

Exceptional Women and Classical Borrowings:

A Case Study of Margaret Cavendish's Use of Latin and Greek Loanwords

Rasmus Knuuttila

Bachelor's Thesis

English, Bachelor of Arts, Department of English

School of Languages and Translation

Faculty of Humanities

University of Turku

March 2025

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Bachelor's Thesis

English, Bachelor of Arts, Department of English

Rasmus Knuutila

Exceptional Women and Classical Borrowings: A Case Study of Margaret Cavendish's Use of Latin and Greek Loanwords.

Number of pages: 32 pages, 2 appendices

This study examines Margaret Cavendish's use of Latin and Greek loanwords in five of her texts, published between 1655 and 1671. The examined texts cover a number of different genres, including, among others, a play, two letter collections and an essay collection. Previous historical sociolinguistic research into women's lexical usage in the Early Modern period is very scarce. The aim of this study is to contribute to filling this gap.

The study is quantitative in nature. It examines the frequencies and the relative frequencies of 26 lexemes in Cavendish's texts. 16 lexemes are of Latin origin and 10 lexemes are of Greek origin. The results from Cavendish's texts are compared with the more male-dominated *Early English Books Online* (EEBO) corpus.

Overall, the results from Cavendish's texts follow the same trends that can be observed in the EEBO-corpus, although some trends appear more clearly than others. Of the examined Latin loanwords, seven occur in Cavendish's texts. The only one of the examined Greek loanwords to occur is *alphabet*.

The results of this study tentatively suggest that previous research has too categorically assumed that women were behind men in borrowing in the Early Modern English period. However, much more future research is required to confirm the conclusions drawn in this thesis. On a more methodological level, this study concludes that the different genres and types of texts selected for this kind of research can also have a marked effect on the occurrences of individual lexemes in the dataset. This is something that future studies should also take into account.

Key words: Loanwords, Historical Sociolinguistics.

Table of contents

1	Introduction	6
2	Background and Theory	7
3	Methodology	10
3.1	Selection of Data	10
3.2	Selection of Lexemes	12
3.3	Examination of Texts	13
4	Results	15
4.1	The studied texts – overall remarks	15
4.2	Individual texts	15
4.2.1	The World's Olio (1655)	16
4.2.2	Playes (1662)	16
4.2.3	CCXI Sociable Letters (1664)	17
4.2.4	Philosophical Letters (1664)	17
4.2.5	Nature's Picture (1671)	18
4.3	EEBO-Corpus	19
5	Discussion	23
6	Conclusion	26
	List of References	27
	Primary sources	27
	Secondary sources	27
	Appendices	31
	Appendix 1 Irregular forms of the studied lexemes	31
	Appendix 2 Total frequencies and relative frequencies of the studied lexemes	31

TABLE 1 THE DATASET	11
TABLE 2 STUDIED LEXEMES WITH LANGUAGES OF ORIGIN	13
TABLE 3 FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN THE WORLD'S OLIO (1655)	16
TABLE 4 FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN PLAYES (1662)	16
TABLE 5 FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN CCXI SOCIABLE LETTERS (1664A)	17
TABLE 6 FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN PHILOSOPHICAL LETTERS (1664B)	17
TABLE 7 FREQUENCIES AND RELATIVE FREQUENCIES OF GREEK LOANWORDS IN PHILOSOPHICAL LETTERS (1664B)	17
TABLE 8 FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN NATURE'S PICTURE (1671)	18
TABLE 9 FREQUENCIES AND RELATIVE FREQUENCIES OF GREEK LOANWORDS IN NATURE'S PICTURE (1671)	18
TABLE 10 TOTAL FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN THE EEBO CORPUS	19
TABLE 11 TOTAL FREQUENCIES AND RELATIVE FREQUENCIES OF GREEK LOANWORDS IN THE EEBO CORPUS	21
TABLE 12 IRREGULAR FORMS OF THE STUDIED LEXEMES	31
TABLE 13 TOTAL FREQUENCIES AND RELATIVE FREQUENCIES OF LATIN LOANWORDS IN CAVENDISH'S TEXTS	31
TABLE 14 TOTAL FREQUENCIES AND RELATIVE FREQUENCIES OF GREEK LOANWORDS IN CAVENDISH'S TEXTS	32

FIGURE 1 RELATIVE FREQUENCIES OF THE STUDIED LATIN LOANWORDS IN THE EEBO-CORPUS	20
FIGURE 2 RELATIVE FREQUENCIES OF THE STUDIED GREEK LOANWORDS IN THE EEBO-CORPUS	22

1 Introduction

English gained ca. 30 000 new lexemes over the Early Modern period (Davies [2005] 2013, 29). Görlach attributes this increase to speakers' need to discuss new concepts and the desire to make English more rhetorically viable ([1991] 2012, 137–138). According to Van Gelderen, ca. 50% of the new words that came to the language between 1500 and 1660 were *loanwords* (2016, 179). Loanwords can be defined as words that a language integrates into its own vocabulary from another language in a process known as *borrowing* (Campbell 2013, 56). The highest number of loanwords was borrowed from Latin (Kastovsky 2006, 256). With the revived interest in classical cultures, Greek loanwords also became more common, although they were much more infrequent than Latin loanwords (Nevalainen 2006, 53–54) and mostly belonged to specialist domains of vocabulary (Nielsen [2005] 2012, 178). This study focuses on these borrowings of Greek and Latin origin (henceforth *classical loanwords*). Many of these borrowings were not easily understandable for those who had not received a classical education (Nevalainen 2006, 39). This led to many of these people feeling linguistically insecure, leading to *hard-word dictionaries* being published (ibid.). They only included *hard words*, i.e. recently borrowed words that had not yet fully established themselves in the language (Dixon 2018, 48).

Women were especially affected by the above-described linguistic insecurity (Nevalainen 2006, 39), as they often did not have access to classical education and, thus, lacked the knowledge of Latin (Nevalainen 2006, 136). There were, nonetheless, some exceptionally well-educated women (Nurmi 2012, 53). However, their lexical usage has – as far as the author of this paper is aware – mostly been left unstudied by historical sociolinguists. This study aims to contribute to filling this gap. More specifically, this study aims to explore how the lexical usage of these exceptional women differs from their male peers. This is achieved through examining how the exceptional 17th century author and Duchess of Newcastle Margaret Cavendish uses classical loanwords in five of her works. She was greatly interested in science (Quinsee 2001, 2). This, in addition to her life being well-documented (Quinsee 2001), makes her a great subject for this study. Although there has been some scholarly interest in her works (Quinsee 2001, 3–4), her lexical usage has been studied very little. The study is done by comparing the frequencies of selected lexemes in her texts with the *Early English Books Online* corpus. Next, some more theoretical points will be discussed.

