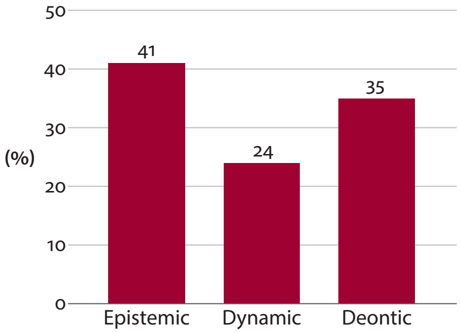


Figure 1. Distribution of modal categories in CHET

Figure 1 illustrates the distribution in percentages of modal categories in the history texts under examination. The most frequent type is epistemic, and this is followed by deontic and dynamic modals, in this order. Within epistemic modality, we have included those forms designating epistemic necessity, even if we hold the position that inferentiality is somewhere between epistemic meaning, i.e. asserting some degree of truth concerning the proposition hedged, and evidentiality, i.e. communicating the author's involvement in the elaboration of the proposition (Cornillie & Delbecque, 2008). Be it as it may, we also argue that there may be functional overlapping between epistemic and evidential resources as the result of the reader's pragmatic interpretation and evaluation of the status or nature of the evidence the speaker has to claim *P* (cf. Alonso Almeida, 2015a). In this context, reliability is put at stake along with an estimation of the writer's degree of confidence and commitment towards the information framed by these modals.

The results shown in Figure 1 are not a surprise. As put forward in Alonso-Almeida & Mele-Marrero (2014), scientific prose of the earlier periods is characterised by the search of a steady balance between authority and academic modesty. In this vein, traditional and new ideas brimmed printed book pages, and these are told with confidence or pretended confidence avoiding mere sequences of directives verging on imposition. This delicate equilibrium among what is (dynamic and deontic), what must or should be (deontic), and what is eligible for actualisation (epistemic) is gained by a strategic use of modal forms in the texts. The pragmatics of their use are varied, being the most evident a desire for acceptance and respect within the authors' academic social network.

Table 2, below, presents the modal verbs found to mean epistemic, deontic and dynamic meanings. Epistemic necessity modals are also included.

Table 2. Modal meanings

Modal meaning	**Modal verbs**								
	can	**could**	**may**	**might**	**must**	**will**	**would**	**shall**	**should**
epistemic	*	*	*	*		*	*	*	
inferential		*		*	*		*	*	*
deontic	*	*	*		*	*		*	*
dynamic	*	*	*			*	*	*	

In what follows, we exemplify and describe these uses and their pragmatic gains.

4. Modal verb categories

In this section, we describe our findings concerning the use of modal verbs in the *Corpus of History English Texts* (CHET). Even if separating data according to genre and genre variables could be of much interest, we will ignore these aspects in the present chapter so that we may concentrate on providing some general considerations of modality and leave the study of modals from a gendered and a genre perspective as prospects for future studies. We will comment on modal categories in CHET in the order resulting from the analysis: epistemic, deontic and dynamic.

4.1 Epistemic modals

As disclosed in Section 3.2. above, epistemic modals are the most frequent type of modals in CHET with the distribution shown in Figure 2:

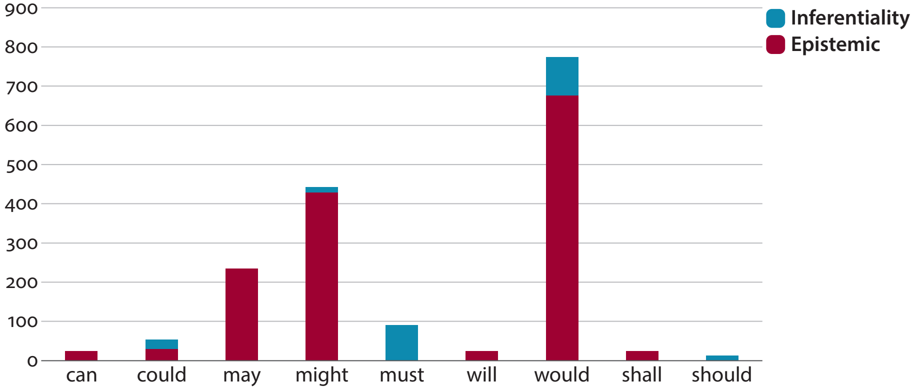


Figure 2. Distribution of epistemic modals, including inferential modals

The forms *would, might* and *may* are the most frequent epistemic modals in the history texts analysed. The modals *can, could, will* and *shall* show low frequencies. Inferential meaning is entailed in the forms *must, could, should* and *might*. The form *should* is by far the commonest of this group.

The form *would* in the following examples refer to hypothetical situations:

(2) Unfortunately, those who accepted Christianity did not retain the remembrance that they had hereditary enemies, who *would* not fail if opportunity offered savagely to attack them. (Kingsford, 1887: 202 [159 (9864)])

(3) It *would* seem as if by the Irish *Lismacuan* or the fort on the harbour, was meant Galway, the parsonage of which, for a long period before the establishment of the wardenship by the Pope, belonged to that monastery; if so, the circumstance *would* tend much to strengthen Baxter's ingenious conjecture.
(Hardiman, 1820: 37 [146 (3305)])

The meaning of *would* in (2) and (3) in the apodosis of the conditional sentences relates to unreal situations which can take place providing the circumstance given in the prodosis. The use of *would* in *it would seem* in (3) conveys the idea of probability concerning the author's evaluation of events introduced in the proposition starting *seem as if…* The pragmatic function of *would* in this case is the avoidance of imposition, and so the author shields his public face from criticism, and also protects the reader's face from reaction.

A similar situation is expressed with the use of *might* in (4):

(4) The result, he supposed, would be, that he should be recalled to settle it, or that at all events he *might* thus preserve himself and the royal family, and by the assistance of Ireland, Scotland, and Louis, be hereafter in a condition to return to it. (Smyth, 1840: 62 [150 (4433)])

The use of the modal verb in (4) is epistemic, as it indicates the probability that the actor, i.e. *he*, can achieve the objective of preserving himself and the royal family. This probability relies on the fulfilment of the condition stated earlier in the text, and this is also reinforced by the particle *thus*, indicating logical consequence. The presence of *might* preceding this logical item mitigates the illocutionary force of the proposition hedged. This mitigating effect is backed by the use of the attributive device, i.e. *he supposed*, given earlier in the text.

Along with *might, may* is used to indicate degrees of probability, as in (5), and possibility, as in (6), below:

(5) The Church of *Inchennan* before the Reformation belonged to the Knights-Templars, who had all their Lands, within the Shire of *Renfrew*, erected into a Regality, called the *Regality* of Greenend. And becauſe ſome *may* have the Curioſity to know ſomewhat concerning the Templars, I ſhall furniſh my Reader with the Hiſtory of them, hoping he'll pardon the Digreſſion.
(Crawfurd, 1710: 75 [124 (283)])

(6) and there is this remarkable of *Donelly*, Sheriff of *Tyrone*, who ſerv'd Two Years for that County, that he had not one Foot of Freehold in the Shire; and for his Honeſty the Reader *may* gueſs at it by this Story.
(Oldmixon, 1716: 60 [125 (4710)])

The use of the modal verb in (5) reports on the likelihood that readers show some interest in the topic of the Templars, and this is taken as the reason to introduce the topic. In (6), *may* describes a situation likely to happen, i.e. guessing Donelly's honesty, as the reader is provided with information to come to this conclusion. The function of this modal in the two examples is to attenuate the meaning of the propositions in order to avoid blatant imposition on the reader. This mitigating effect of *may* is either reinforced by adverbs like *possibly*, or otherwise by stance adverbs like *safely, perhaps, truly, fairly, naturally, properly, reasonably*, and *easily*. All these adverbs often collocate with cognitive verbs, such as *consider, conclude, assume, suppose*, and *imagine*, for example, as shown in Examples (7) to (9), below:

(7) Firmin, *we may naturally ſuppoſe*, was carefully inſtructed by his pious parents in all moral and religious duties. (Cornish, 1780: 2 [137 (225)])

(8) From hence, therefore, it *may be safely concluded*, that the city of the Auterii, mentioned by Ptolemy, was not the ancient town of Galway. (Hardiman, 1820: 36 [146 (2780)])

(9) and, if walls have ears, we *may fairly assume* that the stones in the first Protestant church in Ireland have been listening to both sides of the question for centuries. (Bennett, 1862: 26 [155 (2302)])

The following instances contain examples of *can, could*, and *will*, which are rarely used with the epistemic sense exhibited in these examples:

(10) If, in corroboration of the foregoing reasonings, recurrence shall be had to the testimonies of Tacitus, and also of our native historians, already mentioned, relative to the commerce of Galway, at the very time that Ptolemy describes Nagnata as the most considerable place on the western coast of Ireland, very little doubt *can* remain as to their identity. (Hardiman, 1820: 37 [146 (3431)])

(11) But how *could* society hold together if no barrier was to be raised against the influx of paid propagandists, and the flood of sedition which came pouring in from France?. (Burrows, 1895: 172 [161 (1062)])

(12) Every thing that *could* possibly be spared was set on fire, and soon after the ten thousand Greeks proceeded on their retreat. (Sewell, 1857: 260 [153 (10081)])

(13) And I hope all my good Countrey-Men, whenever occaſion offers, *will* make grateful acknowledgments, and returns of theſe favours, to that Reverend and Learn'd Body. (Anderson, 1705: 60 [123 (2157)])

Example (10) contains an example of *can* with epistemic meaning. In this particular case, this modal verb entails epistemic probability as suggested by the context. The whole expression *very little doubt can* may be interpreted in terms of certainty, as the whole expression refers to the implausibility that doubt may arise from the statement concluding that Nagnata designates contemporary Galway. Examples (11) and (12) present cases of epistemic *could*, and this modal verb might be also accompanied by the adverb *possibly* reinforcing its meaning. The use of *will* in (13) appears to report on the actualisation of the event described in the modal. The presence of *I hope* suggests the idea of hesitation, and so does the condition manifested in *whenever occasion offers*.

4.1.1 *Inferentiality*

Inferentiality is expressed by *would, could, should, must* and *might*. The following instances have been excerpted from CHET:

(14) He took Delight in very idle and cruel Diverſions, and was wrought upon by *Naraskin*, to believe that his Favourite *muſt* needs have been concerned in ſome of the Enterprizes of his Couſin. (Bancks, 1740: 25 [130 (6751)])

(15) Greatly apprehenſive of the fatal effects that *muſt* inevitably have proceeded from a rupture with a nobleman whoſe power and popularity were ſo firmly eſtabliſhed, he made overtures to Robert for an accommodation; and, after much difficulty, obtained a perſonal interview, in which he apologized for his paſt intentions, and ſolemnly ſwore, in the preſence of the Archbiſhop of Rouen, that he would never again form any deſign againſt the perſon or liberty of the Earl. (Gifford, 1790: 182 [139 (2722)])

(16) That William had a perfect right to be considerably out of humour, cannot be doubted; and if he had not expressed his own sentiments at a proper juncture, and given the weight of his decision to the arguments and expostulations of the Whigs, it is impossible to say how long and how preposterously the Tories *might* have persevered in their most impracticable opinions, and again, how long the moderation and caution of the Whigs *might* have been able to sustain itself, and might have continued to maintain the peace of the community; in other words, whether a civil war *might* not have been the result, or at least the return of James. (Smyth, 1840: 69 [150 (6485)])

(17) Yet he resolved to prosecute the siege of Grenada with vigour, but in the meantime sent to remind Abu Abdalla of the promise he had made to deliver the city into his hands as soon as he *should* have subdued the strong places held by Zagal. It was now that Abu Abdalla, being delivered from his rival on the throne, resolved really to defend his faith. (Callcott, 1828: 226 [147 (3429)])

All the instances quoted indicate reasoning through the use of the modal verb followed by the perfective. The use of one modal or the other depends on the degree of the author's commitment to the truth of the propositions hedged. Thus, the modal verb *might* represents more theoretical supposition than *must*, which indicates more practical truth following from logical reasoning. Actually, this inferential meaning of *must* in Examples (14) and (15) is intensified by the use of the adverbials *needs* and *inevitably*. The modal form *should* also denotes some kind of obvious necessity that the designated action took place to be able to explain the attested historical events. A further function of inferential modals is to indicate discourse organisation, as shown in (16). The modal *might* is deployed to introduce a sequence of ideas in order to support the author's stance. These are given after the matrix *it is impossible to say…how long… the Tories might… how long the moderation and caution of the Whigs might… whether a civil war might*. In short, inferential modals allow writers to elaborate information as deduced from their own critical thinking based on historical evidences.

4.2 Deontic modals

Deontic modals represent 35% of modal cases in CHET. The distribution of forms is given in Figure 3 below:

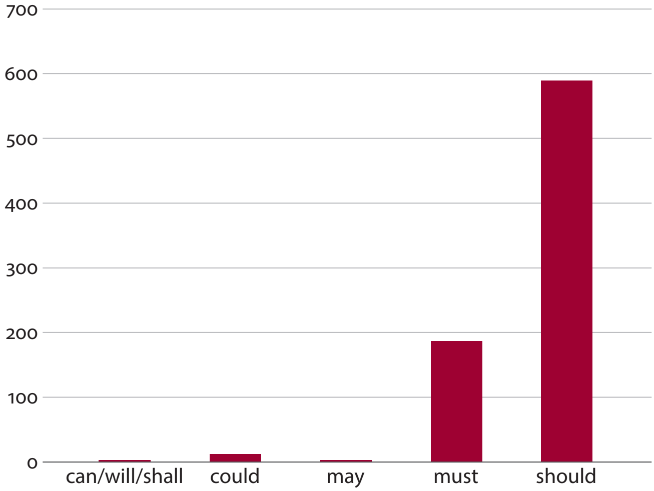


Figure 3. Distribution of deontic modals

The forms *should* and *must* are the preferred ones in the texts followed in the far distance by *could*.

(18) Who foraſmuch as he had deſpiſed the Condemnation of *Martin Luther*, and of his Books and Works, together with the Publication, Declaration and Monition thereof, and had kept in his Poſſeſſion, the Books and Works of the ſaid *Martin*, by which he was involved and intangled in the Sentence of the greater Excommunication, by the Authority of Pope Leo X. of happy Memory,

and for other juſt and lawful Cauſes, the ſaid Reverend Father inhibited and interdicted the ſaid *Forman*, that hereafter he *ſhould* not celebrate Maſs, nor preach publickly before the people, until he ſhould otherwiſe be diſpenſed with, under the Pain of Law. (Strype, 1721: 81 [126 (3006)])

(19) The Privy Council of England recommended that the county *should* be divided into two shires – Youghal to be the capital of the one, and Cork city of the other. (Bennett, 1862: 24 [155 (1860)])

The deontic forces of *should* in the examples above show two different motivations. In (18), the source is external, and prohibition is due to "lawful Cauſes", and insubordination will incur in penalty. That means that this minister is deprived of his evangelising and catechetical functions, and so he cannot do them, because an authority imposes lack of freedom of action (see Lyons, 1977: 827). The use of *should* in Example (19) follows from advice, as indicated in the matrix "The Privy Council of England recommended that…". As seen, the modal verbs are properly contextualised, and these contexts indicate the degree of imposition from the speakers. Similarly, *must* is deployed to clearly imply external obligation to perform an action, as in (20), below:

(20) Thus it was that this Country, before without a Name, fell to the *Muſcovites*, who call it *Siberia*, a Word ſignifying a Priſon, becauſe hither are ſent many Criminals, ſome to perpetual Baniſhment, and ſome for a Term of Years, who are obliged to ſhoot for their Livings or ſtarve, and to bring in a certain Quantity of Furs Weekly, or be ſeverely puniſhed. They *muſt* take particular Care that the Furs are without Holes or Stains of Blood, which make them very dextrous in ſhooting with a ſingle Ball at the Head of the Creature. (Bancks, 1740: 32 [130 (9501)])

In this example, criminals in Siberia are forced to kill animals for their living and give their skin in return. The skin should show almost intact to avoid punishment. In this context, the use of *must* indicates a requirement, and hence an imposition.

The verbs *can, will, shall* appear once with deontic nuances, and *may* appear three times with this value. In Example (21), the form *can* seems to communicate freedom of actions, and this is also contextually signalled by the expression "as they ſhall think convenient", which could be seen as somehow granting permission for free will:

(21) For Winter, when they go a Journey, they have what they call a travelling Wagon; in which they put their Beds, and Bedding. They *can* either ſit upright, or lie along, as they ſhall think convenient. They generally take good Store of ſtrong Liquor, Tongues, Hung-Beef, or any Thing that is ported: For there is but bad

> Entertainment upon the Road: They travel Night and Day. This is what I have been told; for I never went a Journey while I was there.
> (Justice, 1739: 20 [129 (4535)])

4.3 Dynamic modals

The modals *could, can, will* and *shall* are the most frequent forms with dynamic meaning. Less common is *may*, and this is followed by *would*, as shown in Figure 4, below:

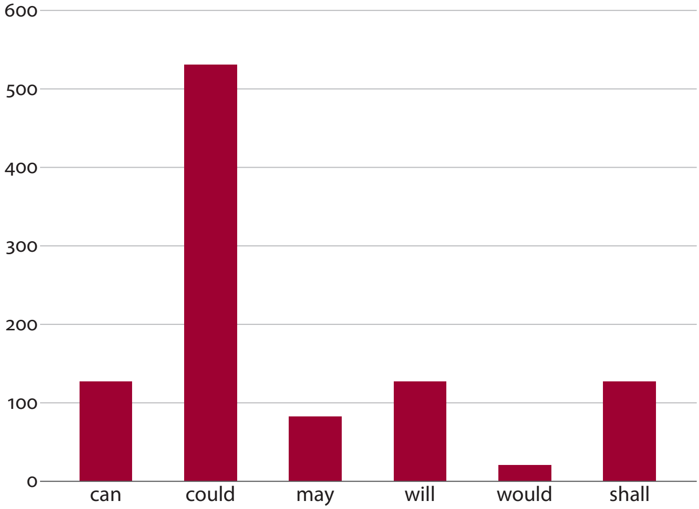


Figure 4. Distribution of dynamic modals

The following examples illustrate cases of *could*:

(22) They imagined the existence of a plot to destroy their race, and that the Jesuits were the magicians who were to effect it. The claim to supernatural influences which ran through the Jesuit teaching encouraged this idea, and the Indians *could* not understand that such power could only be beneficently exercised.
(Kingsford, 1887: 200 [159 (9299)])

(23) But, while I was in *Ruſſia*, there was one built, by Mr. BROWN, allowed by every Body to be as fine a Ship as *could* fail upon the Ocean.
(Justice, 1739: 36 [129 (8422)])

The form *could* in (22) shows the Indians' lack of ability to understand that the supposed supernatural powers of the Jesuits belong to the realm of beliefs, and so those powers are only hypothetical. This use of modality refers to extrinsic motivation, as the Indians have not been able to learn from the essentials of Christianity. In (23), the form *could* refers to the actual potentiality of the ship to sail the ocean. Rather than simply circumstantial, this modal verb clearly indicates an intrinsic feature concerning the design of the ship that makes it effectively fulfil the act of navigating the ocean.

The form *can* appears to behave in the same way as *could*, as shown in (24):

(24) To have attained such wonderful success in the development of his country's resources, and now to see the whole object of his life torn away from him by war, was a prospect the horror of which *can* only be measured by those who have made a study of the work he performed during the first period of his administration. (Burrows, 1895: 169 [161 (339)])

In this example, the form *can* indicates some restriction concerning the possibility that the event "a prospect of the horror…be measured" could be realised, as this depends on external factors. The same may be also true of *may* in (25), at least partially. First, this modal refers to the speaker and readers' potential to deduce the type of government King James had. Obviously, gaining this deduction requires that information on the topic has been previously attained by the conceptualiser to allow for this cognitive process. Secondly, this form might be also adopted to refer to the intrinsic potential of human beings to memorise and process information to elaborate new information.

(25) We *may* imagine what kind of Government King *James* deſign'd, when he was attended with ſuch a Council; and yet it is certain, ſome of thoſe who were Proteſtants wou'd have been turn'd out, if they had not abſented themſelves, and declin'd appearing at the Board. (Oldmixon, 1716: 61 [125 (4984)])

It seems evident from the analysis of these dynamic modals in CHET that they are context-sensitive (Vetter, 2015), and so their meanings depend on how the text is interpreted. Historians rely on these dynamic senses in order to describe properties of people, events and objects, which may be either internal or circumstantial. Essentially, dynamic modality does not seem to own a fully modalising effect, as it often denotes factuality. The modulating force of dynamic devices might be found in the authors' interest to avoid a potential face-threatening act, and thus these modals can be also labelled as negative-politeness strategies. The use of *may* in (25), for instance, makes even more patent this avoidance of imposition, as the writer selects this form rather than *can*, and this is probably because of the strong association of *may* with the expression of politeness, as described in Section 4.1.

5. Conclusions

In this paper, we have described the use of modal verbs in CHET. Our analysis of the texts has revealed that historians tended to mainly choose epistemic modals in order to hedge the information they present. Without further considerations on the nature of inferentiality, we have also detected that there are many modal verbs conveying epistemic necessity, following the terminology in van der Auwera & Plungian (1998). This modal sense is useful for these writers, as they are able to communicate deductions based on their reading-around of the topic, their own knowledge, and the documented evidence at hand. In this way, they can make a claim while self-protecting their public face, as these modulating devices appear to function as a negative politeness strategy.

Deontic modals have been registered to be the second most frequent modal type in CHET. The meanings of these verbs are frequently obligation, necessity and imposition, either internally or externally inspired, with *must* and *should*, while permission and freedom of action are expressed by means of *can, could* and *may*, for instance. Dynamic modality represents a quarter of the modal meanings in CHET, and these concern senses of disposition and potentiality. Writers make use of these senses to talk about internal and external conditions of participants, events, and objects without fully committing to its truth. These modal verbs are context-sensitive, and so their meanings depend on particular readings of the text. In general, dynamic modality has this obvious descriptive force, but it also seems to have a mitigating function, which is highly appreciated and strategically exploited by writers in their elaboration of historical discourse.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Allen, Cynthia. (1975). Old English modals. In Jane Grimshaw (Ed.), *Papers in the history and structure of English* (91–100). University of Massachusetts: University of Massachusetts Occasional Papers in Linguistics 1.

Alonso-Almeida, Francisco. (2015a). On the mitigating function of modality and evidentiality 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). Propositional epistemic and evidential strategies in Spanish and English texts on Computing. In Juan R. Zamorano et al. (eds.), *Thinking Modally: English and Contrastive Studies on Modality* (383–408). Newcastle upon Tyne: Cambridge Scholars Publishing.

Alonso-Almeida, Francisco & María-Luisa Carrió-Pastor. (2015). Sobre la categorización de seem en inglés y su traducción en español. Análisis de un corpus paralelo. *Revista Signos. Estudios de lingüística*, 48(88), 154–173. https://doi.org/10.4067/S0718-09342015000200001

Alonso-Almeida, Francisco & Margarita Mele-Marrero. (2014). Stancetaking in seventeenth-century prefaces on obstetrics. *Journal of Historical Pragmatics*, 15(1), 1–35. https://doi.org/10.1075/jhp.15.1.01alo

Auwera, J. van der & Vladimir A. Plungian. (1998). Modality's Semantic Map. *Linguistic Typology*, 2.1, 79–124.

Biber, Douglas; Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Campbell, Alistair. (1959). *Old English Grammar*. Oxford: Oxford University Press.

Chafe, Wallace. (1986). Evidentiality in English conversation and academic writing. In Wallace Chafe & Johanna Nichols (Eds.), *Evidentiality: The Linguistic Coding of Epistemology* (261–272). Norwood: Ablex.

Collins, Peter. (2009). *Modals and Quasi-modals*. Amsterdam and New York: Rodopi. https://doi.org/10.1163/9789042029095

Cornillie, Bert & Nicole Delbecque. (2008). Speaker commitment: Back to the speaker. Evidence from Spanish alternations. *Belgian Journal of Linguistics*, 22, 37–62. https://doi.org/10.1075/bjl.22.03cor

Denison, David. (1993). *English Historical Syntax*. London: Longman.

Haspelmath, Martin. (1994). Functional categories, X-bar-theory, and grammaticalization theory. *Sprachtypologie und Universalienforschung*, 47, 3–15.

Hoye, Leo. (1997). *Adverbs and Modality in English*. Essex: Longman.

Hoye, Leo. (2009). Modality in Discourse: The Pragmatics of Epistemic Modality. In Anastasios Tsangalidis & Roberta Facchinetti (Eds), *Studies on English Modality. In Honour of Frank Palmer* (99–131). Bern: Peter Lang.

Kemenade, Ans Van. (1992). The History of English Modals; a Reanalysis. *Folia Linguistica Historica*, 26, 143–166.

Lightfoot, David W. (1979). *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Lyons, John. (1977). *Semantics*, Volume 2. Cambridge: Cambridge University Press.

Palmer, Frank R. (1986). *Mood and modality* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Palmer, Frank R. (2001). *Modality and the English modals*. London: Longman.

Plank, Frans. (1984). The modals story retold. *Studies in Language*, 8.3, 305–364. https://doi.org/10.1075/sl.8.3.02pla

Roberts, Ian. (1985). Agreement parameters and the development of English modal auxiliaries. *Natural Language and Linguistic Theory*, 3, 21–59. https://doi.org/10.1007/BF00205413

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

Vetter, Barbara. (2015). *Potentiality. From Dispositions to Modality*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198714316.001.0001

Willet, Thomas. (1988). A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language*, 12, 51–97. https://doi.org/10.1075/sl.12.1.04wil

CHAPTER 9

A corpus-based study of some certainty adverbs in the *Corpus of History English Texts*

María José Esteve Ramos and Inés Lareo
Universitat Jaume I / Universidade de Vigo

1. Introduction and objectives

This chapter deals with an aspect of scholarly discourse that seems especially relevant in history writing from the late Modern English period: how to convince the reader of the veracity of facts and events using adverbs to show the writers' truthfulness. Historians tried to be like witnesses of events that wanted to be reliable and that said to the readers, in some way, "trust me! this is the truth", like witnesses when they testify at trials. History is in many ways a truth-based discipline; we rely on historical texts because we expect them to be rigorous and truthful. However, the evolution of the practice of writing history reflects how in the past things were very different. Moskowich (see Chapter 3, this volume) deals with the history of the discipline and reminds us that its very name in fact used to mean "a relation of incidents, either true or false, when it was first introduced in English in the 14th century". However, by the time of the late Modern period this had changed, and by 1756 the difference between history and story was attested in the *OED* (Moskowich, 2019). Thus, fictional or invented episodes would no longer be considered as part of the discipline, this coinciding with our current understanding of the subject. Moskowich emphasises this idea:

> Nineteenth-century scholars applied the scientific method previously described by John Locke (1690) as the "plain historical method" (Stromberg, 1951), and contemporary authors such as Humboldt (1822) corroborated such an approach when expressing his belief that History should in fact be exact, impartial and critical. This was precisely the origin of the present-day assumption that History is based on a collection of true and verified facts (Black, 1926; Stromberg, 1951).
> (Moskowich, 2017: 88)

If such a distinction was present in the scholarly community, then we can infer that the pursuit of truth in historical writing might also have been a relevant objective

https://doi.org/10.1075/z.225.09est

© 2019 John Benjamins Publishing Company

for writers. Persuading and convincing readers of the truthfulness of the facts presented would have been a much more notable feature of discourse than it is today. However, we might recall that this is still a time in which facts and fiction might coincide in historical literature, since history in this period "is the interpretation of the past and it is, therefore, subjective and not directly verifiable" (Moskowich, 2019). In this sense, different authors, conditioned by a variety of contexts and influences, may have made use of elements, such as certainty adverbs, in their discourse. By focusing on a study of adverbs of certainty, then, we may shed some light on the perception of historical discourse at the time.

Studies on the evolution of adverbs have adopted a variety of different perspectives and have analysed different periods, fields and types of texts (Alonso-Almeida, 2012; Conrad & Biber, 2000; Downing, 2002; Moskowich & Crespo, 2014; Simon-Vanderbergen & Aijmer, 2007). In Simon-Vanderbergen & Aijmer (2007) the semantic field of modal certainty is explored, and the evolution of the semantic and pragmatic use of these adverbs is considered. The research is based on corpora, although not historical ones. The authors explain that their criterion for choosing these adverbs was "that the adverbs should express the speaker's strong commitment to the truth of the proposition" (2007: 69). Our present aim is to present the attempts of writers of history to transmit this sense of truthfulness to readers, and hence we have also focused on this type of adverb. However, we will limit ourselves to the classification proposed by Quirk *et al.* (1985), as explained in the methodology section, below.

Adverbs contribute to assertions, and assertiveness is fundamental in convincing readers of the reliability of an author's point of view. These days we have numerous examples -especially in political discourse and in the media- where seeking the truth has become a job for experts. But how were certainty adverbs used in the late Modern Period? Moreover, how might this have been different in, or have otherwise affected, a wide range of specialised discourses in English? In this chapter, we will attempt to explore the use of some adverbs of certainty, and to ask whether the intentions of the writer may also be reflected in this specific element of discourse. Our approach here will be to use one of the most potentially useful tools for such a purpose, the *Coruña Corpus of English Scientific Writing* (CC), and more specifically, the subcorpus of History, the *Corpus of History English Texts* (CHET).

CHET, like the CC, includes a section with data on sociolinguistic variables such as sex, age and place of education. Thus, a second aim here is to see if there is a difference between writers of both sexes in the use of certainty adverbs. We might assume that women, who were at the time seeking to become part of the community of historians, may have used more adverbs of certainty than men, in order to be more assertive and successfully convince the reader of the veracity of what they were saying. In this period, women were pushing for acceptance in that exclusive group of writers, and such discourse devices may have afforded them a means of showing assertiveness.

A third aim is to explore the use of these adverbs according to different genres, a parameter which is also included in the corpus (Moskowich & Crespo, 2016). Not all genres are expected to contain an equal number of adverbs of certainty. Among the different possibilities, we would expect lectures to show a higher number of tokens than treatises, for instance. Lectures were written with the ultimate purpose of being read aloud by the author before an audience, a context in which the use of devices to give weight to the author's sense of certainty, such as adverbs, may well have been particularly appropriate.

Finally, it would be also interesting to explore the age factor, in both the eighteenth and nineteenth centuries, to see if this is a significant variable in the use of certainty adverbs.

To summarise, this chapter will explore the use of some adverbs of certainty in the *Corpus of History English Texts* (CHET) by focusing on variables such as time, sex, age and text-type. By considering the two centuries included in CHET, we will be able to see whether eighteenth-century historians are more or less assertive than their nineteenth-century counterparts, and to confirm whether our expectations regarding the use of these adverbs in history texts are correct.

2. Data and methodology

As noted above, our data will be drawn from the *Corpus of History English Texts* (CHET), one of the subcorpora of the *Coruña Coruña* (CC) (see Moskowich, this volume; Crespo & Moskowich, 2015, 2016; Moskowich & Parapar, 2008). The beta version of CHET under analysis contains 404,424 words extracted from eighteenth century texts (201,938) and nineteenth century texts (202,486). Each century includes 20 samples of texts, written by 20 different authors. The authors have different origins and the types of texts are also varied. The number of female historians included in the nineteenth century is six but in the eighteenth century it is only two. The difference here between the two centuries also reflects differences in society at the time, especially in the world of scholarship that the corpus seeks to represent.

As a theoretical basis for the study, we have used Quirk's *et al.* (1985: 615–627) classification of certainty adverbs, the so-called *disjuncts*. These are divided into two different types: *style disjuncts* and *content disjuncts*. The former are used to convey the speaker's comment on the style and form of what is being said, and the latter are used to make observations on the actual content of an utterance and its truth conditions. According to Quirk *et al.* content disjuncts are themselves divided into two subtypes:

A. “Expressing the degree of or conditions for truth of content”
B. “value judgment of content”.

Following this classification, we can find different groups within each type:

Type A. ‘Degree of truth’. These comment on the conditions for truth and state the sense in which the speaker judges what he says to be true or false.

Group (i) Express conviction, either as a direct claim (e.g. *undeniably*) or as an appeal to general perception (e.g. *evidently*). The adverbs included in this group are:[1] *admittedly* (0), *assuredly* (0), *avowedly* (5), *certainly* (43), *decidedly* (0), *definitely* (1), *incontestably* (0), *incontrovertibly* (1), *indeed* (105), *indisputably* (0), *indubitably* (0), *surely* (9), *unarguably* (0), *undeniably* (0), *undoubtedly* (18), *unquestionably* (1), *clearly* (12), *evidently* (13), *manifestly* (2), *obviously* (4), *patently* (0), *plainly* (16), *undeniably* (0).

Group (ii) Express some degree of doubt. The list of this type of adverbs includes: *allegedly* (0), *arguably* (0), *apparently* (17), *conceivably* (0), *doubtless* (5), *likely, maybe* (0), *most likely, perhaps* (76), *possibly* (20), *presumably* (2), *purportedly* (0), *quite likely, reportedly* (0), *reputedly* (0), *seemingly* (3), *supposedly* (0), *very likely* (30).[2]

Group (iii) State the sense in which the speaker judges what he says to be true or false. There is a reference to the “reality” or lack of “reality” in what is said. Here Quirk *et al*. group this type of adverbs according to three different senses: the ones that assert the reality: *actually* (16), *factually* (0), *really* (30); the adverbs that express a contrast with reality: *apparently* (17), *formally* (12), *nominally* (3), *theoretically* (1), *hypothetically* (0), *ideally* (0), *officially* (0), *ostensibly* (0), *outwardly* (0), *superficially* (0), *technically* (0); or the ones that claim that what is being said is true in principle: *basically* (0), *essentially* (6), *fundamentally* (1).

Having considered this classification and the adverbs listed therein, we have focussed our attention on *type A group (i)*, that is, adverbs expressing conviction. These are, in our opinion, the forms that best represent and reflect the objective of this study: to be assertive about the truth of what is being related. Did historians in the eighteenth and nineteenth centuries use adverbs of this type widely as a means of commenting on the conditions for truth and to show that what he or she was writing was really true?

---

[1] This sequence includes all the adverbs listed in each group and type by Quirk *et al.* (1985). The numbers in brackets represent the number of tokens (instances) found in CHET.

[2] Although Quirk *et al.* distinguish four variants of this type, that is, without pre-modification and pre-modified by *very, most* and *quite*, all the variants are included within these 30 tokens.

