Univerzita Karlova

Filozofická fakulta

Ústav anglického jazyka a didaktiky

Filologie – Anglický jazyk

DISERTAČNÍ PRÁCE

Mgr. Veronika Volná

Lexical Obsolescence and Loss:

The Case of Early Modern English (1500–1700)

Lexikální mortalita v rané moderní angličtině (1500–1700)

Vedoucí práce: Mgr. Ondřej Tichý, Ph.D.

2023

Acknowledgments

I would like to express my sincere appreciation and gratitude to my supervisor, Dr. Ondřej Tichý, for his guidance throughout the entire process of this dissertation.

I extend my heartfelt thanks to my professors and fellow students for their support and encouragement throughout my academic journey. I am truly grateful for the opportunities, knowledge, and growth that the Department of English Language and ELT Methodology has provided me.

Finally, to my husband, friends, and family, thank you for the five years' worth of encouragement and support.

Declaration

Prohlašuji, že jsem disertační práci napsala samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 28. 6. 2023   Mgr. Veronika Volná v.r.

Abstract

Literature dealing with the development of the English lexicon has been concerned primarily with new additions to the language's vocabulary, such as borrowings, coinages and word formation patterns, while the topic of lexical obsolescence and loss remains under-researched. In the Early Modern period, the rate at which new lexical items appeared in the English language was unprecedented, especially in the years 1590-1620, as documented in the Oxford English Dictionary's online Timelines feature. In tandem with the rapid expanding lexicon, there was a portion of the vocabulary that was undergoing obsolescence or complete disappearance. Over the course of the Early Modern period, English lost a significant portion of its word- stock, including those short-lived coinages or borrowings which had entered the language only several decades or centuries earlier.

Using authentic examples from the EEBO (Early English Books Online) corpus, this dissertation provides an insight into the role of lexical obsolescence and loss in the development of Early Modern English. Based on frequency data, a list of candidates for obsolescence has been generated featuring words such as *sacerdote* (lat.), *travalier* (fr.), and *breastlap* (en.), all of which were common at the onset of the Early Modern period but had disappeared by the end of the 17th century, likely due to competition with another already existing, more widespread and perhaps more native-sounding synonym.

Although near synonymy and polysemy are identified as the basic catalysts for obsolescence, other relevant factors include the disappearance of the word's original referent, weakening of emphasis through overuse, and political correctness. The most promising candidates were selected and the conditions accompanying the decline of these forms analyzed and discussed. Building on the examples drawn from the EEBO as well as previous classifications by Visser (1949) and Görlach (1991), several potential systems of classification are proposed for lexical obsolescence, taking into account form and function, language-internal and -external processes, as well as the real-world manifestation of obsolescence.

Key words

Early Modern English, lexical obsolescence, EEBO, classification of obsolescence

Abstrakt

Literatura zabývající se vývojem anglické slovní zásoby se obecně zaměřuje především na nové přírůstky v podobě výpůjček, novotvarů a slovotvorby, zatímco téma obsolescence či zastarávání a ztráty lexika zůstává málo prozkoumané. V období raného novověku se nová slova v anglickém jazyce objevovala nebývalým tempem, a to zejména v letech 1590-1620, jak dokládá online slovník Oxford English Dictionary pomocí nástroje Timelines. Současně s rychlým rozšiřováním slovní zásoby docházelo k zastarávání nebo úplnému zániku jiné její části. V průběhu raně moderního období ztratila angličtina značnou část své slovní zásoby, včetně těch novotvarů, frází a výpůjček, které se do jazyka dostaly teprve před několika desetiletími či staletími.

Na základě skutečných příkladů z korpusu digitalizovaných anglických knih z raného novověku EEBO (Early English Books Online) poskytuje tato disertační práce pohled na roli, kterou ve vývoji raně moderní angličtiny zastává obsolescence a ztráta lexikálních prostředků. Na základě frekvenčních údajů byl vytvořen seznam slov, která podle během tohoto období pravděpodobně zanikla. V tomto seznamu kandidátů jsou uvedena slova jako *sacerdote* (lat.), *travalier* (fr.) a *breastlap* (en.), která byla na počátku raného novověku běžná, ale do konce 17. století se z korpusu vytratila, pravděpodobně v důsledku konkurence s jiným již existujícím, rozšířenějším a možná i domáčtěji znějícím synonymem.

Ačkoli za základní katalyzátory obsolescence jsou většinou považovány synonymie a polysémie, k dalším relevantním faktorům patří například zánik původního referentu slova, oslabení emfáze v důsledku nadužívání a politická korektnost. Následoval výběr nejslibnějších kandidátů a analýza s diskuzí o podmínkách, které mohly vést k úpadku těchto slov. Na základě příkladů čerpaných z korpusu EEBO a předchozích klasifikací Vissera (1949) a Görlacha (1991) bylo navrženo několik potenciálních systémů klasifikace lexikální obsolescence, které zohledňují formu a funkci, jazykově vnitřní a vnější procesy a také reálné projevy obsolescence v praxi.

Klíčová slova

raná moderní angličtina, lexikální obsolescence, zastarávání slov, EEBO, klasifikace obsolescence

Table of Contents

List of figures

List of tables

1   Introduction

For the most part, literature dealing with the development of the English lexicon has been primarily concerned with new additions to the language's vocabulary including borrowings from other languages, productive word-formation patterns and so forth (for instance in Plag 2018; Bauer L., R. Lieber and I. Plag 2015; Lipka 1992), whereas the topic of lexical obsolescence and loss remains under-researched. If we were to take the example of Early Modern English, the focus of this thesis, there is an evident disbalance between the numbers of studies explaining additions to the vocabulary and semantic change (to cite several, Stockwell and Minkova 2001; Haugen 1950; Miller 2012), and those attempting to classify the process of word loss (the most prominent of these sources being a lecture by Visser 1949 and a short chapter in Görlach 1991: 139-143).

Although the rate at which new words emerged in the Early Modern period (most notably in the years 1590-1620) was vast and indeed unprecedented, it tells only one part of the story of English vocabulary. A much less known fact is that we can simultaneously witness the ongoing obsolescence or loss of words, which affected a large portion of the English word-stock, including those short-lived coinages or borrowings which had entered the language only several decades or centuries earlier. It is this second, hitherto underrepresented aspect of lexicology, that will be explored in depth within the scope of this thesis.

With the aid of authentic examples found in the EEBO (Early English Books Online) corpus, I will attempt to illustrate the process of lexical obsolescence and loss in Early Modern English, supplemented by a more detailed analysis of selected examples. Based on the cases found in the EEBO corpus as well as those cited in previous literature, this thesis proposes several possible classification systems for lexical obsolescence and follows with a discussion of their respective merits and pitfalls. Since the practice of current English dictionaries shows a relative lack of systematic labelling of obsolete forms, it is hoped that the proposed classification may find its use in contemporary lexicography.

Chapter 1 serves as an introduction to the concept of *lexical obsolescence,* including an overview of the terminology used throughout this study, the current state of research, and the motivations for the selection of this topic. This includes an explanation of what is understood by the term *lexical obsolescence,* a discussion of its possible definitions based on previous research and a further clarification of the key terms *obsolescent, obsolete,* and *lost,* which will be used throughout this thesis.

Chapters 2 and 3 offer a general overview of the state of the English language in the Early Modern period, and deal with grammar and vocabulary, respectively. The main points of interest regarding Early Modern English grammar can be found summarized in chapter 2, with emphasis on those features which could conceivably have a role to play in lexical obsolescence and loss.

The areas of focus are the emergence of a standard (i.e., standardized usage) in English, ongoing shifts in phonology (Great Vowel Shift), orthography (spelling standardization and attempts at reforms), morphology (levelling and loss of endings) and syntax. Chapter 3 describes the changes taking place in the English vocabulary, namely the influx of new borrowings from (predominantly) other Indo-European languages, the motivations for these borrowings and existing word-formation patterns.

The mechanisms of language change are described in chapter 4, which provides an account of language-internal processes as well as language-external ones (where relevant). From the viewpoint of language typology, the shift towards a more analytical model constitutes the most salient internal cause of language change, whereas the external mechanisms include various political, religious and socio-linguistic motivations.

The methodology used in this thesis is outlined in chapter 5. This section includes a description of the EEBO corpus (Early English Books Online), the process of building and tagging the corpus, the structure of the data and the possible obstacles that this might bring. The automated process of extracting the candidates (i.e., potentially obsolete forms) is described in detail, as well as the subsequent manual sorting and evaluation of the forms returned by the automated process.

Chapter 6 contains a detailed analysis of the obsolete forms which were found in the EEBO data and made it through both the automatic and manual filtering process. The qualitative analysis is supported by the OED entries for the given words. In closing, the obsolete forms are grouped in categories based on the circumstances of their decline.

A classification of obsolete forms can be found in chapter 7. Beginning with a detailed account of earlier classifications by Visser (1949) and Görlach (1991), groundwork is laid down for a more comprehensive system. Several systems for classification are proposed based on the examples found in the EEBO as well as those from the earlier classifications, after which follows a discussion of their drawbacks and potential applications (e.g., in dictionaries).

Chapter 8 features a general discussion of the methodology used in this thesis, weighing its merits and pitfalls. In this section I further address the limitations of the EEBO corpus, namely the questions of representativeness and spelling normalization. Finally, several improvements are proposed, as well as possible ways of expanding the scope of this study.

Chapter 9 summarizes the key findings of this thesis and in conclusion offers some final thoughts on the study of lexical obsolescence and loss.

1.1 Lexical obsolescence

1.1.1 Defining obsolescence

In the most general terms, *lexical obsolescence* is understood as the process which constitutes a word's decline. That is, it tells the story of a word which had previously appeared in the language above a certain frequency threshold (see chapter 4.4), but has begun to progressively lose its popularity and is moving in the direction towards extinction. Obsolescence is a natural process in language and affects not only the lexical aspect, but also, for instance, grammatical constructions. In her investigation on the topic of grammatical obsolescence, Rudnicka works with the following definition, which is largely applicable to the *lexical* obsolescence also:

> "Grammatical obsolescence describes a situation in which a previously popular and productive construction is, often gradually, losing its productivity and popularity over time until the construction disappears or there are only residues or fossilised forms left." (Rudnicka 2019: 4)

By extension, the related terms *obsolescent* and *obsolete* also require some clarification. As defined by Rudnicka, in order that a grammatical (in our case lexical) item may be considered *obsolescent,* it is not necessary for it to be below a certain frequency threshold, but the process of obsolescence should be evidently ongoing, in other words "there should be a visible negative correlation between the time and the frequency of use" (2019: 6). The adjective *obsolete* describes the final stage of obsolescence, i.e., the word's complete disappearance from active use in both spoken and written language. That is, the state where a word, which had previously been extant in the language in question, is no longer used and only appears in the context of meta language, for example in dictionary entries or, for that matter, in this thesis.

In order to find a set of candidates with which to illustrate and classify the numerous causes of *lexical obsolescence,* in chapter 4 of this thesis I search the EEBO corpus for evidently *obsolete* items, i.e., words which had completely disappeared from the corpus by the end of the Early Modern period. Needless to say, no historical language corpus, not even one as extensive as the EEBO data set, is an exhaustive representation of language production in the given time period. Therefore, it must be acknowledged that simply because a given word is absent from the language data sample does not necessarily mean that it is obsolete in the language outside the scope of that data sample. However, since this study operates within a corpus linguistics framework, it is permissible to conclude that "if a form does not appear in a corpus, it does not exist in the language the corpus represents" (Tichý 2018a: 83) while allowing for a margin of error since the corpus, however large, is still a sample of the actual language. In a corpus the size of Early English Books Online, a frequency of 0 has a confidence interval with an upper limit of 4 ($\alpha = 0.05$ significance level), meaning that technically we cannot rule out its existence in the language entirely. The central theses and methodology of this study build on the work of Tichý and will therefore be operating with his definition of *obsolete* as being dependent on the

available data sample, in this case the EEBO language corpus. In light of this, an *obsolete* word is defined as "lost, that is from a corpus linguistic perspective either not present in the data or indistinguishable (in frequency) from errors" (ibid.).

1.1.2 Lexical obsolescence in the English language context

From the 15ᵗʰ to the 18ᵗʰ century, the English language underwent a number of changes in all of its facets. Among these, we count the obsolescence or complete disappearance of a significant portion of the English word-stock. A crucial portion of the words lost was of Latin and French origin, and had entered the language only several centuries earlier (OED 2014). By the end of the Early Modern Period (early 1700s), the following examples were no longer present in the language based on their entries in the OED (2014):

> (1a) *sacerdote* (lat.) priest
> (1b) *pabulation* (lat.) feeding, fodder
> (1c) *pacated* (lat.) pacified, calmed
> (1d) *quicquidlibet* (lat.) whatever one pleases, anything whatsoever

It stands to reason, and certainly in the above examples it is easily conceivable, that the most readily available explanation for a word's disappearance is its competition with another already existing, more prevalent, and perhaps more native-sounding, synonym. Although near synonymy and polysemy are indeed recognized as the basic catalysts for obsolescence, Görlach (1991: 140) mentions several other factors, both internal and external to the language:

1. disappearance of the word's original referent
2. weakening of emphasis through overuse
3. political correctness/euphemisms
4. homonym/homophone conflict
5. phonic inadequacy
6. word formation patterns no longer productive

Ad 1. an example of this might be an industrial tool becoming obsolete, and the word denoting it going out of use as a result. Ad 2. such was the case of many intensifiers such as *wondrous,* which lost their charge due to frequent use. Ad 3. for example, when a taboo word was replaced by a more acceptable euphemism and subsequently disappeared from the language. Ad 4. words sharing the same written form or at least the same pronunciation may lead to confusion, which may serve as motivation for the replacement of one of these words. Ad 5. words containing sounds which are foreign to the language or difficult to pronounce may be in danger of becoming obsolete. Ad 6. for example, when the suffix *-ish* was no longer used to form words with the sense of "pertaining to", already existing words with this structure, such as the obsolete adjective *soulish*, fell out if use too (Görlach 1991: 143). A detailed overview of causes, including the above factors, can be found in chapters 3 and 6.

1.1.3 Issues with identifying obsolescence

The potential causes of lexical obsolescence are numerous, and many have been referred to by both Görlach (1991) and Visser (1949). However, this abundance can result in difficulty in separating the individual causes from one another. As Görlach himself concedes: "the number of concurrent factors involved [in lexical obsolescence] often makes it difficult or impossible to reconstruct the specific causes that have led to the loss of an individual word" (1991: 140).

Furthermore, Görlach draws attention to the fact that not all instances of word loss represent true cases of obsolescence, such as the disappearance of coinages which were entered into dictionaries but were never accepted by the speech community. The most well-known are *nonce words*, coined by a native speaker and occurring only in that one instance, and *latent words*, derivatives smuggled into the dictionary by the lexicographers themselves (Read 1978: 95-6). Forms which arise due to misspellings or mispronunciations and are incorrectly recorded in dictionaries form a separate category, dubbed *ghost words* by Skeat in 1886, who described them as "words which had never any real existence, being mere coinages due to the blunders of printers or scribes, or to the perfervid imaginations of ignorant or blundering editors" and lexicographers should never have recorded them in the first place (Read 1978: 95). These items cannot be counted among cases of obsolescence, due to the fact that they were never fully integrated into the lexicon.