2 Background and Theory

As discussed above, the English lexicon grew substantially over the course of the Early Modern English period (Davies [2005] 2013, 29), although there is some debate about when the growth peaked, see (Lancashire 2012, 641) and (Görlach [1991] 2012, 137). Many loanwords were also infrequent in usage (Durkin 2014, 308). Lists of loanwords are provided by Kastovsky (2006, 258–259) and Serjeantson ([1935] 1968, 264–270). Lancashire characterises the English lexicon as a *split lexicon* with, on the one hand, an older common core and, on the other hand, newer borrowed vocabulary (2012, 638). He also highlights the role of specialised vocabulary in the growth of the lexicon (Lancashire 2012, 642–643). Although Latin lost some of its prestige after the Restoration, specialist vocabulary continued to be borrowed from Latin (Nevalainen 2006, 54). Many Latin loanwords were bookish in nature (Nevalainen 2006, 53). Barber discusses contemporary attitudes to language extensively, including the Inkhorn Controversy (1997, 43–102). For more theoretical discussion on borrowing, see Kay and Allan (2016) and Campbell (2013, 56–90).

When discussing gender, it is important to, firstly, state that, while it is a social construct, this study uses an individual's biological sex to determine their gender, as Nevalainen and Raumolin-Brunberg also have done (2017, 110). Secondly, it is important to remember that the significance of different social variables depends on the historical period (Kiełkiewicz-Janowiak 2012, 307). Thus, a good understanding of the examined society is also needed (Nevala 2016, 286). Nonetheless, according to the *uniformitarian principle*, the causes for linguistic change remain the same throughout all periods (Trousdale 2010, 313). Other social variables besides gender are also important to account for (Kiełkiewicz-Janowiak 2012, 313). Kiełkiewicz-Janowiak especially focuses on the interrelationship between an individual's gender and their social class (2012, 318–320). Often, differences between women can be greater than those between men and women (Kiełkiewicz-Janowiak 2012, 313). Gender stereotypes also have to be taken into account, as they can not only have an effect on an individual's usage but also on how a researcher interprets their data, for example small differences can be interpreted as marked contrasts (Kiełkiewicz-Janowiak 2012, 317–318). The Early Modern period is also the first period in which women's language is explicitly criticised (Trousdale 2010, 85).

On a more concrete level, *change from above* and *change from below* are important concepts. Whereas change from above refers to developments in the language that the speakers are

aware of, such as borrowing, change from below refers to changes that the speakers themselves are not aware of (Raumolin-Brunberg 2012, 718). In the Early Modern period, women were only leaders in change from below (ibid.). Nevalainen and Raumolin-Brunberg provide an extensive account of previous studies into gender's effects on language usage in Early Modern England (2017, 116–118). They, for example, mention that women were behind men in borrowing (2017, 117). They also discuss the effects of gender on many grammatical developments of the period (2017, 118–131).

In addition to more general patterns and theories of how gender affects an individual's language use, it is also important to understand how an individual's *idiolect*, i.e. their unique way of using language (Campbell 2013, 191), can change. A useful framework for studying the language use of an individual is *micro-sociolinguistics* and within it the *micro-model* (Anipa 2012, 176). This model places the focus on the individual, as understanding the language use of individuals is essential for understanding larger-scale phenomena (ibid.). According to this model, the behaviour of an individual is greatly affected by their previous experiences (ibid.). These also influence how individuals experience the world around them (ibid.). This model also emphasises the lack of regularity in an individual's language use (Anipa 2012, 178). As Anipa explains, the written data used in historical sociolinguistic research has the benefit of providing the model with a finite amount of data on an individual's language use (2012, 177). However, written data cannot represent the entirety of an individual's language use (ibid.).

To gain a proper understanding of education in Early Modern England, it is important to conceptualise education as more than just attending school (Charlton 2002, 4). For a detailed description of the education system in Early Modern England, see Brink (2001, 99–103). In terms of literacy, Charlton estimates that most women belonging to the middle and upper classes were literate (2002, 4). Whitehead, however, claims that 90% of women were illiterate (1999, x). For this period, the most common metric for measuring literacy is an individual's ability to write their signature (Charlton and Spufford 2003, 27). However, there are many issues with this method, including, for instance, that women are underrepresented in the data (Brink 2001, 97). In the period, upper-class girls could receive tutoring (Nurmi 2012, 53) or they could be sent to live with another family, to be educated there (Charlton 2002, 8). Women were mostly denied access to Latin, which prevented them from advancing their education (Eskin 1999, 110). There were exceptions amongst women of the royalty and the nobility (Eskin 1999, 120).

To conclude this section, the author Margaret Cavendish (1623–1673) is briefly discussed. Overall, she published 13 books over her lifetime and has been linked to the literary tradition of the Early Modern female authors who came before and after her (Fitzmaurice 2011, n.p.). Fitzmaurice also adds that, when Cavendish had her writing published, it enabled future female authors to more readily publish their works (2011, n.p.). At the time, it was exceedingly rare for women to have their book published or have financial motives behind their writing (Ezell 2002, 79–80). Fitzmaurice also remarks upon the diversity of genres in which Cavendish wrote (2011, n.p.). These include drama, scientific texts, letters as well as prose fiction (*ibid.*). She even published an autobiography (Quinsee 2001, 1). Cavendish was also greatly interested in science and corresponded with many notable philosophers (Quinsee 2001, 2).

3 Methodology

This section is organised into subsections outlining the selection of the data and the lexemes as well as the data analysis process.

3.1 Selection of Data

Margaret Cavendish was chosen to be the subject of study for this paper for multiple different reasons. Firstly, her life is very well documented. This is important for a case study in the field of historical sociolinguistics, because, as Anipa states, an individual's behaviour is always affected by their previous experiences (2012, 176). Because so much is known about Cavendish, it is possible to take into account how her previous life experiences may have influenced her use of classical loanwords. Her exceptionality as a character also adds to her interest as a subject of historical sociolinguistics study. Her exceptionality can better highlight how she differs from other women, because often there are more differences between women than between men and women (Kiełkiewicz-Janowiak 2012, 313). Cavendish was also significant in paving the way for the Early Modern female authors that came after her (Fitzmaurice 2011, n.p.), which also adds to her interest. In addition, her interest in science (Quinsee 2001, 2) means that she is more likely to write texts that contain the more specialised lexemes that this study examines. Equally importantly, many of her works are available in the *Early English Books Online* database (henceforth EEBO).