Examples (1)–(13) below show the 13 types from *type A group (i)* in Quirk's *et al.* classification attested in CHET:

(1) THIS Exception againſt the ſuppoſ'd Charters of Edgar, and the Duplicated Charter of William Rufus of England, is ſo *clearly* and fully voucht, that more need not be ſaid. (Anderson, 1705: 74 [123 (4935)])

(2) The laying any stress upon dreams, unless *evidently* calculated to answer some valuable purpose, only serves to prop up the old rotten cause of superstition, every degree of which may prove a source of uneasiness to some honest and good minds. (Cornish, 1780: 20 [137 (3472)])

(3) He was *obviously* and *avowedly* forced into war at last by the nation, and had no choice. (Burrows, 1895: 168 [161 (130)])

(4) But with the greateſt appearance are counterfit, by ſome obvious circumſtances *plainly* related by the moſt, if not all the Engliſh Hiſtorians, Ancient and Modern, who treat of thoſe times. (Anderson, 1705: 70 [123 (4239)])

(5) Permanently to occupy an office the duties of which were so *manifestly* incompatible with the devotion of his time and thoughts to the primary objects of his public life, had never probably been the intention of the primate. (Aikin, 1833: 375 [148 (8762)])

(6) Yet the calmer reflection of a prince would *ſurely* ſuggeſt, that the ſame acclamations were applied to every character and every reign: and if he had riſen from a private rank, he might remember, that his own voice had been the loudeſt and moſt eager in applauſe, at the very moment, when he envied the fortune, or conſpired againſt the life, of his predeceſſor. (Gibbon, 1788: 112 [138 (9612)])

(7) One part of this relation, is therefore *undoubtedly* falſe; and what is there to ſupport the reſt of it?. (Chapman, 1750: 44 [132 (7207)])

(8) The constitutional question as to the right of taxation, to any extent, not infringing the above rule, is considered as *definitely* settled. (Gray, 1872: 350 [156 (168)])

(9) The patent seems to have been procured by corruption, and the issue of copper coins was *unquestionably* quite too large: but the detriment was absurdly magnified. (Killen, 1875: 237 [157 (5211)])

(10) But their not conſulting before-hand with the other Governments was *certainly* a great overſight. (Penhallow, 1726: 85 [127 (1738)])

(11) I muſt here inſert a Remarkable Tranſaction, which ſhould have been mention'd in the beginning of this Year, when *indeed* it happen'd. (Tyrrell, 1704: 955 [122 (2837)])

(12) though a man, who was poſſeſſed of ſuch uncommon fortitude and honeſty as, in the preſence of a ſovereign ſurrounded by his adherents, to aſſert the ſupremacy of his rival, and boldly accuſe his principal ſupporter, is *indiſputably* deſerving of the higheſt praiſe an author can beſtow, and ſhould be tranſmitted to poſterity as an object of imitation to every member of a national aſſembly. (Gifford, 1790: 189 [139 (9919)])

(13) it is sufficient to observe, that no fact can be more *incontrovertibly* established, than that the Irish committed to writing, in their native language, immediately after the introduction of Christianity, not only the laws, bardic historical poems, &c. (Petrie, 1839: 38 [149 (3243)])

Having established our object of study, the next step was to search the corpus for the adverbs listed by Quirk *et al*. for this type and group using the *Coruña Corpus Tool* (*CCT*). The examples found were checked, and the senses which we were interested in were confirmed using the *Oxford English Dictionary Online* (*OED*). It is crucial in a study of this kind to verify not only the presence of these adverbs in the corpus (and hence that they were part of the language at the time), but also that the specific senses that we are interested in were in use. The *OED* includes evidence for all of these 13 types in the period 1700–1900.

3. Analysis

The corpus findings are set out in Table 1, which includes information on the variables that we are going to use in this study. The first of these, time, is in the first column; following columns provide information on sex, the name of the authors, the age of the author when she or he wrote the work, genres, examples found distinguishing different types and the number of tokens, and, finally, the total number of tokens and their normalised figures (nf) per sample.[3]

Table 1. Data and findings of certainty adverbs in CHET including nf

Date/ words	Sex	Author	Age	Genre	Type/Tokens	Tokens/nf 10,000
1704 10089	M	Tyrrell, James	62	Treatise	Indeed 3	3 2.97

---

[3] Figures have been normalised to 10,000 (nf) as a means of making the comparison possible and ensuring more reliable results.

Table 1. (*continued*)

Date/ words	Sex	Author	Age	Genre	Type/Tokens	Tokens/nf 10,000
1705 10066	M	Anderson, James	43	Essay	Clearly 5 Evidently 1 Plainly 5	11 10.99
1710 10111	M	Crawfurd, George	29	Treatise	Certainly 1	1 0.98
1716 10077	M	Oldmixon, John	44	Treatise	Surely 1 Indeed 1	2 1.98
1721 10088	M	Strype, John	78	Treatise	Indeed 1	1 0.99
1726 10192	M	Penhallow, Samuel	61	Treatise	Indeed 1 Certainly 2	3 2.94
1732 10135	M	Horsley, John	47	Treatise	Clearly 1 Plainly 4 Indeed 1 Certainly 2	8 7.89
1739 10005	F	Justice, Elizabeth	36	Travelogue	Indeed 7	7 6.99
1740 10057	M	Bancks, John	31	Treatise	Indeed 2 Certainly 1	3 2.98
1745 10006	M	Hooke, Nathaniel	Unknown	Treatise	Plainly 1 Surely 1 Indeed 1	3 2.99
1750 10187	M	Chapman, Thomas	33	Essay	Evidently 1 Manifestly 1 Plainly 2 Undoubtedly 3 Indeed 3 Certainly 1	11 10.79
1760 10059	M	Birch, Thomas	55	Essay	Certainly 1	1 0.99
1762 10114	F	Scott, Sarah	42	Treatise	Indeed 1	1 0.98
1770 10070	M	Adams, Amos	41	Lecture	Clearly 1 Indeed 4 Certainly 1	6 5.95

(*continued*)

Table 1. (*continued*)

Date/ words	Sex	Author	Age	Genre	Type/Tokens	Tokens/nf 10,000
1775 10020	M	Anderson, Walter	52	Treatise	Clearly 1 Evidently 2 Certainly 3	6 5.98
1780 10035	M	Cornish, Joseph	30	Biography	Evidently 1 Surely 1 Undoubtedly 1 Indeed 6 Certainly 4	13 12.95
1788 10014	M	Gibbon, Edward	51	Treatise	Surely 1 Indeed 1	2 1.99
1790 10319	M	Gifford, John	32	Treatise	Plainly 1 Indeed 2 Certainly 1	4 3.87
1795 10116	M	Adams, John	45	Treatise	Plainly 1 Undoubtedly 1 Indeed 1 Certainly 1	4 3.95
1800 10178	M	Stock, Joseph	60	Treatise	Surely 1 Undoubtedly 1 Indeed 7 Certainly 1	10 9.82
Total **18th c**		**201,938**				**100** **4.95**
1802 10158	M	Adolphus, John	34	Treatise	Avowedly 2 Undoubtedly 1 Indeed 1 Certainly 1	5 4.92
1805 10214	F	Warren, Mercy Otis	77	Treatise	Avowedly 1 Indeed 5	6 5.87
1810 10065	M	Bigland, John	60	Treatise	Undoubtedly 2 Indeed 4	6 5.96
1814 10017	M	Britton, John	43	Dictionary	Evidently 2 Indeed 1 Certainly 2	5 4.99
1820 10255	M	Hardiman, James	38	Treatise	Clearly 1 Evidently 1 Indeed 1	3 2.92

Table 1. (*continued*)

Date/ words	Sex	Author	Age	Genre	Type/Tokens	Tokens/nf 10,000
1828 10332	F	Callcott, Maria /lady	43	Treatise	Indeed 3	3 2.90
1833 10022	F	Aikin, Lucy	52	Treatise	Evidently 1 Manifestly 1 Indeed 5 Certainly 1	8 7.98
1839 10117	M	Petrie, George	49	Lecture	Clearly 1 Evidently 2 Incontrovertibly 1 Obviously 2 Indeed 7 Certainly 3	16 15.81
1840 9993	M	Smyth, William	75	Lecture	Clearly 1 Plainly 2 Surely 3 Indeed 5 Certainly 7	18 18.01
1844 10008	M	D'Alton, John	52	Treatise	Undoubtedly 1 Indeed 2	3 2.99
1855 10166	M	Masson, David	33	Textbook	Indeed 4	4 3.93
1857 10037	F	Sewell, Elizabeth Missing	42	Textbook	Indeed 7	7 6.97
1860 10043	F	Freer, Martha Walker	38	Treatise	Indeed 1	1 0.99
1862 10005	M	Bennett, George	40	Treatise	Undoubtedly 2 Indeed 1 Certainly 2	5 4.99
1872 10045	M	Gray, John Hamilton	58	Treatise	Definitely 1	1 0.99
1875 10083	M	Killen, William Dool	69	Treatise	Avowedly 1 Evidently 1 Obviously 1 Unquestionably 1 Certainly 1	5 4.95

(*continued*)

Table 1. (*continued*)

Date/ words	Sex	Author	Age	Genre	Type/Tokens	Tokens/nf 10,000
1884 10057	M	Breese, Sidney	84	Treatise	Evidently 1 Undoubtedly 1 Indeed 1	3 2.98
1887 10041	M	Kingsford, William	68	Treatise	Indeed 2 Certainly 1	3 2.98
1893 10730	F	Cooke, Alice M.	26	Article	Clearly 1 Undoubtedly 5 Indeed 1 Certainly 3	10 9.32
1895 10158	M	Burrows, Montagu	76	Treatise	Avowedly 1 Obviously 1 Surely 1 Indeed 12 Certainly 3	18 17.72
Total 19th c		**202,486**				**130** **6.42**
TOTAL		**404,424**			**Types 13/ 230**	**5.68**

In terms of the first variable, time, a slight difference between the use of these certainty adverbs is seen between the eighteenth (4.95) and nineteenth (6.42) centuries. Clearly, some changes in the development of scholarly discourse may have affected this trend and may explain why the use of certainty adverbs has increased here. The need for debate, the establishment of a more structured division of disciplines and an increase in the number of writers dealing with this discipline may have triggered a more assertive discourse.

The second variable, sex, was also analysed over time. Table 2 shows the *nfs* for male and female authors for the two periods:

Table 2. *Normalised figures* of certainty adverbs according to "time" and "sex"

NF/10,000	Subperiod	Sex	*nf*/10,000
5.79	Eighteenth century	Male	5.06
	Nineteenth century	Male	6.73
5.27	Eighteenth century	Female	3.97
	Nineteenth century	Female	5.70

The first column shows the normalised figures for the whole two-century period for male (5.79) and female (5.27) authors. The fourth column includes the normalised figures for each century. The similar levels of difference between the normalised figures for men and women in both centuries (5.06 – 3.97 = 1,09; 6.73 – 5.70 = 1,03) confirm the general findings for this variable. The tendency seen here was also observed by Moskowich & Crespo (2014), although they looked at a wider variety of adverbs in three different nineteenth-century subcorpora of the CC. Therefore, our first hypothesis, that women needed to be more assertive and to show certainty more overtly or more often than men, is clearly not supported by our findings.

In fact, the use of these adverbs by men is slightly higher than in the case of women (see Table 2 and Figure 1). Perhaps assertiveness may be connected with authority and visibility, something that has already been explored in other sub-corpora of the CC:

> In previous works on involvement in late Modern English scientific writing by women (Crespo & Moskowich, 2015), the idea was posited that that there is some kind of power asymmetry, in Lakoff's (1990) terms, together with the one which a priori assigns a more involved or less informational style to female writers than to male ones. In other words, women are less detached than men (Argamon *et al.*, 2003).
> Moskowich (2016: 75)

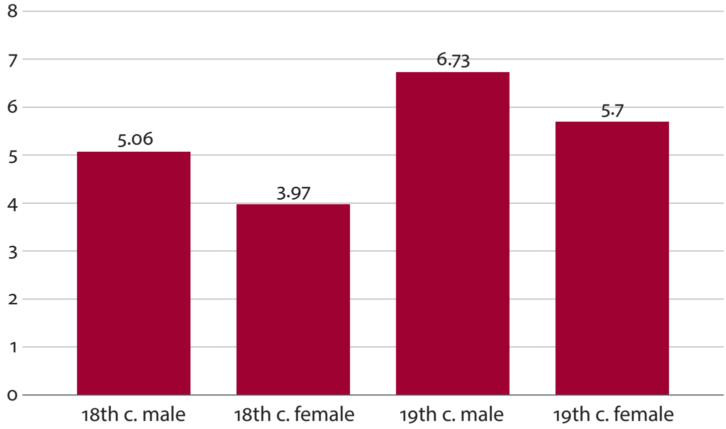


Figure 1. Distribution of certainty adverbs according to "sex" and "time" variables

In Moskowich's *Genre and Change in the Corpus on History Texts*, she points out that women "often worked in the shadows" (2017: 91). In comparing different subcorpora, she explains that the presence of female authors varies, and that a possible explanation for such variation might be found in social factors, for example in the role of women in society at the time. In the case of the Astronomy subcorpora, which lists only two female authors, she reminds us that "it was seen as inappropriate for women to observe the sky at night" (Herrero, 2007; Moskowich, 2012).

In the case of History, the role of authority clearly lay with men; however, women use this resource less frequently. It is true that the number of samples in the analysis may have conditioned these findings, but it seems unarguable that women historians revealed themselves to be much more neutral and restrained in terms of their language.

As a means of assessing the variable "age" we have grouped the authors by decade. Figure 2 sets out age ranges and normalised figures per 10,000 words for each age range.

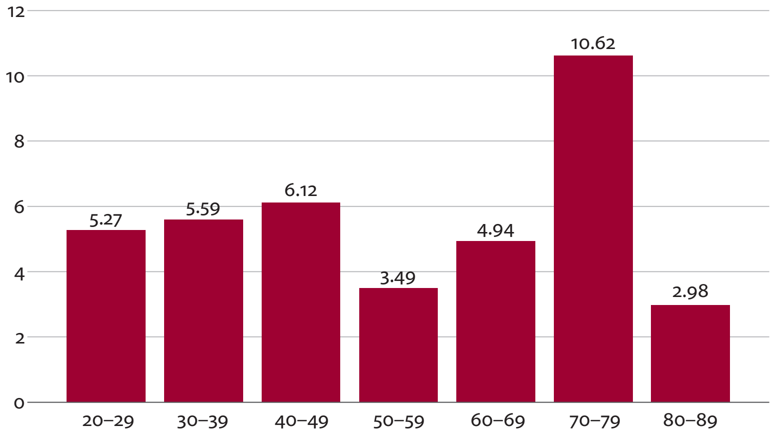


Figure 2. Distribution of certainty adverbs according to the "age" variable

Although our initial thought was that younger writers might need to be more assertive than mature authors, the findings do not support such a premise, and in fact seem to show the contrary. The figures increase with age until the 50–59 age range, when it decreases notably. From their fifties onwards, writers in the eighteenth and nineteenth centuries could be considered elderly people, in that life expectancy was significantly shorter than it is today. Do our findings imply that writers within this age range no longer need to be as assertive as before? Another possibility is that they do not need to be so assertive, in that their reputations (related to the factor of age) allow them to be more neutral and less emphatic and persuasive. However, we see that in the following two decades the numbers increase once more, only to fall away dramatically (2.98) for the oldest writers (aged 80–89). This requires a broader exploration with other disciplines and a larger set of data before any clear conclusions can be drawn.

The results of the fourth variable, "genre", are shown in Figure 3:

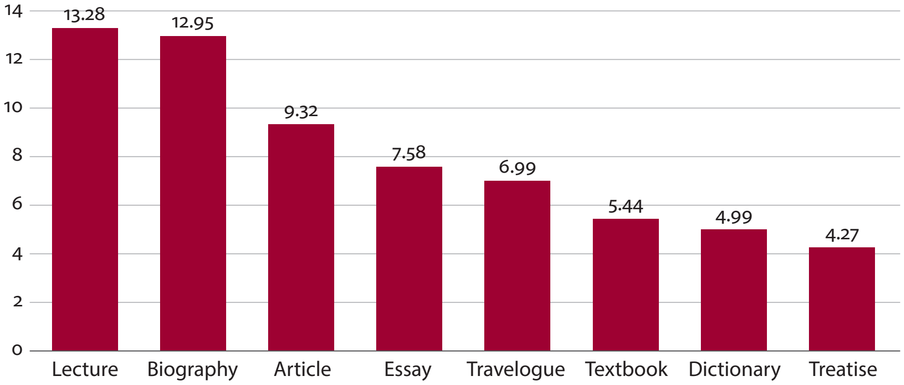


Figure 3. Distribution of certainty adverbs according to "genre" variable

These results take frequency into account, and hence the genres with higher frequencies are shown first. As we noted in the Introduction, we have assumed that lectures might present a higher number of adverbs of certainty than treatises, in that the features of these two text-types are completely different, as are also their intended audiences. The findings here support our premise. Lectures, along with biographies, are the two text-types with the highest numbers of results. An interpretation these results is enriched if we reflect on the observations of Kohnen:

> Text types may be thought of as dynamic patters of communication combining aspects of function, context and form. For the purposes of historical analysis, the basic sameness of a text type in the course of the centuries may be maintained by providing a minimal functional-situational definition (…) (Kohnen, 2001: 198)

Lectures, Essays and Articles are more directly associated with the academic sphere. This may mean that in such contexts – where a priori discourse would tend to be more neutral – there is a need to convince members of the audience, reassuring them of the truth of the facts presented. On the contrary, Treatises have the lowest findings, although we might add that these texts do not represent the same genre that we would label a treatise today. As Moskowich (this volume) observes, the term was used at that time to refer "almost to any piece of writing of some extension", not necessarily having anything to do with the academic life, and therefore not necessarily requiring the reader to be convinced or shown a high degree of certainty in the statements made. If we take a look at Figure 3 again, even the genre dictionary generates a slightly higher number of cases than treatises.[4] "Treatise" cannot be understood as what we nowadays label as a treatise. Consequently, its place at the end of the scale, implying that this genre is the one with the lowest number of certainty adverbs, is the expected result.

---

[4] In CHET, this genre represents, as Moskowich (this volume) describes, "half a catalogue and half a dictionary of technical terms".

Another interesting aspect of our analysis is concerned with which adverb types are used most frequently. These data are presented in Figure 4.

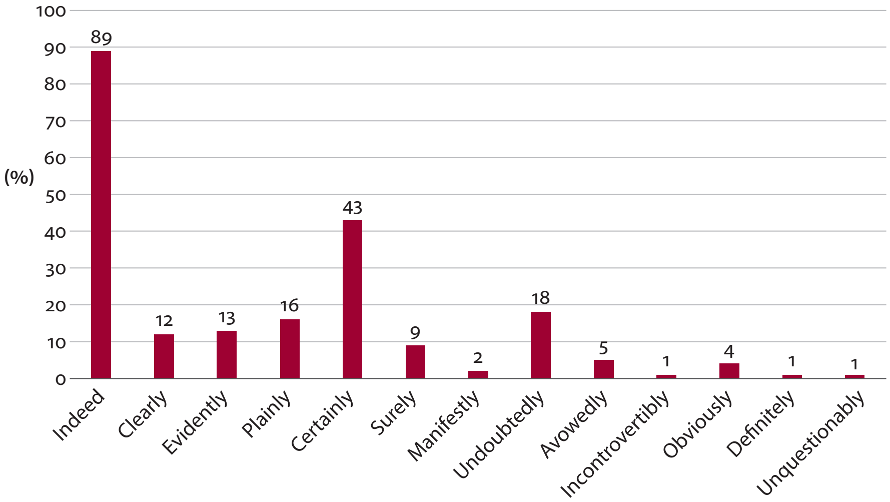


Figure 4. Distribution of certainty adverbs according to adverb-type

As these findings show, the three tokens *indeed*, *certainly* and *undoubtedly* significantly outnumber all others. This would confirm our expectations in the sense that these three forms express a particularly strong sense of conviction. Also of note here is that in Moskowich & Crespo's study (2014), although they analysed more subcorpora and more adverbs and did so for the nineteenth century only, the ranking of the adverbs is the same. Our two first results, *indeed*, and *certainly*, are similarly placed in their study, if we omit those types which do not figure in our analysis here. However, apart from this quantitative aspect, it is also interesting to focus on some of the examples extracted from the material (shown earlier already) in order to test our premises:

(10) But their not confulting before-hand with the other Governments was ***certainly*** a great overfight. (Penhallow, 1726: 85 [127 (1738)])

(7) One part of this relation, is therefore ***undoubtedly*** falfe; and what is there to fupport the reft of it?. (Chapman, 1750: 44 [132 (7207)])

In the case of these Examples (10) and (7), the use of *certainly* and *undoubtedly* clearly shows the intention of convincing the reader of the truth of what is being said. In Example (11), the use of *indeed* (the most frequent form) shows the same intention:

(11) I muſt here inſert a Remarkable Tranſaction, which ſhould have been mention'd in the beginning of this Year, when ***indeed*** it happen'd.
(Tyrrell, 1704: 955 [122 (2837)])

In the case of *incontrovertibly*, shown in (13) below, although the frequency of this example is very low, the intentionality shows similar intensity:

(13) it is sufficient to observe, that no fact can be more ***incontrovertibly*** established, than that the Irish committed to writing, in their native language, immediately after the introduction of Christianity, not only the laws, bardic historical poems, &c. (Petrie, 1839: 38 [149 (3243)])

A more detailed analysis of the individual examples would enhance and extend the findings from the current quantitative results, allowing for a greater understanding of the topic. This, however, will be the matter of future research.

4. Concluding remarks

In early periods, one of the scientific-academic fields in which the writing itself is the only evidence for the facts presented is history. Historians of the pre-contemporary era dealt exclusively with written materials and/or objects, with no recording of visual evidence or oral messages to help new generations to verify the facts. The role of the historians consisted not only in writing about the past, by rephrasing and quoting other authors' words, but also in trying to convince the reader about the truth of the facts stated. Beyond mere description, they were also prone to include evidence, making value judgments on past events. The use of references and other sources helped them to reinforce their message and to convince the audience of the validity of their writing. So, do history writers of this period resort to adverbs of certainty to make their claims more trustworthy? We have tried to answer this question by using CHET, since historical corpora have recently been shown to be powerful tools for such undertakings, something noted by Taatvisainen & Jucker:

> The availability of historical corpora and other electronic resources together with the increasing sophistication of corpus-linguistic tools has opened up a whole range of new research paradigms, and it has become possible for historical pragmaticists to probe into research questions that could not have been answered before.
> (Taatvisainen & Jucker, 2015: 12)

We will conclude by restating our premises and whether these were validated in the analysis, as well as noting what conclusions we can draw from this.

As a general observation, it is clear that the mechanism of persuasion will not be marked exclusively through the use of certainty adverbs, yet it is indeed interesting to analyse the use of such adverbs in the late Modern English context, and thus to enrich our understanding of how scholarly English, in this case history writing, operates. A future comparison with other periods, especially contemporary English, would constitute another interesting line of research.

The findings reported in this study have shed some light on the behaviour of these adverbs, in line with some of our initial assumptions. However, other findings have run contrary to our expectations. The analysis of the first variable, time, showed an increase in the use of these adverbial devices from the beginning of the eighteenth century to the end of the nineteenth century. This might be evidence of growing rivalry between history writers. The increase in the number of authors writing about history and the general evolution of the discipline, with the emergence of a differentiation between story and history in the middle of the eighteenth century, might also have been one of the reasons underlying this increased use. Concerning the sex variable, the findings revealed that men used certainty adverbs of this type more than women, contrary to our expectations, and showing that women writers-scholars did not need to be more persuasive or assertive than men to be recognised by readers as part of the authoritative group of writers. On the other hand, our assumptions about the variable text-type have certainly been confirmed in the analysis. Lectures and Biographies showed the highest number of examples, as we had predicted. As for age, no tendencies can be discerned with certainty, and it is necessary to widen the study to other subcorpora and larger data sets.

This study has also shown that several aspects here require qualitative research to complement and broaden the current findings. The present study has, at least, provided the first step towards uncovering what is true and what is not true in history writings in late Modern English. And adverbs, *indeed*, will help us to find this out.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P and *Rede de Lingua e Literatura Inglesa e Identidade III* (ED431D2017/17). These grants are hereby gratefully acknowledged.

References

Alonso-Almeida, Francisco. (2012). Sentencial evidential adverbs and authorial stance in a corpus of English computing articles. *Revista Española de Lingüística Aplicada. Monograph: conversing ways of applying Corpus Linguistics*, 15–31.

Conrad, Susan & Douglas Biber. (2000). Adverbial marking of stance in speech and writing. In Susan Hunston & Geoff Thompson (Eds.), *Evaluation in Text: authorial stance and the construction of discourse* (56–73). Oxford: Oxford University Press.

Crespo, Begoña & Isabel Moskowich. (2015). A Corpus of history texts (CHET) as part of the *Coruña Corpus* Project. *Proceedings of the international scientific conference Corpus linguistics-2015*. St. Petersburgh: St. Petersburgh State University.

Crespo, Begoña & Isabel Moskowich. (2016). At close range: prefaces and other text types in the Coruña Corpus of English Scientific Writing. *Revista de lenguas para fines específicos*, 22/1, 213–237.

Downing, Angela. (2002). "*Surely you knew*!" *Surely* as a marker of evidentiality and stance. *Functions of Language*, 8/2, 253–285.

Herrero, Concepción. (2007). "Las mujeres en la investigación científica". *Criterios*, 8, 73–96.

Kohnen, Thomas. (2001). On Defining Text Types within Historical Linguistics: The Case of Petitions/Statutes. *European Journal of English Studies*, 5/2, 197–203. https://doi.org/10.1076/ejes.5.2.197.7308

Moskowich, Isabel. (2012). "A smooth homogeneous globe" in CETA: Compiling late Modern Astronomy Texts in English. In Vázquez, N. (ed.), *Creation and Use of Historical English Corpora in Spain*. Cambridge: Cambridge Scholar Publishing.

Moskowich, Isabel. (2016). Pronouns as stance markers in the *Coruña Corpus*: an Analysis of the *CETA, CEPhiT* and *CHET*. In Francisco Alonso (Ed.), *Stancetaking in Late Modern English Scientific Writing. Evidence from the Coruña Corpus. Essays in Honour of Santiago González y Fernández-Corugedo* (73–91). València: Universitat Politécnica de València.

Moskowich, Isabel. (2017). Genre and change in the Corpus of History English Texts. *Nordic Journal of English Studies*, 16/3, 84–106.

Moskowich, Isabel. (2019). *An Introduction to CHET, the Corpus of History English Texts*. In Isabel Moskowich, Inés Lareo, Paula Lojo Sandino & Estefanía Sánchez Barreiro (Comps.). *Corpus of History English Texts (CHET)*. Accessed [15/03/2019] from http://hdl.handle.net/2183/21849

Moskowich, Isabel & Begoña Crespo. (2014). Stance is present in scientific writing, *indeed*. Evidence from the *Coruña Corpus of English Scientific Writing*. *Token: A Journal of English Linguistics*, 3, 91–114.

Moskowich, Isabel & Begoña Crespo. (2016). Classifying communicative formats in *CHET, CECheT* and others. *EPiC Series in Language and Linguistics*, 1, 308–320. https://doi.org/10.29007/kvjx

Moskowich, Isabel & Javier Parapar. (2008). Writing Science, compiling Science. *The Coruña Corpus of English Scientific Writing*. In María Jesús Lorenzo Modia (Ed.), *Proceedings from the 31st AEDEAN Conference* (531–544). A Coruña: Universidade da Coruña.

Quirk, Randolph; Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik. (1985). *A Comprehensive Grammar of the English Language*. New York: Longman.

Simon-Vanderbergen, Anne-Marie & Karin Aijmer. (2007). *The Semantic Field of Modal Certainty*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110198928

CHAPTER 10

How intimate was the tone of female History writing in the Modern period?

Evidence from the *Corpus of History English Texts*

Begoña Crespo
Universidade da Coruña

1. Introduction

This paper is concerned with female authors and the attitude to their topics, as conveyed through their writing. Late Modern as well as twentieth- and twenty-first century scholarly writing has been said to be broadly objective in tone (Atkinson, 1999; Hyland, 2002; Dontecheva-Navrátilová, 2013), showing little or no trace of personal involvement. As such it was intended to offer the reader reliability and trustworthiness, basic pillars of academic objectivity; indeed, scholarly and scientific writing has been characterised by "clarity, economy, rational argument supported by evidence, caution and restraint" (Bennett, 2009: 52). In recent times, however, it has been claimed that such discourse is not as objective and detached as we might suppose (Hyland, 2002) and that authors tend to express their thoughts and feelings through their texts and make some kind of connection on an emotional level with their readership. Also, female writing in particular is said to present a more intimate tone than male writing (Argamon *et al.*, 2003; Biber and Burges, 2000; Palander-Colin, 2006).

In this paper I will explore the extent to which the tone of Modern female history writing was intimate or involved, especially in comparison to non-scholarly works by women written at the same time. Intimacy and emotion conveyed through language involves both the addressor and the addressee in the communicative process; in writing this corresponds to the author and reader, in line with the interpersonal function in Systemic Functional Linguistics (Halliday & Matthiessen, 2004).

I will examine some of the linguistic strategies that have been recognised as indicators of authorial presence (Biber, 1988; Quirk *et al.*, 1985) and the reactions or feelings that can be aroused in the readership. These linguistic strategies will

https://doi.org/10.1075/z.225.10cre

© 2019 John Benjamins Publishing Company

be further examined from the point of view of two variables: time and genre or communicative format.

The initial hypothesis is that there will be no significant differences between scholarly and non-scholarly writing here, in the sense that our focus is on history, one of the soft sciences, and also because women's language in general tends to be more personal and to favour interaction more than in the case of men (Lakoff, 1990). As Hyland noted: "Experienced writers select rhetorical options for projecting authority and engaging with readers that reflect the epistemological assumptions and social practices of their fields, with more explicit authorial involvement in the soft disciplines" (2002: 1098).

To this end, the study will be divided into four sections. Section 2 will present the author-reader interaction and the way in which writers manifest their own viewpoints and opinions. Section 3 will describe the data used, which are taken from the *Corpus of History English Text* (CHET), a subcorpus of the Coruña Corpus of English Scientific Writing (CC), with data on Modern texts by women drawn from the *Penn-parsed Helsinki Corpus of Modern British English* (PPCMBE), (Kroch *et al.*, 2016). Section 4 will set out the analysis of these data and a discussion of the findings. Some concluding remarks will then be presented in the fifth and final section.

2. Authorial presence in scientific discourse: Identity and interaction

Impersonality in writing has traditionally been interpreted as the most direct and persuasive strategy the author has at her disposal in order to communicate academic content to the reader without being present in the text (Hyland, 2001). However, it has recently been shown (Kuo, 1999; Harwood, 2005; Hyland, 2008; Dontecheva-Navrátilová, 2013; Pagliawan, 2017) that the opposite tendencies, represented by subjectivity and interpersonality in the objective reporting of findings to a discipline-specific community, are manifested through signs of authorial identity and the writer-reader relationship.

In linguistics, involvement is most commonly related to the notion of subjectivity, even though the two are sometimes treated separately (Pander Maat & Degand, 2001; Simó, 2012: 79). As Finegan has argued, "subjectivity concerns the involvement of a locutionary agent in a discourse, and the effect of that involvement on the formal shape of discourse – in other words, on the linguistic expression of self" (1995: 1).

Intimacy or involvement is linguistically represented in written texts by features such as first and second person pronouns, private verbs, evaluation mechanisms such as hedges, amplifiers, modality through possibility modals, causative subordination, emphatics, discourse particles, and the pronoun *it*. These are some of the features used by Biber (1988: 101–121) to determine the degree of involvement or detachment of texts using factor analysis within the framework of his Multidimensional Analysis, particularly in terms of Dimension 1: Involved vs informational. This list of features will be the starting point for the analysis of data in the present study. Exploring the visibility of the author and her interaction with the reader is based on Bakhtin's (1986) idea of the dialogic nature of language use and the dialogic overtones of individual utterances. A couple of years earlier, Widdowson (1984: 59) argued that "the writer has to conduct his interaction by enacting the roles of both participants" and that in this way the text is built up as a series of writer responses to the reader's anticipated reactions (Thompson, 2001: 58). The linguistic structures used are simply ways in which writers can bring the underlying dialogue to the surface (Thompson, 2001: 59). This dialogic nature endows the writing with a touch of closeness and intimacy between author and reader. Showing involvement is a consensus-seeking mechanism used by members of an epistemic community, with authors seeking readers' acceptance of and commitment to their proposals.

From the seventeenth century onwards, under the influence of Empiricism, science began to be conveyed in an object-centred manner, leaving aside the author-centred model. The authors' main task was now to show reliable results that attracted the reader's attention in a clear and simple way, seeking an air of credibility and trustworthiness. Yet after nearly two and a half centuries of moving away from any inclusion of emotional traces in their work, the discourse of scholars today seems to be undergoing a pendular swing towards the acceptance of subjectivity, to the point that the notion of authors positioning themselves in their own writing is not only feasible but is in fact recommended as a means of conveying authority. As noted above, this is what present-day research on academic discourse tells us. During the eighteenth century, a detached style had become the goal in most forms of public discourse, and particularly in scientific-scholarly writing (Puente-Castelo, 2017). This need for objectivity brought about the gradual emergence of the passive, which in linguistic terms frees the description of an experiment from the personal experience of its author. The following two centuries saw the continuation of this objectivising trend in scientific discourse, partly supported by the Romantics' insistence on an absolute genre division between "Poetry and Matter of Fact, or Science" (Wordsworth, 1802: 69). In fact, it has been claimed that this development dates from the nineteenth century (Halliday, 1988: 166; Taavitsainen, 1994: 339). González-Álvarez & Pérez-Guerra (1998) have shown that scientific writing between the fifteenth and the seventeenth centuries moved from a more involved to a more informational style. Atkinson (1999) also claimed that the involved versus informational dimension in articles from the *Philosophical Transactions* of the Royal

Society was geared towards detachment, in consonance with Biber and Finegan's (1997) earlier work in which they concluded that scientific writing was more elaborate and impersonal after the seventeenth century, and that impersonality extended to the twentieth century (Monaco, 2017).

As well as seeking to confirm or reject the presence of intimacy in female scholarly writing, this study will attempt to compare academic and non-academic discourse in the same sense. To this end, some of the linguistic features will be here described in detail, looking at women's language as well as at the changes implied in applying time and communicative format as variables in the analysis.

3. Materials and methodology

Two corpora have been used: texts written by women from the beta version of the CHET subcorpus of the CC, and texts written by women in the PPCMBE. Both contain samples of texts produced between 1700 and 1900.