1.2   Current state of research

There is not an abundance of research available on the subject of lexical obsolescence and loss, perhaps partly due to the fact that it is not always simple to determine the status of a word – extant or obsolete? The majority of existing studies pertains to specific subsections of the lexicon, for example the obsolescence of English loanwords in Italian (Pulcini 2008) or is focused on lexical loss within the scope of a dialect (or group of dialects), for example lexical obsolescence in the Arabic dialects (Kleinberger 2011). Moreover, these studies examine word loss in languages other than English and are related to more recent cases of obsolescence. The research in English obsolescence that is of most relevance to this thesis both in scope and methodology is that of Tichý (2021, 2018a) and Rudnická (2019, 2015), although both of these studies are concerned with the Late Modern English period and their focus reaches beyond individual lexical items.

Nevertheless, it is likely that many aspects of this process are universal and shared across languages, and thus the general definitions, approaches and methodologies used in these studies may be adapted to fit the requirements of research dealing with the Early Modern English lexicon.

1.2.1   Methodology for a corpus-based study of obsolescence

Previous research includes Tichý's (2018a) study on lexical mortality which covers the period from Late Modern to Present-day English (1700–2000) and is the most relatable to this study in terms of its scope and method. The corpus-driven nature of this study as well as the use of a frequency-based algorithm for determining potential cases of lexical obsolescence are the main features shared with this thesis. Working with the *Google Books* dataset, Tichý (2018a: 85) processed n-gram strings ranging from unigrams to 5-grams by setting a filter for n-grams which were among the 40,000 most frequent in any of the decades and at the same time have a maximum observed frequency of 0.03 p.p.m. (p.p.m. = parts per million words, elsewhere i.p.m = instances per million words) in the last decade. However, Tichý observes that "not all the words in [the 0.03 p.p.m.] frequency band are necessarily considered to be lost: some highly technical or specialised vocabulary may, for example, be found in this frequency band and yet be considered neither lost nor obsolete in any sense" (ibid.: 89). After filtering out false positives – proper names, Optical Character Recognition (OCR) errors and variety-specific forms (ibid.: 90-91) - Tichý was able to identify two primary categories of obsolete unigrams: forms which have been replaced by a different corresponding form, and those which have not:

> "The forms that have a clear substitute were usually either replaced by a formally related word or they belong to a specialised terminology that is often well structured – largely because otherwise we would not have been able to spot the substitute. In case of a replacement by a related form, the loss of the form in question is a part of a larger process: spelling standardization, morphological analogy or change in word-formation strategies" (ibid.: 97).

The results suggest that these processes - spelling standardization (*oxyd -> oxide*), morphological analogy (*shew/shown -> show/shown*) and change in word-formation strategies (*acetous -> acidic*) (ibid.) - are all language-internal, while the cases of lexical obsolescence with no obvious substitution are likely to be linked to language-external causes.

Tichý's methodology for extracting obsolescence candidates from the corpus seems highly effective, given that the obsolete n-grams (lexical bundles) covered a wide variety of types based on their respective syntactic functions, semantic fields, and the most probable explanations for their decline, for example, scientific terminology (*vitriolic acid*, *nitrous air*), appellations (*His Czarish Majesty*), countability and accommodation (*letters patents*), complex verb phrases (*be made appear*), etc. (Tichý 2021: 117-123).

1.2.2   Grammatical obsolescence

Beyond the realm of individual lexical items there is the obsolescence of entire grammatical constructions. Rudnicka's (2019) study follows the decline of the purpose subordinators *in order that, in order to, lest, so as to,* and the rise of the *to*-infinitive, drawing upon several Late

Modern English corpora of American English. The results of this research indicate that the decline of the constructions cannot be strictly attributed to their competition with one another, but rather their potential for obsolescence being dependent on higher levels of the language where change – whether internal or external – triggered changes on the constructional level. In this case, it was the level of the purpose subordinators. What is understood by higher-level processes is the sum of changes above the constructional level, which can constitute anything from standard sentence length or changes in punctuation usage, to loss of inflectional morphological categories (ibid.: 175). Rudnicka's inquiry into the COHA (Corpus of Historical American English) returned results which strongly suggest a correlation between sentence length and the distribution of purpose subordinators, i.e., with steadily decreasing sentence length there was a decline in the use of the aforementioned purpose coordinators, complemented by a rise in use of the *to*-infinitive. Competition on a higher level can also be seen in the overall decline in finite subordinate (adverbial) clauses, which correlates with, and presumably triggers the decline of the purpose subordinator *in order that* which, as a rule, introduces a finite subordinate clause (Rudnicka 2015). Furthermore, the study concluded that a given construction's frequency of use, for example in the available language corpora, may predict the direction of its future development, i.e., a lower frequency could indicate potential obsolescence in the given construction's future (Rudnická 2019: 221-222).

Another study of grammatical obsolescence by Petré (2010) documents the decline of the Old English *weorðan* (*become, be*) and the subsequent loss of *wearð*. From the onset of the Middle English period, *weorðan* was already on a decidedly downward trajectory frequency-wise, and was being replaced in usage by alternative means of expressing the same grammatical function, namely copulas of change such as *become*, *be* or *wax* or phrases containing *begin (to + V)* and *be made/done + XP* (Petré 2010: 458). Furthermore, the use of *weorðan*, which was strongly linked to inverted word order, a feature of the bounded Old English system. The decline of this system meant that *weorðan* would soon follow as a result of the changing ways in which grammar was used to structure narratives (Petré 2010: 480).

1.2.3 Other surveys of obsolescence

The phenomenon of lexical obsolescence has also been studied in dialects and, of course, a variety of languages. When considering the possible causes, the findings offer examples of a range of factors including language contact, analogical change, and homonymic/synonymic conflict.

Lexical loss in Medieval Spanish and the effects of homonymy specifically can be seen in Dworkin (1995) where, in order to test the hypothesis that Old Spanish was averse to homonymy and near-homonymy, the circumstances leading to the loss several Old Spanish verbs (*acender*, *puñar*, *puñir*, and *punir*) are explored by closely examining their historical development in Spanish as well their cognates in other Romance languages. It is concluded that the obsolescence

of *acender* was indeed most likely caused by its homonymy with *ascender*. Similarly, the homonymy of *puñar* and *puñir*, as well as their near-homonymy with *punir* "may have substantially contributed to [their] demise" (Dworkin 1995: 538).

Rini (1990) provides a detailed account of the loss of the Old Spanish pronominals *connusco* and *convusco* (in Modern Spanish, the paradigm is *con nosotros* and *con vosotros*) and explores the reasons behind the survival of the remaining elements of the paradigm, namely the forms *conmigo, contigo,* and *consigo*. Beginning with Early Latin and following the forms' phonological and morphemic development, Rini concludes that the loss appears to have been sparked by a series of analogical changes, beginning with the forms *nosco, vosco* becoming *connosco, convosco* by analogy to *contigo,* and later the change from *connosco, convosco* to *connusco*, *convusco*. The resulting forms were so detached from their original phonology that they were no longer recognized for the pronouns (*nos, vos)* that they represented. In contrast, the morphemes -mi-, -ti-, -si- still corresponded with the other tonic pronouns, *mi, ti, si*, and therefore *conmigo, contigo,* and *consigo* remained an easily recognizable part of the paradigm. An additional consideration is that "they were not flanked by the nearly identical con- ... -co. These two major differences between the forms *conmigo, contigo, consigo and connusco, convusco* might indicate why the former survived while the latter disappeared" (Rini 1990: 63)

In a study of obsolescence in Turkish-Ottoman vocabulary, Kleinberger (2011) looks for patterns of the obsolescence affecting 61 selected Turkish-Ottoman words by surveying how they are understood in various communities speaking Arabic dialects. The data collection was carried out in the form of 253 linguistic questionnaires, wherein the informants demonstrated their knowledge of the 61 words. Unsurprisingly, the age of the informants was the key factor which showed in apparent time which of the words were obsolescent. All of the remaining parameters (gender, geography, education and religion) had ultimately little to no effect on the informants' familiarity with the words. The Turkish-Ottoman words which had Arabic morphology were perceived as Arabic in origin, and those which were successfully assimilated into the Arabic dialect did not show any signs of obsolescence. It appears that semantic fields do have a role to play, as the results of the survey show that culinary words generally tend to be stable, whereas military terms are in danger of becoming obsolete or subject to semantic shifts.

The process of lexical obsolescence was explored in Berber, a language whose lexicon was greatly influenced by language contact with Latin, Spanish, French and, most notably, Arabic. Authors Chaira and Hamada (2018) compare the degree of lexical loss (the term *erosion* is used) in the Berber dialect Tacawit across the Aurès Massif, Occidental Aurès and Oriental Aurès regions. This is done on a limited semantic field, for which purpose animal names were selected, namely *bird, fish, cat, bee, pigeon,* and *female goat*. Participants were asked to write down the Berber term for each of the words, and the degree of lexical erosion was found to be highest for *bird* (95.62%), and moving through *pigeon* (68.72%), *fish* (52.34%), *cat* (49.88%), *female goat* (44.83%) down to the most widely recognized term *bee* (37.71%). The results also show a

dependence of the degree of erosion on region size, i.e., "the less maintained an item is, the narrower will be the area it is used in. The first lexical variable bird, for instance, is used across a territory that is the narrowest compared to all other items. The lexical variable bee, on the other hand, occupies a territory of use that is wider than those of the other five lexical items" (Chaira and Hamada 2018: 69).

Lexical obsolescence has also been studied in creolized languages, for example Gilman's (1979) research on lexical loss in Cameroon Pidgin. Following the principles of lexicostatistics, Gilman compares Cameroon Pidgin to four other languages (English, German, French and Spanish) using a 200-word list which contains core vocabulary items specifically selected to be as culturally independent as possible (1979: 174). The lists were then compared to determine how many words had been lost between the languages; the only language missing any of the 200 meanings from its repertoire was Cameroon Pidgin, which was missing 3 words. These words were *freeze, snow* and *green,* and the likely causes of their obsolescence are: "the first two may be by meteorological conditions in West Africa, the third by a restructuring the color-name system in CP" (Gilman 1979: 175). The rest of the study was concerned with determining which type of lexical replacement took place and found that all cases of replacement were the result of either borrowing, coinage, semantic shift, or semantic extension.

1.2.4   Obsolescence in Early Modern English

Literature dealing with the development of the English lexicon for any given period is primarily concerned with the new additions to the language's vocabulary – borrowing words from other languages, the coining of new words and various word-formation strategies (affixation, compounding, clipping, etc.) – and yet it is rare to encounter the question of what becomes of those lexical items which have grown redundant or outdated. The Early Modern English period is one of rapid vocabulary growth, especially in the years 1530-1660 (Görlach 1991: 136), with most of the new words introduced as borrowings and new coinages. Simultaneously, obsolescence of lexical items was not uncommon, most noticeably in the 15th century, when the emerging standard led to the stigmatization of a number of words from peripheral dialects, both regional and social (ibid.: 139).

Naturally, word loss continued, albeit to a lesser degree, in the later decades and centuries as a result of a number of factors, both internal and external: weakening due to overuse, disappearance of the original referent, ambiguity resulting from polysemy or homonymy, etc. (Görlach 1991; for a full list see chapter 6.1). Furthermore, with the enormous influx of vocabulary, many new words were introduced to the language as a result of redundant borrowing as well as native production, resulting in an array of synonymous expressions. Taking into account the general tendency towards economy in the lexicon, the chances that such words would remain in the language were very slim from the very beginning. This phenomenon in native production can be illustrated using the synonyms *disfaithful (†1530) and unfaithful,* which

were both formed by derivation from the adjective *faithful* and differ only in their synonymous affixes. Not surprisingly, only one variant survived. A similar case is the synonymous nouns *disadventure* and *misadventure,* of which only one remains extant in standard usage.

Among the possible causes of formal obsolescence in Early Modern English we may count the ongoing changes in inflectional paradigms, which resulted in a number of redundant forms which would never become fully established, e.g., the past tense *wrote* had the variants *writ* and *wrate.* Finally, many cases of formal obsolescence are attributable to the process of standardization which spelling underwent in the Early Modern English. Following standardization, usually only one spelling variant remained, e.g., in the case of *royal,* which won out over the forms *royalle* and *royel,* etc., which then became obsolete.

1.3   Thesis objectives

Using authentic examples from the EEBO corpus, this thesis seeks to provide an insight into the role of lexical obsolescence and loss in the history of the English language, namely in the Early Modern period. In extension to earlier research, I propose several viable systems of classification for obsolete forms– both in terms of the degree of their obsolescence (based on their frequency and distribution in the corpus, as well as citations in the OED) and in terms of the conditions and circumstances of their decline. In a comparative conclusion, the typological reshaping and later standardization of English will be discussed in connection to obsolescence.

The motivation for choosing the Early Modern period to study the process of lexical obsolescence was twofold: firstly, the previous research on the topic of lexical obsolescence covers the periods of Old English, Middle English, and Late Modern English (Tichý 2018a, Tichý 2021, Tichý and Čermák 2015), leaving the thus far uncharted territory of the Early Modern period. Secondly, the quantity and availability of data representing the period made it possible to employ a corpus-driven approach, which meant that there could be continuity with Tichý's (2018a) research as far as methodology was concerned. The language corpus Early English Books Online (EEBO) was chosen as the primary source for cases of formal obsolescence in Early Modern English. The EEBO is a corpus of English written texts and contains almost 800 million words from more than 25 000 digitalized books which were published between 1420 and 1710, thus covering the entirety of the Early Modern period.

Currently, there is no existing framework in place which would serve to categorize obsolete and obsolescent words. Although usage labels are not uncommon in contemporary dictionaries of English, they each have their unique system with emphasis on different aspects of usage, for example, Merriam-Webster has labels for *obsolete* and *archaic* words[1], while Macmillan Dictionary is more focused on style and attitude labels[2]. To date, the most finely-grained

---

[1] https://www.merriam-webster.com/help/explanatory-notes/dict-usage
[2] https://www.macmillandictionary.com/live/labels.html

classification can be found in the OED (2014) but it is not systematic – though one might intuitively recognise what some of these labels mean, there are no definitions of these available and it is difficult to imagine precisely that they mean in relation to the degree of obsolescence (the labels currently used in the OED are *Obsolete, Historical, Archaic, Disused, Irregular, Rare* and *Regional*).

Moreover, the OED labels only address the degree of obsolescence, but there are no specifications as to the causes. Outlined in detail in chapter 6.1, Görlach (1991) and Visser's (1949) lists of possible causes of lexical obsolescence lay no claims to being systematic or exhaustive, which further goes to show how under-researched this topic is. One can only speculate as to how philosophy, ideology and attitudes towards language may have affected the preference of one form over another, thus inevitably leading to the latter form's disappearance from the language.