As many of her works are available in the EEBO-database, it was important to narrow down the selection to make sure that this study is manageable within the scope of a BA-thesis. Five of her texts were selected. This number should provide enough data on her writing while still being manageable. In the selection process, texts not available as transcribed TCP-versions were first excluded. TCP-versions are needed, because they make it possible to easily search for lexemes. After this initial delimitation, there were still many texts left. To provide this study with a sufficiently large and representative dataset, the five longest texts – based on the word counts provided on EEBO – were selected. Table 1 provides a summary of the dataset.

Table 1 The Dataset

A table of the studied texts and their years of printing. Book titles are given here in a shortened form. For full forms, see the *List of References*. Word counts are those that are given on EEBO. The total word count is 810318

Text	The Year of Printing	Word count
The World's Olio	1655	97825
Playes	1662	287689
CCXI Sociable Letters	1664	107085
Philosophical Letters	1664	154829
Nature's Pictures	1671	162890

The World's Olio (Cavendish 1655) is a collection of essays (Fitzmaurice 2011, n.p.). *CCXI Sociable Letters* (Cavendish 1664a) and *Philosophical Letters* (Cavendish 1664b) are letter collections. *Nature's Pictures* (Cavendish 1671) is a collection featuring narratives in verse form as well as short stories (Fitzmaurice 2011, n.p.).

There are several reasons for using the EEBO-corpus as a point of comparison. First of all, the corpus includes 755 million word-form tokens (English Corpora n.d.), which is in itself a great advantage. The interface facilitates this study by enabling sorting of occurrences by decade and lemmas while still providing access to the language of Cavendish's male peers. As women wrote under two per cent of the texts published over the Early Modern period (Nevalainen 2006, 136), a corpus consisting of only printed texts (English Corpora n.d.) is bound not to include many female authors. In addition, having one's book printed in the Early Modern period was expensive (Quinsee 2001, 2), meaning that the EEBO-corpus will automatically overrepresent the usage of the upper classes. This eliminates a significant amount of social variation, providing this study with access to individuals who could be considered Cavendish's peers. It is also notable that, in this period, most books were printed in London (Nevalainen 2006, 135), which also largely eliminates the effects of regional variation. One potential issue with the EEBO-corpus is that it may not accurately represent the usage of Cavendish's male peers, i.e. affluent individuals active in London social circles. This risk is, however, greatly reduced by the large amount of data. A second more practical issue with this corpus is the accuracy of the *group by lemmas* function which is discussed in section 3.3. *Examination of Texts*.

3.2 Selection of Lexemes

The lexemes were selected from lists provided in Kastovsky (2006, 257–259). Kastovsky discusses Latin and Greek loanwords separately (*ibid.*). His list of Latin loanwords is divided by the time and type of borrowing into four subgroups (Kastovsky 2006, 257–258). The subgroups consist of loanwords borrowed in the 15th and 16th centuries, those borrowed in the 17th century, those borrowed in the 18th century, and loan translations. The first subgroup was selected for this study, as it enables this study to focus on lexemes that may not have been in common usage at the time the studied texts were written. If the other subgroups had been included, this study would have had to account for more general changes in the lexicon. Kastovsky categorises his list of Greek loanwords based on whether they were mediated through Latin or whether they were directly borrowed from Greek (Kastovsky 2006, 259). Both categories were included in this study.

As the selected lists contain 55 lexemes in total, not all of them could be included in this small-scale study. Hence, the lists were combined into one list and alphabetised, and, then, every other lexeme was selected for this study. This enabled this study to select lexemes in as unbiased manner as possible without overly favouring some letters of the alphabet. Thus, the list was narrowed down to 26 lexemes. 16 are borrowed from Latin and 10 from Greek. Some of the lexemes are more specialised, but, as Cavendish also wrote scientific texts (Fitzmaurice 2011, n.p.), her texts probably also feature some technical vocabulary. The selected lexemes can be found in Table 2 below.

Table 2 Studied Lexemes with Languages of Origin

Lexeme	Origin	Lexeme	Origin
Alphabet	Greek	Hostile	Latin
Archive	Greek	Innuendo	Latin
Axis	Latin	Instruct	Latin
Caesura	Latin	Junior	Latin
Catastrophe	Greek	Militia	Latin
Corona	Latin	Narcosis	Greek
Crisis	Greek	Pathos	Greek
Dismiss	Latin	Philander	Greek
Elastic	Greek	Popular	Latin
Epiglottis	Greek	Produce	Latin
Exit	Latin	Psyche	Greek
Fungus	Latin	Sinus	Latin
Genus	Latin	Virus	Latin

3.3 Examination of Texts

Due to its small scale, this study can only include the base form of the lexeme and the inflected forms, such as plural forms. Including derived forms would add unnecessary complications, because, then, even the size of word-families could influence the results. Cavendish's spelling was deemed to be standardised enough to only include the standard orthographical forms of the studied lexemes. This obviously means that, if either the author or the printer has chosen to use nonstandard spellings, this can slightly affect the results. However, as the spelling of printed English had become largely standardised by 1650 (Nevalainen and Tieken-Boon Van Ostade [2006] 2012, 290), the likelihood of non-standard spellings is low. Table 12 in *Appendix 1* outlines any irregular forms – mostly irregular plural forms – that are included in this study. These forms are based on the *Oxford English Dictionary* (Oxford University Press 2024). The texts were examined one-by-one. Each of the studied lexemes was searched for using the web browser's search function – the most efficient method available. Then, the results were gone through manually to exclude any occurrences that were not within the scope of the study. Lastly, the occurrences were marked on a spreadsheet for later analysis. No distinction was made between the base and the inflected forms, because this study is not concerned with such grammatical variation. The relative

frequencies of each of the studied lexemes in each individual text as well as in the entire dataset of Cavendish's works were also calculated and marked in the spreadsheet.

With the corpus, the *group by lemmas* function was used to include all the forms of a lexeme. This makes the results of this study dependent on the accuracy of the function. If the search results, for example, do not include all the inflected forms or include many instances of Latin lexemes, this will affect the results. However, as the total word-form token count of the corpus also includes these Latin passages, they have not been excluded, as this would make the relative frequencies provided by the corpus invalid. The relative frequency data is significant, because it allows for a comparison between Cavendish's texts and the much larger dataset of the EEBO-corpus. The total number of occurrences in the corpus was marked in a spreadsheet along with the number of occurrences from three decades: 1650s, 1660s, and 1670s. Examining these decades separately enables this study to compare how Cavendish's usage compares to her male peers who wrote over the same decade. Cavendish's examined texts were all printed over those three decades. The relative frequencies – both for the entire corpus and the three decades – were also marked in a spreadsheet.

4 Results

First, the results for the studied texts are discussed generally as a whole. Then, each of the studied texts is discussed individually in its own subsection. Lastly, the results for the EEBO-corpus are outlined. Those lexemes with zero occurrences in a given text have been omitted from the tables. Frequencies are used to discuss the occurrences of lexemes within a single text, whereas relative frequencies, given per million words, are used to compare the different texts and datasets.