CHET represents late Modern history writing and the female authors compiled in it (see Table 1) are the following:

Table 1. Samples from CHET

Author	Title of work	Publication date	Communicative format	No. of words
Justice, Elizabeth	*A Voyage to Russia: describing the Laws, Manners, and Cuſtoms, of that great Empire, as govern'd, at this preſent, by that excellent Princeſs, the Czarina. Shewing the Beauty of her Palace, the Grandeur of her Courtiers, the Forms of Building at Petersburgh, and other Places: with ſeveral entertaining Adventures, that happened in the Paſſage by Sea, and Land.*	1739	Travelogue	10,005
Scott, Sarah	*The History of Mecklenburgh, from the Firſt Settlement of the Vandals in that Country, to the Present Time; including a Period of about Three Thouſand Years.*	1762	Treatise	10,114
Warren, Mercy Otis	*History of the rise, progress and termination of the American revolution. Interspersed with Biographical, Political and Moral Obſervations. In three volumes. Vol. I.*	1805	Treatise	10,032

Table 1. (*continued*)

Author	Title of work	Publication date	Communicative format	No. of words
Callcott, Maria Lady	*A Short history of Spain. In two volumes. Vol. II*	1828	Treatise	10,333
Aikin, Lucy	*Memoirs of the Court of King Charles the First. In two volumes. Vol. I*	1833	Treatise	10,022
Sewell, Elizabeth Missing	*A first history of Greece*	1857	Textbook	10,057
Freer, Martha Walker	*History of the reign of Henry Iv. King of France and Navarre.*	1860	Treatise	10,102
Cooke, Alice	*The Settlement of the Cistercians in England. The English Historical Review, Vol. 8, No. 32. (625–648)*	1893	Article	10,761
	TOTAL			**81,347**

Women's texts in the PPCMBE represent non-scientific discourse. The list of authors can be found in Table 2:

Table 2. Samples from PPCMBE

Author	Title of work	Publication date	Communicative format	No. of words
Austen, Jane	*Private letters from Chapman, R. W. (ed.). 1952. Jane Austen's letters to her sister Cassandra and others. 2nd ed. London: Oxford University Press.*	1805–1808	LETTERS_PRIV	9,650
Davys, Mary	*The northern heiress: or, the humours of York. A comedy.*	1716	DRAMA_ COMEDY	10,294
Montagu, Mary Wortley	*Letters of the Right Honourable Lady M--y W---y M----e: written, during her travels in Europe, Asia and Africa, to persons of distinction. Dublin, 1763.*	1718	LETTERS_PRIV	9,344
Montefiore, Lady Judith	*Private journal of a visit to Egypt and Palestine by way of Italy and the Mediterranean. London: Rickerby. Reprinted 1975 (Jerusalem: Hebrew University).*	1836	TRAVELOGUE	10,195

(*continued*)

Table 2. (*continued*)

Author	Title of work	Publication date	Communicative format	No. of words
Nightingale, Florence	*Vicinus, Martha, and Bea Nergaard (eds.). 1990. Ever yours, Florence Nightingale. Cambridge, MA: Harvard University Press.*	1888–1889	LETTERS_PRIV	3,302
Nightingale, Florence	*Vicinus, Martha, and Bea Nergaard (eds.). 1990. Ever yours, Florence Nightingale. Cambridge, MA: Harvard University Press.*	1890	LETTERS_PRIV	6,201
Reeve, Clara	*The champion of virtue: a Gothic story. Colchester.*	1777	FICTION	9,432
Turner, Sharon	*The history of the Anglo-Saxons, from their first appearance above the Elbe, to the death of Egbert: with a map of their ancient territory. London.*	1799	Treatise*	8,743
Victoria, Queen of England	*Buckle, George Earle (ed.). 1926. The letters of Queen Victoria. A selection from Her Majesty's correspondence and journal between the years 1862 and 1878. Second series, vol. 1: 1862–1869. London: Murray.*	1863–1865	LETTERS_PRIV	9,368
Yonge, Charlotte Mary	*Simmons, Clare A. (ed.). 2001. The clever woman of the family. Broadview Press.*	1865	FICTION	9,329
	TOTAL			**85,858**

* Classified as history by the original compilers, but that is not a communication format classification for our purposes here. Applying the criteria used for classification in the CC, this text is a treatise.

The overall word count of all the samples used in the analysis is 167,355, with 81,497 words from CHET and 85,855 from PPCMBE. In order to achieve more rigorous results in the analysis, figures have been normalised to 10,000 words.

Two different retrieval tools have been used: the Coruña Corpus Tool (CCT) for retrieving data from CHET, and AntConc (Laurence Anthony's freeware corpus analysis toolkit for concordancing and text analysis) for PPCMBE.

The set of lexico-grammatical features to be searched for, and the corresponding specific items, are presented in Table 3:

Table 3. Set of linguistic features

Set of features	Individual items
1st and 2nd p. pronoun	*I, me, mine, you, yours, we, us, ours*
Hedges	*at about, something like, more or less, almost, maybe, sort of, kind of*
Amplifiers	*Absolutely, altogether, completely, enormously, entirely, extremely, fully, greatly, highly, intensely, perfectly, strongly, thoroughly, totally, utterly, very*
Private verbs	*Anticipate, assume, believe, conclude, decide, demonstrate, determine, discover, doubt, estimate, fear, feel, find, forget, guess, hear, hope, imagine, imply, indicate, infer, know, learn, mean, notice, prove, realize, recognize, remember, reveal, see, show, suppose, think, understand*
Emphatics	*for sure, a lot, such a, real+ADJ, so+ADJ, DO+V, just, really, most, more*
possibility modals	*can, may, might, could (cou'd)*
causative subordination	*because*
discourse particles	*well, now, anyway, anyhow, anyways*
pronoun it	*it*

I have combined automatic information retrieval systems with manual disambiguation. On occasions, some of the forms may represent distinct lexical categories, as in the case of private verbs like *doubt, fear, notice* which, by a process of conversion or zero derivation, can also play the role of nouns (see Example (1)).

(1) But the distress and annoyance which it could not fail to occasion to the puritan clergy, – all, to a man rigid Sabbatarians, was no doubt the real inducement to this measure; and it fully answered its unwise and uncharitable purpose. (Aikin, 1833: 356 [148 (4664)])

Other forms are not lexically but pragmatically ambiguous: their role in fact depends on the communicative intention of the writer and on the context. This is the case with *now*, in Example (2), which can be either a time adjunct or a discourse particle.

(2) Enter Sir. Loobily in a Piss-burnt Periwig, a great Riding Coat, and dirty Linnen.
Wel. to. Gam. Now, Gamont, look and tremble.
I Am. I was afraid, Sir. Loobily, we should not have had your Company, and that you were so taken up with your Horses, you could not find Time to see your Mistress. (Davys, 1716: 45)

I am aware that this is not a complete multi-factorial analysis, since only separate features are analysed. However, even if my results cannot be claimed to be conclusive in Biber's sense, they might nevertheless reveal patterns present in late Modern scientific writing, as well as indicate trends worth pursuing in future research.

The following section will present the data used in the analysis together with an interpretation of these data.

4. Analysis and discussion

As mentioned above, my total data set amounts to some 167,355 words, of which 81,497 are from eighteenth- and nineteenth-century history samples representing scholarly writing, and 85,858 words are from non-scholarly texts written by women during the same period, these latter from the PPCMBE corpus. I am conscious of the fact that it is not an especially large sample from which to extract data, but we might bear in mind that this scenario, representing a scarcity of women writers, mirrors well the social reality of the late Modern English period. All in all, then, the data can give us a picture of how women are mediated in their own works, whether scholarly or not.

The set of features to be investigated are listed in Table 4, below, including their occurrence in the texts, in both raw and normalised figures.

Table 4. Raw figures for linguistic features

Features	**Frequency**	**NF/10,000**
1st and 2nd p pronoun	4,120	246.18
Hedges	98	5.86
Amplifiers	624	37.29
Private verbs	1,913	114.31
Emphatics	1,014	60.6
Possibility modals	810	48.4
Causative subordination	49	2.93
Discourse particles	47	2.8
Pronoun 'it'	1,488	88.91

All these features contribute to the creation of a sense of proximity between writer and reader. In general, the features that stand out are first and second person pronouns (246.18), private verbs (114.31) and the pronoun *it* (88.91). Emphatics (60.6) and amplifiers (37.29) come fourth and fifth, respectively, with the remaining features present at very low frequencies of occurrence. As graphically shown in Figure 1 below, the overall count for involvement features is 6% of the total number of words analysed in all the samples of late Modern female writing.

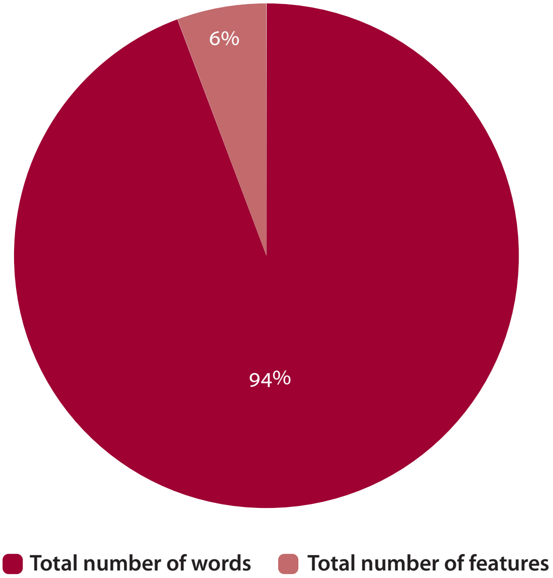


Figure 1. Involvement features in CHET and PPCMBE

Within this percentage, first and second person pronouns are the most common features. First person pronouns help authors to build a coherent self-image. They are used by writers as mechanisms to make explicit their opinions, beliefs, judgements and can act as markers that help organise the structure of the text by playing several communicative roles (see Tang and Johns, 1999, for their continuum of authorial presence). First person pronouns are clearly the most evident manifestation of the self.

As Hyland has pointed out: "Authorship in academic writing in English both carries a culturally constructed individualistic ideology and places the burden of responsibility for the truth of an assertion heavily on the shoulders of the writer" (2002: 1110). He adds that "the authorial pronoun is a significant means of promoting a competent scholarly identity and gaining acceptance for one's ideas". These considerations typify contemporary academic writing. The low frequency of such pronominal forms in late Modern texts could be interpreted as an indication that few or no traces of identity, acceptance or responsibility are present. Example (3) below illustrates the use of first person pronouns to reveal authorial identity and opinion.

(3) After this, we came to the Church: The Outſide of which did not tempt me to take a View of the Inſide; but, being deſired, could not refuſe the Civility of the Stranger; and I found it much beyond my Expectation.
(Justice, 1739: 4 [129 (762)])

Second person pronouns are reader-reference pronouns, useful to get the reader involved in the content of the text by assigning him/her an active, participatory role. Real interaction takes place when writers include their readership in their works as part of the community of practice. Emotions expressed by authors create connections between the audience, the author and the subject, ultimately building a relationship between the audience and the text. The inclusion and negotiation of knowledge facilitate credibility and scholarly success. Some of the reactions used in this kind of interaction are approval, acceptance, collaboration, involvement or persuasion, all of these being manifestations of the interpersonal function, as seen in Example (4):

> (4) As I have told you, that the Water all Winter is froze into a fix'd Subſtance, to bear ſo great a Weight upon it; you will imagine that they are at a Loſs for Water: But, indeed, they are not. (Justice, 1739: 21 [129 (4930)])

Private verbs indicate "intellectual states such as belief and intellectual acts such as discovery. These states and acts are 'private' in the sense that they are not observable" (Quirk *et al.*, 1985: 1181). Biber (1988: 105) observes that there is a strong correlation between high counts of private verbs and the degree of involvement in a text. *Understand* in (5) is one such example:

> (5) IF we would rightly understand the conditions of Cistercian work in England at the time when the order was taking root in this country, we must look to the early history of the movement of which it formed a part. From this we must endeavour, to some extent, to reconstruct both the system under which that work was carried on, and the position to which the order had attained in Europe at large at the time. (Cooke, 1893: 625 [160 (12)])

The relationship of the pronoun *it* with involvement lies in the general reference of the pronoun, meant to indicate either an animate being or an abstract concept (Biber, 1988: 225–226: Mischke, 2005: 101) which has been already mentioned, and which is thus known to both the writer and the reader. It brings written texts closer to orality, and high occurrences tend to show a closer sense of contact or intimacy. An instance of this can be seen in (6):

> (6) Being one even plain, the stormy ocean which environs makes it a perilous habitation; it has therefore been surrounded by a strong dyke, to propel the assaulting waves, and to give to its natives the sleep of security. Its vicinity to the continent invests it with a charm superior to the attractions of the other islands and islets, which swarm round Jutland. It contains three or four parishes, and about as many villages. (Turner, 1799: 27)

The comparison between scholarly and non-scholarly writing shows that only 23% of the different linguistic features under analysis here are employed by late Modern female history writers, as opposed to 77% in the samples from PPCMBE.

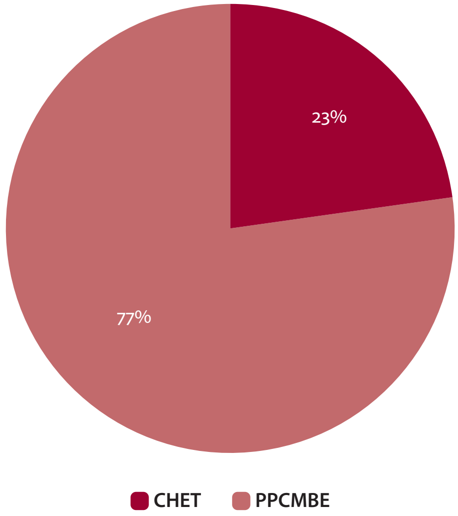


Figure 2. Comparing involvement in CHET and PPCMBE

The frequency of occurrence of involvement markers in non-scholarly samples triples that of scholarly texts (see Figure 2 above), this clearly going against my initial hypothesis that a large difference between these two kinds of prose was not expected because history is in fact one of the soft sciences.

In what follows two variables will be applied to the analysis of involvement features in the overall data set.

4.1 The time variable

Because samples of texts have been collected from the eighteenth and nineteenth centuries, a diachronic analysis of linguistic forms here is possible. Figure 3 below illustrates the use of these features in each century and each kind of writing, scholarly (CHET) and non-scholarly (PPCMBE).

Normalised figures confirm that the low frequency of involvement in the history samples is even lower in the nineteenth than the eighteenth century. Features in non-scholarly language are more frequent and their use increases over time (747.09–1033.2), contrary to what happens in the language of history writing, where a decrease is seen (484.11–228.09). As we can appreciate from Figure 3, the gap between CHET and PPCMBE samples is bigger with the passing of time. These findings fit well into the objectivising trend that characterises scientific-scholarly writing since the eighteenth century. We can confirm, then, that this specialised discourse follows a path towards impersonality and detachment, which would eventually consolidate in the twentieth century.

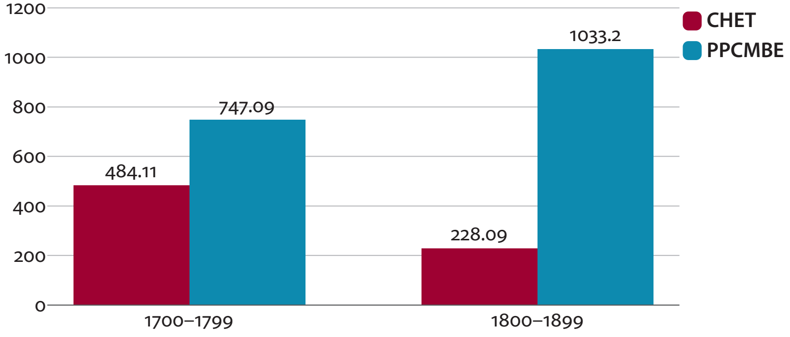


Figure 3. Involvement features over time (nf)

While in the eighteenth century there are some noticeable peaks in the use of certain involvement markers, such as first and second person pronouns (147.12), private verbs (102.39) and the pronoun *it* (90.46), there is a far lower and steady use of these markers from 0 to 16.7 in the nineteenth century, this reflecting the object-centred reporting of facts (see Figure 4 below). Previous claims by authors such as Biber and Finegan (1997), González-Álvarez & Pérez-Guerra (1998) and Atkinson (1999), then, are confirmed here.

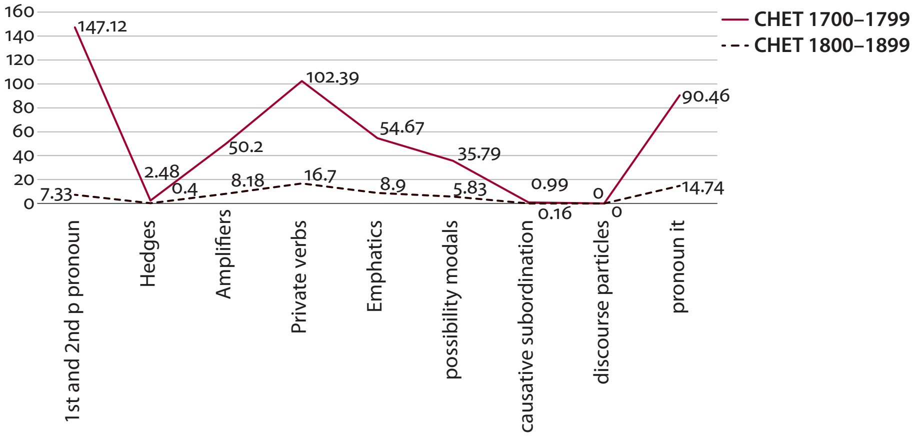


Figure 4. Involvement features in CHET in two periods

Figure 5 shows that in our non-scholarly data the frequency of occurrence of these features is quite homogeneous in the two centuries, the highest peak being that for personal pronouns (499.22–448.74) and the lowest for discourse particles (2.7–5.81).

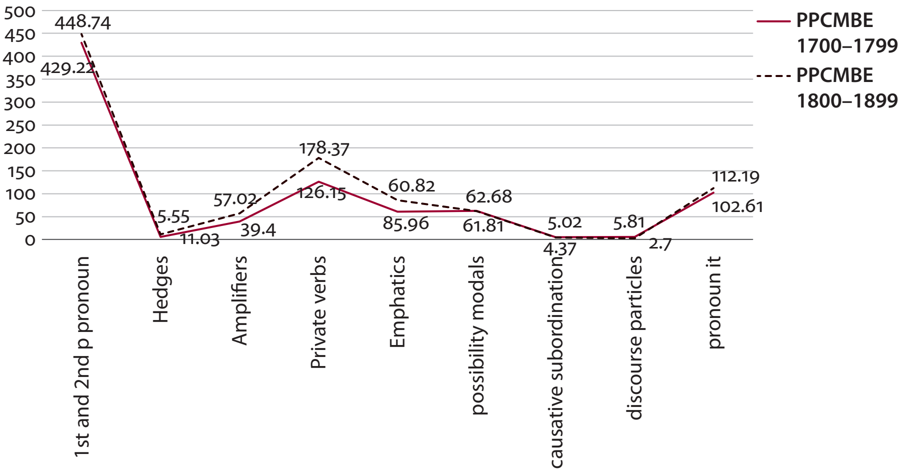


Figure 5. Involvement features in PPCMBE in two periods

So, data from these two corpora illustrate that whereas similar frequencies of occurrence of these linguistic resources are seen in non-scholarly discourse in the eighteenth and the nineteenth centuries, there is a remarkable change of frequency in the language of scholarship in these periods.

The second variable to be examined is communicative format.

4.2 The communicative format variable

Following the proposal in Moskowich & Crespo (2016: 2), instead of using the terms *genre* or *text-type* I will use the recently coined expression 'communicative format' to refer to the interdependence of form and function in a text:

> It is undoubtedly the case that texts are produced with a clear function, in that the main aim of human language is to achieve some kind of response on the part of the receiver. However, depending on the kind of response the sender/addressor envisages, that is, the function of the text, form will vary. Hence, there is no absolute independence of form and function, and texts adopt forms depending on the function they perform (telegram, advertisement, treatise…). This mutual dependence means that form and function can be seen as a whole, one which ultimately cannot be wholly divided. For this reason, we believe that the symbiosis between the form and function of a given communicative act deserves a new name: communicative format.

Since the samples currently under scrutiny serve to contrast scholarly and non-scholarly language, the written communicative formats in each set of samples are different. In the corpus of history texts, we have four: treatise, textbook, travelogue and article; there are also four in the PPCMBE texts: private letters, drama-comedy, travelogue and fiction. The information about the frequency of occurrence of involvement features in each format will now be explored, allowing for a more fine-tuned description of the writer-reader interactive relationship.

In the CHET texts, travelogue (689.66) has the highest frequency of occurrence (normalised figures are shown in Figure 6). Travelogue is defined in the *OED* as "an (illustrated) lecture about places and experiences encountered in the course of travel". This very definition seems to invoke an intimate tone.

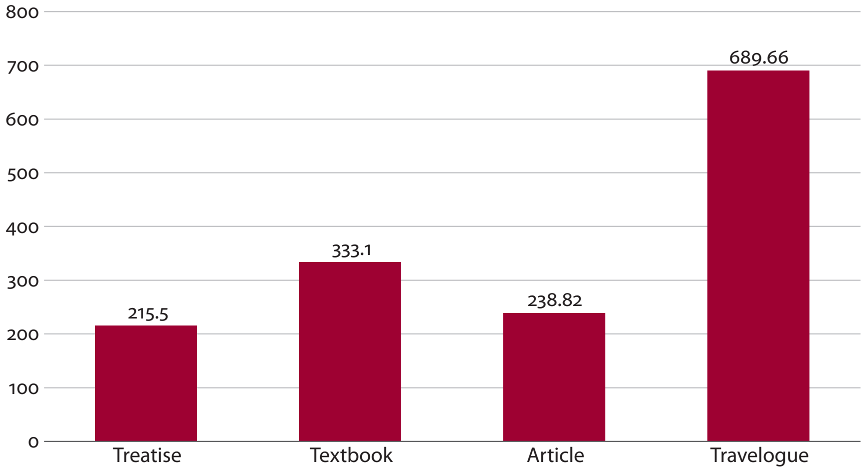


Figure 6. Involvement and communicative formats in CHET

The second in terms of frequency if occurrence is textbook (333.1) which is defined in the *OED* as "a book used as a standard work for the study of a particular subject; now usually one written specially for this purpose; a manual of instruction in any science or branch of study, esp. a work recognized as an authority". In that its function is mainly instructive, authors need to make some attempt to approach the reader and to write convincingly.

The interpersonal tone is realised by first person plural pronouns, with both nominative and objective cases (*we*, *us*). These fulfil an inclusive role which brings the reader closer to the text and thus creates a friendly atmosphere (Yang Xinzhang, 2017: 349). In an instructive work of this sort, successful transmission of knowledge requires the engagement of the student/reader.

Treatises (215.5) contain the lowest amount of involvement features. This is in line with the trend towards objectivity, and the informational nature and use of specific formats to convey scientific-academic knowledge.

In the PPCMBE texts, representing non-scholarly language (Figure 7), drama-comedy is the format containing the highest frequency of occurrence of involvement features per 10,000 words (1420.24). It is followed by private letters (1060.08) which, in principle, should yield a high degree of author intimacy. Curiously enough, and in contrast with what was seen in CHET, travelogue exhibits the lowest number of occurrences (449.36). The category 'travelogue' could perhaps be considered as a relatively common or general format within history writing, that is, not especially scholarly in tone. Hence, it is perhaps of little surprise that its frequency of linguistic forms denoting authorial presence is closer to those found in CHET than in other PPCMBE formats.

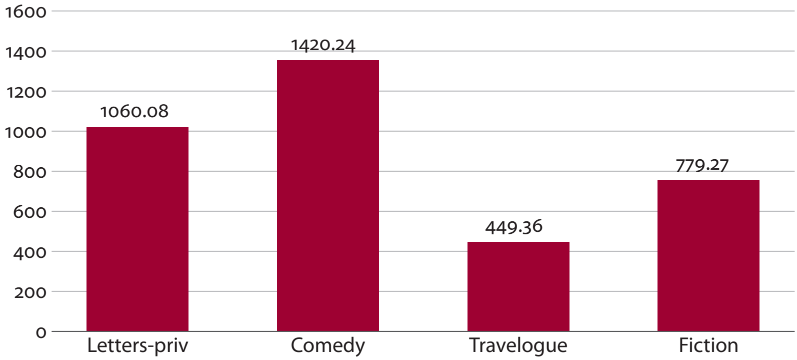


Figure 7. Communicative formats in PPCMBE

In this case, the communicative formats with the highest frequency of use are those which represent an informal, colloquial register or everyday conversation and those that express emotions or any kind of personal feelings.

In what follows I will present the most frequently used features denoting intimacy and authorial presence, paying special attention to how the use of these can be interpreted in late Modern scientific works written by women.

4.3 Detailed analysis of the most frequent involvement features

Considering the individual frequency of occurrence of each set of involvement markers in the two groups of texts, I have noticed further differences which can help us identify more characteristics of female scientific-scholarly writing in late Modern English. Personal pronouns, private verbs and the pronoun *it* exhibit the highest frequency in non-scientific texts, whereas private verbs, the pronoun *it* and emphatics are the three most frequently used features in scientific-scholarly samples. The findings are set out in Figure 8:

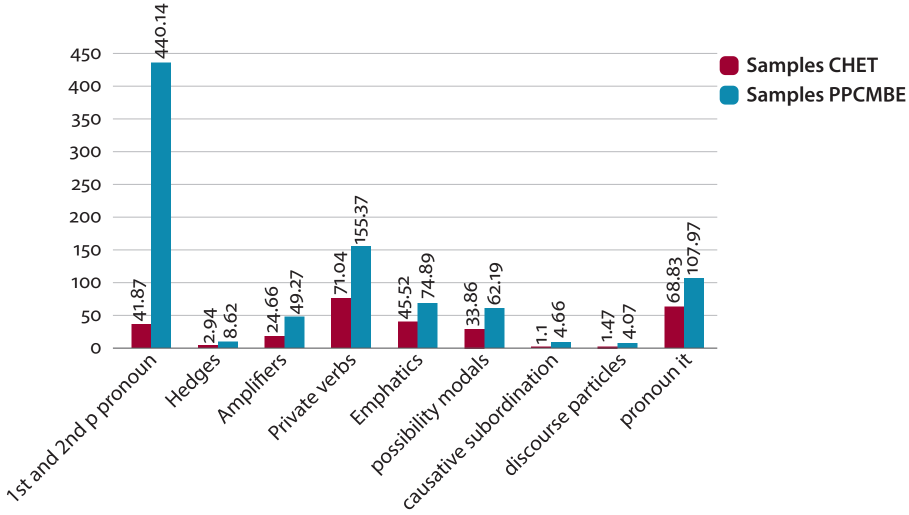


Figure 8. Comparison of individual involvement features in CHET and PPCMBE

First and second person pronouns are only fourth in the involvement cline in history texts.

As can be inferred here, the basic difference lies in the frequency of use of personal pronouns. To explain this, it might be interesting to look at the different occurrences and uses of first person pronouns, on the one hand, and second person pronouns, on the other. These results expressed in normalised figures (nf/10,000) confirm the scant presence of personal pronouns in scientific samples, especially when compared to non-scientific prose. All in all, reference to the authorial self through first person pronouns predominates over second person pronouns in both kinds of writing. In fact, the author's identity is thus referred to in history texts at a frequency of 38.77 per 10,000 words, with second person pronouns hardly present (1.96). The same tendency, but on a far greater scale, is attested in non-scientific prose (323.2 for first person vs 116.94 for second person). Figure 9 represents these proportions graphically.

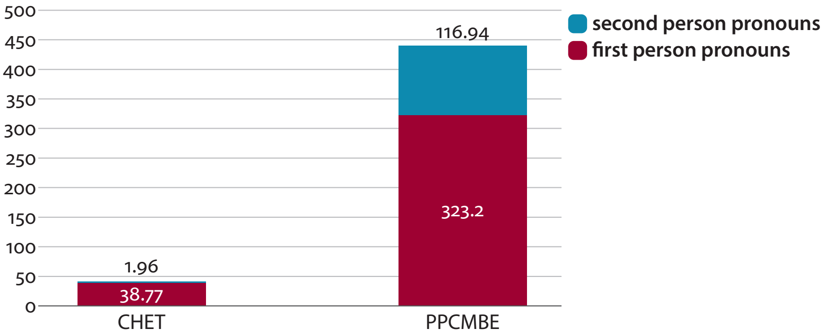


Figure 9. First and second person pronouns, nf = 10,000

Some uses of first person pronouns are illustrated in (7) to (9) below:

(7) But I cannot ſay there was any thing remarkable, more than a very fine new Pew in the Church for the Mayor and Aldermen, and ſome Relicks of an old Monaſtery, a ſmall Diſtance from the Church, which ſtands upon a very high Precipice. (Justice, 1739: 2 [129 (301)])

(8) We were bleſt with fine Weather, but ſmall Gales of Wind. We were a Fortnight in our Paſſage to Elſinore, a Place in the King of Denmark's Dominions, half Way to Ruſſia, where the Captains are obliged to go a Shore; and as the Wind was not fair, we had an Opportunity of going with him. (Justice, 1739: 4 [129 (624)])

(9) The movement itself was but one in the great series of revivals within the Benedictine order which were the outcome of the mighty, but ever-changing, influence of the idea of extreme renunciation. We are not, however, here concerned with the shape in which the Cistercians received this idea. (Cooke, 1893: 625 [160 (152)])

As we can see in Figure 10 below, a dramatic decrease in the use of first person pronouns can be attested in the history samples (from 134.7 to 7.33). This reinforces the idea that women's scientific writing had adapted to the patterns of the period, these dictated and assimilated by the scientific community of practice. The absence of self-reference, then, reflects the consolidation of an informational register.

As for second person pronouns, there are few addresses to the reader in the eighteenth (7.95) and none in the nineteenth century. There is a clear decline in the use of pronouns which, we can assume, is in line with the changing discursive practices of the discipline, itself brought about by the birth of nineteenth-century historiography.

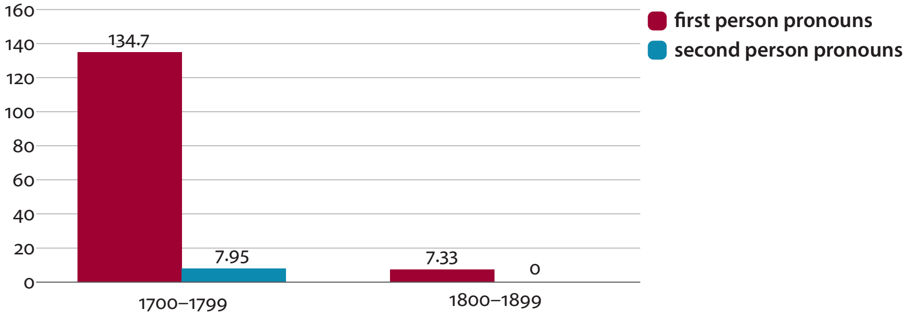


Figure 10. Pronouns in CHET over time

As Moskowich (2017: 81) has noted:

> History or historiography (…) was heavily influenced by the Positivist ideas of Auguste Comte (1798–1857) throughout the 19th century, and the objective description of facts tended to be the primary concern of writers. Perhaps in order to be respected by other discourse communities, authors of history had to adopt the supposedly objective perspective that had been so successful in other fields.

In Example (10) below, the author appeals to the target audience so as to engage them and to urge them to take on her message. She establishes a relationship of closeness with the readership so that they can build a picture of the author's communicative intention.

(10) As I have told you, that the Water all Winter is froze into a fix'd Subſtance, to bear ſo great a Weight upon it; you will imagine that they are at a Loſs for Water: But, indeed, they are not. (Justice, 1739: 21 [129 (4910)])

Results in CHET are not replicated in PPCMBE which, once again, may be a valid means of outlining female history writing in the Modern period. Figure 11 illustrates the data obtained in common speech:

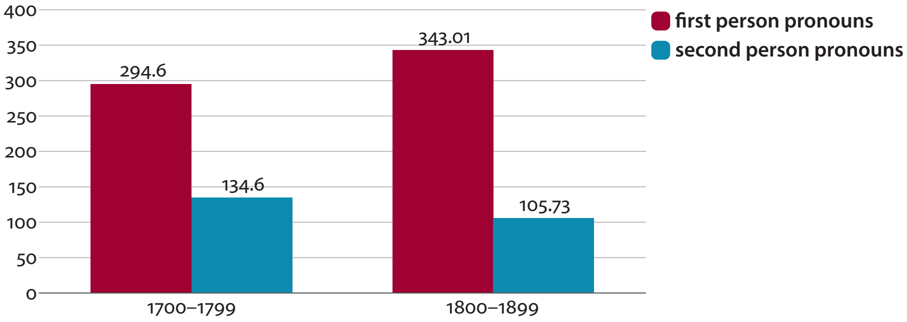


Figure 11. Pronouns in PPCMBE over time (nf)

A slight increase in the frequency of occurrence of first person pronouns here (294.6–343.01) is probably explained by the communicative formats used by authors (predominantly private letters). This same reasoning could apply to the reduction of second person pronouns (134.6–105.73) in the nineteenth century. Focusing on the occurrence of these pronouns according to the communicative format in which they appear, we note that in non-scientific samples, comedy (1420.24) and private letters (1060.08) exhibit the highest number, as might be expected.

As for communicative formats (see Figure 12), it is comedy in which most involvement markers are found, probably due to the oral-like nature of this communicative format.

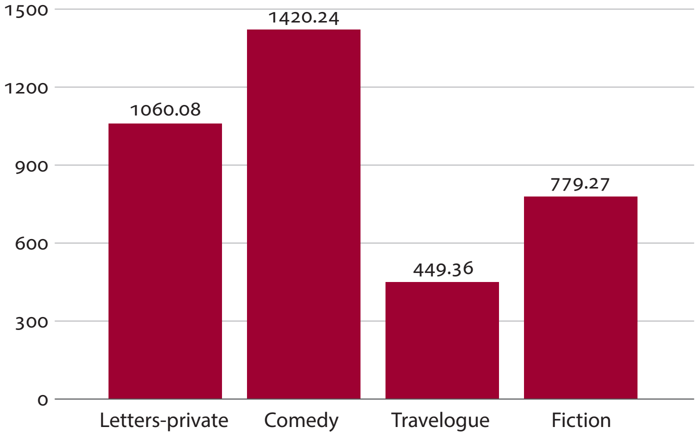


Figure 12. Pronouns per communicative format in PPCMBE

In the case of scientific texts we find a different tendency, as shown in Figure 13.

Travelogue within scholarly writing presents a very large comparative frequency of occurrence (287.85 forms per 10,000). It was only after the eighteenth century that writing about past events in which the author had taken part began to be considered as falling within of the realm of science and scholarly writing. The critical examination of sources, the search for authentic materials, and the creation of objective narratives about past stages of the human condition endowed the discipline of history with its modern scientific character. In the eighteenth century, under the influence of Rationalism, the historian resorts to evidence she thinks is relevant and then writes an account which she believes to be accurate and illuminating. This kind of discourse may conform to the canons of a literary genre, as well as to those of a more scholarly inquiry. Well into the century, such writing may even contain a thesis or theory to explain the actions of the past in some domain (Cook, 1988). Stromberg (1951) claimed that this form of historicism did not come to maturity until the following century: "to the romantics these

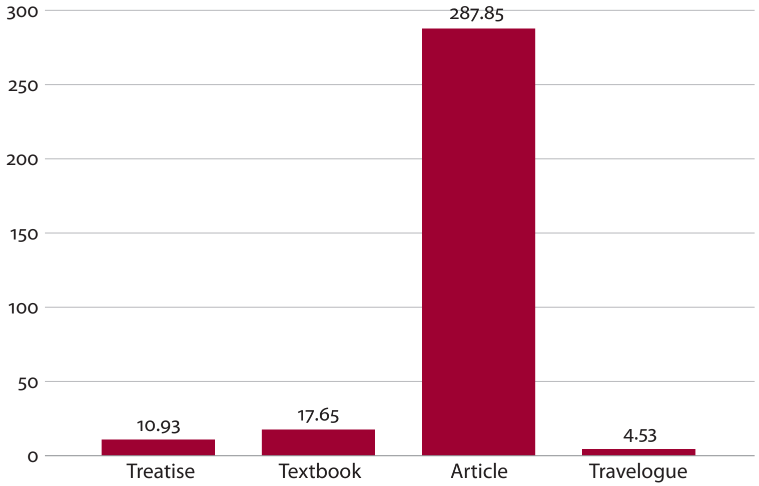


Figure 13. Pronouns per communicative format in CHET

eighteenth-century rationalists seemed completely useless: unreal and scholastic" (Stromberg, 1951: 301). Indeed, as mentioned in previous chapters, it would be in nineteenth-century rationalistic historiography that the scientific method was finally applied (Porter & Ross, 2003). This socio-external consideration on the shift from history to historiography might well account for the very notable occurrence of pronouns in the travelogue format which, incidentally, is one of the samples here from the early eighteenth century.