The aim is to design a scheme which could be utilized in contemporary lexicography and English dictionaries specifically. For such a scheme to serve its purpose, it is necessary to achieve a balance between complexity and clarity. The system should be clearly structured, unambiguous, and equipped to address even the most complicated cases of obsolescence without being confusing or cumbersome – a tall order, since there is great potential for overcomplication, which becomes clear in the following chapters as I analyse and attempt to classify the data retrieved from the EEBO.

2 *Characterizing Early Modern English*

The objective of section is to document the state of the English language during the Early Modern period. Beginning with a short description of the varieties of English spoken at the time and the inevitable development of standard usage, this chapter goes on to briefly describe contemporaneous developments in English phonology, orthography, morphology, and syntax. Finally, a characterization of Early Modern English Vocabulary is provided, including attitudes toward the etymological structure of the lexicon and means of enriching the language using both native and non-native strategies.

2.1 *The fragmented nature of the English vernacular*

The need for a standard originated from the chaos and lack of structure that was prevalent at the onset of the Early Modern period and was proving unsustainable for one single nation. The Renaissance was a period in English when the language was extremely varied, leading to a fragmentation or "brokenness" that was commented upon by contemporary scholars and aptly dubbed the "Babel" of the vernacular, for example, by 17th-century lexicographer Thomas Blount, who believed that England was "a 'self-stranger' nation – one growing alien to itself through the diversity of available forms" (Blank 2006: 214). The divisions that existed were primarily based on education, region, and trade.

English speakers were divided based on class; the language which was used by the elite was virtually impenetrable to the less educated, lower classes. On the opposite end of the social spectrum, emerging from the criminal underworld, was the *Thieves' cant.* This variety was a mixture of English, Latin, Dutch, French, Spanish, and many original coinages, used by thieves and hustlers with the purpose of being free to speak of their activities in a language that would be unintelligible to anybody outside that exclusive group. The *Thieves' cant* thrived predominantly in London at that time, since the large city offered anonymity and the "steady flow of new and unsuspecting victims" provided lucrative business opportunities for the *coney-catchers* (*coney* = "a fool; a dupe") (Coleman 2012: 120).

The language barrier in English provided the motivation for the first English-to-English dictionaries as a means of social reform, since the exclusivity of educated, elitist language put many at an unfair disadvantage. Hence the efforts to create dictionaries of 'hard words,' and technical terms from various fields, in order to make the vocabulary accessible to any literate speaker of English, thus "distributing the wealth of new words to the disadvantaged" and creating the space for one unified variety of English. One of the first was the *Table Alphabeticall* compiled by Robert Cawdrey in 1604, where it was explicitly stated that it the intended audience were "Ladies, Gentlewomen and any other vnskillful persons" (Blank 2006: 232).

In addition to the class-based division among English speakers, the geographical factor was of huge significance, the greatest divide existing between the Northern, Western, and Southern

varieties. In the North, the language developed under the influence of Old Norse under the Danelaw (Stockwell and Minkova 2001: 33) while retaining its modest share of local Scots words, resulting in a vocabulary vastly divergent from that of the South, where the influence of French-speaking nobility following (especially) the Norman conquest set an entirely different linguistic scene. Much like the class- and education-based sociolects, the diverse regional varieties also saw the stigmatization of speakers from more rural regions. Of the three key dialects, the Western (Somerset, Devon, Cornwall) was the most peripheral, seen by London writers as barbarous' 'rustic' and wholly unintelligible. However, the standing of the Northern dialect was not a great deal better, as it was also considered provincial, and a relic of sorts, since many northern words were often confused with archaisms or obsolete English words (Blank 2006: 216-220).

The full extent of the linguistic divide between the North and South is best illustrated by the classic "no *egges,* just *eyren*" anecdote, recorded in Caxton's prologue to the *Eneydos*. In the text, Caxton tells the story of a group of merchants travelling from the North and seeking to purchase eggs from a Southerner. When asked for *egges* (a word of Old Norse origin), the woman replied that she did not speak French, but understood as soon as another member of the group clarified that it was *eyren* (Old English in origin) that they wanted (Smith 2006: 122-3). Although we can never know for certain whether "I can speak no French" was the woman's sassiness or a genuine sign of misunderstanding, this anecdote drives home the fact that when speakers of the Northern and Southern dialects met, they might as well have been speaking different languages.

2.2   The emergence of a standard

Standard language is defined as having "maximal variation in function and minimal variation in form" (Nevalainen 2006: 29). Maximal variation in function means that the standard language facilitates successful communication in a wide range of situations, mediums and for many levels of formality. The same language could be used in the printed medium, in the royal court, as well as in the homes of the working classes. Minimal variation in form means that the norms (whether official or unspoken) governing spelling and grammar usage would be shared by more or less the entire language population. As soon as the first hints of a standard begin to emerge, one can expect that the population will gravitate towards the more prestigious dialects available. In the case of the English standard this was the language spoken by the educated classes based in London and arguably representing the lifestyle and level of education towards which most other speakers aspired.

In the Late Middle English period, the variation of form in English was extraordinarily rich, and particularly when it came to orthography. To give just one example, as many as 500 variants of the word *through* are attested in the *Linguistic Atlas of Late Mediaeval English* (Nevalainen 2006: 30). Over the course of the Early Modern period, English usage becomes increasingly

standardized. Among the first catalysts leading to the decline in orthographic variation were the scribes in King Henry V's Signet Office, who travelled with him on his campaigns (Nevalainen and Tieken-Boon van Ostade 2006: 274) and gave basis for a variety of English that came to represent an authority for standard usage. Soon to follow were the Westminster-based offices of the Chancery, where government clerks produced official documents, which were then disseminated throughout England. This was where the Chancery Standard originated, a decidedly southern variety of English but with its fair share of northern dialect features (Nevalainen 2006: 29-30).

Alongside the Chancery variety, there were three other standards emerging in the early 15th century; the Wycliffite variety, Chaucer's dialect and the Greater London variety (Nevalainen and Tieken-Boon van Ostade 2006: 275). Chaucer, as well as other popular writers at the time, wrote in English and arguably contributed to the rising popularity of the vernacular. The selection of a standard variety, it seems, was not so much determined by markers of prestige (such as education) as it was by support from the government and administration of King Henry V, and "it is significant that the selection of the variety which was to develop into what is generally referred to as the Chancery Standard originated with the king and his secretariat: the implementation of a standard variety can only be successful when it has institutional support" (Nevalainen and Tieken-Boon van Ostade 2006: 274). Another influential institution was the Royal Society of London, one of whose side projects was to be the creation of a universal language under which all would be united, which is a testament to the fragmentation of English at the time (Blank 2006: 237). Founded in 1660 to promote science and, in part, to improve the state of the English language, the Society prescribed a clear, "native prose" style which was to be exact and easily understood not just by the scholars, but all literate people (Nevalainen and Tieken-Boon van Ostade 2006: 291-2).

It was only thanks to the advent of the printing press that these language norms could spread as the printed texts were distributed throughout the population at a speed hitherto unimaginable. When William Caxton first set up his printing press in 1476, the language variation was still considerable and one of his chief concerns was determining the most appropriate variety for his English translations. As a businessman, strived to produce books that would be read by the largest possible demographic, and so he was looking for the most universal variety that would appeal to speakers of all the dialects of English. In the end, Caxton settled on "the variety used by his intended audience, educated people and those belonging to the higher regions of society," (Nevalainen and Tieken-Boon van Ostade 2006: 278) perhaps unsurprisingly since this was the variety with the most prestige and institutional support.

2.3   *Early Modern English grammar*

When writing about English grammar of the Early Modern period, we have two types of sources to lean on. The most important source of information regarding language use is the actual written

English of the time; be it letters, legal documents or works of fiction, Early Modern English has a wealth of preserved publications, available by way of online corpora and library archives, where we can observe authentic examples of language and grammar being used in the 16th and 17th centuries. The second source is grammars that were written at the time, publications whose authors explicitly described both the current state of the language as well as the desired state that the language should aspire to.

The first documented grammar of English was William Bullokar's 1586 *Pamphlet for Grammar,* written in English, which was a rarity for any grammar book at the time (Linn 2021: 65). These first grammars were prescriptivist in nature, since their primary purpose was teaching students how to use the language "correctly," and authors of these books or pamphlets often made efforts to liken English to Latin, because they wanted to show that English had the same grammatical categories and strict adherence to rules as Latin. These "teaching grammars" are less dependable, since they often described the language as it should be and not as it really was. On the other hand, the 1653 *Grammatica Linguæ Anglicanæ* by John Wallis was more on the descriptive side, offering insights on the true state of the English language as it was used by contemporary native speakers. Below are some of the key characteristics of Early Modern English grammar from the point of view of variability as well as their possible contribution to the lexical obsolescence that ran rampant at the time.

2.3.1 Phonology

When considering the key phonological changes that took place from Middle to Early Modern English, the Great Vowel Shift comes in at the top of the list. Simply put, it was a process which took place over the course of several centuries and affected all long vowels in English. By all accounts, it began with the high vowels /iː/ and /uː/ which were diphthongized into /aɪ/ and /aʊ/ respectively, setting off a chain reaction by freeing up slots in the vowel space. Once /iː/ had shifted, the high-mid front vowel /eː/ moved to take its place, in turn leaving an empty slot for open-mid /ɛː/ (most of which then moved further up to /iː/ in later phases) and so forth (Lass 2006: 82).

The shift took place in several phases, so as a result the whole picture is messier than one would like to imagine – for a more detailed account, see the section on vowels in Lass (2006: 81-91). To further complicate matters, in addition to the phases outlined over the course of the centuries, we also have the regional factor to consider. In fact, Smith (2007) speaks of two separate vowel shifts occurring simultaneously: "As well as a 'full' Shift affecting both the long front and long back vowels of Middle English, characteristic of southern varieties, there was also a distinct Shift, affecting primarily long front vowels, which is found in northern accents" (Smith 2007: 127).

There is no clear answer to what triggered the Great Vowel shift, rather there is a number of factors, both sociological and linguistic, all of which likely contributed to the process. To name a

few, Smith discusses a) the rise of a standardized form of English to which many speakers aspired, b) the presence of French vocabulary (and, presumably, pronunciation) which was seen as prestigious, c) the growth of London, which saw an influx of immigrants from the countryside and was seen as the focus for the standard that many gravitated toward, and d) the loss of final <-e> which resulted in an unusually large quantity of monosyllabic words in the core vocabulary and prompted the phonemization of vowels affected by Middle English open syllable lengthening (2007: 129).

As for consonants, the most crucial phonological change occurring at this time was the system-wide addition of phonemic /ŋ/ and /ʒ/ to the inventory of Early Modern English phonemes. Aside from this, most of changes the at the time involved the loss or substitution of a consonant, such as a) the disappearance of /t/ and /d/ from consonant clusters with /s/, e.g., *castle* and *handsome,* b) the loss of word-initial /g/ and /k/, as in *gnaw* and *knight,* or c) the assibilation of clusters /sj/, /zj/, /tj/, and /dj/ to /ʃ/, /ʒ/, /tʃ/, and /dʒ/ respectively, e.g., *ocean, seizure, fortune* and *soldier* (Millward and Hayes 2012: 247).

Although the Great Vowel shift and changing pronunciation in general are among of the most salient processes taking place in Early Modern English, their effects on formal obsolescence are minimal if not non-existent. We may speak of select phonological or phonetic forms such as the disappearance of the palatal realizations in words like *night*, however those are outside the scope of this thesis.

2.3.2   Orthography

It is the cooccurrence of the Great Vowel Shift and Caxton's printing press that we most often blame for the puzzling orthography of present-day English. In older varieties of English, the correspondence between phoneme and grapheme was relatively straightforward and, as a result, the written language of Old and Middle English may be considered an accurate reflection of contemporaneous pronunciation. This is attested by the variability of regional spellings, for example two copies of the same text, separated by a decade at best; *Ancrene Riwle* (Text A) and *Ancrene Wisse* (Text B), created by scribes in Worcestershire and north-west Herefordshire respectively, differ in several aspects, most notably words beginning with *u* in Text A (*uikelares* and *uorme*) begin with *f* in Text B (*fikeleres* and *forme*). The spelling systems observable in these and many other texts are regionally divergent and are believed to reflect local pronunciations (Corrie 2006: 89).

The divergence of spelling and pronunciation took place gradually over the centuries, for the most part owing to the ongoing sound changes and the conservative nature of spelling which struggled, and ultimately failed, to catch up. When the printing press was introduced to England in the late 15th century, the Great Vowel Shift was already underway and as phoneme departed from grapheme, there came a sense that orthography was no longer as adaptable. With the wide dissemination of print, the written form was resistant to changes which would mirror the gradual

shifts in pronunciation. By 1650, English spelling had become fixed in the printed media (Nevalainen 2006: 32), and reflected the sounds of London and southern dialects.

Further contributing to the divergence of text and sound was the authoritative status of French and Latin, which meant that these languages were consulted when it came to choosing the correct spelling of a word of those origins, and by "recreating connections between words borrowed from Latin and French, [those responsible for creating norms for the written language] introduced etymological spellings into English" (ibid.).

As a result, there were numerous efforts to bring about a reform of spelling, making it phonemic once more. The qualms of spelling reformers such as John Cheke and John Hart were that "superfluous letters occurred in words such as *authorite* (the letter <h>), *condempned* (<p>), *eight* (<g>) and *people* (<o>), and unnecessary variation was found in homophones like *sunne* 'sun' and *sonne* 'son'" (ibid.). The extent of formal obsolescence brought about by changes in spelling is very apparent at first glance. Take for example this quote by Francis Bacon and take note on how many orthographic forms are present that have been lost since the year 1605 when the text was written: "Schollers in Vniuersities come too soone, & too vnripe to Logicke & Rhetoricke; Arts fitter for Graduates then children, and Nouices: For these two rightly taken, are the grauest of Sciences, beeing the Arts of Arts, the one for Iudgement, the other for Ornament" (in Nevalainen 2006: 67).

2.3.3 Morphology

All developments in morphology between the Old English and Early Modern period can be summarized as the levelling and loss of inflectional endings as the language became increasingly analytic. These systemic changes had led to the loss of a large number of forms, for example the dative *kinges* and nominative plural *eyen* (eyes) and may be attributed to the increasingly reduced pronunciation in unstressed syllables, followed by the inevitable disappearance of grammatical suffixes, and leading in turn to the "increasing obligatoriness of pronouns" (Görlach 1991: 87). By the EModE period, the system of verbal inflections had already been reduced to such a degree that only the 2nd person <-(e)st> and 3rd person <-(e)th> or <-(e)s> remained (Görlach 1991: 88), and by the middle of the Early Modern period, the 3rd person <-(e)th> had already been pushed to the brink of extinction by <-(e)s> (Nevalainen 2006: 89). By this time, the regular <-ed> suffix was productive in most preterite and past participle verbs, and "many verbs, in particular rare ones or those confined to poetic registers, disappeared or were used with regular (weak) tense forms" (Nevalainen 2006: 90-1). That these lesser-known verbs would start being used with the regular <-ed> ending is logical, since the lack of familiarity would lead to doubts as to which of the irregular endings to use, and speakers would be more likely to opt for the safe, regular variant, whereas those irregular verbs which were used on a daily basis would not cause confusion or qualms of this sort. This period also saw the rise of auxiliaries *do/does* and *have/has,* while their respective archaic forms *doth* and *hath* fell out of use over the course

of several short decades, as shown by Gelderen (2006: 168-9) using the example of several plays from the years 1580-1620.