4.1 The studied texts – overall remarks

The loanwords of Latin origin are much more common in Cavendish's texts. Of those, *produce* is the most frequent with 359 occurrences and the total relative frequency of 443.04. The second most frequent is *instruct* and the third most frequent is *exit*. The relatively high frequency of *exit* can mostly be attributed to its use in stage directions in *Playes* (Cavendish 1662). The other Latin loanwords are much less frequent. *Popular*, for example, occurs only five times and has the total relative frequency of 6.17. Some lexemes, for instance, *axis* have no occurrences. In total, seven of the sixteen Latin loanwords – 44 per cent – occur in the selected texts. They have 554 occurrences in total and the combined relative frequency of 683.68.

The only one of the Greek loanwords to occur in Cavendish's texts is *alphabet*. It has only two occurrences and the relative frequency of 2.47. The infrequency of Greek loanwords can most likely be attributed to their specialised nature not matching with the subject matter and the contents of the texts. For example, *Playes* (Cavendish 1662) is a collection plays – a genre in which dialogue often plays a central role. This makes it unlikely to feature the more specialised Greek loanwords. The infrequency of these loanwords is discussed further in the *Discussion*-section. In total, the examined classical loanwords have 556 occurrences, giving them the total relative frequency of 686.15. Eight of the studied lexemes, i.e. 31%, occur in Cavendish's texts. Tables with the total frequencies and relative frequencies of the studied Latin and Greek loanwords can be found in *Appendix 2*.

4.2 Individual texts

The texts are discussed in order of publication.

4.2.1 The World's Olio (1655)

First, the results for the Latin loanwords will be discussed. They are outlined in Table 3 below.

Table 3 Frequencies and Relative Frequencies of Latin Loanwords in The World's Olio (1655)

Lexeme	Frequency	Relative Frequency
Instruct	15	153.34
Militia	3	30.67
Popular	2	20.44
Produce	53	541.78

The studied Latin loanwords have relatively few occurrences in this text. The most frequent lexeme is *produce* with 53 occurrences. The second most common lexeme is *instruct* with 15 occurrences. *Militia* occurs three times and *popular* only twice. Overall, only four of the sixteen studied Latin loanwords occur in this text. In total, the lexemes have 73 occurrences. The relatively frequency of all the studied Latin loanwords combined is 746.23. None of the studied Greek loanwords occur in this text.

4.2.2 Playes (1662)

Table 4 below outlines the results for the Latin loanwords.

Table 4 Frequencies and Relative Frequencies of Latin Loanwords in Playes (1662)

Lexeme	Frequency	Relative Frequency
Exit	90	312.84
Instruct	41	142.52
Produce	43	149.47

Again, most of the studied Latin loanwords do not occur in this text. The most frequent lexeme is *exit* with 90 occurrences. The high number of occurrences can mostly be attributed to the fact that stage directions often use this verb. This phenomenon will be further discussed in the *Discussion*-section. *Produce* has 43 and *instruct* 41 occurrences. However, it should also be noted that, although *instruct* only has 15 occurrences in *The World's Olio* (Cavendish, 1655), its relative frequency is much higher in that text due to the notable length difference between these two texts. Overall, three of the studied lexemes occur in this text, having 174 occurrences in total. Their combined relative frequency is 604.82. This text has no occurrences of the studied Greek loanwords.

4.2.3 CCXI Sociable Letters (1664)

Table 5 outlining the results for the Latin loanwords can be found below.

Table 5 Frequencies and Relative Frequencies of Latin Loanwords in CCXI Sociable Letters (1664a)

Lexeme	Frequency	Relative Frequency
Instruct	11	102.72
Produce	22	205.44

Only *instruct* and *produce* occur in this text. *Instruct* has 11 occurrences and *produce* has 22 occurrences. *Produce* also has a relatively high relative frequency: 205.44. In total, the two lexemes have only 33 occurrences. Their combined relative frequency is 308.17. None of the Greek loanwords occur in this text.

4.2.4 Philosophical Letters (1664)

Table 6 below outlines the results for the Latin loanwords.

Table 6 Frequencies and Relative Frequencies of Latin Loanwords in Philosophical Letters (1664b)

Lexeme	Frequency	Relative Frequency
Instruct	1	6.46
Junior	1	6.46
Produce	181	1169.03

Again, only a few of the studied Latin loanwords occur in this text. *Instruct* and *junior* both occur only once, and, thus, also have very small relative frequencies. *Produce*, on the other hand, occurs 181 times and has more occurrences than any other lexeme in a single text. Overall, three of the studied Latin loanwords occur in this text, with 183 occurrences in total. Their combined relative frequency is 1181.95.

Table 7 below outlines the results for the Greek loanwords.

Table 7 Frequencies and Relative Frequencies of Greek Loanwords in Philosophical Letters (1664b)

Lexeme	Frequency	Relative Frequency
Alphabet	1	6.46

Alphabet is the only Greek loanword to occur in this text. It has a single occurrence. Although both the frequency and the relative frequency of this lexeme are extremely small, its occurrence is still significant, because this is the first of the examined texts where any of the

studied Greek loanwords occur. In this text, the author uses it as a part of a metaphor where she compares remembering something to a writer going through the alphabet when they are attempting to write a rhyme (Cavendish 1664b, 32). This can be seen as an instance where the contents of the text and the examined Greek loanwords happened to match each other.

Overall, the studied lexemes have 184 occurrences in this text, giving them the total relative frequency of 1188.41.

4.2.5 Nature's Picture (1671)

The results for the studied Latin loanwords are found in Table 8 below.

Table 8 Frequencies and Relative Frequencies of Latin loanwords in Nature's Picture (1671)

Lexeme	Frequency	Relative Frequency
Dismiss	3	18.42
Instruct	25	153.48
Popular	3	18.42
Produce	60	368.35

Again, many the studied Latin loanwords do not occur in this text. *Produce* has the highest frequency with 60 occurrences, giving it the relative frequency of 368.35. The second most frequent lexeme is *instruct* with 25 occurrences and the relative frequency of 153.48. *Dismiss* and *popular* both have three occurrences in the text and the relative frequencies of 18.42. In total, four of the Latin loanwords occur in this text. Their combined frequency is 91 and their combined relative frequency is 558.66.

The results for Greek loanwords are outlined in Table 9 below.