Another of the most frequent linguistic features is that of private verbs (so called by Quirk *et al.*, 1985: 203). These verbs refer to mental activities (*believe, think*) which in turn refer to sensations, cognition, emotions, states of perception, and states of bodily sensation. In this sense, authors using private verbs clearly reveal their involvement with the subject under discussion, especially so, as we will see, when they are preceded by first person pronouns (see Example (11) below). Figure 14 below sets out the frequency of occurrence of this feature in each corpus.

As can be seen, the presence of private verbs in the PPCMBE corpus (155.37) more than doubles their occurrence in CHET (71.04). Despite history being one of the soft sciences, hence more likely to use involvement features, this set of attitudinal markers seems to fall behind in comparison with non-scholarly texts. Yet this is the most common feature (71.04) expressing involvement in the scientific corpus here, and the second most common in the non-scientific one (155.37).

The use of private verbs denoting intimacy and closeness is reinforced by the parallel presence of first person pronouns. Thus, involvement is expressed through

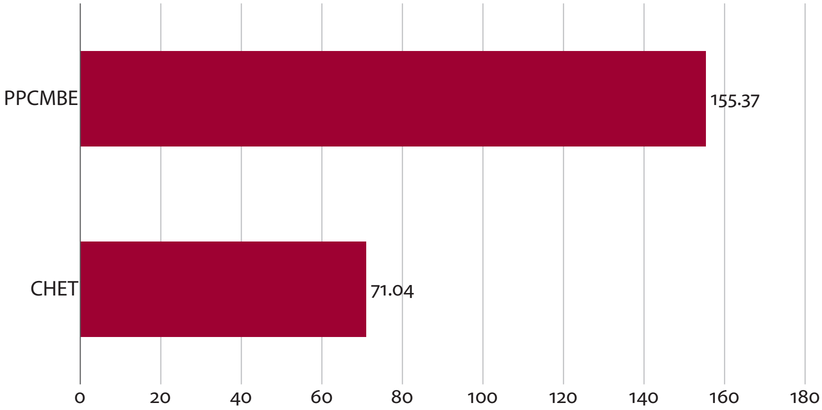


Figure 14. Overall frequency of private verbs

mental states, cognitive acts, feelings or perceptions as symptoms of the author's presence in the texts.

(11) In the Winter, they have very fine Illuminations, ſuch as, I believe, there is not the like in any Place. They are Four Times a Year: Upon her Majeſty's Birth-Day, the Day ſhe was named, that of Her Coronation, and New-Year's-Day; the Yearly Expence of which is Fifty Thouſand Pounds. (Justice, 1739: 22 [129 (5136)])

Moreover, we note that the frequency of occurrence of private verbs follows opposite trends over time in the two corpora (see Figure 15 below).

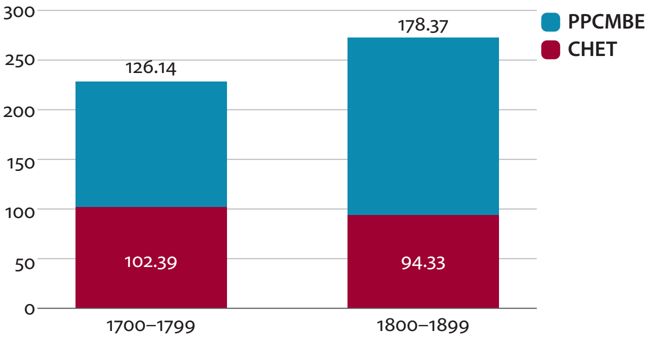


Figure 15. Private verbs in CHET and PPCMBE over time

An increase in the occurrence of these forms can be attested in nineteenth-century non-scientific material whereas a slight decrease is observed in scientific material, once again this being in line with the objectivising patterns in science writing.

The normalised figures in CHET do not seem to indicate an emphasis on interpersonal relations in the sense of avoiding communicative exchanges or associations between interlocutors (Mischke, 2005). However, we can note that, as in Example (12), their combination with personal pronouns can reinforce the meaning of these: inclusion is highlighted (Example (13)) and a discourse marker function is also attested.

(12) We have seen before that Cyrus had been connected with the affairs of Greece, especially with those of the Spartans, of whose bravery and talents he had in consequence formed a high opinion. When he began to plan his rebellion, his thoughts naturally turned to the Greeks; for he felt that with their help he might succeed in dethroning his brother, but without it he could scarcely hope to do so. (Sewell, 1857: 252 [153 (7691)])

(13) The story passes immediately from the recall of Robert to Molême to the election of Stephen. Migne, [ib]. [p]. 995, and Tissier, [Biblioth]. [Pat]. [Cist]. [i]. 13, follow Manriquez in thinking that a portion of the manuscript has been lost here. [endnote] but we are given to understand that the inevitable effects of the legislation which precedes were severely felt, and that the monks were sorrowful because their number did not increase. (Cooke, 1893: 630 [160 (2658)])

Following Hyland's (2002: 1092) proposal, my data here leads me to claim that academic writing is an act of social identity like any other form of communication, conveying a representation of the writer in command of her discipline-specific content (Ivaníč, 1998). Our discursive choices align us with certain values and beliefs that support particular identities. In the samples under analysis in the present study, scientific identity seems to have priority over female identity, in that the linguistic choices made by women seem to conform to the scientific canon of the times.

Private verbs and emphatics are used to seek attitudinal solidarity, and the frequency of use of private verbs is lower in CHET than in PPCMBE, as shown in Figure 16. The communicative formats Travelogue (109.94) and Textbook (106.39) show the highest frequencies here. This can be explained on the grounds of the evolution of history and historiography in the former case, and on the instructive goal and communicative purpose of the text in the latter.

There is a remarkable difference in frequency of use between the eighteenth (109.94) and the nineteenth century travelogues (43.83) which confirms the establishment of history as a scientific-academic discipline.

The pronoun *it* is another of the most frequent features denoting involvement in both history and narrative writing. According to Quirk *et al.* the pronoun *it* can have both a referring and non-referring function: "The neuter or nonpersonal pronoun it ('REFERRING it') is used to refer not only to inanimate objects […], but also to noncount substances […], to singular abstractions […] and even to singular collections of people […]:" (1985: 347).

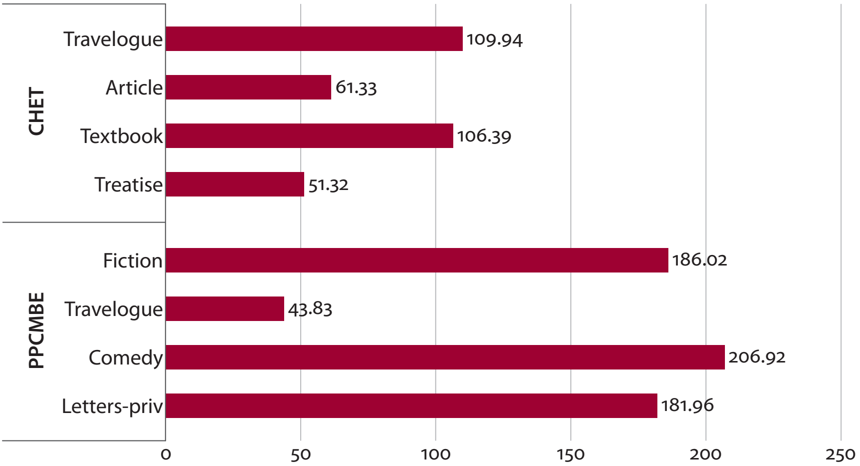


Figure 16. Private verbs per communicative format

Biber (1988: 226) assumes a more frequent use of the pronoun *it* in spoken language and a less frequent use in contexts that are more informational, due to the fact that *it* "can be substituted for nouns, phrases and whole clauses".

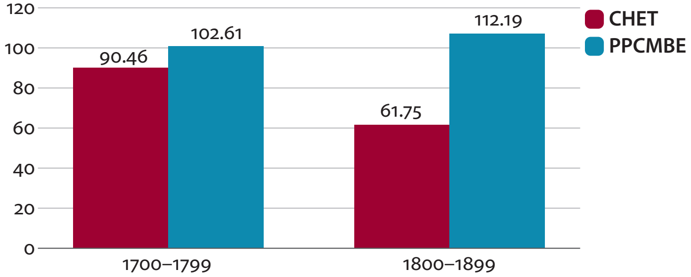


Figure 17. Pronoun *it* over time

It, as well as first and second person pronouns, are non-attitudinal involvement features but ones which nonetheless transmit closeness. The use of *it* (see Figure 17 above) has declined over time in the history corpus (90.46–61.75); in contrast, a small increase has been detected in the PPCMBE texts (102.61–112.19). The frequency of the pronoun *it*, as was the case with other involvement features, follows divergent paths in the two corpora. These results match with the analysis of pronominal forms and private verbs, which also show a decrease of involvement features in the nineteenth century as following the pattern of objective reports and limiting traces of authorial identity (see Example (14)).

(14) Theſe Princes ſettled the points in diſpute to the ſatisfaction of both parties; and the Margraves engaged, that if any other occaſions of quarrel aroſe, inſtead of endeavouring to decide it by arms, they would refer it to the ſame arbitrators. (Scott, 1762: 129 [134 (1074)])

As for communicative formats, Travelogue (117.94) stands out in scientific writing, and comedy (131.14) does the same in non-scientific prose, which is reminiscent of what was seen with private verbs in the same situation. The use of *it*, a non-attitudinal feature, requires that the participants in the communicative act know what is being referred to, that is, shared knowledge is essential for successful communication. Thus, the use of this deictic element underlines the reader and writer's reliance on mutual knowledge to code and decode utterances.

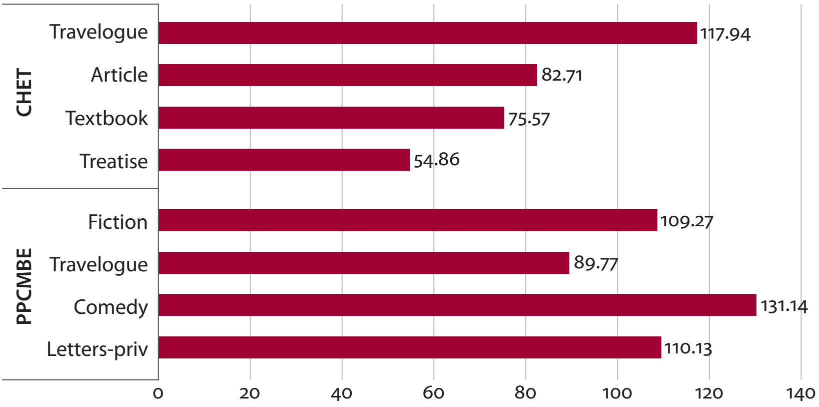


Figure 18. Pronoun *it* per communicative format

Formats with the highest counts for *it* are those demanding the greatest degree of sharing of information, hence the author's assumption of implicit knowledge on the reader's part.

Emphatic forms are the third most frequent group of items found in CHET (45.52) though this normalised figure is still lower than that for PPCMBE (74.89). Emphatics include expressions and lexical items that could be further classified as boosters, maximisers, emphasisers or intensifiers. Their role is defined by Quirk *et al.* (1985: 583) as having a reinforcing effect on the truth value of the clause or part of the clause to which they apply.

Holmes (2001: 287) identifies emphatics as "features which may boost or intensify a proposition's force" and which highlight the assertiveness of a statement and help the speaker to be accepted and to gain credibility, to be taken seriously, especially in the case of women's discourse. The fact that emphatics constitute a distinctive feature in female scientific production reveals the nature of female interaction patterns or purposes, in which they seek solidarity and connection with the target readership (see Examples (15) and (16) below).

(15) His plays were full of wit, and amused the people extremely, yet they had a really serious meaning, and were intended to show the folly and sinfulness of many of the Athenian customs.. (Sewell, 1857: 247 [153 (6565)])

(16) and they resolved to send out everything that might render Hispaniola a permanent establishment, and to furnish Columbus with such a fleet as might enable him to go in search of those new and rich countries. (Calcott, 1828: 252 [147 (9760)])

The descending line in the use of emphatics in history texts (Figure 19) confirms, once more, the objectivising trend of nineteenth-century science which discarded any signs of authorial presence.

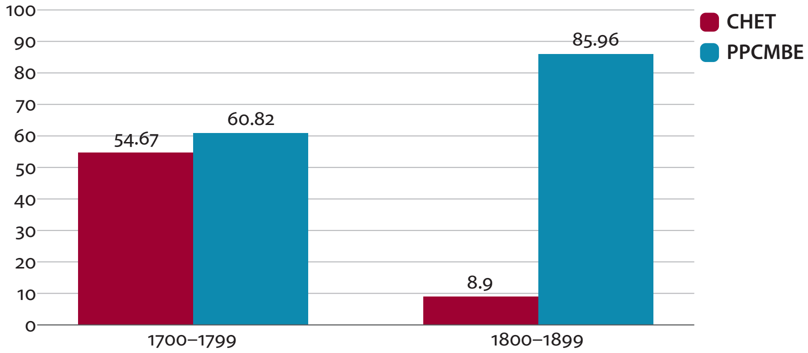


Figure 19. Emphatics over time

As for communicative format in scientific prose, Travelogue (49.97) and Textbook (42.59) are the ones containing most emphatic forms. Figure 20 shows that fiction (100.69) and private letters (9.32) use emphatics more frequently than comedy (65.09) and Travelogue (50.02). The occurrence of forms is nearly the same in the scientific and non-scientific samples of the Travelogue category (49.97 vs 50.02), which might point to a rather intimate tone of Travelogue in the history corpus.

On the one hand, we should recall that the Travelogue in CHET was written during the first half of the eighteenth-century and its tone approaches more closely that of a story than of history *per se*. On the other hand, the ultimate goal of the Textbook was instructive, thus manifesting the intentional attitude of persuading the community of practice.

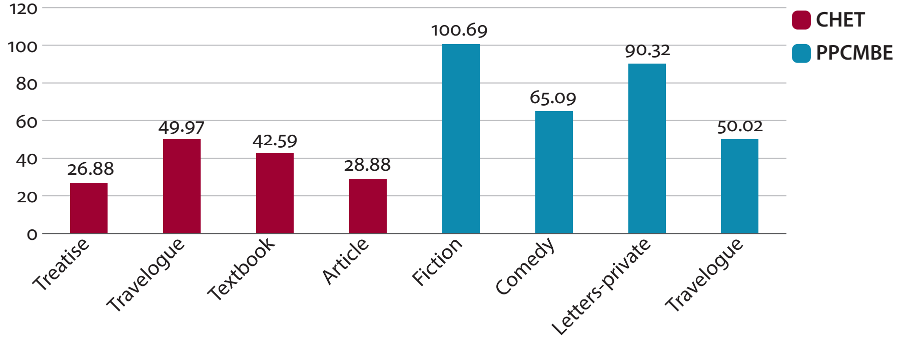


Figure 20. Emphatics per communicative format

5. Concluding remarks

The analysis in the previous pages indicates that samples of scientific and non-scientific writing follow diverging tendencies, both in the frequency of occurrence of these features in general and over the two centuries analysed, as well as in each of the formats considered. The results, then, contradict the initial hypothesis, which was that since history is one of the soft sciences, and since the samples examined were written by women, the features analysed should in principle follow similar trends. What my findings in fact seem to show is that female history writing follows the same writing patterns as other scientific disciplines, such as those of the hard sciences, which in principle are not suspicious of using the characteristics of common language. And, in this sense, female authors follow the writing trends imposed in the community of practice. This is confirmed by the individual analysis of each feature.

Personal pronouns, private verbs and the pronoun *it* exhibit the highest frequencies of use in non-scientific texts, whereas private verbs, the pronoun *it* and emphatics are the three first most frequently used features in scientific samples. In the history corpus, first and second person pronouns occupy the fourth position.

Pronouns are the most obvious forms of conveying authorial identity, values and engagement. They are also a direct way of addressing the readership in search of acceptance and in showing commitment to the message. If we recall Atkinson's (1999) suggestion that there is a movement in scientific discourse from subject-centred to object-centred writing, this is precisely what the low use of first and second person pronouns in the history corpus demonstrates. The eighteenth-century travelogue written by Justice (1739), although only one single text, is indeed the one in which most of these pronouns were found. The inclusion of another nineteenth-century travelogue in the PPCMBE corpus, and the scant presence of pronouns and other features denoting involvement in it, also supports the evolution of history as a discipline that fits into the patterns of scientific writing in this period. From a diachronic point of view, the decreases in the use of first and second person pronouns (134.7 to 7.33 for first person, 7.95 to 0 for second) are certainly very remarkable. A similar behaviour has been attested in the case of private verbs, the pronoun *it* and emphatics: a decreasing frequency of these features in history texts over the two centuries under survey, this being the opposite of the general trend in the language, and with travelogue and textbook (except in the case of the pronoun *it*) being the communicative formats containing the highest number of involvement forms in normalised figures.

This movement towards objectivity in the subcorpus of history texts (CHET) written by women moves in the same direction as scientific writing, hence the same direction as male-authored science writing. Female texts seem to fit into the general patterns of scientific writing, shifting towards the depersonalised discourse of modern science in line with what has been claimed by authors such as Halliday (1988: 166) and Taavitsainen (1994: 339), who date this form of discourse to the nineteenth century. The influence of Enlightenment thought, the manifestation of philosophical currents such as Positivism rooted in seventeenth-century Empiricist movement, underlie these discursive practices. Yet despite the popularisation of science at the end of the nineteenth century, women could not yet play an active role in scientific life (Knight, 1986). What is also interesting here, then, is that it seems possible to confirm that, as Prelli (1989) and Besnier (1994) have argued, such objectivity may reflect ideological constructs, that an increasing detached form of expression reflects a particular cultural ideology (Besnier, 1994: 284), and that female authors probably adhered to these ideological constructs in search of affiliation and recognition.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Argamon, Shlomo, Moshe Koppel, Jonathan Fine & Anat Rachel Shimoni. (2003). Gender, Genre, and Writing Style in Formal Written Texts. *Text*, 23 (3), 321–346. https://doi.org/10.1515/text.2003.014

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

Bennett, Karen. (2009). English academic style manuals: A survey. *Journal of English for Academic Purposes*, 8(1), 43–54. https://doi.org/10.1016/j.jeap.2008.12.003

Besnier, Niko. (1994). Involvement in linguistic practice: An Ethnographic Appraisal. *Journal of Pragmatics*, 22, 279–299. https://doi.org/10.1016/0378-2166(94)90113-9

Biber, Douglas. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511621024

Biber, Douglas & Jená Burges. (2000). Historical change in the language use of women and men: Gender differences in dramatic dialogue. *Journal of English Linguistics*, 28, 21–37. https://doi.org/10.1177/00754240022004857

Biber, Douglas & Edward Finegan. (1997). Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen & Leena Kahlas-Tarkka (Eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen* (253–275). Helsinki: Societe Neophilologique. (Reprinted in Conrad & Biber (Eds.) (2001), 66–83.)

Cook, Albert. (1988). *History/Writing*. Cambridge: Cambridge University Press.

Dontcheva-Navrátilová, Olga. (2013). Authorial presence in academic discourse: functions of author-reference pronouns. *Linguistica Pragensia*, 23(1), 9–30.

Halliday, Michael Alexander Kirkwood. (1988). On the Language of Physical Science. In Mohsen Ghadessy (Ed.), *Registers of Written English: Situational Factors and Linguistic Features* (162–178). London: Pinter.

Halliday, Michael Alexander Kirkwood & Christian Matthiessen. (2004). *An Introduction to Functional Grammar*. London: Arnold.

Harwood, Nigel. (2005). We Do Not Seem to Have a Theory … The Theory I Present Here Attempts to Fill This Gap: Inclusive and Exclusive Pronouns in Academic Writing. *Applied Linguistics*, 26(3), 343–3. https://doi.org/10.1093/applin/ami012

Holmes, Janet. (2001). *An introduction to Sociolinguistics*. Longman.

Hyland, Ken. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091–1112. https://doi.org/10.1016/S0378-2166(02)00035-8

Hyland, Ken. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Ivaníč, Roz. (1998). *Writing and Identity. The Discoursal Construction of Identity in Academic Writing*. Amsterdam and Philadelphia: John Benjamins. https://doi.org/10.1075/swll.5

Knight, David (Ed.) (1986). *The Age of Science. The Scientific World-View in the Nineteenth Century*. Oxford: Basil Blackwell.

Kroch, Anthony, Beatrice Santorini & Ariel Diertani. (2016). *The Penn Parsed Corpus of Modern British English (PPCMBE2)*. Department of Linguistics, University of Pennsylvania. CD-ROM, second edition, release 1. Accessed [21/03/2018] from http://www.ling.upenn.edu/ppche/ppche-release-2016/PPCMBE2-RELEASE-1

Kuo, Chih-Hua. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English for Specific Purposes*, 18(2), 121–138. https://doi.org/10.1016/S0889-4906(97)00058-6

Lakoff, Robin T. (1990). *Talking power: The politics of language in our lives*. New York: Basic Books.

Mischke, Gertruda E. (2005). *Analysing 'involvement' in Distance Education Study Guides: An Appraisal-based Approach*. The University of South Africa.

Monaco, Leida Maria. (2017). A Multidimensional analysis of Late Modern English Scientific Texts from the Coruña Corpus. Unpublished PhD Diss. Universidade da Coruña.

Moskowich, Isabel. (2017). "Pronouns as stance markers in the Coruña Corpus: An analysis of the CETA, CEPhiT and CHET". In Francisco Alonso-Almeida (Ed.), *Stancetaking in Late Modern English Scientific Writing. Evidence from the Coruña Corpus* (73–91). Colección Scientia [Applied Linguistics]. Valencia: Servicio de Publicaciones de la Universidad Politécnica de Valencia.

Moskowich, Isabel & Begoña Crespo. (2016). "Classifying communicative formats in CHET, CECHeT and others". *EPiC Series in Language and Linguistics*, 1, 308–320. https://doi.org/10.29007/kvjx

Pagliawan, Dominador L. (2017). Feature Style for Academic and Scholarly Writing. *Academic Journal of Interdisciplinary Studies*, 6(2), 35–41. https://doi.org/10.1515/ajis-2017-0004

Porter, Theodore & Dorothy Ross (Eds.) (2003). *The Cambridge History of Science. Vol. 7. The Modern Social Sciences*. Cambridge: Cambridge University Press.

Prelli, Lawrence J. (1989). *A Rhetoric of Science: Inventing Scientific Discourse*. Columbia, SC: University of South Carolina Press.

Puente-Castelo, Luis. (2017). On Conditionality: A Corpus-based study of conditional structures in Late Modern English Scientific Texts. Unpublished PhD Diss. Universidade da Coruña.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Simó, Judit. (2012). Involvement and Subjectivity in the Linguistic Description of Counterfactual Scenarios in Chess Commentaries, 2012. *Selected Papers from UK-CLA Meetings*, Vol 1, 78–93. Accessed [21/03/2018] from http://uk-cla.org.uk/proceedings

Stromberg, Roland N. (1951). History in the Eighteenth Century. *Journal of the History of Ideas*, 12(2), 295–304. https://doi.org/10.2307/2707520

Taavitsainen, Irma. (1994). On the Evolution of Scientific Writings between 1375 and 1675: Repertoire of Emo-tive Features. In Francisco Fernández et al. (Eds.), *Papers from the 7th International Conference on English Historical Linguistics, Valencia, Sept. 1992* (329–342). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.113.32taa

Tang, Ramona & Suganthi John. (1999). The 'I' identity: Exploring Writer Identity in Student Academic Writing Through the First-Person Pronoun. *English for Specific Purposes*, 18, 23–39. https://doi.org/10.1016/S0889-4906(99)00009-5

Thompson, Geoff. (2001). Interaction in Academic Writing: Learning to Argue with the Reader. *Applied Linguistics*, 22(1), 58–78. https://doi.org/10.1093/applin/22.1.58

Yang, Xinzhang. (2017). Linguistics Features of Knowledge Construction in Chemistry Textbooks. In Jonathan J. Webster & Xuanwei Peng (Eds.), *Applying Systemic Functional Linguistics: The State of the Art in China Today* (337–350). London: Bloomsbury Publishing.

CHAPTER 11

Neither *I* nor *we*

Inexplicit authorial voice in eighteenth century academic texts

Margarita Mele-Marrero
Universidad de La Laguna

1. Authorial voice

The importance of the self is nowadays undisputable in scientific writing: the visual or the intellectual approach, self-observation, have been vital for the historical evolution of science, a science that cannot be separated from its scientists. It has not always been like that, though. The idea of the prevalence of reason in the eighteenth century created expectations for more object-centred writings that would elude the presence of the author in the text, however, as stated by Elliot (2010: 2): "English scientific culture developed as part of the creation of new public spheres of rational debate founded upon autonomous individuals". It is precisely the manifestation of these individuals in their works our main interest here. Specifically, we will be concerned with the use and, mainly, the absence of the first-person pronoun as forms of stance.

The theoretical foundations followed in this chapter are those present in Hyland (2001, 2002, 2004, 2005; Sancho-Guinda, 2012) concerning *stance* and *engagement*. These are understood as, on the one hand, an authorial voice evident through features which allow writers to "convey their judgements, opinions, and commitments […] the ways that writers intrude to stamp their personal authority onto their arguments or step back and disguise their involvement". Engagement, in turn, refers to how writers establish a more direct relationship with their readers, "pulling them along with their argument, focusing their attention, acknowledging their uncertainties, including them as discourse participants, and guiding them to interpretations" (Hyland, 2005: 176). These *attitudinal* and *alignment* interactional macro-functions, again in Hyland's terms, have in common the use of first person pronouns as textual indicators (see Figure 1, 2005: 177). The inclusion of first-person pronouns is a clear form of manifesting authority and conviction on what is being said, of course,

https://doi.org/10.1075/z.225.11mel

© 2019 John Benjamins Publishing Company

certain modifiers might also mitigate the strength of this, but initially the use of *I* introduces directly the writer's presence, as well as the use of *we,* in its inclusive form, involves the reader in the written discourse.

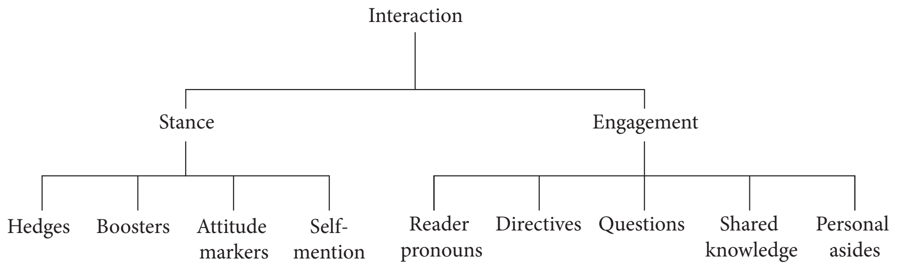


Figure 1. Key resources of academic interaction (adapted from Hyland, 2005: 177)

The author's position within his work can be modulated by means of *hedges* that appear to lessen the degree of commitment (*suggest, may, might, perhaps*…) as opposed to *boosters* (*clearly, obviously, certainly*…) which contribute to underline the author's assertions. In the same line, attitudinal markers (*appropriate, logical, unfortunately*…) establish their personal "affective" relation with the content of their discourse. The way authors engage with their readers is also revealing about their own voice; not only the inclusive first person *we*, but *you*, can inform about the extent to which the author feels involved in his social community. *Personal asides* and *appeals to shared knowledge* (*as you know, familiar*…) also inform about the vision the writer has of his readers and how he/she wants to be seen. *Directives* eliciting action or guiding the reader through different imperative forms (*you must, let's*…), as well as the *rhetorical questions* can complete the picture when we analyse an author, and make the reader feel more or less comfortable and confident about the text he/she is reading.

Although Hyland considers first person pronouns and possessive adjectives as stance and engagement markers, in this paper we will concentrate especially on subject forms since, from the evidence in our texts, they represent a primary force for self-reference. Other elements of stance and engagement will be observed whenever necessary to shed light on our main purpose.

Hyland's prolific work (2001, 2002, 2005, & Bondy 2006, & Sancho-Guinda 2012) on authorial voice has attempted to demonstrate the importance of the type of discipline in the choice and use of self-mention. In 2001, he considered discipline was a determining factor, indicating that first person occurred more often in humanities and social sciences than in hard sciences. In turn, Sword (2012) found that the opposite was more common. This current paper derives from a previous research (Mele-Marrero, 2017a) in which I tried to establish if the differences concerning the use of first person pronouns observed by Hyland versus the opinion of Sword, had a replica or precedent in some of the eighteenth century academic subcorpora that belong to the *Coruña Corpus of English Scientific Writing* (CC, henceforth): CETA and the beta version of CHET. That research revealed that, in absolute terms, first person singular pronouns were more frequent in History works (CHET) than in Astronomy works (CETA), but when stance functions were clearly delimited, authorial voice was stronger in CETA than in CHET. That work also evidenced a considerable variation among authors in their use of first person and gaps by the early second half of the century for both CETA and CHET (1761; 1760, 1775 respectively) where three authors showed no use of these pronouns, that is neither *I* nor *we*. One corresponded to CETA and two to CHET. The purpose of this paper is to explore the reasons for this exclusion, which was not generalised in the rest of the authors analysed (see Figures 2 and 3), and locate, if possible, alternative forms of stance and engagement that may have been used by these authors to make themselves apparent in their works.

In the following sections, we will find general information about the authors and works which concern this study, as obtained from CC metadata or other indicated sources. Section 3 will present quantitative data, stating number of personal pronouns found, modals and other stance and engagement markers that can be of relevance in the representation of the author's voice, following basically those already mentioned in Hyland's reproduced Figure 1 above. These data will be commented individually first and then interrelated, trying to find any connections, if extant, basically among the three authors but also including other texts and writers that might be of help in this final comparison. To conclude, I will try to expose any reasons that connect these authors in stance terms or if theirs is a particular choice within their social community.

2. Authors in CETA and CHET

The CC covers various scientific disciplines that conform its different subcorpora with works which had been directly written in English by English-speaking authors, and published between 1700 and 1900. This structure allows us to make a first distinction between soft and hard sciences, but we may also monitor singular cases by means of the metadata offered, which facilitate knowledge about each author, especially about their education and other relevant aspects.

In the corpus information, the genre of the texts is also identified providing the researcher with an anticipation of the parameters implied. Nevertheless, the compilers state (Moskowich & Crespo, 2012: 15):

> Delimiting genres is complicated and it represents a problem that, according to Monzo (2002: 141) can be solved if we assume that genres may be sub-classified in variants attending to cultural criteria. She proposes the following:
>
> – Paragenre: genre belonging to one professional community
> – Diagenre: analogous genre in a culture that can be identified and recognised by the territory it occupies
> – Idiogenre: genre reflecting a particular author's idiosyncrasy in his/her texts in a constant way.

Thus, following this classification, the CC main aim is to procure texts within a paragenre.

My own selection (Mele-Marrero, 2017a) included CETA's 18 male English authors of the eighteenth century (177,865 tokens) and CHET's 15 male English authors (Moskowich et al. 2019) within the same period of time (151,566 tokens). As can be appreciated in Figures 2 and 3 below, nearly all the authors use subject forms for the first person pronouns or at least one of them (*I* or *we*), except Stewart in CETA, and Birch and Walter Anderson in CHET for whom the lines drop completely.

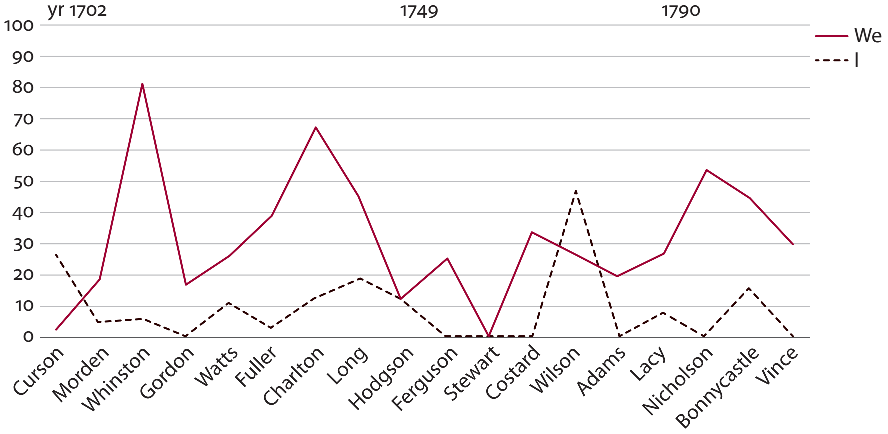


Figure 2. Authorial variation in CETA

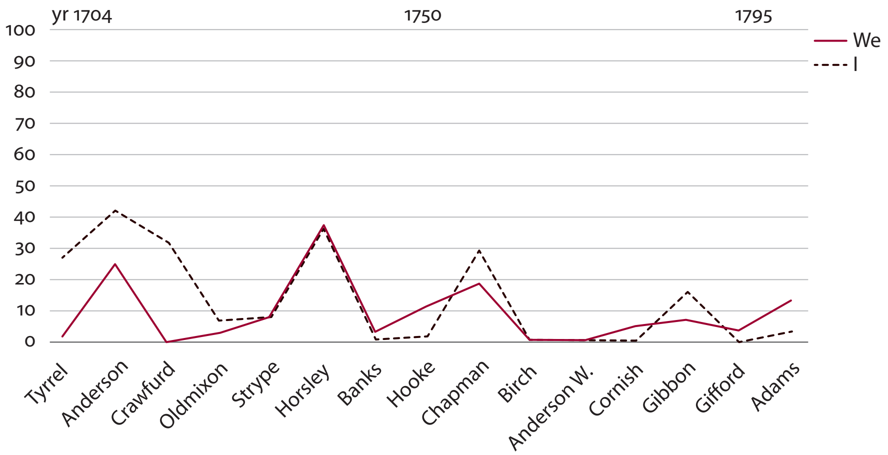


Figure 3. Authorial variation in CHET

Peaks previous and following the work of our “neither nor” authors, do not seem to point to a generalized tendency. It is true that numbers appear to drop in CHET but, Adams’s (1795) use of *we* is higher than his predecessor’s, Oldmixon (1716). This makes the absence of first person pronouns more exceptional and that is why we have isolated these authors for a particular analysis.

2.1 Thomas Birch (1705–1766)

Birch’s career was ecclesiastical, he had no higher education at university even if at 48 he was created D.D. of the Marischal College, Aberdeen, and of Lambeth (1753). His connections with the Church of England and patronage of Sir Philip York, helped him to obtain promotion within the religious institution. His obligations did not prevent him to be active in other cultural areas, becoming a member of the Antiquarian Society or Secretary of the Royal Society and devoting good part of his life to historical writings and biographies. Among these, he participated in the compilation of the English lives in the *General Dictionary, Historical and Critical* (1734–41) and wrote *The Life of Henry Prince of Wales* (1760) (CETA metadata). His works had detractors, though his person had supporters, the first complained about the “wearisome minuteness of detail and their dullness of style” of his writings whereas those who appreciated him, especially on the tory wing, seem to have valued Birch’s religious and political views (*DNB*).