The system of nominal declensions had already been reduced from Old English's four cases to two – the nominative and genitive, and the productive way of forming the plural was reduced to a single <-s> ending, written with an apostrophe (Nevalainen 2006: 74). This meant that singular noun *woman,* for example, could take on two forms; *woman's* in the genitive when denoting attribution or ownership, and *woman* in all other usages. The function of inflectional endings, which would have been used in the dative and accusative cases, was taken over by prepositions whose frequencies steadily rose.

The system of pronouns also became simplified, beginning with the merger of the accusative and dative, following which all objects had the same forms of pronoun (sg. *me/you/thee/her/him/hit/him* and pl. *us/you/them)* (Gelderen 2006: 166-7). The second person singular still distinguished between the familiar *thou* (*thee*) and polite *yee* (*you*) pronouns; when addressing friends and family members, one would use *thou*, whereas the use of *you* signalled respect. Over the course of the Early Modern period, the preference for the formal "you" was so overwhelming that the informal pronouns *thou/thee* fell out of use completely. Another development in the realm of pronouns was the addition of the neuter genitive *its,* whose function had been expressed by the pronoun *his* up until the Early Modern period*,* e.g., in instances such as "the tree and his fruit". It is generally agreed that this form was created analogically to genitive pronouns *hers*, *yours*, etc. (Gelderen 2006: 167).

2.3.4   Syntax

Essentially, the syntactic properties of Early Modern English are very similar to that of the Present-day system; in contrast to Old and Middle English, by the Early Modern period word order was already fixed due to the levelling and loss of the most inflectional endings over the course of the previous centuries. The most notable changes in syntactic structure in this period include the obsolescence of double definite determiners, the rise of auxiliary *do*, the gradual stigmatization of double negation, and S-V inversion in declarative sentences.

The system of determiners was very similar to the Present-day one, except for the possibility of combining two definite determiners in front of a noun (Nevalainen 2006: 104), for example as used in 1623: "And this my wicked Mistris may reclame […]" (EBBO, A00012). This became replaced by the *of*-construction over the course of the Early Modern period.

In the Early Modern period, auxiliary *do* became an obligatory element in a number of contexts, especially in questions, "especially in negative interrogatives, *do* became the rule by the end of the seventeenth century" (Nevalainen 2006: 108), for example "Darest thou behold thy happinesse?" (EEBO, A00968) and in negative affirmative sentences. The affirmative *do*, which today has an emphatic function in positive sentences, was used widely even in non-

emphatic contexts (ibid.: 109). The increased use of the auxiliary *do* is likely to be linked to the fixation of word order in the Early Modern period (Görlach 1991: 119).

The disappearance of multiple negation primarily affected the General dialect, while in other dialects it is extant to this day. In Early Modern English constructions with *not*, non-assertive forms *ever/anything* began to gradually take the place the negative forms *never/nothing*. There are certain trends which can be observed in the process, such as that "those who used multiple negation most in the Early Modern period came from social ranks below the gentry and professions. Women also used it more than men throughout the period" (ibid.: 113).

In declarative sentences with an adverb occupying the first position, the inversion of subject and verb was common in the 16th century, especially if the verb phrase included an auxiliary, or in the case of a heavy subject. This disappeared over the course of the Early Modern period, and today is present only in focusing constructions, as in *Here is/comes NN* (Nevalainen 2006: 113-4)

2.4 *Early Modern English vocabulary*

At the beginning of the Early Modern period, English was in a precarious position due to the dominance of classical Latin, the international language of culture and learning which reached far beyond the borders and whose influence permeated most spheres of language including religious discourse, legislature, and education. With its highly inflected and regularized grammar, rich vocabulary and perceived rhetorical beauty, the prestigious position of Latin was indisputable. Nevertheless, perhaps owing to national pride, there was a strong desire shared by many authors such as John Cheke and Thomas Chaloner to write in English, and the importance of education in the vernacular was propagated by many early humanists such as George Puttenham and Thomas Wilson (Knowles 1997: 69, 81). The only matter standing in the way of the vernacular's becoming a fully-fledged language of writing was the evident lack of terminology in English, whose vocabulary was, to the minds of contemporaneous scholars, severely underdeveloped in most specialized semantic domains. Therefore, the chief concern of English scholars in the 16th century was with the translation of Latin terminology into the vernacular, resulting in the emergence of a substantial number of new words in the language.

In the Early Modern period, the rate at which new lexical items appeared in the English language was unprecedented, a fact which is documented in the OED's interactive application *Timelines*, according to which the most productive decades were 1590-1620 (OED Online, 2014). In the decade 1590-1599, as many as 21,991 new words entered the language, followed by 21,225 in 1600-1609 and 20,775 in the years 1610-1619. At the turn of the 18th century, there was a significant decrease in the incoming vocabulary, most notably in the years 1730-1739 as seen in Figure 1 below. This is largely attributed to a methodological problem with the gathering

of data, during which quotation slips written by American readers were lost and never reached the editor's scriptorium (OED 2020).
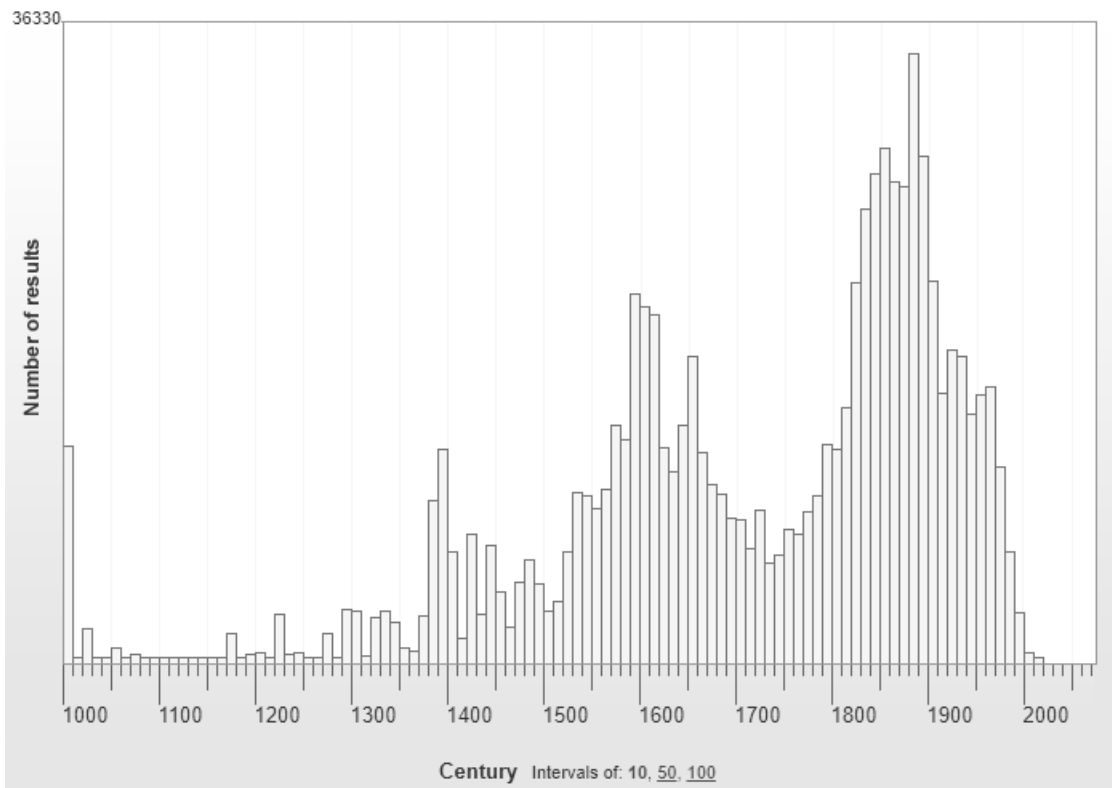
Figure 1: OED's *Timelines* visualization of new vocabulary in English per decade

However, two key issues must be addressed regarding the accuracy of the dates provided by the OED. Firstly, the data which was initially used for compiling the OED was incomplete, which has inevitably led to various discrepancies between the first recorded date available and that which appears in the OED. For example, Görlach demonstrates that such is the case of the word *gaietie*, which appeared as early as 1582 in Richard Mulcaster's *Elementarie* but its earliest use in the OED was cited as 1634 (Görlach 1991: 137), although it must be said that the OED has undergone many revisions since then, and in the case of *gaiety* the year 1573 has now been added as the first citation. Secondly, the first entry of any given word reflects, at best, the year of its first written record, and there is much uncertainty as to the word's prior usage, i.e., for how many decades or centuries the word existed in the spoken language before being written down (ibid.) Nonetheless, although the specific dates attributed to new additions to the lexicon are potentially problematic in these respects, they do provide us with a rough idea of the situation in a wider context of English. It should be noted, however, that the issue of fuzzy timelines is generally not pertinent to specialized and "learned" vocabulary, whose first usage is more likely to correspond with its first written record.

The categories of incoming vocabulary in the final decade of the 16th century, which bore witness to the arrival of almost 22,000 new words, confirm the general trend of borrowings and coinages at the time. The vast majority of new words pertain to scientific concepts (2635 new words) among which are items such as *allheal* (n., 1597-present; created as a compound of two existing English words; denoting a range of plants used in medicine), *rabulane* (n., 1593; supposedly derived from a Middle French form *rave*; meaning turnip or radish), *olorina* (n. 1596-1796; derived from Latin; denoting a type of plant), *defensative* (n. and adj., 1563-present; borrowed from Latin with the addition of a suffix; meaning protective or having the property of defending). The categories and the number of entries in each of them can be seen in Figure 2 below.
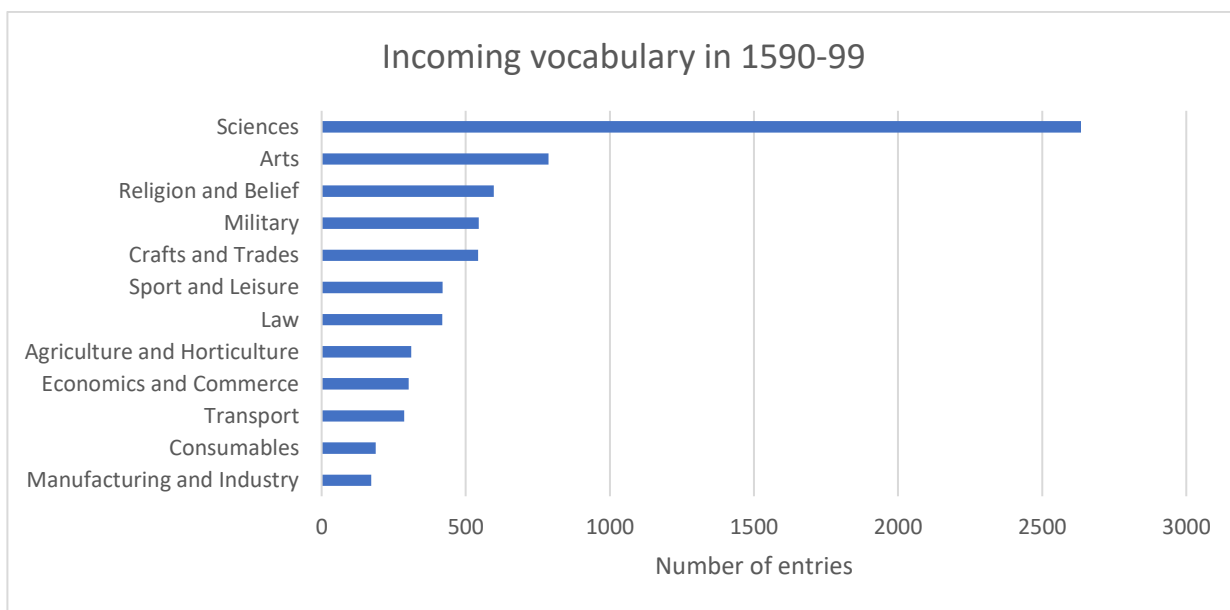
Figure 2: Semantic domains of incoming vocabulary as categorized by the OED (2014)

2.4.1 Lexical Borrowing

Due to the ongoing language contact between English and many (not only Indo-European) languages, it comes as no surprise that the bulk of new vocabulary in the 16th and 17th centuries was in the form of lexical borrowings. The two principal motivations for borrowing can be described as need and prestige, which Durkin (2009: 142) summarizes thus: "borrowing for need is necessary borrowing, because there is a lexical gap, and borrowing or prestige is unnecessary borrowing, because an adequate means of expressing the same concept already exists." In the case of need, the Early Modern period was witness to the emergence of many new concepts and discoveries, especially but not limited to the natural sciences, including *satellite* (1645), *optical* (1610) and *plenilunary* (1646). Borrowing for prestige at a time when Latin and French were

considered to be more sophisticated than English resulted in borrowings such as *decurt* (v., 1631; shorten or cut down), a process which gave English its typical richness of stylistic variability.