Table 9 Frequencies and Relative Frequencies of Greek Loanwords in Nature's Picture (1671)

Lexeme	Frequency	Relative Frequency
Alphabet	1	6.14

Again, *alphabet* is the only one of the studied Greek loanwords that occurs in the text, and, again, it only has a single occurrence. In this text, it is used in a figurative manner by a character who is apologising for the crudeness of his language (Cavendish 1671, 345). This is another instance where the contents of the text and the examined lexemes happened to match each other. Overall, the studied lexemes have 92 occurrences, giving them the total relative frequency of 564.80.

4.3 EEBO-Corpus

Table 10 below provides the total frequencies and the total relative frequencies for the studied Latin loanwords.

Table 10 Total Frequencies and Relative Frequencies of Latin loanwords in the EEBO Corpus
The lexemes are sorted by frequency. Figures are provided by the EEBO-corpus

Lexeme	Total Frequency	Total Relative Frequency
Produce	30342	40.19
Exit	18659	24.71
Instruct	17848	23.64
Popular	12752	16.89
Genus	9066	12.01
Militia	7603	10.07
Axis	2809	3.72
Hostile	2805	3.72
Corona	2151	2.85
Sinus	2127	2.82
Junior	1723	2.28
Dismiss	1488	1.97
Virus	471	0.62
Innuendo	224	0.30
Fungus	186	0.25
Caesura	37	0.05

As can be seen in Table 10, some studied lexemes are much more frequent than others. *Produce* is the most frequent with 30342 occurrences and the relative frequency of 40.19. *Exit* is the second most frequent with the frequency of 18659 and the relative frequency of 24.71. The third most frequent is *instruct* with 17848 occurrences and the relative frequency of 23.64. *Militia*, *popular* and *genus* also have relative frequencies over ten. Some lexemes, however, have notably few occurrences. For instance, *fungus* only has 186 occurrences and the relative frequency of 0.25. The Latin loanword with the smallest number of occurrences is *caesura* with 37 occurrences and the relative frequency of 0.05.

Cavendish's texts examined in this study were published over three decades: 1650s, 1660s and 1670s. Figure 1 gives the relative frequencies of the studied Latin loanwords over those decades in the EEBO-corpus as well as the averages for the whole period.

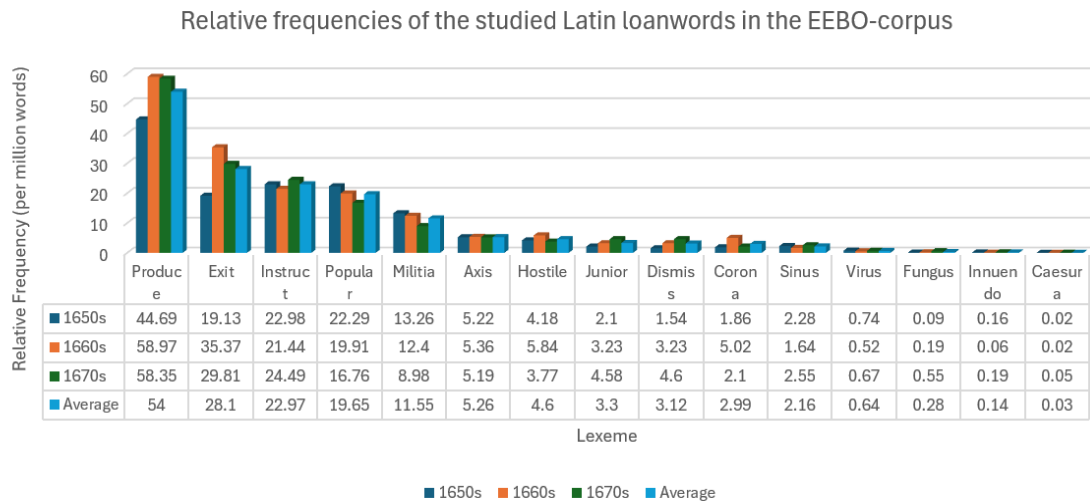


Figure 1 Relative frequencies of the studied Latin loanwords in the EEBO-corpus

Over all the studied decades, the most frequent and the least frequent lexeme remain the same. The most frequent is *produce* and the least frequent is *caesura*. Over the entire examined period, *produce* also has a much higher relative frequency than the other lexemes. However, there are some notable changes in the relative frequencies of the other lexemes. The relative frequency of *exit* is 19.13 in the 1650s and 35.37 in the 1660s, but it drops back to 29.81 in the 1670s. In fact, *exit* is the second most frequent of the studied lexemes in the corpus in the 1660s and 1670s, relinquishing *instruct* to the position of the third most frequent lexeme. The relative frequency of *popular* – the third most frequent lexeme in the 1650s – also shows a downward trend in the corpus. The relative frequencies of *dismiss* and *junior*, on the other hand, seem to be on the rise.

Table 11 below shows the frequencies and the relative frequencies of the studied Greek loanwords in the entire EEBO-corpus.

Table 11 Total Frequencies and Relative Frequencies of Greek loanwords in the EEBO Corpus
The lexemes are sorted by frequency. Figures are provided by the EEBO-corpus

Lexeme	Frequency	Relative Frequency
Alphabet	370247	490.39
Crisis	1219	1.61
Philander	1064	1.41
Catastrophe	1055	1.40
Psyche	980	1.30
Epiglottis	179	0.24
Elastic	59	0.08
Pathos	59	0.08
Archive	27	0.04
Narcosis	4	0.01

Overall, in comparison to the Latin loanwords, the frequencies are notably low. However, *alphabet* has the exceptionally high frequency of 370247, giving it the relative frequency of 490.39 which is higher than that of any of the Latin lexemes discussed above. On the other hand, some of the Greek loanwords only have a very limited number of occurrences. Even the second most frequent lexeme, *crisis*, only occurs 1219 times, giving it the low relative frequency of 1.61. Perhaps even more notably, *narcosis* only has four occurrences in the entire corpus, giving it the diminutive relative frequency of 0.01.

Figure 2 shows the relative frequencies of the studied Greek loanwords over the three decades from the 1650s to the 1670s as well as the average relative frequencies.