2.2 Walter Anderson (?-1800)

Anderson studied at the University of Edinburgh where he obtained a MA in 1742. He became a minister of the Church of Scotland at the village of Chirnside in the Scottish borders. There he produced his historical works whose publication was probably afforded by the sale of one of his houses (CHET metadata). One of them is, *The history of France*, in three volumes, partly included in CHET. His style and sources were censured by Smollett in *The Critical Review* (1769: 170) observing he insisted excessively in the bloody and cruel parts "often told in a descriptive manner" and not having consulted relevant state-papers.

2.3 Matthew Stewart (1717–1785)

Included in CETA, Stewart was born and educated in Scotland. In 1745 "he was ordained minister of Roseneath, Dumbartonshire, which charge, however, he resigned on being elected professor of mathematics at Edinburgh University" (*DNB*). He wrote mainly on Geometry and planetary motion, the work selected is *Tracts, Physical and Mathematical. Containing an Explanation of several important points in Physical Astronomy; and a new Method for ascertaining the Sun's distance from the Earth, by the Theory of Gravity* (1761) (CETA, metadata). He had also written in Latin *Pappi Alexandrini collectionum mathematicarum libri quarti propositio quarta generalior facta, cui propositiones aliquot eodem spectantes adjiciuntur* (1754) and later *A solution of the problem involved in Kepler's second law of planetary motion* (1756). In *The Monthly Review* edited by Griffiths, (1762: 458), he is considered an "ingenious author having explained the most difficult things in the doctrine of the Centripetal Forces", referring precisely to the work included in CETA.

3. First person pronouns and other stance/engagement markers

Because in the previous work only subject forms had been searched, here the scope was extended to object and possessive forms of the first person when these could make some difference, such as *us* in *let us* or *let's* expressions. Given this paper's interest in the author's voice, examples of *one* referring to the writer were also considered, as well as second and third person subject forms if they were indicative of some sort of engagement between the author and the reader (see Mele-Marrero & Alonso Almeida, 2011). I also searched for modals that could express some personal involvement – *may, might, should, must, can, could* – and verbs indicating some commitment like *think, believe*, as well as other attitude markers like *unfortunately, appropriate, logical, remarkable, important…*

Tables 1 to 3 below display the results obtained per author and work with further commentaries and examples in each case to describe the data before comparison.

3.1 Birch's authorial voice

Table 1. Results in Thomas Birch (1705–1766) "The Life of Henry Prince of Wales" (1760)

BIRCH (10,059 tokens)

1st ps.	
P. pr. subject forms sg.	_
P. pr. subject forms pl.	_
P. pr. object forms	1
P. pr. poss. forms	2

	2nd ps.	3rd ps.
P. pr. subject forms sg.	_	151
P. pr. subject forms p.	_	10
One	23*	

Modals	31
Other markers	12

Birch's text does not present any first-person singular except when quoting, but we find two cases of the possessive *our* (Examples (1) and (2)) and a single object form *us* engaging the reader in the historical facts narrated. The rest are third person forms referring to the historical characters in the text. The use of *one*, appears as numeral or indefinite but it does not represent the author. Concerning modals, the most frequent is *should*, 19 items, but these correspond to paraphrases of other's utterances (3), the same is evidenced in the 9 examples of *might* found, except for (4), which shows the author's opinion, and another three cases of *could*. The remaining markers are basically some hedges like *perhaps* or *at least*, and other attitude markers such as *remarkable* or *important* (5). The preface and dedication (not part of CHET's selection) include the form *I* nine times (6).

(1) Edmund Howes, *our* English chronicler of that age.
(CHET, Birch, 1760: 32 [133 (5746)])

(2) natural knowledge, very little cultivated in *our* country at that time.
(CHET, Birch, 1760: 34 [133 (6147)])

(3) He adds, that if he *should* find by any token, that his performance was acceptable to the Prince, he *should* be excited to publish soon some greater works under the auspices of his Highness's patronage.
(CHET, Birch, 1760: 47 [133 (8599)])

(4) It is not improbable, that his tutor, [Mr] Newton, *might* have a considerable hand in this and other letters of the Prince.
(CHET, Birch, 1760: 36 [133 (6611)])

(5) on the death of Matthew Earl of Lennox, and filled that *important* office but thirteen months. (CHET, Birch, 1760: 10 [133 (2080)])

(6) In these respects *I* flatter myself, that my researches have enabled me to do justice in some degree to the memory of a Prince, who has long been the object of a general admiration, by a detail of particulars not known to, or defectively related by, former historians. (Birch, 1760: 2–3 [133 (PREF)])

3.2 Anderson's authorial voice

Table 2. Results in Walter Anderson (?–1800) "The History of France" (1775)
ANDERSON (10,020 tokens)

1st ps.	
P. pr. subject forms sg.	_
P. pr. subject forms pl.	_
P. pr. object forms	_
P. pr. poss. forms	_

	2nd ps.	3rd ps.
P. pr. subject forms sg.	_	144
P. pr. subject forms p.	_	16
One	15*	

Modals	37
Other markers	25

The search for first person pronouns in CHET rendered no results in Anderson, we do find third person subject forms, when alluding the historical characters, and including a considerable amount of *it* forms, 30, that basically introduce passive constructions, as in Example (7). Anderson presents more modal and other markers variety as compared to the previous author, thus: *might, may, should, must, could* or *seem* (Examples (8) to (10)) Nevertheless, not all the examples refer directly to the author's opinion but to the characters described:

(7) *It* was agreed that Matthias should be chosen governor.
(CHET, Anderson, 1775: 109 [136 (528)])

(8) Even by the lowest of them the political use of pageantry of his name *might* well be comprehended. (CHET, Anderson, 1775: 129 [136 (6383)])

(9) Gregory's name was justly affixed to it, and *must* remain in the annals of ages.
(CHET, Anderson, 1775: 124 [136 (5040)])

(10) while that of the other memoirs of the times *seems* both obsolete and tedious.
(CHET, Anderson, 1775: 140 [136 (9731)])

3.3 Stewart's authorial voice

Table 3. Results in Matthew Stewart (1717–1785) "Tracts, Physical and Mathematical. Containing an Explanation of several important points in Physical Astronomy; and a new Method for ascertaining the Sun's distance from the Earth, by the Theory of Gravity" (1761)

STEWART (12,180 tokens)

1st ps.	
P. pr. subject forms sg.	_
P. pr. subject forms pl.	_
P. pr. object forms	_
P. pr. poss. forms	_

	2nd ps.	3rd ps.
P. pr. subject forms sg.	_	18
P. pr. subject forms p.	_	_
One	25*	

Modals	28
Other markers	89

Stewart appears as a very particular instance, he does not present any type of first person pronoun, nor any other pronoun than *dummy it* forms that introduce passive structures (11); the use of *one* appears as a numeral or a reciprocal pronoun (12) but not as an indefinite pronoun that could make a reference to himself, as an author, or to the reader. Furthermore, in the preface to this work (not included in CETA) the personal references are reduced to *the author* and *the reader* establishing certain distance by using third person sg. pronouns (see Example (13)). When talking about colleagues, he does not include himself explicitly in the group (14).

(11) the same way *it is shown*, that if the moon had described the circle.
(CETA, Stewart, 1761: 341 [12 (251)])

(12) are to *one* another nearly as the distances of the moon from the earth.
(CETA, Stewart, 1761: 340 [12 (5771)])

(13) There are several propositions particularly in the second and third tracts, the use of which will not easily appear to *the reader*. They were intended to explain several things in the lunar theory which *the author* had in view, but which, as the book has swelled to a greater size than was at first expected, he was obliged to defer at present. (Stewart, 1761: vi [12 (PREF)])

(14) astronomy, and several solutions of it have been given by *mathematicians* of the first rank. *Their* methods of solution require a considerable knowledge in the more difficult parts of mathematics 3 but the method here given, requires only the knowledge of the elementary parts, and of plain trigonometry. As *the reader* is supposed to be well acquainted with Euclid's Elements, there are no references made to them. (Stewart, 1761: vii [12 (PREF)])

When considering other forms of stance, in spite of raw numbers, Stewart does not seem very prolific either. Of the modal verbs searched, he uses *may* 27 times and just once *can*; *may* is used in formulas, not really expressing personal doubt or possibility (15), the same occurs with *can* (16). He avoids the use of hedges like: *seem, suggest*, *appear*, *maybe*; but also, boosters like: *clearly*, *obviously*, *truly*, or verbs like *think*, *believe*, *demonstrate* or *reveal*. He does use *will* (203 tokens, of these 200 with the verb *to be*) but for mathematical derivations and results (17).

(15) that the rectangle FHA *may* be triple the rectangle FCA. (CETA, Stewart, 1761: 351 [12 (2552)])

(16) GH: of all the lines that *can* be drawn from the point G to any point in the semiellipse. (CETA, Stewart, 1761: 382 [12 (8861)])

(17) Because KT is great in comparison| of TA, fix times the rectangle KTD *will* be nearly equal to the square of TA; therefore TA *will* be to TD nearly as fix times KT to TA. (CETA, Stewart, 1761: 341 [12 (152)])

Regarding engagement forms, we find directives through the imperatives *let* (84 tokens) and *suppose* (5 tokens) ((18) and (19), respectively), but he never uses them as inclusive forms, that is: *Let's* or *let us suppose*. Once more these forms obey more to the mathematic propositions than to intended interactions with the reader.

(18) *Let* TK represent one third part of the distance of the sun from the earth at T, and *let* TL, represent two different distances (CETA, Stewart, 1761: 340 [12 (37)])

(19) *suppose* a body acted upon by a centripetal force. (CETA, Stewart, 1761: 357 [12 (3522)])

In the same line, personal asides and expressions indicating explicit shared knowledge are absent. Nevertheless, it must be said that in the preface (not included in CETA) he relies on the reader's knowledge, as shown in (20):

(20) As *the reader is supposed to be well acquainted with Euclid's Elements*, there are no references made to them; but as his knowledge of conic sections may not be so extensive, there are *references made to the excellent treatise* on this subject, by that great geometrician Dr Robert Simon Professor of Mathematics in the university of Glasgow. (Stewart, 1761: vii [12 (PREF)])

4. Cross-case analysis

It is not easy to find common factors among the three authors that could justify their stance choices. Those that would appear obvious like origin and education, period, discipline or genre do not seem significant motives once evaluated.

Even if our three scholars happen to have been educated in Scotland, so were other contemporaries who made a considerable use of the first person in their works, thus, James Ferguson (1756) and Alexander Wilson (1773) in CETA and George Crawfurd (1710) in CHET (see Mele-Marrero, 2017a). The connections with the Church of Scotland and a more effaced style could be plausible, but I cannot find clear links among the authors' education and/or their religious options, especially in non-religious texts. On the other hand, we must be aware that although the church was losing full control of education at the time, its influence was still important. Therefore, even if other authors did not conclude their studies becoming vicars, they would still be under that same influence of their classmates.

> After 1707 the Scottish universities developed along their own lines without intervention from the state. […] The universities continued to train the Presbyterian clergy, and for students aiming at a clerical career there was a presented four-year curriculum based on six subjects: Latin, Greek, mathematics, logic, moral philosophy and natural philosophy (roughly equivalent to physio, but with a mathematical rather than experimental emphasis). Completing the course was enough, and formal graduation in arts became unusual […] Sixteen or seventeen was the typical age of entry by the 1860s, though younger boys might come direct from parish schools. (Anderson & Wallace, 2015: p. 268)

In addition, and although the Kirk might have its idiosyncrasies, we have other authors out of Scotland who were also ordained and have again an ample use of first person in their works: Isaac Watts (1674–1748), George Costard (1710–1782) (CETA) and Joseph Cornish (1750–1823) (CHET).

Expanding the period of study in each field, if we consider previous and later writers, for Stewart, a mathematician, we find predecessors, like Hodgson (1709–1772, CETA), who use first person singular and plural, or Newton himself (1643–1772), so admired and followed by his contemporaries, who as seen in the examples below (Examples (21) to (23)) found space for self-mentioning in his work (Newton, c. 1665-c.1672: 81^{r}):

(21) The Radius DC of which circle *I* may therefore call the radius of Circular motion or velocity about that axis EF. (l.24)

(22) *I* find such a point P from which the perpendic{la}rs. (PK, PH) let fall to those axes bee reciprocally proportional to the angular velocitys about those axes (l.31)

(23) And these four rules *I* gather thus: The whole velocity. (l.51)

Authors in the nineteenth century sampled in CETA all use *I* or *we* pronouns. Even a Scottish minister of religion, mathematician and astronomer (CETA metadata) like Small (1732–1808) uses the first-person singular in the introduction to his work (1804) and plural forms in the sampled parts in CETA.

Regarding our male authors in CHET, previous and later historians, make use of first person in the eighteenth century. The next century, though, presents two authors with no *I* or *we* in CHET's selection, the Irishmen:[1] Hardiman (1782–1855) and D'Alton (1789–1867). Hardiman makes use of *our* (24) as the CCT indicates; searching further in the full text, examples of *we* can be found (25), whereas *I* appears only in the dedication and as part of quotations which are very frequent. For D'Alton, the CCT does not render any results for first person pronouns, in the original text we find *I* forms in the preface and inside quotations, but in the conclusion, in the manner of Stewart, D'Alton refers to himself as *the author* (see Example (26)):

(24) Considering, therefore, all the foregoing testimonies and reasonings, and the conclusions drawn from them, *our* entire concurrence is given to the opinions of the learned and respectable writers above quoted, that the ancient town of Galway. (CHET, Hardiman, 1820: 38 [146 (3628)])

(25) This circumstance, though unobserved by Vallancey, very forcibly corroborates this opinion, both names evidently agreeing in meaning and derivation, and each serving to illustrate, and very satisfactorily to explain, the origin and signification of the other: when, therefore, *we* consider the weak foundation of traditional report, and the fabulous complexion of the story, attributing the name to the woman. (Hardiman, 1820: 3)

(26) In closing this history […] To that nobleman, and to those who have confidingly honoured this work with their names as individual subscribers, *the Author* has but to express a hope, that if they shall be satisfied with its execution, they will cheer *him* with an early assurance of their continuing co-operation, in such similar productions as *he* may be induced to undertake. (D'Alton, 1844: 466)

Just one more author in the mid nineteenth-century section of CHET does not seem sufficient to allege a decreasing tendency in the use of the first person, although first person plural might be more frequent (Mele-Marrero, 2017b).

---

[1] My initial selection only included English male authors, but I decided to include here these Irish authors from the nineteenth century because their results were comparable and D'Alton appeared as the only similar case.

A claimed, and proved, important factor in different stance positioning is discipline, but for these three authors that difference does not seem valid. Divergent fields of study do not interfere with their 'neither *I* nor *we*' option. We might then move to the particular genre of their texts.[2] Stewart develops propositions and formulas which are addressed to his peers, in his text he exposes his knowledge rather than teaching. In the case of Birch, he presents a biography of Henry, Prince of Wales with no textbook structure; the same could be said of Anderson when writing about the history of France. The three texts have in common a descriptive function but this should not imply a necessary lack of personal involvement. We find first person forms in other biographies and treatises. Stewart and Birch would coincide with the "Essay" classification of CC (see Moskowich & Crespo, 2012: 19) but so does James Anderson (not to be confused with Walter Anderson) who in turn would be writing a treatise.

From a superior level, considering the epistemological distinctions indicated as initial aim in the CC (Moskowich & Crespo, 2012: 15), our authors would be positioned in a highest or high epistemological level and this would apparently confront them with other texts of medium epistemic levels such as textbooks. The lack of explicit personal involvement in these texts would contrast with that genre connection that I have found in other authors for obstetric manuals in the seventeenth century (see Mele-Marrero, 2011 and Alonso-Almeida & Mele-Marrero, 2014) where the use of first person was more prominent. Most of CHET works even though informative cannot be considered and therefore are not classified as textbooks; for CETA, there are textbooks where the use of first person is certainly high, (see Whinston or Charlton in Figure 2 above) but this also happens in treatises and, particularly, in the sample by Wilson (see also Figure 2) whose article, of highest epistemological level, presents a major use of first person singular, *I* forms. Hyland has found in modern texts the prominent use of inclusive *we* both in academic writing and in more popularizing type of texts like science journalism. In the first this inclusive *we* tends to relate to a scientific community whereas in the second it is more an appeal to align reader and writer "perhaps against an institutional 'they'" (2010: 225). Most of CETA and CHET scholars accommodate to Hyland's findings, but not our "particular" authors.

My initial thesis facing these authors' results was that they had their own *idiogenres,* that is, their stance was part of a "more" particular choice within the century and a text-type within a conforming discipline. In order to assert this, I had to previously analyse the possibilities referred to in the former sections and then examine the author's full texts and others they might have written. This was not an easy task since it implied working with "raw" material outside the corpus, but on the other hand, it would also help endorsing the *Coruña Corpus* results.

---

[2] A discussion on *genre* is not within the limits of this article. My main objective, subject personal pronouns and stance, makes me consider differences between disciplines and text-types/genres, already pre-established by the corpus, from which I will distinguish, mainly and for my purpose, between manuals with a clear teaching purpose and other texts of scientific dissemination to be shared with either specialised scholars or a lay public.

Below are summarised the findings obtained after considering our three writers' works.[3]

Anderson's text read in full includes first person plural *we* out of quotations, at least 31 tokens, but no examples of *I* were found out of quotations. He was also the author with a higher variety of modals. It appears that his choice of no direct self-mentioning appeared in the chapters selected in CHET (Book I viii and Book II ii). It must also be said that the use of *we* made by Anderson is nearly always a form of engagement to include the reader in his narration of historical facts, quite similar to that indicated by Hyland for scientific journalism mentioned before:

(27) In the conference at St Brix, *we* are not distinctly informed, upon what special footing the Queen-mother treated with Henry. (Anderson, 1775: 170)

(28) it is with regret that *we* find an odious and despicable view of it subjoined, by the historians. (Anderson, 1775: 445)

Given the above examples, we cannot consider Anderson a real instance of absence of direct self-mentioning. His text, a very lengthy piece of about 536 pages, is not very consistent in this sense. This could justify the criticism his style received, involving the reader in his opinions. Nevertheless, I think this higher presence of authorial voice in his case was made apparent in the variety of modals and other markers already traceable in CHET's extract.

The results for Birch are different, in the 554 pages of his *Life of Henry Prince of Wales*, including appendices and index, we only find the use of first person in the Dedication and Preface, 11 pages in total, where he uses *I* to express experiences, feelings and procedure – as expected in these parts of the book – but here, he does not declare any scientific claims or findings. The rest of the text corresponds to the data exemplified in this beta version of CHET: neither *I* nor *we* out of the recurrent quoting of full letters and discourses that save Birch form presenting his personal view on the facts. This agrees with the absence of verbs like *seem*, *think* or *believe* in CHET's extract and conforms the line of his personal "style". In one of his later works, *Life of Archbishop Tillotson* (1762), Birch follows the same procedure, we can find four to eight pages (e.g. 1762: 222–229) devoted to the reproduction of documents in full. There are first person pronouns (about 7, counted manually), these appear before the mentioned quotes to justify their inclusion:

---

[3] Complete works and their prefaces have additionally been accessed through open online platforms like Archive org. or Google books. Digitized data in these texts, that do not form part of CC, are not parsed and therefore close reading was the only way to evaluate them.

(29) But this fact will be best represented in the Dean's own words in his letter to Lady Russel, which I shall insert here at length, as well as some others, since a mere abstract would be an injury to the reader. (Birch, 1762: 202)

Or in order to indicate opinions, comments on his search of documents as in:

(30) But that was not the case, as I find by the registers of that Church. (Birch, 1762: 152)

For Birch we cannot talk about an idiogenre completely exempt of first person, taking into account the items extant in this later work, however we can see a clear intention in producing books which are closer to compilations which allow for an explicit lack of commitment. It is also curious how even in the *Life of Archbishop Tillotson*, Birch includes a final appendix by John Beardmore (pupil of the Archbishop) which contrasts greatly with his own work, this appendix being full of stance markers including first person pronouns and many attitudinal markers.

I do not think Stewart's explicit "absence" from the text is simply connected with mathematics as a discipline; as mentioned before, other relevant mathematicians use first person at least to mark procedure and Stewart does not even use it in his preface nor in the rest of the text (1761) excluded in CETA. In a previous work published in Volume I of *Essays and Observations, Physical and Literary*, published by the Philosophical Society of Edinburgh, Stewart preferred Latin to English. Certainly, the topic, the extension of one of Pappus' theorems, *Pappi Alexandrini collectionum mathematicarum libri quarti propositio quarta generalior facta, cui propositiones aliquot eodem spectantes adjiciuntur,* could sound more appealing in this language, but it also added some distance from the common reader, since all the other essays published in the same volume (1754) are written by his peers in English, leaving Stewart as the only dissonant note. Two years later, in the second volume of the above-mentioned *Essays*, Stewart writes in English "A Solution of Kepler's problem", but again avoids the use of first person and the main form of engagement found is directives using the imperative form: *Let* or *suppose* (but not *let's suppose*). In 1763 he published two works, one in Latin *Propositiones Geometricæ more veterum demonstratæ,* with some didactic purpose as he states in the reader's preface (no self-mention in it), and *The Distance of the Sun from the Earth determined by the Theory of Gravity*. In this last work, Stewart once more refers to himself in the preface as "the author" and uses third person singular pronouns:

(31) *The author* of the following sheets was early of opinion that this method would fail; and therefore, for many years past, he has employed. (Stewart, 1763: v)

Nothing indicates that this unsigned part of the text was written by anybody else, it just appears that Stewart feels more comfortable adopting an external position that allows him to comment on his own work. Nevertheless, his direct absence does not mitigate the strength of some assertions:

(32) *he thought proper* to consider this subject *himself*; and *has ascertained* the solar force affecting the gravity of the moon to the earth, and from that has calculated *very accurately* the mean distance of the sun from the earth.. (Stewart, 1763: vi)

In the final part of this book (of 103 pages), Stewart finds space for the use of *we* as the exclusive plural form. He uses it three times (see Examples (33) to (35) extracted from pp. 73, 74, 81) to comment "directly" what he thinks about a quotation of another author, Prof. Machin,[4] who attempted the same type of calculus as his, and in a face-saving action to avoid criticism towards Newton and Bernoulli.[5]

(33) which agrees very nearly to what *we* have above demonstrated it to be. (Stewart, 1763: 73)

(34) The distance *we* have assigned is nearly a mean between both. (Stewart, 1763: 74)

(35) *We* will not pretend to give a positive decision which of their observations were best chosen. (Stewart, 1763: 81)

These *we*s correspond to *I*s, they could point to a timid change in Stewart's stance that could be due to a personal evolution and greater confidence, or, I am more inclined to believe, a need to be more evident but still with certain estrangement from the text. This need could derive from the personal issues among the mathematicians he was commenting on which required a more subjective positioning. In the rest of the text and specially in the one included in CETA, Stewart's stance is concerned with his porisms or theorems and no engagement with the reader or other authors' is found. His interest is, as stated by Platfair (1788: 62), to introduce into "the higher parts of mixed Mathematics the strict and simple form of ancient demonstration". Thus, it seems to me that Stewart's choices coincide with the idea portrayed by Hyland (2015: 41) when he states that "there is a tension between using genre conventions to index membership and so claim similarity and gaining a reputation by taking a different stand using a distinctive set of genre options". This is clearly what he does, to maintain a differentiated style and even language to gain a reputation among his peers what, apparently, worked quite well for him.

---

[4] Machin, John (d. 1751), professor of astronomy in Gresham College, enjoyed a high mathematical reputation, until his not well received attempt to rectify Newton's lunar theory in his 'The Laws of the Moon's Motion according to Gravity' (*DNB*).

[5] Johann Bernoulli (1667–1748) was a Swiss mathematician who "defended the cause of G. W. Leibniz in the dispute with Isaac Newton over who had originated calculus. His text in integral calculus appeared in 1742 and his differential calculus shortly afterward. During his last years he worked mainly on the principles of mechanics. His works were published in *Opera Johannis Bernoullii*, 4 vol. (1742)" (*Britannica* s.v.).

5. Final remarks on exceptional voices

It is clear from general results obtained that CC's aim at producing an even collection of authors and texts within one professional community is fulfilled.

The results from the analysis of samples by particular authors in the corpus concord basically with other full works by them. Nevertheless, when considering quantifiable data, as in any corpus, the researcher must be aware that there are other qualitative elements that cannot be left unattended as simple variations. Here, the diversity of modals found in Anderson's selection was a clue to find out more stance elements in the rest of the text and thus exclude him from the "neither *I* nor *we*" group.

Birch and Stewart are exceptional within the eighteenth-nineteenth centuries in their stance taking, since both deviate from the typical authorial voice representation in their disciplines. Soft and hard sciences share in these authors a lack of the clear, direct sound of first person pronouns and other elements that would allow us and their contemporary readers to "hear" them out loud. Nonetheless, they created a discipline-related authorial image almost independent of themselves, a particular silent stance, which is clearly the result of their choices. Whereas for Birch his "absence" did not bring the success he might have searched for, Stewart achieved recognition. This could be due to the extreme reached by Birch, as compared to the "step further" taken by Stewart, a consequence of their quality as writers in their disciplines. On the other hand, a second implication is that stance taking in academic writing can be quite idiosyncratic: it does not always obey the rules but, as a consequence, obtains different possible receptions on the side of one's peers. It is therefore not strange that not many authors dare to surpass the lines.

As the compilation of the whole *Coruña Corpus* advances it will be easier to establish if authors' tendency to conform to the rules is more or less frequent depending on the discipline and if, as in the case of Birch and Stewart, deviation could be linked to degrees of success.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Alonso-Almeida, Francisco & Margarita Mele-Marrero. (2014). Stancetaking in seventeenth century prefaces on obstetrics. *Journal of Historical Pragmatics*, 15(1), 1–35. https://doi.org/10.1075/jhp.15.1.01alo

Anderson, Robert & Stuart Wallace. (2015). The Universities and National Identity in the Long Nineteenth Century c.1830–1914. In Robert Anderson, Mark Freeman & Lindsay Paterson (Eds), *The Edinburgh History of Education in Scotland* (265–286). Edinburgh: Edinburgh University Press.

Anderson, Walter. (1775). *The History of France*. London. Accessed [11/04/17] from *Internet Archive* at https://archive.org/details/bub_gb_clj4P6s2Z6EC

"Bernoulli, Johann" (30 Jan. 2017). *Encyclopædia Britannica*. Accessed [11/06/17] from https://www.britannica.com/biography/Johann-Bernoulli

Birch, Thomas. (1760). *The Life of Henry, Prince of Wales: Eldest Son of King James I*. London: Millar. Accessed [11/04/17] from https://books.google.es/books?id=TUJAAAAQAAJ&printsec=frontcover&hl=es#v=onepage&q&f=false

Birch, Thomas. (1762). *The life of John Tillotson, Lord Archbishop of Canterbury*. London: J. & R. Tonson. Accessed [11/04/17] from *Internet Archive* at https://archive.org/details/lifeofjohntillot00birc

Elliot, Paul. (2010). *Enlightenment, Modernity and Science: Geographies of Scientific Culture and Improvement in Georgian England*. London: IB Tauris.

Griffiths, Ralph (Ed.) (1762). Review of *Tracts, physical and mathematical: containing, an explication*. *The Monthly Review*, 25, 458–460.

Hyland, Ken. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific purposes*, 20, 207–226. https://doi.org/10.1016/S0889-4906(00)00012-0

Hyland, Ken. (2002). Authority and invisibility: authorial identity in academic writing. *Journal of Pragmatics*, 34, 1091–1112. https://doi.org/10.1016/S0378-2166(02)00035-8

Hyland, Ken. (2004). Engagement and Disciplinarity: the other side of evaluation. In Gabriella Del Lungo (Ed), *Academic Discourse: new insights into Evaluation* (13–30). Amsterdam: Peter Lang.

Hyland, Ken. (2005). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7 (2), 173–192. https://doi.org/10.1177/1461445605050365

Hyland, Ken. (2010). Metadiscourse: mapping interactions in academic writing. *Nordic Journal of English Studies. Special Issue on Metadiscourse*, 9 (2), 125–143.

Hyland, Ken. (2015). Genre, discipline and identity. *Journal of English for Academic Purposes*. 19, 32–43

Hyland, Ken & Marina Bondi (Eds.) (2006). *Academic discourse across disciplines*. Berlin: Peter Lang. https://doi.org/10.3726/978-3-0351-0446-2

Hyland, Ken & Carmen Sancho-Guinda (Eds). (2012). *Stance and voice in written academic genres*. New York: Palgrave Macmillan. https://doi.org/10.1057/9781137030825

"Machin, John" (1893). In Clerck, Agnes Mary, *Dictionary of National Biography* 1885–1900, Volume 35. Accessed [11/06/17] from https://en.wikisource.org/wiki/Machin,_John_(d.1751)_(DNB00)

Mele-Marrero, Margarita. (2011). Self-mentioning: Authority, authorship or self-promotion in 17thC prefaces to manuals on obstetrics?. *Revista de Lenguas para Fines Específicos*, 17, 147–166.

Mele-Marrero, Margarita. (2017a). The evolution of self-mentioning in 18thC. CETA and CHET Coruña Corpus. In Francisco Alonso-Almeida (Ed.) *Stancetaking in Late Modern English Scientific Writing. Evidence from the Coruña Corpus* (57–70). Estudios de Lingüística Aplicada 2. Valencia: Universitat Politècnica de València.

Mele-Marrero, Margarita. (2017b). Exclusive-we in the 17th-18thC self-mentioning and stance in scientific works from *Coruña Corpus* (CETA), paper delivered at the 41st AEDEAN Conference.

Mele-Marrero, Margarita & Francisco Alonso-Almeida. (2011). The role of the pronouns he and she in seventeenth-century obstetric directives. *Studia Neophilologica*, 83, 169–181. https://doi.org/10.1080/00393274.2011.614113

Monzó Nebot, Esther. (2002). La profesió del traductor juridic i jurat: descripció sociológica del professional i análisi discursiva del trasgenere. Unpublished PhD Dissertatíon. Castelló: Univetsitat Jaume I.

Moskowich, Isabel & Begoña Crespo (Eds.) (2012). *Astronomy 'playne and simple': The writing of science between 1700 and 1900*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.173

Moskowich, Isabel; Inés Lareo, Paula Lojo Sandino, & Estefanía Sánchez-Barreiro, Estefanía (Comps.) (2019). *Corpus of History English Texts*. A Coruña: Universidade da Coruña. Accessed [31/03/19] from https://ruc.udc.es/dspace/handle/2183/21849

Newton, Isaac. (c. 1665–c. 1672). *The Lawes of Motion*. In Iliffe, Rob (Director) *The Newton project*. Accessed [12/08/17] from http://www.newtonproject.ox.ac.uk/view/texts/normalized/NATP00105

Platfair, John. (1788). Account of Matthew Stewart. *Transactions of the Royal Society of Edinburgh*, 1, 57–75. Accessed [11/06/17] from: https://archive.org/stream/transactionsofro01roya/transactionsofro01roya_djvu.txt

Smollett, Tobias George (Ed.) (1769). *The Critical Review: Or, Annals of Literature*, 27, 160–170. London: Printed for A. Hamilton.

Stewart, Matthew. (1736). *The distance of the sun from the earth determined by the theory of gravity…* Edinburgh: A. Millar. Accessed [17/05/17] from: https://books.google.es/books?id=bpXAAAAYAAJ&hl=es&source=gbs_navlinks_s

Stewart, Matthew. (1754). Pappi Alexandrini collectionum mathematicarum libri quarti propositio quarta generalior facta, cui propositiones aliquot eodem spectantes adjiciuntur. *Essays and Observations, Physical and Literary*, 1, 70–141. Philosophical Society of Edinburgh. G. Hamilton and J. Balfour. Accessed [17/05/17] from: https://books.google.mw/books?id=XsIPAAAAQAAJ&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q=stewart&f=false

Stewart, Matthew. (1761). *Tracts, physical and mathematical: containing, an explication…* Edinburgh: Millar & Nourse. Accessed [11/04/17] from *Google Books* at https://play.google.com/store/books/details?id=ns42AAAAMAAJ&rdid=book-ns42AAAAMAAJ&rdot=1

Stewart, Matthew. (1763). *Propositiones geometricæ,: more veterum demonstratæ, ad geometriam antiquam illustrandam et promovendam idoneæ*. Edinburgh: Typis Sands et al., Accessed [17/05/17] from: https://books.google.es/books?id=ylUUAAAAQAAJ&dq=Propositiones+Geometric%C3%A6+more+veterum+demonstrat%C3%A6&hl=es&source=gbs_navlinks_s

Sword, Helen. (2012). *Stylish Academic Writing*. Cumberland, US: Harvard University Press.

CHAPTER 12

Do writers express the same attitude in historical genres?

A contrastive analysis of attitude devices in the *Corpus of History English Texts*

María Luisa Carrió-Pastor
Universitat Politècnica de València

1. Introduction

Writers may communicate their opinion in different ways depending on the expected audience and the genre they are using. In this paper, my hypothesis is that speakers using different genres employ dissimilar ways to communicate their attitude in historical writing. Specifically, I study the use of attitude devices to prove this. Attitude markers have already been analysed by several researchers, such as Hyland (2005), Gillaerts & Van de Velde (2010), Abdollahzadeh (2011), Mur Dueñas (2010, 2011) and Carrió-Pastor (2015), who describe the role of metadiscourse strategies used in academic papers written in English to identify patterns in the utilisation of rhetorical strategies. However, in this paper, I focus on the role of metadiscourse devices as a means to transmit attitude to readers in historical texts written from 1700 to 1850 and compiled in the *Coruña Corpus of English Scientific Writing*.

The corpus selected in this study was extracted from the beta version of the *Corpus of History English Texts* (CHET), which is the third sub-corpus of the *Coruña Corpus of English Scientific Writing* (CC), compiled by the Research Group for Multidimensional Corpus-Based Studies in English (MuStE) and supervised by Dr. Moskowich (Crespo & Moskowich, 2010). The researchers of this group study the development of the English language from a variety of perspectives, focusing both on language change and language variation from the Middle Ages to present-day English.

In this chapter, I argue that English history writers may use different devices in different genres to transmit their attitude to readers. Hence, the main objectives of this study are, first, to study the use of attitude devices in historical texts written

https://doi.org/10.1075/z.225.12car

© 2019 John Benjamins Publishing Company

between 1700 and 1850; second, to compare the frequency of these attitude devices in the genres of essay, lecture and treatise; and, finally, to identify the characteristic usage and polarity of attitude devices in each genre and the possible reasons.

Attitude devices are used, in academic writing, to interact with readers and explain the authors' findings in a specific context, projecting their attitude or personal opinion to potential readers (Koutsantoni, 2004; Hyland, 2005; Carrió-Pastor, 2014, 2015, 2016a, 2016b; Alonso-Almeida & Carrió-Pastor, 2015). These authors argue that unless readers and writers follow the same codes or ways of deciphering meaning (in this case, attitude) in specific texts, communication may break down. As Hyland & Jiang (2016: 261–262) state "Attitudinal stance […] includes [the writer's] evaluations and personal feelings as he or she comments on the material under discussion or on the communication itself".