Based on OED (2014) data, the majority of new words entering English in the 17th century was of European origin (excluding English), and exact numbers are displayed in Figure 3, while Figure 4 zooms in to show the language families subsumed under the largest category of *European languages*. Based on Figure 4, most new words were adopted from Italic donor languages (51,951 items), with German borrowings lagging far behind in second place (15,457), followed by Greek (2,274), Celtic (141), Slavonic (34), Uralic (3) and Armenian (1).
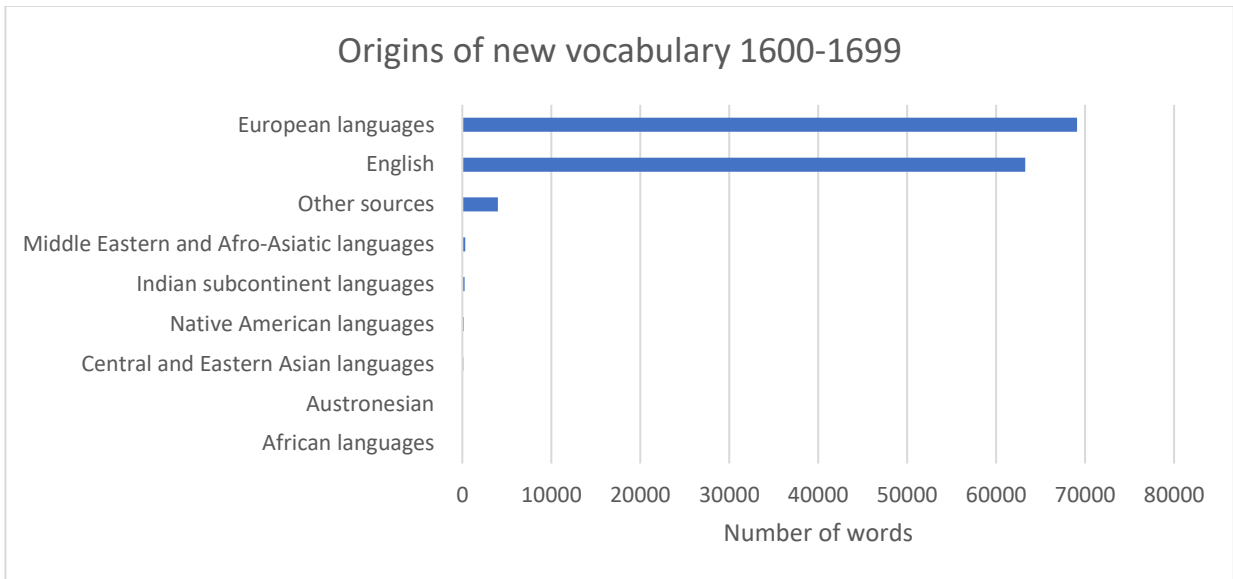
Figure 3: Sorted by origin, the number of new words entering English in the 17th century, based on OED (2014) data
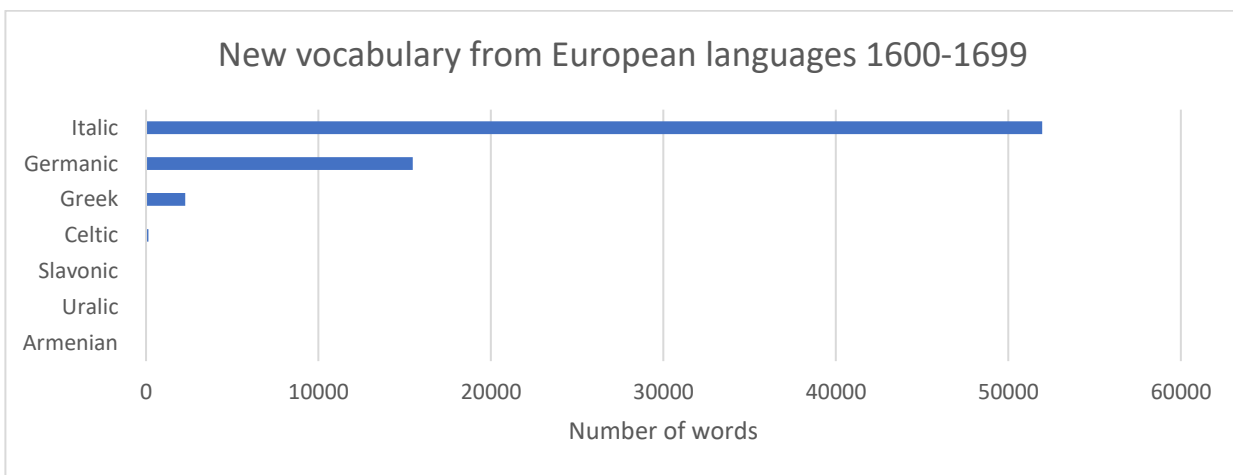
Figure 4: Sorted by origin, the number of new words from European languages entering English in the 17th century, based on OED (2014) data

The most widely accepted typology of lexical borrowing proposes four main types: loanwords, loan translations (calques), semantic loans and loan blends (Durkin 2009: 134-9). Loanwords are the most frequently occurring type of borrowing, in which the word form from the donor language is adopted with little or no change to its original meaning. In fact, the lack of means for expressing the signified referent is among the chief motivations for lexical borrowings in the first place. The loanword may have its pronunciation slightly altered in order to accommodate the phonemic inventory of the target language, and furthermore it is likely to conform to the morphology of the target language by taking the regular <-s> plural form, for example *fresco* (n., 1598 - present; a kind of painting). Loan translations replicate the structure of the word or phrase in the donor language using native elements, which may be either literal translations of the original, or roughly equivalent in meaning. The most notorious example of this sees English in the role of the donor language, with the calquing of the compound *skyscraper* in other languages, for example the Italian calque *grattacielo* (meaning "scrapes-sky") or Czech *mrakodrap* ("claws-at-clouds").

There is a degree of uncertainty as to when we are dealing with a loan translation, and when it is simply a parallel construction created in two or more languages independently (Durkin 2009: 135). Semantic loans do not manifest in the introduction of a new lexical item, but rather the extension of an existing word's meaning in parallel development with a synonymous term in the donor language. Similar to the case of loan translations, it is not always clear whether the parallel extension or change of meaning is purely coincidental or the result of a semantic loan. The category of loan blends includes complex words whose one or more morphs had been replaced by a native morpheme, most commonly an affix. Take for example the nonce word *vagisness* (n., 1604; meaning handsomeness, elegance), from the Italian *vaghezza.* Another example is the commonly occurring substitution of the French nominal suffix <-ier> with the anglicized <-eer> or, in early translations, <-er>, as in *pioneer* (n. and adj., 1517 – present; a soldier employed in digging trenches and clearing the road for the army) from the French *pionnier*.

In the process of borrowing, especially in the case of Latin and French, the new word was adopted and used in its native form, for example in the case of the Ibero-Romance borrowing *adelantado* (n., 1588 - historical use; the title for a governor of a Spanish province), *enamorate* (n. and adj., 1607-1624; a lover) and *gambado* (n., 1625 - historical use; denoting a gaiter attached to a saddle to protect the rider's leg from the weather). Loan blends are another common form of accommodating the foreign morpheme by combining it with a native (or nativized) derivational affix, for example an addition in the field of religious terminology *acolouthite* (n., 1598-1701; denoting a person who attends a priest during mass), whose Greek root entered the English language via Latin *acholitus* and was fitted with the nominal suffix <-ite>, or the term *sassinous* (adj., nonce word from 1632; meaning rocky or stony), which was coined by combining the Latin/Italian lexeme *saxum* or *sasso* with the productive suffix <-ous>.

2.4.2 Word Formation

Next to lexical borrowings, the coinage of new words using elements of predominantly native origin was also widely used throughout the history of English, which is equally true for the Early Modern period, a fact that is attested in the corpora available. The most productive word formation strategies available at the time were compounding, derivation, and conversion (sometimes referred to as zero derivation), which will be described in more detail. The validity of other word-formation strategies such as clipping, acronymization, back-formation and blends has been the subject of debates, in light of the view that word-formation should only refer to those "additive processes and patterns in which words are composed of smaller signs," i.e., those involving various combinations of free and bound morphemes (Görlach 1991: 170).

Compounding involves two free morphemes, *abackstays* (adv., 1625 - 1694; nautical term meaning in line with the direction of the wind). Based on their referent, compounds may be classified as either endocentric or exocentric. The former type is a compound denoting a sub-set of one of the elements, such as *blackbird*, while the latter denotes an object that is external to the meaning of either one of the elements. The productivity of the exocentric type in English is often attributed to French influence following the Norman Conquest, responsible for the appearance of words such as *pickpocket* in the Early Modern period and "a number of formations with *make* are found in the 16th and 17th centuries, such as *makepeace* […] the compounds are frequently paralleled by (and were probably modelled on) verbal phrases, to make peace, to pick someone's purse, etc." (Durkin 2009: 107).

Derivation (or affixation) combines a free morpheme with a bound morpheme. The range of available affixes in Modern English included both native and foreign morphemes, and the general tendency was to avoid mixing native free morphemes with foreign affixes. Native affixes proved to be more flexible in this respect, as in the case of *pragmatical* (adj., 1543 – present; meaning shrewd, also matter-of-fact and practical), a borrowing from Latin *pragmaticus* combined with the English affix <-al>. More conventional derivations attested in the OED include *abandonable* (adj., 1611 - present; capable of being abandoned) derived using the suffix <-able>, and *keenness* (n., 1530 - present; the quality of being keen) derived using the native nominal suffix <-ness>. Lest we forget prefixation, take for example *preadmonish* (v., 1632 – present, rare; to forewarn), where a derivation of the verb *admonish* is created using the prefix <pre->. Görlach (1991: 175) lists the following affixes which were productive in Early Modern English:

1. adj→ n (abstract): -th, -head, -hood, -ness, -ship, -ment, -esse, -ity/acy, -ion, -ure, etc.
2. n/v → adj: -y, -ish, -ful, -ly, -like, -al, -ous, -ic
3. n/adj → v: be- Ø, en- Ø, - Ø, -ize, -(i)fy
4. n/adj → v (privative): - Ø, un-/dis- Ø, dis- ize
5. adj → adj (negation): un-, in-, dis-

A specific type of derivation is zero derivation, which constitutes a free morpheme combined with a zero morpheme, a process which is otherwise interpreted as conversion from one word class to another with no change in the word's morphology. An example of this strategy in the Early Modern word stock is *wadset* (n., 1449 – present; a thing that has been pledged) created though conversion of the existing verb *to wadset.*

At the beginning of the Early Modern period, the insufficient number of prescriptivist forces such as grammarians and dictionaries meant that the control over production was limited and the process of word-formation was much freer, resulting in a multitude of forms whose survival depended on how readily they were accepted by the speech community. Many coinages were in competition with their more established synonymous counterparts, and so many such derivates soon became obsolete, including the forms *yongth* (youth)*, bestnesse, ouermuchnesse,* and *unhonest* (Görlach 1991: 172). In theory, the production of redundant synonymy is being kept at bay thanks to the phenomenon of *blocking*, "by which new formations arc blocked (or pre-empted) by the prior existence of a synonym. There are good reasons for thinking that blocking is an important factor in restricting word formation. Thus, we expect that the prior existence of *difficulty* will block *difficultness* from being formed, and similarly that *coolness* will block *coolth*" (Durkin 2009: 104). However, this presupposes a language community with close ties and a means of effectively sharing information, which would not have always been the case of 16th-century England. In addition, the Early Modern period was a time of linguistic innovation, and thus the principle of blocking was disregarded to some extent, resulting in the existence of the forms *to glad, gladden, englad, engladden,* and *beglad* (Görlach 1991: 172). Such cases of synonymy would ultimately lead to the survival of only one form, unless there was a divergence in meanings, for example *hardship* and *hardness.*

3   Mechanisms of language change

As an entity shaped by the humans who speak it, language with all of its layers finds itself in a perpetual state of change. Arguably one of the most volatile of these layers is the lexicon, least dependent on the internal structures of the language and thus all the more prone to outside influences. At the same time, the lexicon is always at the forefront of language use and so any change is immediately apparent. The following selection of examples taken from Bybee (2015: 98, 124, 189-207) should serve to demonstrate the extent and nature of changes taking place in the English lexicon;

- semantic change of both denotative (*bureau*, originally a "rough cloth" came to denote an *office* through metonymic change) and connotative meaning (unlike present-day English, "to *cause* something" didn't have strictly negative connotations in Early Modern English),
- normalization of spelling, whether through the widespread dissemination of written word or overt changes in spelling conventions,
- loss of inflectional suffixes, both through syntactic change and analogical levelling *(smoke, mourn, reap* all had strong verb inflection in Old English)*,
- grammaticalization, wherein words and morphemes used together form meaningful "chunks" (the future "going to" becomes *gonna*),
- introduction of new words by way of, for example, compounding native morphemes *(lighthouse)*, affixation *(business)*, abbreviation *(laser)*, loan translations *(flea market)* and borrowings from other languages *(coffee, orange)*,
- loanword adaptations, i.e., the alteration of morphology, spelling, and pronunciation to better suit the borrowing language's grammatical structures *(commence)*,
- obsolescence and loss, the process in which a word falls into disuse until becoming an archaic or regional variant *(kin)*, restricted to compounds (*tide* meaning "time" in *Yuletide*) or eventually disappearing from the language altogether (*couth* as in *uncouth*).

3.1   The language-external and -internal dichotomy

The causes of language change are traditionally divided into two distinct categories; external and internal. Simply put, language-external catalysts are the real-world circumstances of a language community which bring about changes in linguistic behaviour, while language-internal causes stem from the inherent structural properties of and processes within languages. In more concrete terms, "internal history may be defined as the study of evolving systems of lexicon, grammar, and transmission (speech- and writing-systems); external history is to do with the ways in which a language is employed over time, for example the shift from script to print, or how particular

languages are associated with particular social functions at particular moments in their history" (Smith 2006: 120).

When tracking the causes of change, it is the external motivators such as language contact, political correctness or institutionally driven standardization that we initially look to for answers. It is only "if external factors are not responsible for the relative occurrence of change types, then the reasons must be sought among internal factors, i.e., these types must be causally connected to structural features of language (in phonology and morphosyntax) or to contingencies of language production (in phonetics)" (Hickey 2012: 406). Take, for example, the voicing of intervocalic consonants, which is arguably very natural from an articulatory point of view, since it requires less effort in the part of the speaker. "Cross-linguistically, the rise of voiced intervocalic consonants is much more common than the reverse, so that a development like /afa/ > /ava/ is more likely than /ava/ > /afa/" (Hickey 2012: 406). This general tendency towards minimal effort is shared across all languages and we can therefore contend that the increase in intervocalic voicing is not driven by cultural (external) factors, but stems from the physiological realities of the speakers (internal).

The practicality of this internal/external dichotomy has been called into question (Smith 2006: 120) with emphasis on how the relationship between internal and external factors is what ultimately creates change. Although these two sources of change are often spoken of in binary terms, but this seems to be rather a matter of convenience, since real-world observations have shown that it is most often external factors that cause internal processes leading to structural changes. As Hickey attests, "considering social motivation as a central factor can improve the understanding of apparently counterintuitive instances of change, or at least of those changes which would not be expected on purely language-internal, structural grounds. In addition, social factors can help to account for the reversal of change and for the important issue of non-change" (2012: 402). Fischer warns that "we may use the dichotomy of internal/external factors as a working method but that we must not fall into the trap of forcing each explanandum (each case to be explained) in linguistic change—which usually concerns a combination of internal and external factors—into the same dichotomous mould" (2007: 32).

Taking the example of loss of inflectional endings in English, we can observe the interplay of internal and external causes. It is generally accepted that the levelling and eventual loss of endings was a product of the gradual shift from an inflectional to an analytical system, an internal process over which the language users had little control. However, it cannot be denied that language contact may have in part also contributed to the loss of inflections, as Millward and Hayes explain that "although Norse and English were similar in many ways, their inflectional endings were quite different. One way of facilitating communication between speakers of the two languages would have been to drop the inflectional endings entirely" (2012: 14). However, external influence is much less clear in the case of the Great Vowel shift (described in more detail in chapter 2.3.1) which can be almost exclusively attributed to internal processes.

Although it cannot be known for certain what triggered it, once the process had begun, i.e., once the high vowels /iː/ and /uː/ were diphthongized, we may argue that what followed was a decidedly language-internal process – the shifting of vowels as a reaction to the newly formed "open" slots in the vowel space is a chain shift that demonstrates "the functional economy of the vowel system: vowels move together to avoid merger and preserve their capacity to distinguish words" (Labov 1999: 117).

3.2 Language-external mechanisms: Sociolinguistic context

External causes of language change can take on a multitude of forms, some of which are outlined in the following chapters, but can for the most part be summarized as the events, circumstances or attitudes that exist in the world and thereby influence how a language is used by its speakers. Perhaps the most easily observed external cause of change is language contact, whose effects on English are well documented by numerous scholars, notably in Miller (2012: 230) where we can trace a great deal of English vocabulary back to the influence of classical languages in the arts and sciences and in Bailey (2006: 342) where we read about how the ongoing exploration and colonization by the British Empire, perhaps to a lesser degree, also influenced the English vocabulary by way of loanwords. Aside from language contact, some of the key motivators of language change include the political, religious and cultural climate. Nevalainen and Raumolin-Brunberg (2003) provide sociolinguistic observations on language change based on data from the period 1410-1681 and pose these four crucial questions;

1. When was the old form replaced by the new?
2. How did the new usage spread in the language community?
3. What was the social status of the people who promoted the new usage?
4. How were the old and new usages evaluated by society and how did this change over time?