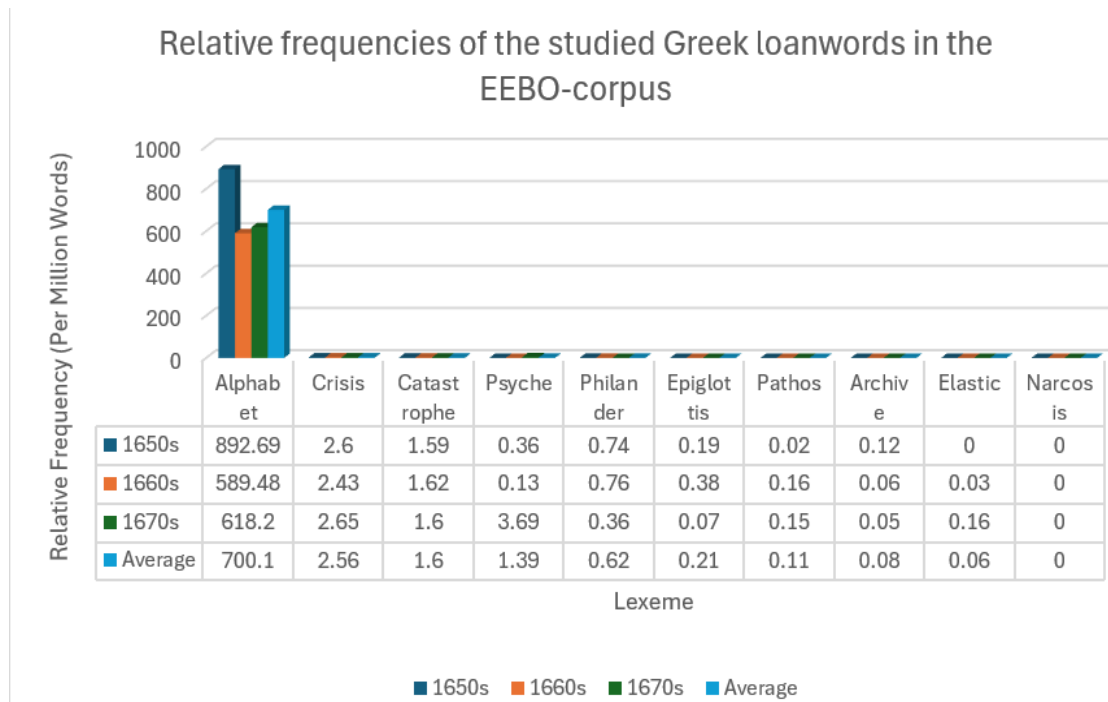


Figure 2 Relative frequencies of the studied Greek loanwords in the EEBO-corpus

As with the Latin loanwords, some notable changes are observable in the relative frequencies. For instance, although *alphabet* clearly remains the most frequent lexeme over all three decades, its relative frequency is somewhat lower in the 1660s than in the 1650s. The relative frequency of *psyche* also grows markedly to 3.69 in the 1670s, making it the second most frequent of the studied Greek loanwords in the 1670s. This also gives it the second highest relative frequency figure for a Greek loanword in any section of the dataset. The relative frequency of *epiglottis* also grows from 0.19 in the 1650s to 0.38 in the 1660s, dropping back 0.07 in the 1670s. It is, in fact, the fifth most frequent of the studied Greek loanwords in the 1660s. It is also notable that there are no occurrences of *elastic* in the 1650s and that *narcosis* has no occurrences over the studied three decades.

5 Discussion

Firstly, it is notable that all the Latin loanwords with a relative frequency of ten or above either in terms of the whole EEBO-corpus or the average for the three decades appear in Cavendish's texts, with the exception of *genus*. I suggest that this exception can be partially attributed to the contents of the texts. It is possible that the author simply did not have the need to use *genus*, and, thus, it does not occur in her texts. It should also be noted that a significant portion of the occurrences of *genus* given by the corpus are in Latin texts, which influences the results for the corpus. It is also notable that, in both datasets, *produce* is much more frequent than the other lexemes, but its relative frequency is much higher in Cavendish's texts. This pattern of higher relative frequencies in Cavendish's texts is also replicated by *instruct* and *exit*. I suggest that this is due to Cavendish's texts discussing many fewer topics than all the books of the EEBO-corpus. The smaller variation in topics is, probably, also going to make variation in lexicon more limited. In addition, the much smaller number of word-form tokens in Cavendish's texts means that the frequencies of individual lexemes have a stronger effect on their relative frequencies. Nonetheless, the main trends are the same for both datasets.

Many of the same phenomena that could be seen in the data for all Cavendish's texts are replicated in the individual texts as well. *Produce* and *instruct* occur in all texts, and, in all texts, with one exception, *produce* is the most frequent lexeme. *Produce* also tends to have a much higher relative frequency than any of the other lexemes. One exception to both statements is *Playes* (Cavendish 1662). There *exit* is the most frequent lexeme with 90 occurrences and a relative frequency more than double that of *produce*. *Produce* and *instruct* also have nearly the same frequencies, 43 and 41 respectively. In the data from the EEBO-corpus, the relative frequency of *exit* does show some growth from the 1650s to the 1660s, rising from 19.13 to 35.37. However, firstly, these figures from the corpus are nowhere near the relative frequency of *exit* in *Playes* (Cavendish 1662), and, secondly, *exit* does not occur in any of the other Cavendish's texts written in the 1660s. This suggests that the high frequency of *exit* is due to the contents of the text and, indeed, all occurrences of *exit* are in stage directions. This explanation is also supported by the fact that, in the EEBO-corpus, *produce* remains the most frequent lexeme over all the decades and has a significantly higher relative frequency than *exit*. In the corpus, *produce* even became more frequent in the 1660s. Another notable phenomenon is the low frequency of *instruct* in *Philosophical Letters*

(Cavendish 1664b). It occurs only once while *produce* occurs 181 times. This drop in frequency is replicated neither by *CCXI Sociable Letters* (Cavendish 1664a), published the same year nor by the EEBO-corpus, where the relative frequency only drops slightly from the 1650s to the 1660s. This suggests that the contents of this text did not favour *instruct*, i.e. the author did not have a need to use the lexeme. The discussion in this paragraph demonstrates how the contents and the type of a text can significantly affect the frequencies of individual lexemes in it.

As for the Latin loanwords of low frequency, some of them – *popular*, *junior*, and *dismiss* – show developments that could be reflections of the trends observable in the EEBO-corpus. For instance, *popular* does not occur in Cavendish’s texts from the 1660s, which matches with the corpus. However, *popular* reappearing in *Nature’s Picture* (Cavendish 1671) does not match with the corpus. In *Philosophical Letters* (1664b), Cavendish uses *junior*. This could be a reflection of the upward trend seen in the EEBO-corpus. *Dismiss* also shows a rising trend in the EEBO-corpus. This trend could be reflected in *dismiss* occurring in *Nature’s Picture* (Cavendish 1671). However, due to the small dataset utilised in this study, I am extremely hesitant to draw any conclusions on whether the developments discussed above are actually reflections of the trends in the corpus. As has become evident earlier, the contents of a text can significantly influence the frequencies of individual lexemes. As such, close-reading would be needed to determine whether Cavendish has used other lexemes where one of these three lexemes could have been used or whether Cavendish did not even have an occasion to use any of them. Such close-reading is beyond the scope of this study.