It should also be noted that the historical period being studied is the Enlightenment, a period that transcends older notions of a passage from traditional to modern conceptions of science. The characteristics of the Enlightenment and the influence of this period continue to arouse controversy, as there has been a strong interest in this intellectual movement. In this chapter, it may be interesting to examine how history writers across Europe expressed attitude and interacted with readers.

Thus, I intend to prove that the attitude of historical writers may change depending on the genre and, accordingly, the expected readers. The research questions of this chapter are the following:

a. What are the lexico-grammatical devices preferred by history writers in the Enlightenment to show readers their position towards a proposition? Why/Why not?
b. Which attitude devices are more frequently used by history writers in the three different genres in the Enlightenment to communicate affective evaluation? Is the frequency of attitude devices linked to the genre and the expected readers?
c. What is the degree of polarity of the attitude devices found in the corpus? Is the degree of polarity different in the three genres under study?

To answer these research questions, firstly, I review the literature on the study of attitude devices and the historical period of the Enlightenment. Secondly, the methodology followed in this paper is explained and, after that, the results extracted from the analysis are shown in tables and figures. Finally, the conclusions of the analysis are explained taking into account the hypotheses, the objectives and the research questions of this study.

2. Attitude devices

In this chapter, the framework of metadiscourse has been chosen to carry out an analysis of writers' attitude in academic writing. Metadiscourse has been defined by Abdollahzadeh (2011: 288) as the way in which "[…] writers tend to convey their personality, credibility, consideration of the reader and the relationship to the subject matter and to readers by using certain devices in their texts". Metadiscursive devices have been of interest to researchers such as Koutsantoni (2004), Dahl (2004), Hyland (2005), Andrus (2009), Abdollahzadeh (2011), Mur Dueñas (2010, 2011) and Swinehart (2012).

Here, I specifically focus on attitude devices, which are included in the interactional metadiscourse categories (Thompson, 2001). As Hyland (2005: 180) explains: "Attitude markers indicate the writer's affective, rather than epistemic, attitude to propositions, conveying surprise, agreement, importance, frustration, and so on, rather than commitment". Writers use different attitude devices to share values and positions with readers in order to convince them of their propositions. Moreover, Mur Dueñas (2011: 3070) describes attitude devices as: "Items which show the writer's affective evaluation of given parameters or entities". Attitude devices have been of interest to researchers such as Abdi, Rizi & Tavakoli (2010), who have studied them from a general perspective, and refer to them as one of the different metadiscourse strategies. Other authors, such as Gillaerts & Van de Velde (2010), Abdollahzadeh (2011), Mur Dueñas (2011), Hyland & Jiang (2016) and Rodgers (2017) have focused on the study of attitude devices taking into consideration the specificity of the discourse analysed, but only Blagojevic (2009) and Mur Dueñas (2010) have dedicated a specific study to such devices, the former a study of academic papers on the humanities and the latter focused on business management research papers.

Some researchers study attitude devices as markers, taking into account a grammatical perspective. In this sense, the best-known classification is the proposal by Hyland (2005: 180). He classifies attitude markers as verbs (*agree, claim, prefer*), sentence adverbs (*unfortunately, hopefully, surprisingly*) and adjectives (*appropriate, logical, remarkable*). Some other authors, such as Mur Dueñas (2010, 2011) and Jiang & Hyland (2016: 18), have also included nouns that show the author's assessment of the proposition, but the authors point out that "The almost complete absence of discourse types of metadiscursive nouns in the sciences shows a reluctance to build claims through reference to the text or to the argument of the contribution of these factors to the matter under study". In this study, nouns have also been included as attitude devices as it has been considered that nouns also show the affective evaluation of writers.

Focusing on a lexico-grammatical analysis, Koutsantoni (2004: 167) states that

> attitude markers comprise: *evaluative adjectives* (such as *significant*, *interesting*, *important*); *evaluative, intensifying, and attitudinal stance adverbs* (such as *significantly*, *considerably*, *unfortunately*, respectively); *obligation and necessity expressions and modals* (*it is necessary*, *must*, *should*) and *discourse-based negative evaluations of previous research*.

Mur Dueñas (2010) also provides a category structure for Spanish and English attitude devices, which has been very useful in designing this study. This author analyses, first, the difference between research-oriented evaluation and topic-oriented evaluation; second, the centrality and polarity of attitude markers and, finally, the evaluative nature of adjectives that was established by Swales & Burke (2003), who divided evaluative adjectives into acuity, aesthetic appeal, assessment, deviance, relevance, size and strength. Mur Dueñas (2010: 63) simplifies Swales & Burke's taxonomy (2003) by proposing three categories: significance (i.e. relevance, importance), assessment (i.e. efficacy, novelty, interestingness, validity, strength, quality) and emotion (i.e. personal, emotional judgements). In this chapter, special attention is paid to the centrality or polarity of the attitudinal devices found in the three genres to identify the values expressed by the authors.

It should be pointed out that few studies have focused on the analysis of attitude devices in different genres in texts dealing with history. I believe that such an analysis is necessary to identify the patterns of rhetorical strategies followed by authors in such texts.

3. The Enlightenment

As this study focuses on the period between 1700 and 1850, the attitude of history writers of that era should also be explored to fully interpret the results of the analysis and to answer the research questions of this study. This period is important as it is a time when a social, cultural and political movement called the Enlightenment took place. It should be thought of as both an intellectual movement and a social phenomenon. Society all across Europe changed during this period and it may be fruitful to know the causes, nature and effects of revolutions and the way this influenced the writing of essays, treatises and lectures. As this chapter is devoted to historical writing, I do not focus on the historical events or revolutions, but instead I specifically pay attention to the writing of sacred, civil and natural history in the Enlightenment.

This movement cultivated the arts and, consequently, learning the arts and music became more widespread due to economic growth. Studies focusing on literature, philosophy, science and the fine arts became more and more important at that time (Edelstein, 2010). Furthermore, the increased consumption of reading materials of all sorts was one of the key features of the Enlightenment. Up until then, people tended to own a small number of books and read them repeatedly (Engelsing, 1969), but the Enlightenment changed this fact and they began to read as many books as they could, as they had access to a greater amount of literature.

Correspondingly, a variety of institutions offered readers access to material without the need to buy books. In general, readers had access to a large variety of books and this fact increased the number of literate people and access to information and culture. Scientific and literary journals also became popular. The first scientific and literary journals were also established during the Enlightenment (Bloch *et al.*, 2005). Although the existence of dictionaries and encyclopaedias dated back into ancient times, the texts changed from simply defining words in a long running list to far more detailed discussions of those words.

In this vein, one of the most important developments that the Enlightenment era brought to science was its popularisation. An increasingly literate population seeking knowledge and education in both the arts and the sciences drove the expansion of the print culture and the dissemination of scientific learning. Economically, the middle and low classes were able to invest money not only in food but also in education, the main focus of the Enlightenment being on making information available to the greatest number of people (Porter, 2003). Collier Frick (1999: 8) explains that "Science and reason seemed to offer the key to the future, to a kind of paradise which would be realised not in the next world, as the theologians asserted, but in this world, here and now".

As a consequence of the aforementioned aspects, it should be noted that during the Enlightenment, the modern development of historiography through the application of scrupulous methods began. As already mentioned in this volume, historiography is basically the study of the development of history as an academic discipline, using particular sources, techniques and theoretical approaches. Furthermore, as Pocock (1996: 9) states: "The succession of past contexts to one another, around which the concept of a civil history was taking shape, became increasingly exposed to conflicting evaluations, and the process of historical change was perceived in correspondingly sophisticated terms". Summing up, all these aspects were part of an Enlightenment movement to systematise knowledge and provide education to a wider audience than just the upper classes. This fact might have influenced the use of certain attitude devices, as history writers were also aware that their works may be read not only by upper-class readers but also by middle- or even low-class ones.

4. Procedure

4.1 Material

The corpus used in this study was retrieved from the *Coruña Corpus of English Scientific Writing*, which is one of the projects currently being carried out by the Research Group for Multidimensional Corpus-based Studies in English (MuStE). As stated on its website (http://www.udc.es/grupos/muste/corunacorpus/index.html) the compilation of the *Coruña Corpus of English Scientific Writing* (CC) has been and is still governed by some of the most common parameters used in Corpus Linguistics, namely, external criteria for the delimitation of dates, sampling techniques, number of words per sample, etc. (Crespo & Moskowich, 2010). In order to avoid repetitions of patterns caused by different authors' idiosyncrasies, one of the principles of the CC is to include only one sample per author in the whole corpus, even when some of them were quite prolific in different fields of knowledge. All the texts form part of a multi-field indexed textual repository which is used by the information retrieval platform accompanying them (*Coruña Corpus Tool*, CCT).

The texts studied in this chapter are part of the *Corpus of History English Texts* (CHET), which is the third part of the *Coruña Corpus*. It has been compiled to represent English History writing in late Modern English (1700–1900), and is used to describe such a tradition both from a synchronic and a diachronic perspective. As is characteristic of the *Coruña Corpus*, each text file in CHET is accompanied by a metadata file providing information about the text sampled and its author's sociolinguistic background.

The corpus chosen for this analysis was composed of a total of nine history texts: three extracted from the genre essay, three texts belong to treatise and three are lectures. The specific data of each of these texts are described below:

TREATISE

a. Period: 1700–1710. Date: 1704. No. of words: 10,089. Author: Tyrrell, James. Title: *The General History of England both Ecclesiastical and Civil: containing the reign of Richard II, taken from the most ancient records, manuscripts and printed historians, with an appendix…* vol III, second part. Printed for W. Rogers. London.
b. Period: 1770–1780. Date: 1775. No. of words: 10,020. Author: Anderson, Walter. Title: *The History of France. From the Commencement of the Reign of Henry III and the Riſe Of the Catholic League; to the Peace of Vervins, and the Establishment Of the famous Edict of Nantes, In the Reign of Henry IV. Together with The moſt intereſting Events in the Hiſtory of Europe, during that Period*. London: printed for the author.

c. Period: 1820–1830. Date: 1820. No. of words: 10,255. Author: Hardiman, James. Title: *The history of the town and county of the town of Galway, from the earliest period to the present time. Embellished with several engravings. To which is added a copious appendix, containing the principal charters and other original documents*. Dublin: printed by W. Folds and sons.

ESSAY

d. Period: 1700–1710. Date: 1705. No. of words: 10,066. Author: Anderson, James. Title: *An historical essay, shewing that the crown and Kingdom of Scotland, is imperial and independent. Wherein the gross mistakes of a late book, intituled, the superiority and direct dominion of the imperial crown and kingdom of England, over the crown and kingdom of Scotland, and of some other books to that purpose are exposed*. With an appendix. Edinburgh: printed by the heirs and successors of Andrew Anderson.
e. Period: 1750–1760. Date: 1750. No. of words: 10,187. Author: Chapman, Thomas. Title: *An essay on the Roman Senate*. Cambridge: printed by J. Bentham.
f. Period: 1750–1760. Date: 1760. No. of words: 10,059. Author: Birch, Thomas. Title: *The Life of Henry Prince of Wales, Eldest Son of King James I. Compiled chiefly from his own Papers, and other Manuscripts, never before publishe*d. London: printed for A. Millar.

LECTURE

g. Period: 1760–1770. Date: 1770. No. of words: 10,070. Author: Adams, Amos. Title: *A concise, historical view of the perils, hardships, difficulties and discouragements which have attended the planting and progressive improvements of New-England; with a particular account of its long and destructive wars, expensive expeditions, &c.: With reflections, principally, moral and religious: In two discourses, preached at Roxbury on the general fast, April 6. 1769*. Boston: printed and sold by Kneeland and Adams.
h. Period: 1830–1840. Date: 1839. No. of words: 10,117. Author: Petrie, George. Title: *On the History and Antiquities of Tara Hill*. Dublin: Royal Irish Academy.
i. Period: 1840–1850. Date: 1840. No. of words: 9,933. Author: Smyth, William. Title: *Lectures on Modern History, from the Irruption of the Northern Nations to the Close of the American Revolution. In two volumes*. Vol II. Cambridge: J. & J. J. Deighton.

The total number of words compiled in my material of Treatise texts was 30,364, the total number in the Essay sub-corpus was 30,312 and the number of words compiled in the sub-corpus of Lecture texts was 30,120. Consequently, the total number of words included in the history corpus used in this analysis was 90,796.

4.2 Method

Once the papers that made up my corpus had been selected from the corpus using the *Coruña Corpus Tool*, the following step was to analyse the texts to look for attitude devices. When this process was finished, the occurrences obtained were carefully analysed in their context to check whether they could be classified as attitude devices and expressed the authors' views.

The classifications and attitude devices provided by Blagojevic (2009), Abdi, Rizi & Tavakoli (2010), Gillaerts & Van de Velde (2010), Abdollahzadeh (2011) and Mur Dueñas (2010, 2011) were also taken into account to detect the attitude devices in this research. First, the attitude devices were categorised following a lexico-grammatical taxonomy. The attitude devices found in the corpus can be seen below in alphabetical order. The taxonomy used in this study comprises adjectives, verbs, nouns and adverbs:

Adjectives: *acceptable, ancient, best, better, confident, contrary, core, critical, dangerous, desirable, difficile, difficult, easy, exorbitant, fundamental, good, great, hard, important, inconsistent, influential, interesting, intriguing, jealous, little, main, major, meaningful, narrow, necessary, new, notable, obvious, poor, powerful, satisfied, serious, strict, sufficient, surprising, true, turbulent, undutiful, unexpected, unfair, unfaithful, unfortunate, unwilling, unreasonable, unfortunate, useful, valid, valiant, valuable, venerable, virtuous, vehement, wise, worthy*.

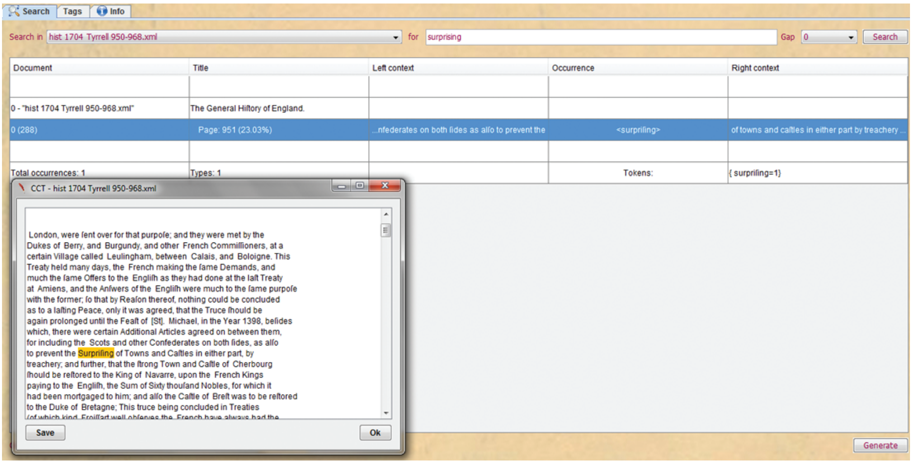
Verbs: *believe, contribute, deny, demonstrate, enjoy, extend, fail, neglect, overlook, support, succeed*.

Nouns: *acceptance, belief, caution, complaint, criticism, crime, credulity, difficulty, dignity, discovery, enemy, failure, faith, importance, integrity, key, limit, peace, power, revenues, scarcity, strength, successful, timidity, value, vanity, violation, violence, wealth*.

Adverbs: *briefly, confidently, considerably, easily, greatly, improbably, indeed, only, particularly, rightly, strongly, tamely, unfortunately, unfrequently, utterly*.

The attitudinal devices were then analysed in context to be sure they indicated attitude and the corpus was checked again with the *The Coruña Corpus Tool* to identify other elements that may be attitude devices. Figure 1 below shows a screenshot of the tool and the process of looking for examples.

Furthermore, *WordSmith Tool 5.0* was also used to identify and revise the total occurrences of the attitude devices. The occurrences of the devices identified were then counted and the frequencies calculated. Moreover, the raw occurrences obtained in the three sub-corpora were normalised to 10,000 words, given that the total number of words in the three sub-corpora was not the same. Examples of the most outstanding attitude devices from the corpus were also discussed.



Search in: hist 1704 Tyrrell 950-968.xml for: surprising Gap: 0

Document	Title	Left context	Occurrence	Right context
0 - "hist 1704 Tyrrell 950-968.xml"	The General Hiſtory of England.			
0 (288)	Page: 951 (23.03%)	...nfederates on both ſides as alſo to prevent the	<surpriſing>	of towns and caſtles in either part by treachery ...
Total occurrences: 1	Types: 1		Tokens:	{ surpriſing=1}

CCT - hist 1704 Tyrrell 950-968.xml

London, were ſent over for that purpoſe; and they were met by the
Dukes of Berry, and Burgundy, and other French Commiſſioners, at a
certain Village called Leulingham, between Calais, and Boloigne. This
Treaty held many days, the French making the ſame Demands, and
much the ſame Offers to the Engliſh as they had done at the laſt Treaty
at Amiens, and the Anſwers of the Engliſh were much to the ſame purpoſe
with the former; ſo that by Reaſon thereof, nothing could be concluded
as to a laſting Peace, only it was agreed, that the Truce ſhould be
again prolonged until the Feaſt of [St]. Michael, in the Year 1398, beſides
which, there were certain Additional Articles agreed on between them,
for including the Scots and other Confederates on both ſides, as alſo
to prevent the Surpriſing of Towns and Caſtles in either part, by
treachery; and further, that the ſtrong Town and Caſtle of Cherbourg
ſhould be reſtored to the King of Navarre, upon the French Kings
paying to the Engliſh, the Sum of Sixty thouſand Nobles, for which it
had been mortgaged to him; and alſo the Caſtle of Breſt was to be reſtored
to the Duke of Bretagne; This truce being concluded in Treaties

Figure 1. An example of one attitude device in *The Coruña Corpus Tool*

Finally, results were shown and analysed, together with examples of attitude devices, which were commented on in context, and the conclusions that were drawn.

5. Results and discussion

The global results of the attitude devices found in the analysis of the treatise, essay and lecture texts can be seen below in Table 1, which shows the raw occurrences and percentages of the genres and the lexico-grammatical categories established:

Table 1. Number of attitude devices found in the treatise, essay and lecture corpora

Attitude devices Lexico-grammatical categories	History corpus Treatise Occurrences (%)	History corpus Essay Occurrences (%)	History corpus Lecture Occurrences (%)	TOTAL Occurrences (%)
Adjectives	248 (14.60)	277 (16.31)	402 (23.67)	927 (54.59)
Verbs	43 (2.53)	18 (1.06)	42 (2.47)	103 (6.07)
Nouns	126 (7.42)	105 (6.18)	190 (11.18)	421 (24.79)
Adverbs	80 (4.71)	69 (4.06)	98 (5.77)	247 (14.55)
TOTAL	**497 (29.27)**	**469 (27.62)**	**732 (43.11)**	**1,698 (100.00)**

It can be observed from the results that history writers that chose the genre lecture used more attitudinal devices in their texts. The characteristics of the genre, which entail an affective attitude of the writer towards the reader, may have been the cause of the higher percentage of occurrences found in the analysis (43.11% vs. 27.62%

and 29.27%). The emphatic expression of affect is relatively infrequent in the genres of treatise and essay. A similar percentage of attitude markers has been observed in both genres and so it seems that their authors do not consider it important to include evaluations and personal feelings in treatises and essays. This may again be due to the characteristics of these genres, since their main purpose is to narrate or explain certain facts rather than being emotional and transmitting affect. Figure 2 shows the comparison of the raw occurrences found depending on the genre used by the history writers. The difference in the use of attitude devices in lectures can be clearly observed:

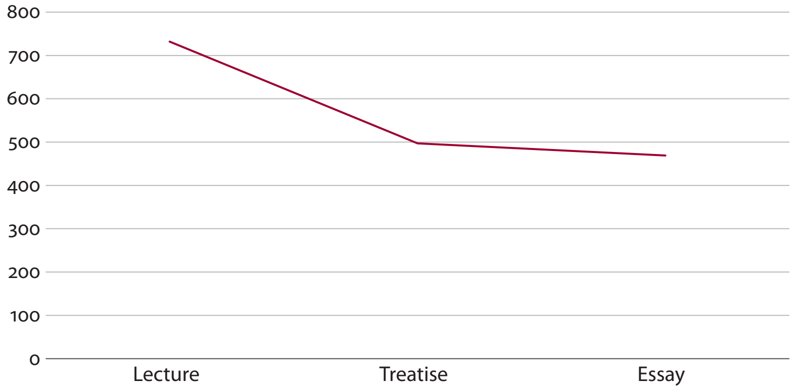


Figure 2. Comparison of the occurrences of attitude devices depending on the genre

The most frequently found lexico-grammatical category in the three genres analysed was that of adjectives (53.70%). This finding coincides with the results of Koutsantoni (2004), Mur Dueñas (2010, 2011) and Hyland & Jiang (2016). The second most frequently used category was nouns (24.39%), followed by adverbs (15.93%) and verbs (5.96%). The results of this analysis of history texts published between 1700 and 1850 differ from those of Mur Dueñas (2010: 56) as, in the case of business papers published in 2002 and 2003, the second most frequent lexico-grammatical category was that of verbs. In the corpus analysed in this chapter, nouns played a role as conveyors of attitudinal meaning. Figure 3 shows the comparison of the occurrences found and the usage of lexico-grammatical categories can be seen very clearly.

After the analysis of the general occurrences of the categories, the results are shown according to the lexico-grammatical category and the different genres. The raw occurrences and the normalised occurrences of the adjectives that act as attitude devices in the corpus analysed can be seen below in Table 2. The results of each sub-corpus are shown normalised to 10,000 words (nf) and ordered from more to less frequent:

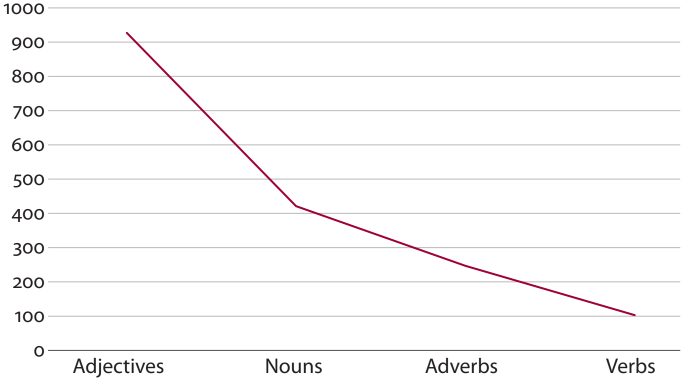


Figure 3. Comparison of the occurrences of the lexico-grammatical categories of attitude devices

Table 2. Raw occurrences and normalised frequencies of adjectives as attitude devices

Attitude devices Adjectives	Treatise raw occurrences (nf)	Essay raw occurrences (nf)	Lecture raw occurrences (nf)
great	62 (20.41)	50 (16.49)	80 (26.56)
ancient	23 (7.57)	36 (11.14)	53 (17.59)
new	16 (5.26)	29 (9.56)	53 (17.59)
good	13 (4.28)	21 (6.92)	14 (4.64)
little	13 (4.28)	11 (3.62)	18 (5.97)
powerful	13 (4.28)	2 (0.65)	5 (1.66)
true	13 (4.28)	34 (11.21)	19 (6.30)
contrary	7 (2.30)	9 (2.96)	4 (1.32)
better	6 (1.97)	2 (0.65)	6 (1.99)
necessary	6 (1.97)	11 (3.62)	20 (6.64)
satisfied	6 (1.97)	0 (0.00)	3 (0.99)
turbulent	6 (1.97)	0 (0.00)	0 (0.00)
best	4 (1.31)	3 (0.98)	5 (1.66)
sufficient	4 (1.31)	1 (0.32)	13 (4.31)
dangerous	3 (0.98)	0 (0.00)	0 (0.00)
hard	3 (0.98)	2 (0.65)	10 (3.32)
important	3 (0.98)	5 (1.64)	22 (7.30)
main	3 (0.98)	3 (0.98)	0 (0.00)
surprising	3 (0.98)	0 (0.00)	0 (0.00)
unfortunate	3 (0.98)	1 (0.32)	3 (0.99)
venerable	3 (0.98)	0 (0.00)	2 (0.66)
jealous	2 (0.65)	2 (0.65)	2 (0.66)
unexpected	2 (0.65)	0 (0.00)	3 (0.99)

(*continued*)

Table 2. (*continued*)

Attitude devices Adjectives	Treatise raw occurrences (nf)	Essay raw occurrences (nf)	Lecture raw occurrences (nf)
unreasonable	2 (0.65)	1 (0.32)	1 (0.33)
confident	1 (0.32)	1 (0.32)	1 (0.33)
core	1 (0.32)	1 (0.32)	0 (0.00)
critical	1 (0.32)	0 (0.00)	8 (2.65)
desirable	1 (0.32)	2 (0.65)	2 (0.66)
difficult	1 (0.32)	0 (0.00)	8 (2.65)
easy	1 (0.32)	1 (0.32)	6 (1.99)
fundamental	1 (0.32)	2 (0.65)	3 (0.99)
influential	1 (0.32)	0 (0.00)	0 (0.00)
interesting	1 (0.32)	5 (1.64)	8 (2.65)
intriguing	1 (0.32)	0 (0.00)	0 (0.00)
meaningful	1 (0.32)	0 (0.00)	0 (0.00)
narrow	1 (0.32)	2 (0.65)	0 (0.00)
strict	1 (0.32)	0 (0.00)	0 (0.00)
undutiful	1 (0.32)	0 (0.00)	1 (0.33)
unfaithful	1 (0.32)	0 (0.00)	0 (0.00)
unwilling	1 (0.32)	0 (0.00)	0 (0.00)
useful	1 (0.32)	0 (0.00)	1 (0.33)
valid	1 (0.32)	0 (0.00)	0 (0.00)
valiant	1 (0.32)	0 (0.00)	0 (0.00)
valuable	1 (0.32)	9 (2.96)	5 (1.66)
vertuous	1 (0.32)	1 (0.32)	0 (0.00)
vehement	1 (0.32)	0 (0.00)	0 (0.00)
wise	1 (0.32)	0 (0.00)	1 (0.33)
worthy	1 (0.32)	5 (1.64)	3 (0.99)
difficile	0 (0.00)	1 (0.32)	0 (0.00)
exorbitant	0 (0.00)	0 (0.00)	1 (0.33)
inconsistent	0 (0.00)	7 (2.30)	7 (2.32)
major	0 (0.00)	2 (0.65)	1 (0.33)
notable	0 (0.00)	2 (0.65)	1 (0.33)
obvious	0 (0.00)	2 (0.65)	1 (0.33)
poor	0 (0.00)	1 (0.32)	2 (0.66)
serious	0 (0.00)	1 (0.32)	1 (0.33)
unfair	0 (0.00)	1 (0.32)	0 (0.00)
acceptable	2 (0.65)	7 (2.30)	2 (0.66)
TOTAL	**248 (81.67)**	**277 (91.38)**	**402 (133.46)**

As stated above, attitudinal adjectives are the most common category in the three genres. The genre in which I found more attitudinal adjectives was lecture with 402 raw occurrences (133.46 per 10,000 words). Essay was the second genre with more occurrences, with a total of 277 (91.38 per 10,000 words), and the least frequent was the genre of treatise, with a total of 248 occurrences (81.67 per 10,000 words). In all, 60 different adjectives denoting attitude were found in the three genres. Many of them were used just once or only a few times in the three genres, although the adjectives *great, ancient, new, good, little, powerful, true, contrary, better, important* and *necessary* were recurrently used. It should be pointed out that these adjectives were not used equally in the three genres. For example, *great* and *ancient* were the most frequent adjectives in the three genres. Examples can be seen in (1) and (2):

(1) "… and to this end, being now under a *great* Conſternation, they diſpatched away the Archbiſhop of York, and Biſhop of London into Ireland, to the King, entreating Him to haſten His Return,"
(Tyrell, 1704: 953. [122 (1327)] Treatise)

The author used the adjective *great* as an evaluative device in the treatise to indicate the consternation of ecclesiastical persons after being accused by the "Wickleffites or Lollards" (members of Parliament who were against the Church, heretics). The writer uses an evaluative adjective to transmit the feelings of the members of the Church in that moment.

(2) "… we commenced, with the map in hand, a second examination of the remains in the order pointed out by the *ancient* descriptions."
(Petrie, 1839: 27. [149 (839)] Lecture)

The historical author, in this case, used the evaluative adjective *ancient* in a lecture text to express the nature of the descriptions. By the use of this adjective the author implies that the map dates from a remote period, it is of great age. In this vein, the synonym *old* could not have been used with the same attitudinal sense.

The third most frequently used attitudinal adjective varies depending on the genre. For example, in the genre Treatise the third most frequently used adjective is *new* (5.26 per 10,000 words). In the genre Essay, the third most frequently used adjective is *true* (11.21 per 10,000 words), and *important* (7.30 per 10,000 words) is the third most frequently used adjective in lectures. Other researchers (Koutsantoni, 2004; Mur Dueñas, 2010; Hyland & Jiang, 2016) have also reported that *important* is one of the top choices in their corpora. It should be highlighted that the most common attitudinal adjectives, excluding *little* and *contrary*, express positive meanings and, in this sense, also transmit the position of history writers who tend to attract readers by using attitude devices that evoke affect from them. But if we observe the case of *little*, history writers used them to express that they would like to know more about a topic, as in Example (3):

> (3) "He is often censured and sometimes corrected, but the fact seems to be, that without his original, and certainly honest account, we should know *little* about the events and affairs he professes to explain."
> (Smyth, 1840: 49. [150 (2989)] Lecture)

In this sense, the meaning of the device should not be interpreted as negative, since the writer explains that he would like to know more about the author mentioned.

Moreover, in the case of *contrary*, the meaning of the attitude device is not negative, but the writer is just reflecting on some limitations, as can be observed in Example (4):

> (4) "But this ſuppoſition of his is quite arbitrary, and deſtitute of any authority to ſupport it. What Mamercus ſaid, as well as did, upon this occaſion, proves alſo very ſtrongly againſt it; and even Zamoſcius himſelf, declares poſitively for the *contrary* opinion." (Chapman, 1750: 55. [132 (8815)] Essay)

It should be noted in Example (4) that the adverb *positively* that precedes the adjective *contrary* lessens the negative meaning of the phrase "the contrary opinion", so we could confirm that the adjectives used as attitudinal devices generally entail a positive meaning.

Nouns are the second most frequent lexico-grammatical category of attitudinal devices in the history corpus analysed. Table 3 compares the raw occurrences and the frequencies normalised to 10,000 words of the nouns that act as attitude devices in the three sub-corpora:

Table 3. Raw and normalised occurrences of nouns

Attitude devices Nouns	**Treatise raw occurrences (nf)**	**Essay raw occurrences (nf)**	**Lecture raw occurrences (nf)**
power	27 (8.89)	56 (18.47)	56 (18.59)
peace	26 (8.56)	6 (1.97)	17 (5.64)
dignity	13 (4.28)	10 (3.29)	0 (0.00)
enemy	13 (4.28)	4 (1.31)	51 (16.93)
revenues	8 (2.63)	3 (0.98)	0 (0.00)
difficulty	4 (1.31)	3 (0.98)	12 (3.98)
faith	4 (1.31)	4 (1.31)	5 (1.66)
importance	4 (1.31)	2 (0.65)	12 (3.98)
complaint	3 (0.98)	1 (0.32)	5 (1.66)
limit	3 (0.98)	1 (0.32)	1 (0.33)
value	3 (0.98)	4 (1.31)	6 (1.99)
timidity	2 (0.65)	0 (0.00)	0 (0.00)
vanity	2 (0.65)	0 (0.00)	0 (0.00)

Table 3. (*continued*)

Attitude devices Nouns	**Treatise raw occurrences (nf)**	**Essay raw occurrences (nf)**	**Lecture raw occurrences (nf)**
acceptance	1 (0.32)	1 (0.32)	0 (0.00)
belief	1 (0.32)	1 (0.32)	5 (1.66)
criticism	1 (0.32)	0 (0.00)	0 (0.00)
crime	1 (0.32)	2 (0.65)	0 (0.00)
credulity	1 (0.32)	1 (0.32)	1 (0.33)
discovery	1 (0.32)	0 (0.00)	0 (0.00)
integrity	1 (0.32)	2 (0.65)	2 (0.66)
key	1 (0.32)	0 (0.00)	0 (0.00)
scarcity	1 (0.32)	0 (0.00)	0 (0.00)
strength	1 (0.32)	0 (0.00)	0 (0.00)
successful	1 (0.32)	0 (0.00)	7 (2.32)
violation	1 (0.32)	1 (0.32)	6 (1.99)
violence	1 (0.32)	1 (0.32)	0 (0.00)
wealth	1 (0.32)	0 (0.00)	0 (0.00)
caution	0 (0.00)	2 (0.65)	4 (1.32)
failure	0 (0.00)	0 (0.00)	0 (0.00)
TOTAL	**126 (41.49)**	**105 (34.58)**	**190 (63.08)**

Twenty-nine attitudinal nouns were identified in my material. Many of them were used just once or only a few times in the corpus analysed. The attitudinal nouns most commonly used by the history writers were *power, peace, dignity* and *enemy*. It should be noted that these devices were not used with the same frequency in the three genres. In the genre treatise, *power* (8.89 per 10,000 words) and *peace* (8.56) were the most frequent. The writers under study used *power* to show the importance of the actions being narrated, as can be seen in Example (5):

(5) "The diſregard of the King, expreſſed by many; the experiments already made of the *power* of that faction, of which he was the head; together with the tameneſs of the government, under the inſults offered to it; conſpired to fill him with hopes of turning the preſent…" (Anderson, 1775: 127. [136 (5686)] Treatise)

In the description of the acts, the writer uses the attitudinal noun *power* to reinforce the strength of the proposition that is followed by the phrase "of which he was the head", which supports the affective claim of the importance of the faction.

In the essays, the most frequently used nouns were *power* (18.47) and *dignity* (3.29). As in the genre Treatise, both nouns were used to infer an affective evaluation upon the texts, as can be seen in Example (6):

(6) “That if the weaknefs of his age would not permit him to do that ſervice to his Majeſty, which he deſired; he would moſt willingly perform what was in his *power*, by daily praying to God to give ſuccefs to his Majeſty’s government, and that he might adminiſter it ſuitably to his own *dignity* and that of his anceſtors, and to the expectations of all his countrymen;…”
(Birch, 1760: 21. [133 (4926)] Essay)

The intention of the writer is to evaluate the Prince positively, thereby expressing his own position to readers and seeking to bring readers into agreement with it. The use of the attitudinal nouns *power* and *dignity* in the same sentence shows the intentional use of the nouns to express the meaning of positive force by the writer.

Among lectures, the most frequently used nouns are *power* (18.59), *enemy* (16.93), *peace* (5.64), *difficulty* (3.98) and *importance* (3.98). In this genre, authors used devices whose meaning is both negative and positive. This combination of negative and positive devices was only found in this genre and it may be caused by the writer’s intention to draw the reader’s attention to the Lecture. An example can be seen in (7):

(7) “And the treaty of Ryſwick, which reſtored *peace* to England and France, did not quickly deliver our frontiers from the Indian *enemy*. Several attacks after this were made on our eaſtern & weſtern borders.”
(Adams, 1770: 32. [135 (6935)] Lecture)

The use of *peace* and *enemy* in the same sentence denotes that the writer of the lecture expresses a stance towards the attacks being suffered (*enemy*) and also tries to align that stance with the interests of the readers of the lecture with the attitudinal noun *peace*.