Answering these questions should bring us closer to identifying the main causes of lexical obsolescence and loss for individual lexical items. For example, the deliberate omission of certain forms from the vocabulary of the upwardly mobile classes could indicate that these forms were perceived as stigmatizing. The following chapters introduce several noteworthy examples of sociolinguistic factors which can be seen as language-external catalysts for change, specifically in reference to the lexicon of the Early Modern period. In addition, the final chapters of this section examine the key ideas and attitudes towards language in Early Modern philosophy and the normative tradition, all of which have the potential to shed further light on the reality of lexical obsolescence.

3.2.1 Language contact and Britain's expansion

English as it is spoken today is a product of centuries of linguistic contact, during which new lexical items from various languages, most notably French and Latin, became part of its core vocabulary. In earlier periods, language contact took place predominantly within England, but with the growth of the British Empire, the points of contact reached far beyond the borders of Britain. In the 11th century, the Norman invasion of England brought with it a new set of words, some of which would gradually take the place of the already established Germanic ones, *army* from French *armé* replaced the Old English *here*, while other borrowings still coexist with their Germanic counterparts. Initially, some 10 000 words entered English in this manner, and three quarters of these have stood the test of time and remain a part of the modern-day English word-stock.

Similarly, several centuries later, a substantial number of words came to English from Latin. The majority of these borrowings entered English in the 16th and 17th centuries as a by-product of the Renaissance, which brought an increased interest in the sciences and ultimately lead to the dominance of classical Latin, the international language of culture and learning, whose influence permeated most spheres of language including philosophical and religious discourse, legislature, and education. According to Görlach (1991), the Early Modern English period was witness to an unprecedented influx of new vocabulary, often resulting in duplicity of meaning, which we can safely argue to be one of the key features directly contributing to lexical mortality later on in that period.

Alongside language contact in the more learned spheres, there was also a great deal of trade, exploration, and colonialism going hand in hand with the expansion of the British Empire. As Leith (1996: 180-1) argues, the process of colonization by speakers of English began centuries before what is generally considered to be the case, i.e., early 17th century. The (predominantly forcible) spread of English in Wales, Scotland and Ireland began as early as the 12th century and continued well into the Early Modern period, leading to extensive contact with Celtic languages and the assimilation of some of their vocabulary into English (*whisky* is a notable example of this). As for colonies beyond the borders of the British Isles, the first were established by the end of the 16th century. Leith (1996: 181) explains that Britain's motivation was manifold; the likely primary incentive was economic, as the government could tax any commodities that were produced or acquired, but there were also political motivations boosted by rivalries among the European states, and lastly social reasons, since the prospect of shipping off any undesirable citizens must have surely been an appealing one.

Political and economic implications aside, the effect of Britain's colonization on English vocabulary was not negligible. Not only did it eventually lead to an array of new varieties of English in the places where new English settlements formed, but also a number of borrowings into English came about from contact with the native languages. In the Early Modern period, the

colonies of North America and Australia were already among the most notable settlements. In these places, "where Europeans largely displaced the precolonial populations, the influence of the original local languages on English was slight," (Leith 1996: 197) and so the loanwords from these indigenous languages were mostly limited to endemic flora and fauna, local customs practised by the native people, and also place names *(racoon, koala, Dandenong)*.

3.2.2 Religious and political factors

In order to gain a deeper understanding of the interplay of external factors leading to changes in the English lexicon, we must take into consideration the effects of the social and political situation in the Early Modern English period. In the 16th and 17th centuries, the tensions running between Catholics and Protestants were felt most acutely and contributed to a series of political upheavals. The events taking place in this period affected all aspects of life on the British Isles, including how language was used; among these we may count the English Civil War of 1642-1651, the short-lived reign of Cromwell's republican government, the restoration of the Stuart monarchy in 1660 and the Glorious Revolution of 1688, which brought England under the rule of new king William III, whose origins can only be described as Dutch.

One of the key events leading to changes in vocabulary usage was the English Civil War beginning in 1642, a period characterized by the breakdown of censorship and the uncontrolled publication of books and pamphlets by whoever had the financial resources to do so, with as many as 700 newspaper titles published in 1645 alone – in fact, it is estimated that 22,000 pamphlets and newspapers were published in the years 1640 – 1660 (Knowles 1997: 97). The lack of censorship made it possible to disseminate texts by a large number of authors, which would have inevitably increased the degree of variation present in printed English at the time, whilst simultaneously making it accessible to a wide reading public.

The effects which the revolutionary years had on the English language were viewed negatively by commentators at the time, and it was said that the turmoil resulted in a chaotic jumble of new words which had been mindlessly coined and adopted: Sprat (1667: 42) argues that the English language generally improved until the Civil Wars when "'it receiv'd many fantastical terms, which were introduc'd by our Religious Sects; and many outlandish phrases'. He sees the possibility of solving the problem: 'set a mark on the ill Words; correct those, which are to be retain'd; admit, and establish the good; and make some emendations in the Accent, and Grammar."

Following the Restoration in 1660, the reign of the Stuart dynasty was somewhat short-lived, as James II was effectively ousted by his Protestant, Dutch nephew/son-in-law in 1688 in what was referred to as the Glorious Revolution (sometimes also Bloodless Revolution). The change of power, incited by those who desired to protect their private property, laid down the foundations of a constitutional monarchy, where the royal power is not unlimited and the true law-making was in the hands of those who owned property (McInnes 1969: 93). The

introduction of the Dutch-born monarch William III, who ruled in the years 1689-1702, resulted in the increased influence of the Dutch language in the English court. William himself demonstrably wrote in Dutch, although it was heavily influenced by French, his first language, and there is good reason to believe that he spoke it, too (Joby 2015: 271). Of course, the presence of Dutch diplomatic activity in England predated the instatement of William III by several centuries and although it was customary for diplomats to communicate in French or Latin, the written correspondence of Dutch diplomats was always predominantly in Dutch, and there is even an instance of an Englishman writing diplomatic correspondence in Dutch (Joby 2015: 235-243). Here we might also mention the prestige that is often associated with the language of the bourgeois and ruling classes, perhaps leading speakers to adopt and spread the prestigious variants, whilst curtailing those that might signal a lower social status. This prestige associated with the language of the court was evident during the reigns of Elizabeth I and Charles II as was attested by Dryden, but much less so following 1688 when England gained a succession of rulers who had no special relationship to English, if they even spoke it at all (Knowles 1997: 120).

3.2.3   The King James Bible

When speaking of prestige and status, what could be more influential than the single most popular work of literature, in arguably the best English translation that had thus far been produced? The King James translation of the Bible was born of the necessity to replace the existing versions which were either seen to contain erroneous translations (such as the Bishops' Bible, which originated during the reign of Elizabeth I) or, as was the case of the Geneva Bible, "notes that were critical of the authority of monarchs." In fact, the initiator of the entire project, John Rainolds, proposed "there might be a new translation of the Bible, because those that were allowed in the reigns of Henry VIII and Edward VI were corrupt, and not answerable to the truth of the original" (Campbell 2010: 34). The ultimate goal of the translation project was to produce an authoritative version which would reinforce James I's power while uniting all religious parties, and the fact that it was translated into a contemporary variant of English meant that anyone who was literate could have relatively easy access to it (Knowles 1997: 94).

The publication of the King James Bible in 1611, often known simply as the Authorized Version, was a great success and quickly became the most widespread version in Britain arguably until the 20th century, when the New English Bible was published in the 1960s (Knowles 1997: 95). It introduced into English a variety of new phrases and metaphors, to quote but one, *"the fly in the ointment."* As a text that was read and absorbed on a daily basis by much of the English-speaking population, one cannot underestimate the extent of its influence on literacy and asserting the role of the written vernacular. As regards other religious texts being circulated in Early Modern England, those of the Protestant faith in particular were in favour of making the Scriptures accessible to speakers of the vernacular, leading to "a popular religious

culture increasingly [becoming] built around cheap print," (Southcombe and Tapsell 2010: 133) which facilitated the dissemination not only of ideas but the vernacular itself, in all its variation.

3.2.4 Politeness

As the political climate changed, the power and even prestige shifted to the middle class, and as a consequence the concept of "politeness" became relevant not only to the courtiers but to the elite in general, as a "means of social classification, demarcating the upper stratum" (Klein 1993: 36-7). It is a concept that gained traction towards the close of the Early Modern period in particular was "politeness," though the phenomenon is still closely linked to British culture to this day. It has somewhat narrowed in meaning; in the late 17th and 18th centuries "politeness" denoted an agreeableness in conversation first and foremost, but it also meant a refined manner of speech, appearance and behaviour. Conversational politeness in particular influenced the choice of vocabulary that a speaker would use, and was defined as easy, soft, polished, natural – "the free, the sweet, the agreeable, the amusing and the open-hearted, open-minded and open-ended" (Klein 1993: 33) as opposed to terms used by the less educated working classes, as well as those deemed overly "bookish," complicated and therefore inaccessible expressions. This understanding of politeness soon broadened and applied to written language, too.

As a result, those words which were considered inferior or unrefined were abandoned in favour of more prestigious alternatives. We may identify social prestige as the driving force behind this tendency, since "a society that interprets variation in speech in terms of "correctness" will understandably give a social evaluation to the variants themselves. Innovations are classed as vulgar or polite, and archaic forms as vulgar or quaint, according to the prestige of the people who use them [and] it is the polite forms that spread from London to middle-class speech in the towns." Knowles (1997: 128) By the same token, what was provincial came to be viewed as unrefined, primitive and vulgar, whereas "the Town" was held in high regard for its politeness (Klein 1993: 40-1).

The emphasis on politeness that can be witnessed throughout the Early Modern period encouraged the propagation of euphemistic language as a means of softening potentially offensive terms. In this sense, politeness and taboo are two sides of the same coin, namely the societal taboos surrounding subjects such as sexuality, bodily functions, illness, and death. Menstruation, a subject enveloped in shame and euphemism (regrettably, such attitudes still persist to this day), was a regular occurrence in the everyday lives of most women, one might think omnipresent verging on mundane, yet it was referred to in euphemistic language such as "flowers" (Read 2008: 8). In fact, medical writers went to such lengths to avoid using the offensive words associated with menstruation, to the extent that they wrote about it in Latin: "Just as, when writing in English, medics would often use Latin to discuss things which might appear sexual, so too we see the use of Latin to discuss sanitary protection in the previous quotations, reinforcing the idea of menstrual blood as an unfit subject for open discourse" (Read

2008: 3). People avoided using direct or explicit language related to such taboo subjects, which not only resulted in semantic shifts and the appearance of new words, but also potentially led to the obsolescence of words originally associated with the undesirable and vulgar concepts.

3.2.5 Learned discourse

The Early Modern period bore witness to a significant shift in how learned discourse was perceived, most saliently the language of science, medicine, and law. The preferred style of writing in the Elizabethan period constituted an elaborate, aristocratic style which mirrored the trends related to lexical borrowing. This extended to scientific writing, where foreign terms were lavishly distributed, and such practises were not without their critics, most notably perhaps Thomas Sprat, whose History of the Royal Society (1667) criticized "the easie vanitie of fine speaking," "this vicious abundance of Phrase, this trick of Metaphors, this volubility of Tongue" (Knowles 1997: 110). This can be understood as the writers hiding behind opulent phrases and ornate language in the hope that their otherwise subpar productivity would not be recognized. Sprat desired to "return back to the primitive purity, and shortness, when men deliver'd so many things, almost in an equal number of words" (Sprat 1667: 111-3). Indeed, the purity and simplicity of discourse was what defined the latter part of the Early Modern period, in the years following the Restoration. Correspondingly, a number of earlier superfluous terms – mostly coinages and borrowings – would go on to become obsolete, as will be shown in later chapters.

The year 1660 can be viewed as a milestone for the written form of the English language, specifically from the stylistic perspective: "Sixteenth-century texts were influenced by Renaissance rhetoric, and beliefs about style and vocabulary. After the revolution, a more utilitarian approach was taken, and the meaning was conveyed in simpler language. At the same time, we can trace the beginnings of a new intolerant approach to language, in particular other people's use of language" (Knowles 1997: 103). Following the year 1660, a heightened interest was taken in science and consequently scientific discourse, which called for a clear and more easily accessible written style. The conviction that there is an indisputably right and wrong way of using language fuelled the rise of the normative tradition in the English language, which is discussed further in chapter 3.2.7. The existence of these new attitudes to the use of English implies a radical process weeding out of superfluous and "outlandish" phrases, a course of action advocated by scholars such as Thomas Sprat and one which may have contributed to many cases of lexical obsolescence. The tendency towards a plain style originated in the domain of scientific writing, but it went on to permeate most other spheres where written English was used and became the norm: "Although the need was specific, the effect was general, and for the next hundred years plainness of style was to be the outstanding feature not only of scientific writing but also of a wide range of text-types from published books to government decrees and private papers" (Knowles 1997: 110).

Along with the growing variability of functions that the vernacular could fulfil and the resulting array of new styles and text types (Görlach 1990: 117), there was a growing availability and circulation of news in print which provided yet more fertile ground for the proliferation of learned discourse. Although the demand for newsprint was undoubtedly boosted by warfare both in Britain and abroad, it was also increasingly culturally motivated – the emphasis on the value of education and the possibility of social mobility in the Early Modern period was unprecedented. Having knowledge of current events and, indeed, having the ability to produce news writing was highly valued, as Southcombe and Tapsell (2010: 133-134) acknowledge, "there was a 'decorum' of news: knowledge became a means of social differentiation that had previously only been visible in dress-codes designed to mark out different strata of society."

A special subcategory of learned discourse is the language of the law, characterized by features such as "legal terms (*manslaughter*), Latin words (*mensrea*), binomials (*null and void*), heavy use of negative and passive structures, and complex subordination in sentences" (Tiersma 1999). In their paper, Lancashire and Damianopoulos (2014) stress the contribution of the legal profession to Early Modern English lexicography. In fact, contrary to popular belief, the publication of the first monolingual dictionary of English should not be attributed to Robert Cawdrey in 1604 but rather John Rastell, who in 1523 "glossed terms of Common Law in the original law-French and separately translated the whole glossary – headwords and explanations – into English" (Lancashire and Damianopoulos 2014: 31). Driven by his belief in the importance of education and accessible legal knowledge, Rastell dedicated himself to the task of making the law of the land understandable to the English-speaking lay public and bringing a number of words into more common use. In fact, Lancashire and Damianopoulos propose that by including an increasing number of headwords in their dictionaries of legal terms, writers and scholars such as John Rastell were in part responsible for expanding the lexicon of the English language (2014: 39).