Of the studied Greek loanwords, only one, *alphabet*, occurs in Cavendish’s texts. It occurs twice and has the modest relative frequency of 2.47. In the data from the EEBO-corpus, *alphabet* is markedly more frequent than any of the other studied Greek loanwords. In the entire corpus, it has the relative frequency of 490.39 and, in the 1650s, the notably high relative frequency of 892.69. Thus, it is hardly surprising that it even occurs in Cavendish’s texts. The second highest relative frequency for the Greek loanwords in the corpus, 3.69, is reached by *psyche* in the 1670s. However, there are many lexemes, for example *archive*, that, even in the 1670s, have extremely small relative frequencies. I suggest that the fact that these less common loanwords do not occur in Cavendish’s texts can be explained by their relative infrequency in the language. In addition, the limited number of word-form tokens in Cavendish’s texts further augments the infrequency of these lexemes. To assess whether Cavendish actually knew and could use these rare lexemes, significantly more data is needed

to compensate for their rarity. Nonetheless, I would tentatively argue that she does follow the same basic pattern as the authors of the EEBO-corpus: the lexeme that has significantly more occurrences in the corpus also occurs in her texts, although with a much lower relative frequency, and the lexemes with low relative frequencies in the corpus do not occur in her texts.

In previous research, it has often been stated that women were not leaders in change from above, such as borrowing (Raumolin-Brunberg 2012, 718). Nevalainen and Raumolin-Brunberg also assert that men were ahead of women in borrowing (2017, 117). However, the data from this study – at least tentatively – suggests that Cavendish's usage follows in many respects the same trends as the somewhat male-dominated EEBO-corpus. Thus, I would be extremely hesitant to categorically state that *all* women were behind men in borrowing and change from above. As Kielkiewicz-Janowiak explains, there are often larger differences between the language use of individual women than between women and men (2012, 313). She also explains that social class affects an individual's usage (Kielkiewicz-Janowiak 2012, 313). Although, for the reasons discussed above, the effects of social class on the results of this study are greatly reduced, its influence cannot be entirely eliminated. Anipa also emphasises that written language can never provide a complete picture of an individual's usage (2012, 177).

6 Conclusion

The results of this study tentatively suggest that, in many respects, Cavendish's usage follows the same trends as the EEBO-corpus, although some trends appear clearer in her works than others. The results of this study also tentatively suggest that previous studies may have too categorically assumed that women were behind men in change from above, such as lexical borrowing, in the Early Modern period. The limitations of this study discussed earlier should still be borne in mind. However, the micro-model, also discussed above, explains that an understanding of a larger phenomenon can be built gradually piece-by-piece (Anipa 2012, 178). This study can be viewed as a first step towards the historical sociolinguistic study of lexical usage.

There are many possibilities for future research. First of all, more research into the lexical usage of these exceptional women as well as women in general is needed. Margaret Cavendish's and her female peers' use of language should also be studied more broadly and then compared with other women of different backgrounds as well as men. It would, for example, be of interest to examine whether different structural changes take place earlier or later in the language of these exceptional women than their less affluent female contemporaries. The effects of other sociolinguistic variables on lexical usage should also be studied. In addition, it would be important to have studies covering different genres and types of texts because, as has been discussed above, the genre or the type of a text can also have an effect on how frequently different lexemes are used in it. Overall, much more historical sociolinguistic interest in the lexical developments of the Early Modern period is needed. This future research can then, hopefully, either verify or disprove the preliminary suggestions made here.

List of References

Primary sources

- Cavendish, Margaret. 1655. *The worlds olio written by the Right Honorable, the Lady Margaret Newcastle*. London: J. Martin and J. Allestrye. Accessed 11 October 2024. ProQuest. Wing / N873.
- . 1662. *Playes written by the thrice noble, illustrious and excellent princess, the Lady Marchioness of Newcastle*. London: John Martyn, James Allestry and Tho. Dicas. Accessed 11 October 2024. ProQuest. Wing / N868.
- . 1664a. *CCXI sociable letters written by the thrice noble, illustrious, and excellent princess, the Lady Marchioness of Newcastle*. London: William Wilson. Accessed 11 October 2024. ProQuest. Wing / N872.
- . 1664b. *Philosophical letters, or, Modest reflections upon some opinions in natural philosophy maintained by several famous and learned authors of this age, expressed by way of letters / by the thrice noble, illustrious, and excellent princess the Lady Marchioness of Newcastle*. London. Accessed 11 October 2024. ProQuest. Wing / N866.
- . 1671. *Natures picture drawn by fancies pencil to the life being several feigned stories, comical, tragical, tragi-comical, poetical, romanical, philosophical, historical, and moral : some in verse, some in prose, some mixt, and some by dialogues / written by ... the Duchess of Newcastle*. London: A. Maxwell. Accessed 11 October 2024. ProQuest. Wing / N856
- The Text Creation Partnership. n.d. Early English Books Online. Accessed 20 September 2024. <https://www.english-corpora.org/eebo/>.

Secondary sources

- Anipa, K. 2012. "The Use of Literary Sources in Historical Sociolinguistic Research." In *The Handbook of Historical Sociolinguistics*, edited by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 170–190. Malden, MA: Wiley-Blackwell. Accessed 29 September 2024. ProQuest.
- Barber, Charles. 1997. *Early Modern English*. 2nd ed. Edinburgh: Edinburgh University Press.

- Brink, Jean R. [2000] 2003. "Literacy and Education." *In A Companion to English Renaissance Literature and Culture*, edited by Michael Hattaway, 95–105. Malden, MA: Blackwell Publishers Ltd. Accessed 15 September 2024. ProQuest.
- Campbell, Lyle. 2013. *Historical Linguistics: An Introduction*. 3rd ed. Edinburgh: Edinburgh University Press. Accessed 22 September 2024. ProQuest.
- Charlton, Kenneth. 2002. "Women and Education." *In A Companion to Early Modern Women's Writing*, edited by Anita Pacheco, 3–21. Oxford: Blackwell Publishers. Accessed 6 September 2024. ProQuest.
- Charlton Kenneth and Margaret Spufford. [2002] 2003. "Literacy, Society and Education." *In The Cambridge History of Early Modern English Literature*, edited by David Loewenstein and Janel Mueller, 15–54. Cambridge: Cambridge University Press. Accessed 4 September 2024. ProQuest.
- Davies, Diane. [2005] 2013. *Varieties of Modern English: An Introduction*. Oxon: Routledge. Accessed 28 August 2024. ProQuest.
- Dixon, R. M. W. 2018. *The Unmasking of English Dictionaries*. Cambridge: Cambridge University Press. Accessed 28 August 2024. Cambridge Core.
- Durkin, Philip. 2014. *Borrowed Words: A History of Loanwords in English*. Oxford: Oxford University Press. Accessed 29 August 2024. EBSCOhost.
- English Corpora. n.d. "Early English Books Online." *English-Corpora*. Accessed 1 October 2024. <https://www.english-corpora.org/eebo/>.
- Eskin, Catherine R. 1999. "The Rei(g)ning of Women's Tongues in English Books of Instruction and Rhetorics." *In Women's Education in Early Modern Europe: A History, 1500–1800*, edited by Barbara J. Whitehead, 101–132. New York: Garland Publishing, Inc. Accessed 1 September 2024. ProQuest.
- Ezell, Margaret J. M. 2002. "Women and Writing." *In A Companion to Early Modern Women's Writing*, edited by Anita Pacheco, 77–94. Oxford: Blackwell Publishers. Accessed 6 September 2024. ProQuest.
- Fitzmaurice, James. 2011. "Margaret Cavendish, Duchess of Newcastle." *The Literary Encyclopedia*. Accessed 17th September 2024. <https://www.litencyc.com>.
- Görlach, Manfred. [1991] 2012. *Introduction to Early Modern English*. Cambridge: Cambridge University Press. Accessed 1 September 2024. Cambridge Core.
- Kastovsky, Dieter. 2006. "Vocabulary." *In A History of the English Language*, edited by Richard Hogg and David Denison, 199–270. Cambridge: Cambridge University Press. Accessed 1 September 2024. Cambridge Core.