Adverbs are the third most frequent lexico-grammatical category of attitudinal devices analysed in the three genres. Table 4 shows the raw occurrences and the normalised frequencies of the adverbs that act as attitude devices in the three categories:

Table 4. Raw and normalised frequencies of adverbs

Attitude devices Adverbs	**Treatise raw occurrences (nf)**	**Essay raw occurrences (nf)**	**Lecture raw occurrences (nf)**
only	45 (14.82)	54 (17.81)	58 (19.25)
particularly	14 (4.61)	5 (1.64)	6 (1.99)
considerably	6 (1.97)	0 (0.00)	2 (0.66)
indeed	4 (1.31)	5 (1.64)	16 (5.31)
utterly	3 (0.98)	1 (0.32)	1 (0.33)

Table 4. (*continued*)

Attitude devices Adverbs	**Treatise raw occurrences (nf)**	**Essay raw occurrences (nf)**	**Lecture raw occurrences (nf)**
briefly	1 (0.32)	1 (0.32)	4 (1.32)
confidently	1 (0.32)	0 (0.00)	0 (0.00)
easily	1 (0.32)	0 (0.00)	2 (0.66)
greatly	1 (0.32)	0 (0.00)	8 (2.65)
strongly	1 (0.32)	0 (0.00)	0 (0.00)
tamely	1 (0.32)	0 (0.00)	0 (0.00)
unfortunately	1 (0.32)	1 (0.32)	1 (0.33)
unfrequently	1 (0.32)	0 (0.00)	0 (0.00)
improbably	0 (0.00)	1 (0.32)	0 (0.00)
rightly	0 (0.00)	1 (0.32)	0 (0.00)
TOTAL	**80 (26.34)**	**69 (22.76)**	**98 (32.53)**

The analysis identified seventeen adverbs expressing attitudinal values in the three genres. The most frequently used adverb in the three genres was *only* – a result that coincides with the data extracted by Mur Dueñas (2010: 57). This device is used to identify some aspects that should be excluded from the proposition, as can be seen in Example (8):

(8) "The next year (1756) our troops made an inactive campaign. Nothing of any great moment happened, *only* the loſs of the Engliſh garriſon at Oſwego, and with it all our communication with the great lakes, or inland ſeas".
(Adams, 1770: 43. [135 (9943)] Lecture)

In the case of the Treatise texts, *particularly* (4.61) and *considerably* (1.97) were also frequently used by the writers. These attitudinal adverbs are used to emphasise the illocutionary force of the proposition, as in Example (9):

(9) "… these accounts, that Ireland, at the very time, held constant communication and traffic with the most formidable parts of the Roman empire, and consequently with Spain, to which Galway lies *particularly* convenient."
(Hardiman, 1820: 34. [146 (1731)] Treatise)

In the genre Essay, *indeed* (1.64) and *particularly* (1.64) were the most frequently used attitudinal adverbs. An example of *indeed* in context can be seen in (10):

(10) "Then, *indeed*, 100 new ſenators were made, and they are ſaid, to have owed their ſeats, to the kings favour." (Chapman, 1750: 35. [132 (5835)] Essays)

The writers of lectures used adverbs more frequently to express attitude than the writers of treatises or essays, specifically, the second and third most frequently used attitudinal adverbs are *indeed* (5.31) and *greatly* (2.65).

The least frequent lexico-grammatical category of attitudinal devices was verbs. The raw occurrences and the frequencies normalised to 10,000 words of the verbs that acted as attitude devices in the three genres analysed can be seen below in Table 5. It can also be observed that the frequencies of the attitudinal verbs in the three genres were quite dissimilar:

Table 5. Occurrences and normalised frequencies of verbs

Attitude devices Verbs	**Treatise raw occurrences (nf)**	**Essay raw occurrences (nf)**	**Lecture raw occurrences (nf)**
support	9 (2.96)	0 (0.00)	4 (1.32)
succeed	9 (2.96)	0 (0.00)	12 (3.98)
enjoy	7 (2.30)	2 (0.65)	2 (0.66)
believe	5 (1.64)	6 (1.97)	9 (2.98)
fail	5 (1.64)	0 (0.00)	6 (1.99)
extend	3 (0.98)	4 (1.31)	5 (1.66)
contribute	2 (0.65)	3 (0.98)	2 (0.66)
deny	2 (0.65)	1 (0.32)	2 (0.66)
demonstrate	1 (0.32)	0 (0.00)	0 (0.00)
neglect	0 (0.00)	1 (0.32)	0 (0.00)
overlook	0 (0.00)	1 (0.32)	0 (0.00)
TOTAL	**43 (13.28)**	**18 (5.93)**	**42 (13.94)**

In the case of Treatise, *support* (2.96) and *succeed* (2.96) are more frequently used, whereas in the genre of Essay, *believe* (1.97) and *extend* (1.31) and, in the genre of Lecture, *succeed* (3.98) and *believe* (2.98) are the most frequently used by history writers. *Support* is also the most commonly used verb in the Business corpus analysed by Mur Dueñas (2010: 56), although the meaning of this verb is not the same in the corpus of history texts analysed here, as can be observed in Example (11):

(11) "hat if they endeavour'd to Revenge it, the King of France might intereſt Himſelf in the Quarrel, and ſend over Forces to join with thoſe in England, to *Support* the King's Authority; ſo that the Duke of Lancaſter had ſome thoughts of ſending to his Sons-in-Law the Kings of Castile, and Portugal, for Aſſiſtance"
(Tyrrell, 1704: 966. [122 (8694)] Treatise)

In this example, *support* is used to state the authority of the King and the loyalty of kings from other countries.

The frequencies of the attitudinal verbs are not very high, but it should be noted that the meaning of all these verbs is positive and so the writers are seeking to bring readers into agreement with the position stated in the texts.

In the analysis of the attitude devices in three different genres in history texts from 1700 to 1850, the values expressed by those devices were also taken into account, following the studies of Swales and Burke (2003) and Mur Dueñas (2010). Accordingly, attitude devices were analysed focusing on the terms of centrality and polarity. As Mur Dueñas (2010: 60–1) explains "… thus, strongly positive or negative adjectives would be polarised, whereas more neutral adjectives would be centralised". In this sense, attitude devices in the corpus were classified as centralised (low degree or unmarked) and polarised (high degree or marked), according to the degree of attitudinal value expressed. The results in raw occurrences and normalised frequencies per 10,000 (in brackets) can be observed in Table 6:

Table 6. Occurrences and normalised frequencies of polarised and centralised attitude devices

Attitude devices	**Treatise centralised**	**Treatise polarised**	**Essay centralised**	**Essay polarised**	**Lecture centralised**	**Lecture polarised**
Adjectives	154 (50.71)	94 (29.64)	203 (66.97)	74 (24.41)	293 (97.27)	109 (36.18)
Nouns	85 (27.99)	41 (13.50)	38 (12.53)	67 (22.10)	111 (36.85)	79 (26.22)
Adverbs	58 (19.10)	22 (7.24)	63 (20.78)	6 (1.97)	81 (26.89)	17 (5.64)
Verbs	22 (7.24)	21 (6.91)	10 (3.29)	8 (2.63)	27 (8.96)	15 (4.98)
TOTAL	**319 (105.05)**	**178 (58.72)**	**314 (103.58)**	**155 (51.13)**	**512 (169.98)**	**220 (73.04)**

It can be observed in the results that in the three genres the writers used more centralised than polarised words. The devices chosen as polarised were identified and analysed in context and the presence of boosters or hedges was also considered a factor of polarisation in order to obtain full understanding of the attitudinal meaning of the three genres. The results extracted here are in line with those of Swales & Burke (2003) and Mur Dueñas (2010). The corpora analysed in these studies are quite different (written and spoken English and academic Business papers, respectively), but the data studied here seem to point to similar results, namely, that history writers do not use as many polarised devices. It may be stated from Table 6 that the writers of lectures used more polarised devices (73.04) than the writers of treatises and essays (58.72 and 51.13 respectively), but if the results are compared with the centralised devices used in the three genres, the preference for centralised rather than polarised evaluative devices is evident from the data.

Examples of centralised attitude devices with a low degree of evaluativeness in the three genres are illustrated by Examples (12–14), while polarised devices with a high degree of evaluativeness are shown in Example (15):

(12) "… and therefore, that he ſhould be extremely ready to acquieſce, in whatever they ſhould think proper to determine, in this *important* affair."
(Chapman, 1750: 2. [132 (409)] Essay)

(13) "… but also in the popular *belief* of the people in several parts of Ireland, to the present time." (Petrie, 1839: 69. [149 (7832)] Lecture)

(14) "… before I conclude this lecture, I *briefly* advert, still remains. The student must never forget that he is at all times to keep his attention fixed…"
(Smyth, 1840: 75. [150 (8701)] Lecture)

(15) "An inſtance of this is given us by Plutarch, in his life of Pompey. This *great* man, who had been honoured with two triumphs, and was now in poſſeſſion of the conſulate…" (Chapman, 1750: 16. [132 (3009)] Essay)

(16) "He about the beginning of September, ſet Sail for Ireland, attended by the Duke of Glouceſter, the Earls of Marche, Nottingham, and Rutland, with a *powerful* Army of about Thirty Thouſand Men.
(Tyrrell, 1704: 952. [122 (902)] Treatise)

(17) "I must take the liberty of considering it, as a monument of national impolicy, and even national want of good *faith* and honour."
(Smyth, 1840: 77. [150 (9157)] Lecture)

In the examples, the contextualised connotation between the meaning of *important* and *belief* in (12)–(13) and the meaning of *great* and *faith* in (15) and (17) should be noted, as history writers used polarised attitudinal devices to influence the opinion of readers by emphasising certain characteristics that were typical of that historical period. Some of the devices extracted from the corpus are used nowadays but it should be noted that most of the attitudinal devices analysed in this chapter entail certain conceptions that can only be interpreted by taking into account the social and economic period of the Enlightenment.

6. Conclusions

The research presented in this paper has focused on the analysis of attitude devices in three different genres of the *Corpus of History English Texts* (CHET) to study the different attitudinal strategies followed by historical writers between 1700 and 1850. The main aim was to identify the variation in the use of attitudinal devices in history writing and to show that genre played a role and that writers adapted their writings to their expected readers. This analysis has thus helped to gain deeper insight into the rhetorical strategies used by history writers, particularly by the community of writers that lived in the Enlightenment, a changing period that entailed fluctuations in multiple sectors.

After the analysis of the results in Tables 2, 3, 4 and 5, it has been found that history writers in the Enlightenment preferred the use of adjectives to show readers their position towards a proposition. This finding is in line with other studies (Koutsantoni, 2004; Mur Dueñas, 2010; Jiang & Hyland, 2016), which noted that adjectives are the preferred lexico-grammatical item of writers in academic English to express attitude.

Focusing on the attitude devices that were more frequently used by history writers to communicate affective evaluation in the three different genres in the Enlightenment, it was observed that the frequent use of devices changed depending on the genre and the lexico-grammatical category, although the most common attitudinal device in the corpus was *great*. But if we take into account their lexico-grammatical category, the two most common attitudinal adjectives in the three genres were *great* (T:[1] 20.41; E: 16.49 and L: 26.56) and *ancient* (T: 7.57, E: 11.14 and L: 17.59). In contrast, in the case of nouns, the most often used item was *power* (T: 8.89; E: 18.47 and L: 18.59), but the second most frequently used varied from one genre to another. In treatises it was *peace* (8.56), while in essays, writers preferred *dignity* (3.29) and in lectures, *enemy* (16.93).

In the case of adverbs the most commonly used item was *only* (T: 14.82; E: 17.81 and L: 19.25), while the second was *particularly* in the case of treatises (4.61) and essays (1.64), and *indeed* (5.31) in the case of lectures.

Concerning attitudinal verbs, the frequencies varied depending on the genre. In the case of treatises, *support* (2.96) and *succeed* (2.96) were the most commonly used, although the frequencies found in the corpus were low. Concerning essays, the most common attitudinal verb was *believe* (1.97) and in lectures the most common item was *succeed* (3.98).

Finally, as can be observed in Table 6, most attitudinal devices were used by authors in a centralised way, given that in the Enlightenment writers tended to be objective. As Pocock (1996: 9) states in the third section of this chapter: "The succession of past contexts to one another, around which the concept of a civil history was taking shape, became increasingly exposed to conflicting evaluations", and so the authors tend to use centralised devices. Concerning the use of polarised devices, their frequency varied depending on the genre. The authors of lectures used more polarised devices (73.04) than the authors of treatises (58.72) and essays (51.13), and the reason for this may be that writers seek to get their readers involved in the lecture and hence more polarised devices are used to achieve this effect.

Overall, then, it could be concluded that history writers from the Enlightenment used similar attitudinal devices to present-day writers of academic English, although it should be noted that genres and the expected readers of each genre caused variations in the use of attitudinal devices. If we analyse the items used in the three genres in detail, it can be observed that the most frequent items are different from those used in academic English due to the specific topic (history) of the texts analysed. Thus, it can be said that the choice of items by the writers is closely connected to the time and the historical events of that period. The intention of the writer is the key aspect when choosing an attitudinal device.

---

[1] T: Treatise texts, E: essay texts, L: lecture texts.

Attitudinal devices have been analysed in this chapter, but I believe it could be interesting to investigate other interactional devices used in history texts from 1700 to 1850, such as hedges, boosters, engagement markers or self-mentions, or from the point of view of stance, evidentiality and presence.

References

Abdi, Reza, Manoochehr T. Rizi & Mansoor Tavakoli. (2010). The cooperative principle in discourse communities and genres: A framework for the use of metadiscourse. *Journal of Pragmatics*, 42(6), 1669–1679. https://doi.org/10.1016/j.pragma.2009.11.001

Abdollahzadeh, Esmaeel. (2011). Poring over the findings: Interpersonal authorial engagement in applied linguistics papers. *Journal of Pragmatics*, 43, 288–297. https://doi.org/10.1016/j.pragma.2010.07.019

Andrus, Jennifer. (2009). The development of an artefactual language ideology: Utterance, event, and agency in the metadiscourse of the excited utterance exception to hearsay. *Language & Communication*, 29, 312–327. https://doi.org/10.1016/j.langcom.2009.05.001

Alonso-Almeida, Francisco & María Luisa Carrió-Pastor. (2015). Sobre la categorización de seem en inglés y su traducción en español. Análisis de un corpus paralelo. *Revista Signos*, 48, 154–173. https://doi.org/10.4067/S0718-09342015000200001

Blagojevic, Savka. (2009). Expressing attitudes in academic research articles written by English and Serbian authors. *Facta Universitatis*, 7–1, 63–73.

Bloch, Jean et al. (2005). *History of ideas. Travel writing. History of the book. Enlightenment and antiquity*. Oxford: Voltaire Foundation. University of Oxford.

Carrió-Pastor, María Luisa. (2014). Cross-cultural variation in the use of modal verbs in academic English. *Sky, Journal of Linguistics*, 27, 153–166.

Carrió-Pastor, María Luisa. (2015). Identification of Rhetorical Moves in Business E-mails Written by Indian Speakers of English. In Erika Daricks (Ed.), *Digital Business Discourse* (226–242). London: Palgrave-MacMillan.

Carrió-Pastor, María Luisa. (2016a). A contrastive study of the hedges used by English, Spanish and Chinese researchers in academic papers. In Francisco Alonso Almeida et al. (Eds.), *Input a word, analyze the world: Selected approaches to Corpus Linguistics* (477–492). Newcastle upon Tyne: Cambridge Scholars.

Carrió-Pastor, María Luisa. (2016b). A contrastive study of interactive metadiscourse in academic papers written in English and in Spanish. In Francisco Alonso Almeida, Laura Cruz García & Víctor González Ruiz (Eds.), *Corpus-based studies on language varieties*. Bern: Peter Lang.

Collier Frick, Carole. (1999). *The Enlightenment*. Los Angeles: University of California.

Crespo, Begoña & Isabel Moskowich. (2010). CETA in the Context of the Coruña Corpus. *Literary and Linguistic Computing*, 25–2, 153–164. https://doi.org/10.1093/llc/fqp038

Dahl, Trine. (2004). Textual metadiscourse in research articles: a marker of national culture or of academic discipline? *Journal of Pragmatics*, 36, 1807–1825. https://doi.org/10.1016/j.pragma.2004.05.004

Edelstein, Ron (Ed.) (2010). *The Super-Enlightenment*. Oxford: Voltaire Foundation. University of Oxford. https://doi.org/10.7208/chicago/9780226184500.001.0001

Engelsing, Ron. (1969). Die Perioden der Lesergeschichte in der Neuzeit. Das statische Ausmass und die soziokulturelle Bedeutung der Lektüre. *Archiv für Geschichte des Buchwesens*, 10.

Gillaerts, Paul & Freek Van de Velde. (2010). Interactional metadiscourse in research article abstracts. *Journal of English for Academic Purposes*, 9, 128–139. https://doi.org/10.1016/j.jeap.2010.02.004

Hyland, Ken. (2005). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7–2, 173–192. https://doi.org/10.1177/1461445605050365

Hyland, Ken & Feng K. Jiang. (2016). Change of attitude? A Diachronic study of stance. *Written Communication*, 33–3, 251–274. https://doi.org/10.1177/0741088316650399

Jiang, Feng K. & Ken Hyland. (2016). Nouns and academic interactions: A neglected feature of metadiscourse. *Applied Linguistics*, published online: 1–25. https://doi.org/10.1093/applin/amw023. [Accessed 12/12/2016]

Koutsantoni, Dimitra. (2004). Attitude, certainty and allusions to common knowledge in scientific research articles. *Journal of English for Academic Purposes*, 3, 163–182. https://doi.org/10.1016/j.jeap.2003.08.001

Mur Dueñas, Pilar. (2010). Attitude markers in business management research articles: a cross-cultural corpus-driven approach. *International Journal of Applied Linguistics*, 19, 50–72. https://doi.org/10.1111/j.1473-4192.2009.00228.x

Mur Dueñas, Pilar. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. *Journal of Pragmatics*, 43, 3068–3079. https://doi.org/10.1016/j.pragma.2011.05.002

Pocock, John G. A. (1996). Classical and Civil History: The Transformation of Humanism. *Cromohs*, 1, 1–34. Accessed [10/10/2017] at http://www.unifi.it/riviste/cromohs/1_96/pocock.html

Porter, Michael. (2003). The Economic Performance of Regions. *Regional Studies*, 37–6, 549–578. https://doi.org/10.1080/0034340032000108688

Rodgers, Elena. (2017). Towards a typology of discourse-based approaches to language attitudes. *Language & Communication*, 56, 82–94. https://doi.org/10.1016/j.langcom.2017.04.002

Swales, John & Amy Burke. (2003). "It's really fascinating work": differences in the evaluative adjectives across academic registers. In Pepi Leistyna & Charles F. Meyer (Eds.), *Corpus Analysis, Language Structure and Language Use* (1–18). New York: Rodopi.

Swinehart, Karl F. (2012). Metadiscursive regime and register formation on Aymara radio. *Language & Communication*, 32, 102–113. https://doi.org/10.1016/j.langcom.2011.05.004

Thompson, Geoff. (2001). Interaction in academic writing: learning to argue with the reader. *Applied Linguistics*, 22–1, 58–78. https://doi.org/10.1093/applin/22.1.58

CHAPTER 13

On cognitive complexity in scientific discourse

A corpus-based study on additive coherence relations

Iria Bello
Ruprecht-Karls-Universität Heidelberg

1. Introduction

Languages resort to diverse linguistic devices to link semantically and pragmatically discourse members, and express simultaneously the relation between them (Pander Maat & Sanders, 2006). This way, speakers may limit the potential ambiguity of the utterances and guide inferential processes in communication, which often leads to less cognitive effort in processing (Blakemore, 1987; Sperber & Wilson, 1995, 2002; Levinson, 2000; Wilson & Sperber, 2002). Discourse markers are one of those functional guidelines (van Dijk & Kintsch, 1983). They provide instructions for information processing and limit the potential ambiguity of utterances. At the textual level, discourse markers help organize discourse and delimit the structure of the information conveyed in texts. Hence, some may offer guidelines on how to structure a text (*on the one hand, on the other hand*), others are useful reformulators (*that is*); there are also connectives that provide guidelines for argumentation (*furthermore, however, besides, in addition, moreover*), whereas others work at the informational level (*even, also, as well, too*).

Additive markers provide instructions on how to connect discourse segments and thus they establish cohesion and coherence relations. Cohesion takes place "when the interpretation of some element in the discourse is dependent on that of another" (Halliday & Hassan, 1976: 4). Coherence is a type of connectedness between discourse segments that creates mental representations that are based on inferences (Sanders & Pander Maat, 2006). It transcends purely linguistic signs and it needs the combination of external and internal input stored in the speaker's minds. Coherence relations may be referential – linguistically encoded in pronouns and demonstratives that connect units with the same mental referent – or relational – encoded in connectives providing relational instructions among discourse segments (Sanders *et al.*, 1992; Sanders & Pander Maat, 2006; Pander Maat & Sanders, 2006; Spooren & Sanders, 2008).

https://doi.org/10.1075/z.225.13vir

© 2019 John Benjamins Publishing Company

De Cesare (2017: 2) addressed the implicitness of additive relations in discourse, asserting that "additive discourse relations can remain unexpressed" and therefore, the inclusion of an additive marker can be seen as a reinforcement of the informative and argumentative structures. For her, additivity is more basic than other logical or rhetorical relations such as contrast, condition or concession and can be considered the default interpretation when combining two adjacent discourse segments. This contrasts with the Causality-by-Default hypothesis proposed by Sanders (2005), according to which listeners/readers by default expect two segments in a discourse to be causally related. According to the classification of coherence relations provided by Sanders *et al.* (1992, 1993, see also Spooren & Sanders, 2008), additive relations fall within the basic operation (the remaining three coherence relations proposed being polarity, source of coherence and order of the segments). They are weak connections based on a relation of conjunction and they contrast with causal relations, where the type of relation is conjunctional as well as implicational, which results in a stronger connection. In acquisitional terms, Spooren & Sanders (2008: 2006) noted that additive relations and connectives are acquired before causal ones (Piaget, 1969; Katz & Brent, 1968; Bloom *et al.*, 1980; Eisenberg, 1980).

The language of science in English, its origins and evolution has been an inexhaustible treasure trove for linguists. Some have focused on prototypical grammatical features of scientific register like passives and nominalizations (Halliday & Martin, 1993; Albentosa, 1997; Albentosa & Moya, 2000; Banks, 2001, 2003, 2005a, 2005b, 2007, 2008; Halliday, 2004; Vázquez, 2006; Lareo, 2011) or specialized vocabulary (Camiña, 2012, 2013; Camiña, Esteve & Lareo, 2012). Others have tried to provide a sociolinguistic perspective to the matter (Atkinson, 1998; Gotti, 2011, 2012, 2013; Beal, 2012; Moskowich & Crespo, 2012; Moskowich & Monaco, 2016). The Scientific Revolution brought a new way of doing and communicating science. Even if we acknowledge that English had already been the language of transmission of scientific knowledge since Old English, during the seventeenth century, the scientific register was readjusted, especially after the publication of the *Philosophical Transactions* (Taavitsainen & Pahta, 1997). According to Barber (1993: 214), the influence of science on language caused not only the expansion of vocabulary, but also the adoption of a plain style written in prose as the most usual way of conveying scientific, and, eventually, any kind of written knowledge: "The rise of scientific writing in English helped to establish a simple referential kind of prose as the central kind in Modern English". This change was motivated by a grandiloquent purpose: to "restore the biblical connection between words and things" (Camiña, 2013: 57), and at the same time to make knowledge accessible to a wider audience by campaigning against redundancy and ambiguity in discourse. The aim of this chapter is to contribute to the description of the scientific register in English by focusing on the simplest type of coherence: additive relations.

The organization of this article is as follows. In Sections 2 and 3, additive focus operators and connectives are discussed. In the following section, the corpus and methodology used for this study are described. This is followed by the analysis of the data obtained after corpus exploitation, which will include an analysis of their frequency and evolution across the two centuries, as well as an analysis of their use according to the gender and place of education of authors and the typology of texts. In the last section of the chapter, conclusions and suggestions for further research are offered.

2. Additive focus operators

Focus operators contribute to the interpretation of focused elements in relation to their background.[1] They therefore work to structure information within a context. They impose restrictions to a set of alternatives and select the most relevant element in a specific context (Huddleston & Pullum, 2002: 594; Gast & Van der Auwera, 2011; Schwenter, 1999). Focus operators like *even, also, as well* or *too* establish a relation of likelihood within a scale and present an additional, focused element as the most informative (Kay, 1990) and the least expected element as non-informative (König, 1991; Gast & Rzymski, 2015). In (1) below, the reader will infer that some palaces are more magnificent than others based on the scalar instruction triggered by *even*. This inference is possible only because pragmatic scales work with pragmatic assumptions shared by speakers (Fauconnier, 1975).[2] In order for this sentence to be understood, readers need to create a mental representation that includes not only the external, visual stimuli they are receiving (the sentence they are reading) but also internal input stored in their minds that includes linguistic information as well as the encyclopaedic knowledge stored in our long-term memory. Consequently, the reader needs to know that the magnificence of palaces may be judged by their gardens and fountains and that, of all palaces in the world, Versailles is (or it was in 1739), the most magnificent.

---

[1] Unfortunately, there is a myriad of different names to refer to these "connective elements [that] often link units smaller than the clause" (Huddleston & Pullum, 2002: 775). Hence, they are usually referred to as focusing subjuncts (Quirk *et al.*, 1985), focus particles (König, 1991; Loureda *et al.*, 2015; Nadal *et al.*, 2017); additive modifiers (Huddleston & Pullum, 2002), scalar (additive) operators (Gast & Van der Auwera, 2011; Gast, 2017), or additive focus adverbs (Andorno & De Cesare, 2017), among others.

[2] Pragmatic scales are related because of the illocutionary meaning of the elements. There is another type of scales, semantic scales, in which the relation of the elements is derived from their propositional meaning, as in the Example (*) below, where the scalarity is provided by the gradation of the adjective:

(*) It's not the only book in the library, and maybe not even the best.

(1) I thought myſelf very lucky in having an Opportunity of ſeeing ſo fine a Place as that really is; […] which I did, many of them walking in the Gardens. This Place is by ſome thought to excell *even* Verſailles in Water-Works. (Justice, 1739: 9 [129 (2072)])

There are two types of additive scales, namely culminative or non-culminative (Portolés, 2007). In (2) below the inference triggered by the focus operator *even* is that of all military ranks explicitly listed in the scale (Colonel, Lieutenent-Colonel, Major and Soldier) it is less likely that the last one (Soldiers) would receive gifts.

(2) Every General had a Gold Medal given him, with the two Czar's Effigies on one Side, and the Princess's on the other, fastened to a Chain of Gold, the whole worth about ten Ducats. Every Colonel, Lieutenant-Colonel, Major, and *even* every Soldier had Gifts in Proportion. (Bancks, 1740: 15 [130 (3176)])

In culminative scales, the last element becomes the focus and also the more informative element in the scale (cf. König, 1991; Jacobs, 1983; Andorno, 2000; Gast & Rzymski, 2011). This is the main difference with the non-culminative *also*, which establishes an egalitarian relationship between all the coordinated elements. Hence, in (3) the scale established by *also* discards the possibility that one of the elements may be more informative than the others. In non-culminative, additive scales, the maximum degree of informativity is achieved only by the sum of all the members in the scale and, following a mathematical commutative property, the elements in the scale can be rearranged without altering the meaning, as in (4). On the contrary, the inversion of the alternatives and the focus in culminative scales containing *even* generates pragmatically awkward statements, as in (5):

(3) [T]he Carthaginians muſt have been a very ſenſeleſs *as well as* ungrateful people […]. (Hooke, 1745: 39 [129 (6042)])

(4) They would burn their Houses and kill their Cattle *as alſo* them. (adapted from Penhallow, 1726: 85 [120 (839)]

(5) #Every Soldier, Lieutenant-Colonel, Major, and *even* every Colonel had Gifts in Proportion. (adapted from Bancks, 1740: 15 [130 (3176)])

3. Additive connectives

Due to their procedural meaning (Blakemore, 1987; Escandell-Vidal & Leonetti, 2011), additive connectives compel the reader to connect two (or more) utterances with the same argumentative orientation. The two connected segments "cooperate together to express the same logical relation in respect to a third proposition" (De Cesare, 2017: 5), that is, the conclusion. Thus, in (6) we see an entailment of arguments (the Romans had good engines; Roman soldiers were animated by their generals), from which the reader can infer the progression of the Romans against the Carthaginians in the battle. Additive connectives enhance discourse coherence by integrating the new information of the last discourse segment into the previous one (processing cycles, cf. Kintsch & Van Dijk, 1978: 368).

(6) [F]or tho' the Carthaginians had the advantage in the lightneſs and ready working of their veſſels, yet the Romans loſt not their aſſurance of ſucceſs in the end; they found themſelves better men when they came to the ſword's point, and they had great truſt in their engines, with which they grappled and boarded the enemy; *beſides*, the ſoldiers were animated by the preſence of their Generals, in whoſe eye they fought, and who themſelves engaged in equal hazard with the reſt. (Hooke, 1745: 32 [129 (2333)])

This reinforces the weight of the preceding argument "by a procedure of argumentative accumulation" (Domínguez García, 2007: 60). The discourse segments linked by additive connectives also form a scale but, unlike scales containing *even*, the last segment does not become argumentative marked. They are therefore non-culminative scales in which elements can be reordered without changing the argumentative strength of the whole set, as in (7) below:

(7) They found themſelves better men when they came to the ſword's point, and; the ſoldiers were animated by the preſence of their Generals, in whoſe eye they fought, and who themſelves engaged in equal hazard with the reſt. *beſides*, they had great truſt in their engines, with which they grappled and boarded the enemy. (adapted from Hooke, 1745: 32 [129 (2333)])

Discourse segments joined together by connectives may be of varying lengths. Therefore, they may be full sentences, as in (6) or smaller units, as in (8).

(8) The soldiers were promised a crown of gold *in addition to* the more solid recompenses they might expect. (Sewell, 1857: 253 [129 (8130)])

4. Methodology

The corpus material for this study was taken from one of the subcorpora of the *Coruña Corpus* (CC) (Moskowich *et al.*, 2012). The CC was designed for diachronic linguistic study and it is made up of several subcorpora of different scientific disciplines. Its timespan covers the eighteenth and nineteenth centuries and each corpus contains two texts per decade written by native English-speaking authors. Each sample text has approximately 10,000 words, which makes up a total of approximately 400,000 analysable words per subcorpus (Table 1). For the present work, CHET, the history subcorpus, was chosen.

Table 1. Number of words in CHET

Century	Word count
Eighteenth century	201,938
Nineteenth century	202,486
Total	**404,424**

This study deals with additive devices. Following Huddleston & Pullum (2002: 586–595) and Quirk *et al.* (1985: 631–647), eight additive devices were chosen: 4 focus operators and 4 connectives. The number of occurrences was automatically extracted with the help of the *Coruña Corpus Tool* (*CCT*), a search engine that has been designed for joint use with the CC.

CHET contains biographical information about the author's gender, age and place of education, and also information about the genre/typology of texts.[3] In CHET there are 8 female and 32 male authors, as a consequence of the unbalanced situation in science at that time. Of all the texts written by women, 2 were published in the eighteenth century. Table 2 shows the distribution of the number of words in texts written by female and male authors.

Table 2. Number of words in CHET according to gender

Gender	Word count
Women	81,497
Men	322,927
Total	**404,424**

---

[3] In an attempt to simplify the terminological maelstrom, in this chapter genre/text-type are understood in the way Taavitsainen (2001) did. She (2001: 140) defined genre as a "mental frame in people's minds which gets realized in texts for a certain purpose in a certain cultural context" and considered that text-types were a linguistic realization of genres.

Concerning the origin of authors in the corpus (see Table 3), the apparent unbalanced distribution reflects the situation of science in English-speaking countries in the period. During the eighteenth century, science was almost exclusively an English enterprise and the American colonies were not only politically and economically, but also scientifically dependent of England (Greene, 1954).

Table 3. Number of words in CHET according to place of education of authors

Place	Word count
England	242,581
Ireland	60,646
Scotland	50,479
U.S.	40,386
Unknown	10,332
Total	**404,424**

Samples in CHET have been divided into eight groups according to text-type, namely, article, biography, dictionary, lecture, essay, textbook, travelogue and treatise. Compilers of the corpus based their text-type decisions in the notion of paragenre, that is, “a genre belonging to one professional community” (Moskowich, 2012: 37). By doing so, they highlighted the sociolinguistic orientation of the corpus. The distribution of text-types in CHET according to number of words is shown in Figure 1. There are 28 samples of academic treatises and dictionaries, travelogues as well as articles are represented with one sample each.

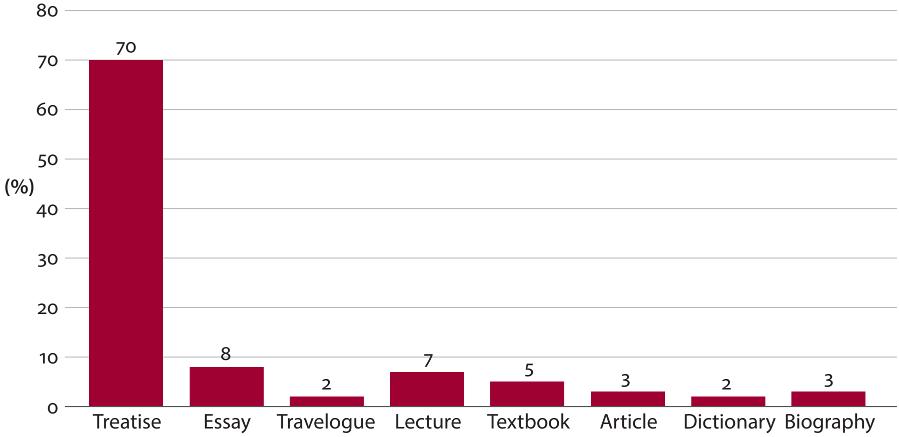


Figure 1. Text-types in CHET

All the possible text-types were then divided into two main groups for this analysis. On one hand, specialised texts (research articles, academic treatises and essays) are aimed at the scientific community. Language use in these texts is *a priori* more sophisticated, whereas non-specialised texts (biographies, travelogues, lectures and text books) are also targeted at either students or the general public. Hence, in non-specialised texts it is more probable to find not only less specialised vocabulary but also less complex linguistic structures. The total number of words for both groups is shown in Table 4:

Table 4. Number of words in CHET according to text type

Text type	Word count
Specialised text	324,044
Non-specialised text	80,380
Total	**404,424**

5. Data analysis

A total of 992 additive markers were found in the corpus. As Table 5 shows, of all the additive markers analysed, focus operators (*also, as well, even, too*) are more frequent. Its normalised frequency (21.35) is higher than that of connectives (2.99) (*besides, moreover, in addition*).[4] Additive argumentative relations are often part of the implicit text (cf. Kintsch & Van Dijk, 1978; De Cesare, 2017) and therefore

Table 5. Frequency of additive devices in CHET
(normalised frequencies per 10,000 words in parentheses)

Additive marker	Frequency
also	439 (10.85)
as well	105 (2.59)
besides	75 (1.85)
even	299 (7.39)
in addition	7 (0.17)
moreover	39 (0.96)
too	28 (0.69)
Total	**992 (24.52)**

---

[4] Huddleston and Pullum (2002) and Quirk *et al.* (1985) also include *furthermore* as an additive connective. However, no occurrences of either this or any equivalent form (*farthermore*, *farther*) were found in the corpus. *Further* had 94 occurrences but all of them were adjectives.

the reader could find it easier to connect two utterances, even if there is no explicit additive connective. This would not apply to focus operators, working at the informational level, which would need to be explicitly included if the writer wanted to delimit the informative structure of discourse.