The first edition of his dictionary contained 165 "obscure and derke termys consernyng the lawis of thys realme" for which Rastell provided not only definitions but practical examples, such as the term "Homage" in which he walks the reader through the procedure itself rather than giving a dry definition. The language of the law was heavily reliant on Latin terminology, much more so than it is today, with a number of Latin phrases originating the Early Modern period, such as *rectus in curia* (1611–present; having full legal rights). The dictionaries and glossaries of the time reflect this, with most of the terms being of French or Latin origin, although words in "plain English, like 'Arest', 'Disceyt', 'Homage', 'Proteccion' and 'Treason'" were no exception. (Lancashire and Damianopoulos 2014: 33).

3.2.6 Philosophy's views on language

With regard to language change and, by extension, lexical obsolescence, external factors such as social networks, language contact, political and economic events are all understood to be vital for the interpretation of possible causes. The changing attitudes towards English played a great role in how the language was used, and Barber identifies "national feeling" as one of the driving forces, igniting "pride in the national language, and […] attempts to create a vernacular literature to vie with that of Greece and Rome" (1997: 45). Seeing as attitudes towards a language and its speakers stand at the centre of all this, it would be an oversight not to include the role of contemporaneous philosophy in the equation. Philosophy and ideology cannot be studied independently and without context, but rather should be viewed as a product of the society in which these ideas originated. Likewise, the state of society can be seen as a reflection of the currently dominant philosophies. Philosophy matters in day-to-day life, as it is concerned with real issues of the current society, and little else could be more immediate than the words that a society's population uses on a daily basis. This chapter examines the key ideas and attitudes towards language in early-modern philosophy, and how these may be applied in order to shed further light on the reality of lexical obsolescence and loss.

Dealing with the workings of the human mind and the process of acquiring knowledge, John Locke's Essay Concerning Human Understanding (1689, edited version by Nidditch 1975) addresses topics such as truth, notions, judgement and reason, and may thus provide invaluable insights into early-modern thought – that is if we choose to understand this text as a product and reflection of the time in which it was created. In Book III, Locke presents his views on language and meaning. As one of the most significant proponents of the ideational approach (Lowe 2005: 99), Locke asserts that using the sounds of one's voice, man is able to articulate "the ideas within his own mind, whereby they might be made known to others, and the thoughts of men's minds be conveyed from one to another" (Locke 1689: 176). In other words, Locke asserts that the primary purpose of language is communication, specifically the communication of the speaker's ideas to the mind of the hearer.

Locke's words are wholly dependent on the understanding between speakers and, "in their primary or immediate signification, stand for nothing but the ideas in the mind of him that uses them" (1689: 178). As one may expect, the arbitrary nature of signification in a language can be a cause of misunderstanding, because the understanding of reality and what a given word signifies varies from speaker to speaker. To remedy this, Locke suggests that in communication one should use only words with clear and concrete ideas associated with them, and also "use words consistently and […] not equivocate". Locke's desire to avoid using ambiguous language ties in with the need for a limited number of precisely defined lexical items. When applied to the situation in Early Modern English, where it was not uncommon for three near-synonyms to exist simultaneously, one can easily understand how unsustainable the situation seemed. If we consider Example 1 from the previous section, *sacerdote* is the Latin word for priest which

appeared in 17th century English. The simultaneous use of these two absolute synonyms would have led to inconsistency and, in the case of speakers not familiar with the non-native borrowing, the obscurity of meaning – in other words everything that Locke was speaking out against.

Language naturally tends towards clarity and economy, which is why naming every individual thing would be an impossible task and it would not aid communication at all, because "it is beyond our power to form and retain separate ideas of all the particular things we meet with: every bird and beast that men have seen, every tree and plant that has affected the senses, couldn't find a place in the most capacious understanding" (Locke 1689: 180). This clear tendency towards the generalization supports the weeding out of unnecessary synonymy from the language and could account for a large portion of now obsolescent words. For instance, Example 2 shows that in the 17th century *pabulation* was a recently introduced Latin synonym for feeding, grazing or fodder. Seeing as the English language had sufficient lexical means of expressing the act of feeding, adding another synonym to the lexicon would have been more superfluous than the language could bear. Even though one word may have several different denotations, Locke postulates that the context alone should suffice when there is no time for the speaker (or writer) to provide definitions and explain his intended meaning, because "the import of the discourse will, for the most part […] sufficiently lead candid and intelligent readers into the true meaning of [the word]" (1689: 223).

It is the case in any language that some words can gradually become emptied of their meaning, either through overuse, or through the disappearance of the original referent from the language. Words are unavoidably ambiguous and definitions are hazy, because although an idea in one speaker's mind corresponds with a specific word, the very same word does not necessarily correspond with the exact same idea in the mind of another (Locke 1689: 204). Locke warns that it is often the case that men learn words as empty sounds without first knowing the concrete ideas or meanings behind them: "[S]o far as words are of use and signification, so far is there a constant connexion between the sound and the idea, and a designation that the one stands for the other; without which application of them, they are nothing but so much insignificant noise" (Locke 1689: 179).

There is a distinct sense of continuity with the writings of Francis Bacon, who distinguished between two types of idolatrous words. Firstly, "'names of things that do not exist' [...]. While these can simply be rejected, the second type is 'obscure and deep-seated' because it concerns everyday and indispensable words whose foundations we do not question. They are names of qualities like 'moist' – 'which do exist but are muddled and vague, and hastily and unjustly derived from things'" (Dawson 2007: 122). This vagueness of meaning is widespread and therefore cannot be eradicated, leaving the speaker once again to rely on knowledge of contexts, or alternatively requiring him to overtly explain his intended meaning.

One of the key points that Locke makes in his Essay is when he comments on the abuse of language, which may come in many forms, but the most salient are the irresponsible coinage of new words and the inconsistent or incorrect use of words (Locke 1689: 208-9). When philosophers or scholars coin a new term in attempts to cover gaps in their theories, it results in a sign with no idea behind it. One may expect coinages, especially those created for a very specific purpose, to be less likely to stand the test of time and are in danger of becoming obsolete within a matter of centuries. Not so much a coinage as a literal translation from Latin, *quicquidlibet* (Example 4), meaning whatever one pleases, anything whatsoever, is an excellent example of this. Superfluous and limited to specific genres and contexts, this was an addition to the lexicon which was destined to fail from its very introduction to the language. The second type of language abuse deplored by Locke is inconstancy in the use of words (with deliberation and for one's own ends) "either applying old words to new and unusual significations; or introducing new and ambiguous terms, without defining either, or else putting them so together, as may confound their ordinary meaning" (Locke 1689: 209). Although it is clear enough that this perspective is mainly concerned with the truthful delivery of a speaker's message, Locke's wording furthermore suggests that he was not a supporter of the semantic shifts taking place in his time. Resistance to semantic shifts, as well as the broadening or narrowing of a word's meaning, is understandable especially owing to one's need for clarity, nevertheless it would appear that this process could not be stopped, in particular due to the influx of Latin loanwords.

In order to be functional, a language must reach an optimal balance between over- and under-specification, which is to say that although words represent ideas in the mind, it is necessary that they cover a much wider, more general sense of the meaning (Locke 1689: 176-7). Furthermore, Locke introduces a very basic theory of etymology, positing that all words originate form sensible ideas with the use of metaphor, in other words there is an abstraction from the primary signification, for example angel originally meant a person who carries news (ibid). Coincidentally, this semantic shift allowed the slot for the original meaning to be filled by the French borrowing, messenger.

When studying obsolescence and word loss, we can witness that the choice between two competing forms is sometimes arbitrary, but in other cases there is a certain predictability – a pattern which may be discerned once we've taken into account all of the intra- and extra-linguistic variables. The preference of one form over another cannot always be simply explained, nevertheless an understanding of society's attitudes towards language and the prevalent ideology of the time may, for example, clarify the preference of a native variant over the more scholarly borrowing. When confronted with a choice between two synonymous forms, one Latin and one English, one of the factors determining which variant wins out will be the attitudes of the speakers and their preferences. To this end, the effects of morphological complexity, word origin, genre and regional use also need to be addressed. Equally vital is for a form's survival is the word-idea-object dynamic, a topic of focus throughout the history of philosophy: 'If at any

time we reason about words, which do not stand for any ideas, 'tis only about those sounds, and nothing else' (Locke 1689: 178-9). Ultimately, this comes down to the word's ability to justify itself in the minds of the speakers, and the strength of its link to the concept which it represents.

In the 18th century and even in the years prior, discourse concerning language change was mostly limited to scholars who addressed the issue as commentators rather than philosophers. Indeed, as we look back into the 16th and 17th centuries and consider the sheer quantity of words which entered the English language, the struggle as one word competed for dominance and the other fought for survival, and the ultimate obsolescence and loss of words which were unable to justify their place in the language, it is a wonder that early-modern philosophers did not give this topic the attention that it indubitably deserves. Locke's attitudes towards and ideas about language can be understood as a representation of at least one facet of 18th century thought. Ultimately, standardisation and normativity are considered necessary within a functioning society and this sense is reinforced by the English normative tradition and a continuation of prescriptivism in the 18th century.

3.2.7 The English normative tradition

Throughout the Middle Ages, Latin was regarded as the ideal, a language in its purest form and although the Renaissance witnessed a shift in attitudes in favour of the vernacular, English, a substantial number of proponents of Latin remained (Görlach 1991: 36). In the 16th century, it was not uncommon to adopt words from Latin without any augmentations and use them in an English text. This practice was not without its opponents, with one group of conservative scholars claiming that only Latin should be used in writing, and on the other hand radical proponents of the English language, who believed that the necessary sentiments could be expressed either by restoring outdated terms or by coining new words using native components (Knowles 1997: 70). These beliefs were put into practice by authors and translators such as Sir John Cheke, whose inventory of native coinages included *crossed* (meaning *crucified*) and *hunderder* (for *centurion*). His words in reaction to Sir Thomas Chaloner's translation of *The Courtier* are a testament to his dislike of borrowings from foreign languages:

> "I am of the opinion that our own tung shold be written cleane and pure, vnmixt and vnmangled with borowing of other tunges […] For then doth our tung naturallie and praisablie vtter her meaning, whan she bouroweth no counterfeitness of other tunges to attire her self withall, but vseth plainlie her own." (Fisiak 1993: 99)

The debate was lively and widespread, with strong sentiments on both sides, and so "neither tradition nor innovation was allowed to emerge the sole victor from the language debates that characterized the mid to late sixteenth century, in which both archaism and neologism were attacked" (Crawforth 2013: 28). In any case, there was the pervading opinion that the borrowing

of Latin terminology into English had gone too far, resulting in the *inkhorne controversy,* "so-called for the horn that held the ink, thus suggesting how such terms were used by people who were highly literate" (Millward and Hayes 2012: 226). Although writers such as Thomas Wilson were open to the use of loanwords and other borrowings, they felt that these words were being used even when an existing equivalent was readily available, and in such a high concentration that the text's meaning was often obscured. Wilson's comment gives a testimony of this, as John Locke had put it, abuse of language: "I dare swear this, if some of their mothers were alive, they were not able to tell what they say" (Millward and Hayes 2012: 226). In reaction to this, there was a heightened interest in the study of Old English, which many 15th- and 16th-century speakers considered to be purer than the status quo, and therefore something to be valued (Considine 2008: 158).

As we move into the latter part of the Early Modern period, there is one remarkable paradox which bears mentioning, and that is the duality of spelling that characterized the written word in the 18th century, standing in stark juxtaposition with the general prescriptivist tendencies of the period. Personal letter-writing, a popular pastime of the literate, had vastly different spelling conventions, or rather lack thereof, in contrast with texts printed by a publisher. Even dictionary author Samuel Johnson is known to have used vastly divergent spelling in his private correspondence as opposed to that found in his Dictionary. And although every printer would have modified the spelling to comply with their house rules, the educated authors were evidently aware of the two systems and conformed to the public one in their formal correspondence. Not only spelling but grammar, too, was subject to this widespread phenomenon, for instance Tieken-Boon van Ostade (2006: 259-260) mentions the incongruent *you was,* the consistent use of *they* in the place of a singular pronoun (*anyone may do as they please*) and the nominative I used where an accusative me was expected (*between you and I*). This was commented on and condemned by normative grammarians at the time, although it appears that they were more successful with some phenomena (*you was*) than others (*you and I*) which persist to this day (Tieken-Boon van Ostade 2006: 255-260).

3.2.7.1 *Eighteenth-century prescriptivism*
By the end of the Early Modern English period, the language reflected the social and political unrest, which was brought by years of civil war, the Reformation, the Scottish and Irish rebellions. There was a general sense that language was in a state of chaos, with too much creativity and unrestrained freedom. This feeling was expressed, among others, by Samuel Johnson in the 1755 Preface to his Dictionary and, in fact, the authority exerted by the Dictionary following its publication was a substitute for the language academy that many of the public were campaigning for (Millward and Hayes 2012: 237). Hickey (2010: 1) describes the 18th century as "the period in which prescriptivism in English established itself," which is to say that there was an increased pressure to use the forms deemed most desirable. Prescriptivism affected all facets of the language, not least of all vocabulary.

A significant number of scholars and intellectuals considered innovation in the English language to be evil and corrupt, and it was an "almost universally" held view that further modification must be stopped (Fisiak 1993: 118-9). Correcting the language was indeed a priority, and rules were stipulated which defined what was to be perceived as correct or incorrect usage, the coexistence of native and non-native words led to a sharper sense of stylistic differentiation, and regional variants led to prejudice against provincial and dialectal speech. Ultimately, with the absence of an official academy, the grammars and dictionaries which "embodied and transmitted codification" came to be regarded as unquestionable authorities, thereby further strengthening the prescriptivist tradition (Beal 2010: 22).

3.2.7.2   Dictionaries of English

The role that lexicographers and compilers of dictionaries played a key role in the Early Modern usage, predominantly among scholars and authors. Not only did they document the meaning, origins and contexts of a myriad words, but they also commented issues pertaining to language usage at the time, leaving a valuable testimony to the problems and attitudes of the epoch. In the late 16th and 17th centuries, when the influx of borrowings and new coinages had reached its peak, the unprecedented variability offered by these new words was perceived as a threat to the purity of English, an attitude shared by a number of antiquarian lexicographers at the time. It fuelled an intense study of Old English manuscripts and the production of glossaries of Old English words, such as Laurence Nowell's *Vocabularium saxonicum* compiled in the mid-16th century and regarded as the first dictionary of Old English. It was later followed by John Joscelyn's Old English glossary, *Dictionarium saxonico–latinum*, "running to about 22,000 entries and subentries, in which Old English words are generally glossed by a Latin equivalent followed by an English one" (Considine 2008: 164-8).

On the other hand, many 16th- and 17th-century lexicographers saw the burgeoning vocabulary as an asset to the English language, as a means by which it could assume a more elevated style and richness than in the preceding centuries. Robert Cawdrey's *A Table Alphabeticall, conteyning and teaching the true writing, and understanding of hard usuall English words* was published in 1604 and is widely considered to be the first English dictionary (Although, as Lancashire and Damianopoulos (2014: 31) assert, that this precedence should rightfully go to John Rastell for his glossary of legal terms published almost a century earlier). Its purpose was to explain, mostly by offering several synonymous expressions, words which, as the title suggests, were to some extent established in the language, but were difficult for the majority of readers. Its secondary purpose was to educate the readers so that they would not only understand the words, but also be able to use them in their own production (Lamb 2014: 133).