- Kay, Christian and Kathryn Allan. 2016. "Change in the English Lexicon." In *The Cambridge Handbook of English Historical Linguistics*, edited by Merja Kytö and Päivi Pahta, 220–236. Cambridge: Cambridge University Press. Accessed 19th September 2024. Cambridge Core.
- Kielkiewicz-Janowiak, Agnieszka. 2012. "Class, Age, and Gender-Based Patterns." In *The Handbook of Historical Sociolinguistics*, edited by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, 307–331. Malden, MA: Wiley-Blackwell. Accessed 3 September 2024. ProQuest.
- Lancashire, Ian. 2012. "Early Modern English: Lexicon and Semantics." In *English Historical Linguistics: An International Handbook*, edited by Alexander Bergs and Laurel Brinton, 637–652. Berlin: De Gruyter, Inc. Accessed 14 September 2024. ProQuest.
- Nevala, Minna. 2016. "Processes of Sociolinguistic and Sociopragmatic change." In *The Cambridge Handbook of English Historical Linguistics*, edited by Merja Kytö and Päivi Pahta, 286–300. Cambridge: Cambridge University Press. Accessed 22nd September 2024. Cambridge Core.
- Nevalainen, Terttu. 2006. *An Introduction to Early Modern English*. Edinburgh: Edinburgh University Press. Accessed 21 May 2024. ProQuest.
- Nevalainen Terttu and Helena Raumolin-Brunberg. 2017. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. 2nd ed. Oxon: Routledge.
- Nevalainen, Terttu and Ingrid Tieken-Boon van Ostade. [2006] 2012. "Standardisation." In *A History of the English Language*, edited by Richard Hogg and David Denison, 271–311. Cambridge: Cambridge University Press. Accessed 12 July 2024. Cambridge Core.
- Nielsen, Hans Frede. [2005] 2012. *From Dialect to Standard: English in England 1154-1776*. Amsterdam: John Benjamins. Accessed 17 September 2024. ProQuest.
- Nurmi, Arja. 2012. "Periods: Early Modern English." In *English Historical Linguistics: An International Handbook*, edited by Alexander Bergs and Laurel Brinton, 48–63. Berlin: De Gruyter, Inc. Accessed 14 September 2024. ProQuest.
- Oxford University Press. 2024. *Oxford English Dictionary*. Oxford: Oxford University Press. Accessed 20 September 2024. <https://www.oed.com/>.
- Quinsee, Suzie. 2001. "Newcastle, Margaret Cavendish, Duchess of, 1624?-1674." In *ProQuest Biographies*. Ann Arbor, MI: ProQuest. Accessed 17th September 2024. ProQuest.

- Raumolin-Brunberg, Helena. 2012. "Early Modern English: Sociolinguistics." In *English Historical Linguistics: An International Handbook*, edited by Alexander Bergs and Laurel Brinton, 715–731. Berlin: De Gruyter, Inc. Accessed 14 September 2024. ProQuest.
- Serjeantson, Mary S. [1935] 1968. *A History of Foreign Words in English*. London: Routledge & Kegan Paul.
- Trousdale, Graeme. 2010. *An Introduction to English Sociolinguistics*. Edinburgh: Edinburgh University Press. Accessed 6 September 2024. ProQuest.
- Van Gelderen, Elly. 2014. *A History of the English Language: A Revised Edition*. 2nd ed. Amsterdam: John Benjamins Publishing Company. Accessed 1 September 2024. ProQuest.
- Whitehead, Barbara J. 1999. "Introduction." In *Women's Education in Early Modern Europe: A History, 1500–1800*, edited by Barbara J. Whitehead, 9–16. New York: Garland Publishing, Inc. Accessed 1 September 2024. ProQuest.

Appendices

Appendix 1 Irregular forms of the studied lexemes

Table 12 below outlines any irregular forms, mostly irregular plurals, that were included in this study.

Table 12 Irregular forms of the studied lexemes

Lexeme	Irregular Form
Axis	Axes
Corona	Coronae
Crisis	Crises
Fungus	Fungi
Genus	Genera

Appendix 2 Total frequencies and relative frequencies of the studied lexemes

Table 13 featuring the total frequencies and relative frequencies of all the Latin loanwords included in this study can be found below.

Table 13 Total Frequencies and Relative Frequencies of Latin Loanwords in Cavendish's Texts

Lexeme	Total Frequency	Total Relative Frequency	Lexeme	Total Frequency	Total Relative Frequency
Axis	0	0	Innuendo	0	0
Caesura	0	0	Instruct	93	114.77
Corona	0	0	Junior	1	1.23
Dismiss	3	3.7	Militia	3	3.7
Exit	90	111.07	Popular	5	6.17
Fungus	0	0	Produce	359	443.04
Genus	0	0	Sinus	0	0
Hostile	0	0	Virus	0	0

Table 14 outlining the total frequencies and relative frequencies of the studied Greek loanwords can be found below.

Table 14 Total Frequencies and Relative Frequencies of Greek Loanwords in Cavendish's Texts

Lexeme	Frequency	Relative Frequency	Lexeme	Frequency	Relative Frequency
Alphabet	2	2.47	Epiglottis	0	0
Archive	0	0	Narcosis	0	0
Catastrophe	0	0	Pathos	0	0
Crisis	0	0	Philander	0	0
Elastic	0	0	Psyche	0	0