Focus operators in non-culminative scales account for 58% of the total. The focus operator *also* (10.85) is the most frequent additive device. The operators *as well* (2.59) and *too* (0.69) function syntactically and semantically in a similar way, as it is shown in (9) and (10), respectively. Their lower frequency may be diaphasically motivated, especially in the case of *too*, which is more commonly used in informal, oral speech. The scalarity triggered by *even* (7.39) is also a useful device to structure information.

(9) This Lady was intruſted with the chief care of his Highneſs's perſon by his Majeſty, to whom ſhe had likewiſe been nurſe: and though the ſeverity of her temper, *as well as* the duty of her office, would not permit her to uſe any indulgence towards the Prince, he ſhewed no leſs affection than reverence for her. (Birch, 1760: 11 [133 (2369)])

(10) They had no money, no colonies nor allies; and the Spartans were watching them with a great deal of jealousy, lest they should find means to regain the power they had lost. The character of the people, *too*, had, in many ways, altered for the worse. (Sewell, 1857: 247 [153 (6431)])

Generally speaking, the evolution of both focus operators and connectives follows the same trend, as Figure 2 shows: after a high number of occurrences in the first third of the eighteenth century, frequency rates decrease drastically – the growth rate in the central years of the century approaches – 60%. Then, at the turn of the century frequencies start to increase. This increase is slow – growth rates never exceed 20% – when compared with the sudden decrease in the previous century but it is also steady, and at the end of the nineteenth century, frequency figures resemble those of 200 years earlier.

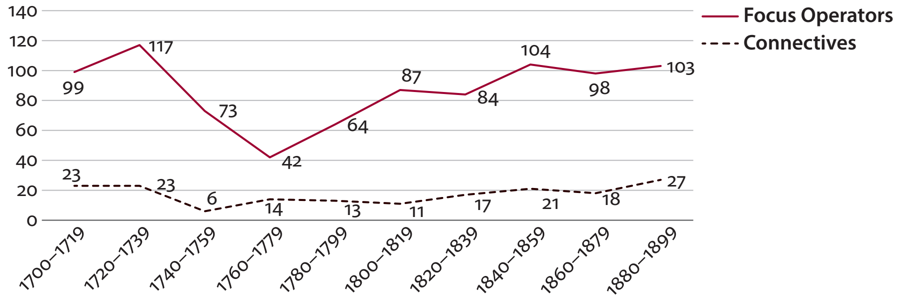


Figure 2. Evolution of additive devices in CHET

This evolution can be explained by the influence of Francis Bacon and the Royal Society and on the new discourse community formed by adherents to the new system. Bacon's fierce criticism of old scholastic models brought a new way of doing – and communicating – science. Bacon expressed the need for a reformation based on the separation between church and science and the adoption of an inductive method based on observation and experiment. Linguistically, he defended the avoidance of abstraction and ambiguity associated with old medieval texts. According to Gotti:

> […] the group of natural philosophers that formed the Royal Society constituted a community of practice and discourse who shared not only methodological aims and research activities but also linguistic conventions and discursive norms. In this way, they characterized themselves not merely as a sociolinguistic group but as a sociorhetorical one (Swales 1990: 24), endowed with its own generic and stylistic features – all contributing to provide a common set of shared practices considered to be basic for the new members by this community. (Gotti, 2013: 282)

This adoption of a whole new set of shared linguistic practices, together with an emphasis on networking and the dissemination of results – the publication of the *Philosophical Transactions* took place in 1666, only four years after the foundation of the Royal Society – contributed to the establishment of a new canon. Parallel to this, the institutionalisation of science in the form of Academies and Scientific Societies during the eighteenth century secured the visibility and prestige of this new standard scientific register. History was not one of the leading disciplines in the Scientific Revolution. Natural philosophers were rather concerned with the explanation of natural phenomena in the field of Physics, Astronomy and Mathematics. If we take into account that the foundation of the Royal Historical Society did not take place until 1868, it might be easier to understand the sudden decrease of additive markers in history texts well into the eighteenth century.[5]

5.1 Gender

According to the data in Table 6, women and men in CHET seem to make a similar use of additive markers. The normalized frequencies for the total of all markers – 22.82 in women's texts and 24.99 in men's – show a very slight difference. The most frequent additive marker employed by both women and man is *also* – 80 occurrences in women's texts and 359 in men's. Women used higher frequencies of some markers (*even, moreover* and *in addition*) and men favoured *also, as well, too* and *besides*. However, the total normalized frequencies of focus markers are very similar – 19.75 in women's vs. 21.98 in men's – and so are those of connectives – 2.94 in women's vs. 3.00 in men's.

---

[5] Other studies (Bello, 2010, 2016b) have proved that even the increase of other linguistic features that would later become standard features in scientific register like nominalisations took place later in those disciplines that were not related to the physical and natural sciences, the first fields of study of the Royal Society.

Table 6. Frequency of additive marker in CHET according to sex of the author (normalised frequencies per 10,000 words in parenthesis)

Additive marker	Women	Men
also	80 (9.81)	359 (11.11)
as well	11 (1.34)	94 (2.91)
besides	8 (0.98)	67 (2.07)
even	66 (8.09)	233 (7.21)
in addition	2 (0.24)	5 (0.15)
too	4 (0.49)	24 (0.74)
moreover	14 (1.71)	25 (0.77)
Total	**186 (22.82)**	**807 (24.99)**

The fact that according to these data there do not seem to be indicative differences in the gender variable is debatable. Women scientists use language in a different way than men. According to Crespo (2015), female scientists resort to persuasion strategies more often than men. They both use different linguistic features to persuade their audience. Crespo based her analysis on the use of different modals and suasive verbs (see also Crespo, 2016). Previous studies (Bello, 2016a) also showed that women tend to use a higher frequency of nominalisations in their texts as a result of their tendency to show involvedness. Perhaps it is interesting to realise that among focus operators, the normalised frequency of the culminative *even* is higher in texts written by women (with a NF of 8.09 in texts written by women compared to the 7.21 of men's texts), whereas men tend to use a higher number of non-culminative devices. The focus operator *even* is context-dependent and, according to Gast & Rzymski (2015), there is an indication of evaluative language in the pragmatic scales they form, so this could be seen as an indication of the tendency to display involvedness that women seem to convey in their texts (Biber *et al.*, 1998; Palander-Colin, 1999; Argamon *et al.*, 2003; Mulac *et al.*, 2013). However, the difference makes this only a tentative (but tempting) conclusion and further research needs to be done to confirm this hypothesis.

5.2 Place of origin

According to the data in Table 7, there are differences in the way American and European authors made use of additive markers in CHET: Not only did American writers use less markers (16.09) and only one type of connectives (*besides*, 1.48) but they also favoured the use of *even* (6.43) over *also* (4.95). In the European countries, the distribution and frequency of the additive markers mirror similar configurations. Irish texts show the highest frequency of focus operators (28.69), whereas the highest frequency of connectives is found in Scottish texts (4.35). The frequencies found in texts written by English authors (22.34 for focus operators and 3.25 for connectives) are also comparable. On the other hand, American texts show the lower frequencies for both focus operators (14.60) and connectives (1.46).

Table 7. Frequency of additive marker in CHET according to place of origin of the author (Normalised frequencies per 10,000 words in parenthesis)

Additive marker	England	Ireland	U.S.	Scotland	Unknown
also	289 (11.91)	76 (12.53)	20 (4.95)	49 (9.70)	5 (4.83)
as well	66 (2.72)	26 (4.28)	9 (2.22)	4 (0.79)	0
besides	51 (2.10)	8 (1.31)	6 (1.48)	10 (1.98)	0
even	171 (7.04)	56 (9.23)	26 (6.43)	36 (7.13)	10 (9.67)
in addition	3 (0.12)	2 (0.32)	0	2 (0.39)	0
moreover	25 (1.03)	2 (0.32)	0	10 (1.98)	2 (1.93)
too	16 (0.65)	4 (0.65)	4 (0.99)	4 (0.79)	0
Total	**621 (25.59)**	**230 (37.92)**	**65 (16.09)**	**115 (22.78)**	**17 (16.45)**

These results may timidly point out that the English scientific register that originated and was developed in England as a consequence of the Scientific Revolution and the institutionalisation of science that followed was adopted in different ways in other English-speaking territories. This may point at the closeness of European authors and the estrangement of American writers, obviously fostered by geographic proximity as well as economic and political (in)dependence.

5.3 Text types

Table 8 shows the frequency of additive markers found in specialised and non-specialised texts. Nfs show that the number of additive markers is considerably higher in non-specialised texts. The nf of all focus operators is 84.22 (677 tokens) in non-specialised texts versus 5.98 (194 tokens) in specialised works. A similar difference can be found in the total frequencies of connectives – 11.19 (90 tokens) in non-specialised texts versus 0.95 (31 tokens) in specialised works. These differences are consistent with the claim made by McNamara *et al.* (1996) that low knowledge readers benefit from coherence marking whereas high knowledge readers benefit from a more implicit text.

Table 8. Frequency of additive markers in CHET according to text type (normalized frequencies per 10,000 words in parenthesis)

Additive marker	Specialised	Non-specialised
also	85 (2.62)	354 (44.04)
as well	27 (0.83)	78 (9.70)
besides	24 (0.74)	51 (6.34)
even	76 (2.34)	223 (27.74)
in addition	5 (0.15)	2 (0.24)
moreover	2 (0.06)	37 (4.60)
too	6 (0.18)	22 (2.73)
Total	**225**	**767**

6. Conclusions and suggestions for further research

According to corpus data, focus operators are more frequent than connectives (see Table 5). As has been explained, this may be due to the workings of implicit additive relations. Markers, such as focus operators, working at the informational level need to be explicitly expressed so that their semantic and pragmatic content can be decoded by the reader. The evolution of both types of additive markers (see Figure 2) seem to be historically motivated. The fluctuation in the frequency of use, with a sharp decrease in the second half of the eighteenth century and a slow, steady increase during the nineteenth century can be seen as the implantation of certain stylistic features associated with the development of scientific register, which in the case of the discipline of History was not as rapid as in other Exact Sciences. Concerning the gender variable, generally speaking men and women seem to use focus operators and connectives in a similar way in CHET (see Table 6). However, the fact that women made a slightly higher number of culminative scales than men could be seen as an indication that women are more prone to use evaluative language. Similarly, there seems to be a timid confirmation of different language use regarding European and American writers (see Table 7). Finally, data in Table 8 confirm that there are differences in the way additive relations are expressed in different text typologies because whereas texts aimed at non-specialised audiences have higher frequencies of additive markers, specialised texts contain more implicit coherence relations.

There are some limitations to this study that should be taken into account. First, this chapter was only concerned with explicit coherence relations marked by focus operators and connectives. In many cases, basic relations like addition are implicitly encoded in discourse, therefore a comprehensive analysis encompassing both explicit and implicit relations would be needed. It would be also desirable to study further the different use of evaluative language in pragmatic scales used by male and female scientists. This would require not only a higher number of texts written by women but also a more detailed analysis of pragmatic scales and culminative focus operators. The study of coherence relations could also be expanded to cover other types of relations like causality or contrast. Similarly, exploitation of other CC subcorpora would also help obtain a more complete picture of the expression of coherence in late Modern English scientific register.

Acknowledgements

The research reported here has been funded by the Spanish Ministerio de Economía, Industria y Competitividad (MINECO), grant number FFI2016-75599-P. This grant is hereby gratefully acknowledged.

References

Albentosa Hernández, José Ignacio. (1997). La sustantivación en el discurso científico en lengua inglesa. CAUCE. *Revista de Filología y su Didáctica*, 20–21, 329–344.

Albentosa Hernández, José Ignacio & Arsenio Jesús Moya Guijarro. (2000). La reducción del grado de transitividad de la oración en el discurso científico en lengua inglesa. *Revista Española de Lingüística*, 30 (1), 445–468.

Andorno, Cecilia. (2000). *Focalizzatori fra connessione e messa a fuoco. Il punto di vista delle varietà di apprendimento*. Milano: Angeli.

Adorno, Cecilia & Anna Maria De Cesare (2017). Mapping additivity through translation. In Anna Maria De Cesare & Cecilia Adorno (Eds.), *Focus on additivity: Adverbial modifiers in Romance, Germanic and Slavic languages (157–200)*. Amsterdam/Philadelphia: John Benjamins.

Argamon, Shlomo, Moshe Koppel, Jonathan Fine, & Anat Rachel Shimoni. (2003). Gender, genre, and writing style in formal written texts. *Interdisciplinary journal for the study of discourse*, 23/3, 321–346.

Atkinson, Dwight. (1998). *Scientific discourse in sociohistorical context: The philosophical transactions of the Royal Society of London 1675–1975*. Mahwah, NJ: Lawrence Erlbaum Associates.

Banks, David. (2001). Analyse des discours spécialisés: Le domaine anglais. *Revue française de linguistique appliquée*, VI (2), 7–16.

Banks, David. (2003). The evolution of grammatical metaphor in scientific writing. In Anne Marie Simon-Vandenbergen, Miriam Taverniers & Louise Ravelli (Eds.), *Grammatical metaphor: views from functional Linguistics* (127–147). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.236.07ban

Banks, David. (2005a). Emerging scientific discourse in the late seventeenth century. *Functions of Language*, 12 (1), 65–86. https://doi.org/10.1075/fol.12.1.04ban

Banks, David. (2005b). On the historical origins of nominalized process in scientific text. *English for Specific Purposes*, 24 (3), 347–357. https://doi.org/10.1016/j.esp.2004.08.002

Banks, David. (2007). L'évolution de la phrase en anglais scientifique. In David Banks (Ed.), *La coordination et la subordination dans le texte de spécialité* (203–221). Paris: L'Harmattan.

Banks, David. (2008). *The development of scientific writing: linguistic features and historical context*. London: Equinox.

Barber, Charles. (1993). *The English language: A historical introduction*. Cambridge: Cambridge University Press.

Beal, Joan C. (2012). Late modern English in its historical context. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy 'playne and simple': The writing of science between 1700 and 1900* (1–14). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bello, Iria. (2010). A diachronic study of nominalizations in Astronomy and Philosophy texts. Paper delivered at the *6th International Contrastive Linguistics Conference*, Freie Universität Berlin, Berlin, 30/09-02/10/2010.

Bello, Iria. (2016a). Nominalizations and female scientific writing in the late Modern Period. *Revista canaria de estudios ingleses*, 72, 35–52.

Bello, Iria. (2016b). Nominalization use in late Modern English scientific texts written by women. Evidence from the *Coruña Corpus*. Paper delivered at the *Sociolinguistics Symposium 21*, 15th–18th June, 2016, Universidad de Murcia (Spain).

Biber, Douglas, Susan Conrad, & Randi Reppen. (1998). *Corpus Linguistics investigating language structure and use*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511804489

Blakemore, Diane. (1987). *Semantic constraints on relevance*. Oxford: Blackwell.

Bloom, Lois, Margaret Lahey, Lois Hood, Karin Lifter, & Kathleen Fiess. (1980). Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. *Journal of child language*, 7, 235–261. https://doi.org/10.1017/S0305000900002610

Camiña Rioboo, Gonzalo. (2012). Accounting for observations of the heavens in the 18th century: new nouns to explain old phenomena. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy 'playne and simple': The writing of science between 1700 and 1900* (93–121). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Camiña Rioboo, Gonzalo. (2013). Word formation in the scientific register of early Modern English: A corpus-based approach. (Unpublished Doctoral dissertation). Available from Repositorio da Universidade da Coruña.

Camiña Rioboo, Gonzalo, María José Esteve, & Inés Lareo. (2012). A study of nouns and their provenance in the astronomy section of the *Coruña Corpus of Scientific English Writing*: Some preliminary considerations. In José Luis Cifuentes Honrubia, Adelina Gómez Gonzalez-Jover, Antonio Lillo Buades, José Mateo Martínez & Francisco Yus Ramos (Eds.), *Los caminos de la lengua. Estudios en homenaje a Enrique Alcaraz Varó* (537–547). Alicante: Universidad de Alicante.

De Cesare, Anna Maria. (2017). Introduction: on 'Addititvity' as a Multidisciplinary Research Field. In Anna Maria De Cesare and Cecilia Andorno (Eds.), *Focus on additivity. adverbial modifiers in Romance, Germanic and Slavic languages* (1–22). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.278.00dec

Crespo, Begoña. (2015). Women writing science in the Eighteenth Century: a preliminary approach to their language in use. *ANGLICA. An international journal of English studies*, 24/2, 103–128.

Crespo, Begoña. (2016). On writing science in the Age of Reason. *Revista Canaria de Estudios Ingleses*, 72, 53–78.

van Dijk, Teun & Walter Kintsch. (1983). *Strategies of discourse comprehension*. New York: Academic.

Domínguez García, Noemí. (2007). *Conectores discursivos en textos argumentativo breves*. Madrid: Arco/Libros.

Eisenberg, Ann R. (1980). A syntactic, semantic and pragmatic analysis of conjunction. *Stanford papers and reports on child language development*, 19, 70–78.

Escandell-Vidal, Victoria & Manuel Leonetti. (2011). On the rigidity of procedural meaning. In Victoria Escandell-Vidal, Manuel Leonetti, & Aoife Ahern (Eds.), *Procedural meaning: problems and perspectives* (81–103). Bingley: Emerald Group Publishing.

Fauconnier, Gilles. (1975). Polarity and the scale principle. In *Papers of the eleventh regional meeting of the Chicago Linguistic Society* (188–199). Chicago, IL: Chicago Linguistic Society.

Gast, Volker. (2017). The scalar operator even and its German equivalents. In Anna Maria De Cesare & Cecilia Andorno (Eds.), *Focus on additivity. Adverbial modifiers in Romance, Germanic and Slavic languages* (201–234). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.278.07gas

Gast, Volker & Jan van der Auwera. (2011). Scalar additive operators in the languages of Europe. *Language*, 87/1, 2–54. https://doi.org/10.1353/lan.2011.0008

Gast, Volker, & Christoph Rzymski. (2015). Towards a corpus-based analysis of evaluative scales associated with *even*. *Linguistik Online*, 71/2.

Gotti, Maurizio. (2011). The development of specialized discourse in the Philosophical Transactions. In Irma Taivitsainen & Paivi Pahta (Eds.), *Medical writing in early Modern English* (204–220). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.012

Gotti, Maurizio. (2012). Variation in academic texts. In Maurizio Gotti (Ed.), *Academic identity traits: A corpus-based investigation* (23–42). Bern: Peter Lang. https://doi.org/10.3726/978-3-0351-0343-4/3

Gotti, Maurizio. (2013). The formation of the Royal Society as a community of practice and discourse. In Joanna Kopaczyk & Andreas Jucker (Eds.), *Communities of practice in the history of English* (269–285). Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.235.16got

Greene, John C. (1954). Some aspects of American astronomy 1750–1815. *Isis*, 45 (4), 339–358. https://doi.org/10.1086/348356

Halliday, Michael A. K. (2004). On the language of physical science. In John Webster (Ed.), *The language of physical science* (162–178). London: Continuum.

Halliday, Michael A. K. & Ruqaiya Hassan. (1976). *Cohesion in English*. Harlow: Longman.

Halliday, Michael A. K. & James Martin. (1993). *Writing science*. London: The Falmer Press.

Huddleston, Rodney & Geoffrey Pullum. (2002). *The Cambridge grammar of English language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316423530

Jacobs, Joachim. (1983). *Fokus und Skalen. Zur Syntax uns Semantik der Gradpartikeln im Deutschen*. Tübingen: Niemeyer.

Katz, Evelyn W. & Sandor D. Brent. (1968). Understanding connectives. *Journal of verbal learning and verbal behavior*, 7, 501–509. https://doi.org/10.1016/S0022-5371(68)80040-4

Kay, Paul. (1990). Even. *Linguistics and Philosophy*, 13, 59–111. https://doi.org/10.1007/BF00630517

Kintsch, Walter, & Teun van Dijk. (1978). Toward a model of text comprehension and production. *Psychological review*, 85, 363–394. https://doi.org/10.1037/0033-295X.85.5.363

König, Ekkehard. (1991). *The Meaning of focus particles: a comparative perspective*. London: Routledge. https://doi.org/10.4324/9780203212288

Lareo, Inés. (2011). Using complex predicates in the writings on astronomy in nineteenth-century English. Exploitation of the corpus of English texts on astronomy. *Revista de lenguas para fines específicos*, 17, 229–252.

Levinson, Stephen. (2000). *Presumptive meanings – The theory of generalized conversational implicature*. Cambridge (MA): MIT Press.

Loureda, Óscar, Adriana Cruz, Martha Rudka, Laura Nadal, Inés Recio, & Margarita Borreguero Zuloaga. (2015). Focus particles in information processing: an experimental study on pragmatic scales with Spanish *incluso*. In Anna Maria De Cesare, & Cecilia Andorno (Eds.), *Focus particles in the Romance and Germanic languages. Experimental and corpus-based approaches. Linguistic Online*, 71/2.

McNamara, Danielle S., Eileen Kintsch, Nancy Butler Songer, & Walter Kintsch. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and instruction*, 14/1, 1–43. https://doi.org/10.1207/s1532690xci1401_1

Moskowich, Isabel. (2012). *CETA* as a tool for the study of modern astronomy in English. In Isabel Moskowich & Begoña Crespo (Eds.), *Astronomy 'playne and simple': The writing of science between 1700 and 1900* (35–57). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Moskowich, Isabel & Begoña Crespo. (2012). *Astronomy 'playne and simple': The writing of science between 1700 and 1900*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.173

Moskowich, Isabel & Leida Maria Monaco. (2016). Linking ideas in women's writing: Evidence from the *Coruña Corpus*. *Revista de lingüística y lenguas aplicadas*, 11, 35–49. https://doi.org/10.4995/rlyla.2016.4660

Mulac, Anthony, Giles Howard, James Bradac, & Nicholas Palomares. (2013). The gender-linked language effect: an empirical test of a general process model. *Language Sciences*, 38, 22–31. https://doi.org/10.1016/j.langsci.2012.12.004

Nadal, Laura, Inés Recio, Martha Rudka, & Óscar Loureda. (2017). Processing additivity in Spanish: *incluso* vs *además*. In Anna Maria De Cesare & Cecilia Andorno, (Eds.), *Focus on additivity. Adverbial modifiers in Romance, Germanic and Slavic languages* (136–156). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.278.05nad

Palander-Collin, Minna. (1999). Male and female styles in 17th century correspondence. *Language variation and change*, 11, 123–141. https://doi.org/10.1017/S0954394599112018

Pander Maat, Henk & Ted Sanders. (2006). Connectives in text. In Keith Brown (Ed.), *Encyclopedia of language and linguistics* (33–41). Amsterdam: Elsevier (Volume 3). https://doi.org/10.1016/B0-08-044854-2/00500-9

Piaget, Jean. (1969). *Judgement and reasoning in the child*. London: Routledge and Paul Kegan.

Portolés, José. (2007). Escalas aditivas. Pruebas del español. *Spanish in Context*, 4/2: 135–157. https://doi.org/10.1075/sic.4.2.02por

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. (1985). *A Comprehensive grammar of the English language*. London: Longman.

Sanders, Ted. (2005). Coherence, causality and cognitive complexity in discourse. *Proceedings/ Actes SEM-05, First international symposium on the exploration and modelling of meaning*, 105–114.

Sanders, Ted & Henk Pander Maat. (2006). Cohesion and coherence: linguistic approaches. In Keith Brown (Ed.), *Encyclopedia of language and Linguistics* (591–595). London: Elsevier (Volume 2). https://doi.org/10.1016/B0-08-044854-2/00497-1

Sanders, Ted, Wilbert Spooren & Leo Noordman. (1992). Toward a taxonomy of coherence relations. *Discourse processes*, 15, 1–35. https://doi.org/10.1080/01638539209544800

Sanders, Ted, Wilbert Spooren & Leo Noordman. (1993). Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, 4, 93–134. https://doi.org/10.1515/cogl.1993.4.2.93

Schwenter, Scott. (1999). *Pragmatics of conditional marking: implicature, scalarity and exclusivity*. New York: Garland.

Sperber, Dan & Deirdre Wilson. (1995). *Relevance*. Oxford: Blackwell.

Sperber, Dan & Deirdre Wilson. (2002). Pragmatics, modularity and mind-reading. *Mind and Language*, 17, 3–23. https://doi.org/10.1111/1468-0017.00186

Spooren, Wilbert & Ted Sanders. (2008). The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, 40, 2003–2026. https://doi.org/10.1016/j.pragma.2008.04.021

Taavitsainen, Irma. (2001). Changing conventions of writing: The dynamics of genres, text types, and text traditions. *European journal of English studies*, 5 (2), 139–150. https://doi.org/10.1076/ejes.5.2.139.7309

Taavitsainen, Irma & Paivï Pahta. (1997). The *Corpus of Early English Medical Writing*: linguistic variation and prescriptive collocations in scholastic style. In Terttu Nevalainen & Leena Kahlas-Tarkka (Eds.), *To explain the present: studies in changing English language in honour of Matti Rissanen* (209–225). Helsinki: *Société Néophilologique*.

Vázquez Orta, Ignacio. (2006). A corpus-based approach to the distribution of nominalization in academic discourse. In María José Luzón, Silvia Murillo & Ana María Hornero (Eds.), *Corpus linguistics: applications for the study of English* (399–416). Bern: Peter Lang.

Wilson, Deirdre & Dan Sperber. (2002). Relevance theory. *Working papers in Linguistics*, 14, 249–287.

Index

A

academic writing 149, 193, 206, 227, 231, 236–237

adverb 65, 77, 87, 95, 97, 107, 115, 120, 123, 129, 131, 133, 150, 158–159, 167–170, 172, 175, 177–182, 237–238, 242–245, 248, 250–252, 255, 261
  stance adverbs *see* stance

alliteration 107–108, 124

Appraisal Theory 25, 36

assertion 140, 168, 193
  assertive 168–170, 176–178, 182
  assertiveness 131–132, 168–169, 177, 208

attitude devices 235–239, 242–245, 247–248, 250, 252–255

authorial presence 185–186, 193, 199, 209, 212

authorial voice 215–217, 221–223, 228, 231

B

balance 2, 45, 155

binomials 103–127

C

centrality 238, 253

certainty 65, 158, 167–169, 172, 176–182

community of practice 194, 201, 209–210, 268

conditional structures 130, 133, 140
  conditional subordinators 130, 132, 138–139
  conditionals 129–145
  inversion conditionals 129–130, 133–139, 141–145

conviction 170, 180, 215

Coruña Corpus of English Scientific Writing (CC) vii, 13, 17, 43–46, 52, 60, 110, 123, 130, 135, 149, 153, 168–169, 177, 186, 188, 190, 217–218, 227–228, 231, 235, 240, 264, 272
  Corpus of English Life Sciences Texts (CELiST) 86, 89, 91–93, 97–98, 130, 135–137, 142–144
  Corpus of English Philosophy Texts (CEPhiT) vii, 12, 44, 51–53, 86, 89–91, 93, 97–98, 130, 135–137, 142–144
  Corpus of English Texts on Astronomy (CETA) vii, 44, 51–52, 86, 89, 91, 93, 97–98, 110–112, 115, 117–121, 123, 130, 135–137, 142–144, 217–220, 223–227, 229–230

Coruña Corpus Tool (CCT) vii, 45, 88, 110, 139, 153, 172, 190, 226, 240, 242–243, 264

D

deontic modality *see* modality

dimension of variation 85–86

dimension score 89

discourse 22, 25, 36, 44, 46, 77, 84, 86, 90–92, 97–99, 129–130, 134, 137–138, 140–141, 144, 149, 160, 164, 167–169, 176, 179, 185–189, 191–192, 195–197, 200, 202–203, 206, 208, 210–211, 215–216, 228, 237–238, 259–261, 263, 267–268, 272

discourse strategies 129

discursive practices 201, 211

dynamic modality *see* modality

E

Empiricism 129, 187

engagement 6, 25, 198, 210, 215–217, 220, 224, 228–230, 256

epistemic modality *see* modality

evaluate 153, 228, 250
  evaluation 25, 30, 150–151, 155, 157, 186, 236–238, 244, 249, 255
  evaluative 131, 238, 247, 253, 269, 271–272

Evans Collection of Early American Imprints 22

evidential 150, 155

F

factor analysis 85, 187

factor loading 85

female authors 53–54, 60, 176–177, 185, 188, 210–211

female writing 185, 192

first person pronouns 193, 200–201, 203–204, 215–220, 222, 226, 229, 231

G

Gaelic 21

genre 1–2, 8, 10, 14, 23, 36, 41, 44–45, 52–53, 58–60, 63, 79, 84, 89–90, 93–94, 96–99, 101, 138–139, 143, 153, 156, 169, 172–180, 186–187, 197, 203, 217–218, 225, 227, 231, 235–238, 240, 243–244, 246–247, 249–256, 264–265

geographical provenance 2

H

historical genres 235

historiography 5–11, 13–14, 17, 42–43, 83, 118, 201–202, 204, 206, 239
History of English 103–104, 124, 152
history writing 5, 10, 90, 97, 118, 167, 182, 185, 188, 195, 199, 202, 210, 240, 254

I
inferentiality 152, 155, 159, 164
interaction 25, 34, 131, 186–187, 194, 208, 216, 224
interpreters 25, 33, 151
inversion 129–130 133–139, 141–145, 152, 262
  inversion conditionals *see* conditional structures

K
Kullback-Leibler Divergence 61

L
language policies 21
late Modern English 1, 5–6, 45, 52, 57, 59–60, 64, 83, 86, 99, 103–104, 110, 124, 167, 177, 182, 192, 199, 240, 272

M
metadiscourse 235, 237
missionaries 10, 25, 31, 33–35
modality 30, 64, 132, 138, 140, 149–152, 155–156, 163–164, 186
  deontic modality 150–151
  dynamic modality 151–152, 163–164
  epistemic modality 150, 152, 155
  modal verbs 129, 145, 149–150, 152–156, 161, 164, 224
Multidimensional Analysis 85, 187
multinomials 105–106, 116, 121

N
narration 6, 24, 28, 31, 36, 65, 67, 83–84, 86, 89, 92, 96–98, 129, 228
  narrative 1, 7, 23, 42, 57, 60, 65, 78, 83–86, 89–98, 145, 206
  narrativity 2, 84, 89
native Americans 22–24, 26, 29, 36
non-scientific discourse *see* scientific discourse

O
objectivity 1–2, 7, 36, 149, 185, 187, 198, 211
of-genitive 69, 78
operator 133–139, 141–142, 145, 152, 261–262, 264, 266–267, 269–272

P
paradigmatic change 130, 138, 145
pattern 2, 13, 36, 46, 57–58, 61–79, 85–86, 93, 106–107, 113–114, 123, 130, 192, 201, 205, 207, 209–211, 235, 238, 240
persuasion 86, 182, 194, 269
  persuade 129, 131, 168, 209, 269
  persuasive 25, 86, 178, 182, 186
Philosophy 5–12, 14, 17, 42, 44–45, 59, 84–86, 89–93, 97–98, 122, 130, 135–136, 138, 142, 145, 225, 239
polarity 152, 236, 238, 253, 260
productivity 58–59, 62–64, 67, 69, 74–75, 77–79, 138

R
reader 6, 24–26, 30, 36, 130–131, 153, 155, 157–158, 163, 167–168, 179–182, 185–187, 192–194, 198, 201, 208, 215–216, 220–221, 223–224, 227–231, 235–237, 239, 243, 247, 250, 253–256, 260–263, 267, 271
  readership 17, 45, 84, 185, 194, 202, 209–210
register variation 84
relative entropy 58, 61–62, 64, 74, 78
reliability 30–31, 155, 168, 185
representativeness 43, 45, 153
Royal Society Corpus (RSC) 58–62, 64–66, 68–75, 77–78

S
Scholasticism 42, 129–131, 138, 145
  Scholastic 11, 129, 138, 204, 268
scientific community 201, 227, 266
scientific discourse 46, 84, 129–130, 138, 186–187, 210, 259
  non-scientific discourse 189
  scientific texts 57–58, 60–61, 64–67, 70–71, 77–78, 108, 149, 152, 199, 203, 210
  scientific writing 1, 18, 45, 57, 130–131, 138, 149, 177, 185, 187–188, 192, 201, 208, 210–211, 215, 260
semantic relations 109
sex 2, 15, 52–53, 55, 60, 139, 168–169, 172, 176–177, 182, 269
stance 22, 25, 29, 59, 140, 149, 160, 215–217, 220, 224–225, 227, 229–231, 236, 250, 256
  silent stance 231
  stance adverbs 158, 238
subjectivity 186–187
surprisal 58, 62–63, 67–69, 71–78

T
to-infinitive 65, 66, 74, 76–78
truth 38, 131, 140, 150, 155, 159, 164, 167–170, 179–181, 193, 208
truthfulness 167–168

W
women 8–9, 14–16, 23, 36, 51, 53, 55, 60, 109, 116, 119, 122, 143, 168, 177–178, 182, 185–186, 188–189, 192, 199, 201, 206, 208, 210–211, 264, 268–269, 272
  women's discourse 208
writer's intention 250

Z
z-scores 89

This volume focuses on the relationship and interaction of language and science between 1700 and 1900. It pays particular attention to English History writing in late Modern English as compiled in the *Corpus of History English Texts* (CHET), a newly released sub-corpus of the *Coruña Corpus of English Scientific Writing*. The chapters cover methodological issues, the period and the status of the discipline itself, as well as pilot studies for the description of scientific discourse using CHET. They embrace topics in several linguistic fields: discourse analysis, syntax, semantics, morpho-syntax. The studies take into account extralinguistic parameters of texts, such as year of publication, sex of the author, geographical provenance of authors and the communicative formats/genres to which the text sample belongs. In the particular case of CHET, the collected samples can be grouped in eight different categories and such categories, as well as the above-mentioned metadata information, can be used to search the corpus.

The book is of interest for scholars specialised in corpus linguistics and historical linguistics, as well as linguists in general. The metadata information used for analysis can also be of interest for historians and historians of science in particular.

The *Corpus of History English Texts* (CHET), accompanied by the *Coruña Corpus Tool* (CCT), purpose-designed software by IrLab, is accessible online at the Repositorio Universidade Coruña at http://hdl.handle.net/2183/21849

Cover image: Tomb of Sir John Moore in the Jardín de San Carlos, A Coruña

© 2015, Merce, CC BY-SA

https://www.flickr.com/photos/merceblanco/20668333459

ISBN 978 90 272 0424 0

9 789027 204240

John Benjamins Publishing Company