Samuel Johnson was opposed to language change of any sort, which was one of his chief motivations for compiling *A Dictionary of the English Language* (1755) – in the hope that usage would become fixed, he did not include the types of words which would be most susceptible to

change, namely words specific to professions and various specialized usages. Unfortunately for Johnson, these words had been recorded a century earlier, as "these exclusions were the principal inclusions of 17th- century dictionaries. Johnson later sneered that his predecessors had spread such words 'with a kind of pompous luxuriance over their productions,' and compiled his own dictionary as an attempt to clear up the mess they had made" (Lamb 2014: 130).

The commentaries accompanying "hard word" lists and dictionaries compiled in the Early Modern period are a testament of the radical change in perspective between the 17th and 18th century attitudes to variety in English, which is to say that, unlike Samuel Johnson, "[seventeenth-century hard words compilers] did not see a disordered language in need of stability, but new and strange regions of language that needed explanation, and to which English speakers required access" (ibid.: 144).

3.2.7.3 *Grammars of English*
The transition from a relatively free usage of English in the 16th and 17th centuries to the prescriptivist zeitgeist of the 18th century is reflected most noticeably in the quantity of grammars created at the time. Only four grammars were published in the 16th century, and thirty-two in the 17th, as opposed to the 18th century, which witnessed the publication of over 200 grammars (Nevalainen 2006: 16). Not only the quantity of publications but also their content says a lot about the change in perspective – while the early grammars were more open to variability and innovation, those written in the later centuries had assumed a strictly prescriptivist tone: "Early Modern English school grammars professed to exclude regional dialects but did not rule out variation in the General dialect. Clearly it was more diffuse than the Late Modern standard variety codified in eighteenth- century grammars" (ibid.).

Furthermore, many of the earlier grammars were published in Latin and were intended for foreign learners of English, while in the later centuries the purpose of grammars and dictionaries was to improve the state of the language among its native speakers and achieve a stability in the language which, as 18th-century grammarians felt, the previous centuries had been seriously lacking. The general tendency then was towards a suppression of optional variety and "the buttressing of restrictive/normative attitudes and pronouncements with logical or […] etymological or historical argument" (Lass 1993: 104). Such attitudes may be gleaned from contemporaneous commentaries on strong verbs, such as James Greenwood's 1711 essay demonstrating attempts at ideological standardization in the preterite forms of strong verbs, in which he identifies as "not proper or usual" a large number of forms even though they were by no means as unusual as Greenwood would have the reader believe, as is attested by sources from both earlier and later decades (Lass 1993: 104-5).

3.3 Language-internal mechanisms

It is suggested in chapter 3.1 that the dichotomy of internally versus externally motivated language change is a fundamentally misleading one since it is evident that language does not exist in a vacuum. After all, a living language requires speakers, who do not come without their social and political idiosyncrasies. However, I would venture that the mechanisms described below lend themselves to the language-internal perspective in that they constitute tendencies leading to language change that is more systemic and long-term, moreover, these mechanisms may be less accessible to speakers on a cognitive level. In fact, they are not limited to the domain of language at all, but rather they are tendencies that permeate all aspects of human cognitive processing (Bybee 2015: 238-9). Below is a brief and by no means exhaustive overview of some of the mechanisms which come into play.

The *reduction of effort* speaks to our natural tendency to exert the least effort, cognitively and physically, whilst still managing to reach the desired goal, in the case of language the effective communication of meaning. It is demonstrated by McMahon (2010) using the example of final /r/ weakening and loss in some varieties of English. Manifesting primarily as an articulatory or acoustic phenomenon, "weakening processes involve the interplay of characteristics of a particular segment, which may be more or less prone to weakening, in a particular environment, which may be more or less conducive to weakening" (McMahon 2010: 105). Weakening predominantly affects unstressed syllables and specifically codas (the final sound of the syllable) and has been shown to be a major factor in the loss of endings in Middle English for example the final [e] in words like *sun (sunn-e), blind (blynd-e)* and *other (other-e)*.

This ties in with the *automation of production* (Bybee 2015: 238) which is especially pertinent to the process of sound change: "As articulatory production is a neuromotor process, it is subject to the reduction and retiming that highly practiced behaviors achieve through repetition." Bybee cites the automation of articulation as one of the leading sources of sound change. This mechanism has been shown to affect mostly high-frequency items, since they are both familiar and frequent enough to be understood regardless of more variation.

The *effects of frequency*, however, work in the opposite direction. It has been observed that high-frequency items are more resistant to change, as "high token frequency strengthens the mental representation of particular items and makes them resistant to change" (Bybee 2015: 238). An example of this would be the weakening of verbal paradigms from Old English onwards, and it is not coincidence that the verbs which retained their strong conjugation are ones which occur with the highest frequencies in everyday language use.

The inclination towards *generalization* can manifest as the preference for productive or frequent patterns not only when assigning them to completely new items, but even going so far as to replace irregular or low-frequency patterns. However, this does not apply in those weak or irregular patterns that happen to have a high token frequency, e.g., when a paradigm pertains to a

small number of words but one or more of those constitute high-frequency words. In Bybee's words, "the operative cognitive mechanism is the response to type frequency; use of a pattern with different items strengthens patterns and builds up a general category which can easily be extended to apply to novel items" (2015: 238).

Analogy effectively results in a simplification of a given pattern or paradigm within the system. When referring to a single pattern, we speak of analogical extension, which is "the generalisation of a morpheme or relation which already exists in the language into new situations or forms" (McMahon 1994: 70). An example of this is the plural ending <-s>, which became the dominant paradigm somewhere between Middle and Early Modern English. The nominative plural endings in Old English were dependent on the paradigm to which the noun belonged, take McMahon's (1994: 71) examples of the words for *sun* (sg. *sunne*, pl. *sunnan*), *stone* (sg. *stān*, pl. *stānas*), and *ship* (sg. *scip*, pl. *scipu*). However, this system began to break down even towards the end of the Old English period, eventually leading to the extension of the <-s> as a widespread plural marker.

Chunking, the tendency to process more than one item together, such as a phrase or collocation, is yet another instance of humans exerting the least cognitive effort. This is especially true of high-frequency chunks, and "as the chunk is used more and more, it tends to undergo more and more internal phonetic reduction and fusion" (Bybee 2015: 124) and this chunking tendency factors into *grammaticalization*, a complex process that simultaneously affects every aspect of the construction in question; its spelling, pronunciation, grammatical function, and meaning. With a high frequency of use, words or phrases which previously had lexical meaning become grammatical markers, leading to their morphological or phonetic reduction. For example, the future construction "going to" is in informal contexts often reduced to "gonna", both in the spoken and written form. Even more relaxed articulation can result in further reduction to [gə̃ɾ̃ə], where the unstressed vowels converge to schwa, and the /n/ becomes an alveolar flap accompanied by the nasalization of the surrounding sounds. Grammaticalization can also involve expansion of meaning or function, for example the auxiliary verb *can* (OE *cunnan),* which was used only with a handful of verbs and had the transparent meaning "to know or have learned a thing", has now expanded to anything from "be able to" to "be allowed" and "have the option."

Finally, there are the two key principles which are often at odds with one another; *iconicity* and *linguistic economy*. While iconicity works in favour of always being explicit for maximum clarity of meaning, linguistic economy tends to make cutbacks wherever possible. Iconicity is the principle dictating that language should reflect the extralinguistic experience as accurately as possible; in morphology, one form should be mapped onto one meaning; in syntax, the events described should be mentioned in the same order in which they took place (McMahon 1994: 85-6). If this principle is ignored, the resulting ambiguity is often such that the economy principle takes the reins and the consequences are nothing short of disastrous for the offending forms, for

example, "the obsolescence of the verb *let* 'hinder' […] might be ascribed to an unfortunate clash with the more common verb *let* meaning 'allow.'" (Hogg and Denison 2006: 39) This might explain the tendency to avoid homonymy and superfluous synonymy. To give another example, due to the uncontrolled borrowing taking place in the Middle English and Early Modern period, the English language contained a sizeable list of groups of three synonyms originating from Germanic, Latin and French, but the only ones that survive into present-day English are those that have been confined to specialized stylistic contexts or shades of meaning (Crystal 2002: 194).

4   Methodology

First, using frequency and distribution as the main criteria, I will attempt to establish an inventory of items which were relatively common at the beginning of the 16th century but became obsolete towards the end of the Early Modern period, e.g., word forms such as *flagitious*, word-formation patterns such as *rivalship*, or parts of inflectional paradigms such as *properest*. In order to find such forms, the structure of English word-stock will be examined to correlate relative frequency of a form to its intuitive commonness from a native speaker's point of view. Next, a corpus-driven methodology (devised to this purpose in earlier research) will be applied to a large set of data for Early Modern English. In previous studies, this was implemented on a corpus of over a hundred billion tokens of English text from the period 1700–2000, while in the case of this thesis I will be working with the 800-million-word corpus Early English Books Online.

In the next chapter, I will provide an overview of the most salient candidates for lexical obsolescence, and the conditions signaling their decline will be selectively analyzed and discussed. Close attention will be paid both to the related changes in language structure and relevant language-external processes.

4.1   EEBO and EEBO-TCP

Early English Books Online (EEBO) is a collection of digitized manuscripts and early print books which was first made available online in 1998; In the preceding years, copies of these published materials had been recorded on microfilms. The University of Michigan began to explore solutions which could provide a searchable version of the digitized texts, a development which would significantly aid work with larger quantities of textual data (Sandler 2003: 47). This prompted the founding of the Text Creation Partnership, which is a unique collaboration between a corporate company ProQuest and over 150 public libraries and institutions (Welzenbach 2012: 4), which co-funded the digitization of over 25,000 printed books in Early Modern English.

As a result, EEBO and EEBO-TCP coexist as two virtually separate projects, and each has its indisputable advantages over the other. The EEBO collection allows users to view digital copies of the early print books, of which there are currently over 125,000 titles included in the collection, while the EECO-TCP project compliments this feature by offering a corpus of searchable texts which are available to members of academic institutions online via a number of platforms, including that of the University of Michigan, Lancaster University, and Charles University. The corpus can be accessed free of charge by registered users, however in most cases this pertains only to the original unlemmatized version of the EEBO corpus. Further reading on digital humanities and TCP in particular, refer to Lavagnino (2012), Baron and Hardie (2012), Martin (2007 and 2006).

Due to the quantity of texts being digitized and the time required to complete this task, the EEBO-TCP corpus exists in two versions. Phase I was completed in 2009 and contains over 25,000 transcribed texts, and the ongoing Phase II has undertaken to digitize an additional 45,000 unique publications. According to the EEBO-TCP web, the number of books encoded in Phase II as of 2014 was 28,466[3]. The documents from Phase II are available at selected institutions, for example Lancaster University's EEBO v3 corpus contains a total of 44,422 texts[4].

4.2 Building the corpus

The EEBO-Text Creation Partnership has created a collection of accurately converted texts which had been carefully transcribed by hand in several phases (Welzenbach 2012: 2-3) and encoded in SGML/XML. Original first editions of the texts are used, which ensures that the resulting corpus should contain an accurate representation of the books exactly as they were printed, typos and all. Although the transcriptions were checked by several editors, there is always a margin of human error to be reckoned with, and as a result we can never know for certain whether the typo can be attributed to the transcriber or whether it was correctly copied as such from the original text. The possibility of progressing from transcribing texts by hand to an OCR-aided method of digitization has been addressed, which would significantly lower the costs but could ultimately result in an increase of incorrectly recognized symbols.

As for manual transcription, the endeavour is not without its challenges, as is attested in Sandler's (2003: 47-48) summary of the problems which the EEBO-TCP transcribers and editors are faced with:

> "Spelling variants abound; hyphenation is not standardized; ornamental letters, illustrations and glosses interfere with the flow of text; macrons are used (or not) to indicate abbreviations or elisions; Greek, Hebrew and other non-roman scripts appear with frequency; letters, words and sometimes pages are lost because the original print strike was uneven, the microfilm image was faulty, the digital image was not correctly optimized, or bleed-through from the verso page obscures print on the recto. Add in alchemical and astrological symbols, diacriticals, superscripts, e-hooks, ligatures, and other oddities of early printing, and you can begin to imagine the challenge facing the keyboarding firms and those reviewing their work against a standard of 99.995 percent character accuracy."

4.2.1 POS tagging

Once the Phase I of the EEBO corpus was completed, some institutions chose to lemmatize or otherwise annotate their own versions. For example, the corpus at Northwestern University was

---

[3] https://www.textcreationpartnership.org/tcp-eebo/
[4] https://cqpweb.lancs.ac.uk/eebov3/index.php?ui=corpusMetadata

annotated using the widespread NLP tool MorphAdorner, which automatically assigned a lemma and a part-of-speech tag to each token (Mueller 2012).

Later when working with the corpus, it is important to keep in mind the fact that there is an error rate to be reckoned with in the assignment of these part-of-speech tags, which should come as no surprise as unerring accuracy cannot be expected from any automated NLP tool (Bauer 2002: 106). Such tools are trained using a set of manually annotated texts from which they create a statistically based language model, which is then used to disambiguate and tag any quantities of text based on the rules that it has learnt from the training data. Even so, the accuracy achieved with modern taggers such as MorphAdorner is relatively high, resulting in approximately 97% correctly tagged words (Mueller 2009: 14). Many factors come into play, for instance the indistinguishability between two identical yet grammatically divergent forms or the variability of word order, which can result in the error rate being as high as 15-18% in some specific cases, as it is for the distinction between the past participle and past tense, which are identical with the exception of a handful of strong words (ibid.). As Mueller suggests, this high error rate is partly caused by the fact that NUPOS tag set was not specifically designed for Early Modern English texts, although many of the incorrectly assigned POS tags and lemmas can be attributed to the variation in spelling and the presence of typos in either the original or digitized texts.

In addition, some inaccuracies are caused by words on the borderline between two (or more!) word classes, for example *however,* which can be classified either as a conjunction or adverb. This is typical especially of grammatical words (prepositions, conjunctions), but it is not uncommon in the case of lexical words, for example *this fountain contains drinking* (adj.) *water – drinking* (n.) *and driving is illegal – he was drinking* (v.) *his coffee.*

The accurate assignment of POS tags is relevant for the purposes of this thesis, since I will be interested in identifying cases where the given form became obsolescent only within the scope of one word class but has remained extant in another, for example the now obsolete *borrow* (n.) or *otherwise* (n.).

4.3   Data structure

The data used in this research has been obtained from Northwestern University with the kind consent of Martin Mueller. The tagged and lemmatized corpus is based on the EEBO's TCP Phase I. The data set which I am working with contains 685,131 types (unique words) which had been mined from the EEBO corpus of approximately 800 million words. The data comprises words which appear in at least five documents; therefore, the analysis was actually conducted on a relevant section of the EEBO corpus. However, for this purpose it was not be necessary to gain access to the corpus in its entirety, due to the need for our obsolescence candidates to have a set minimum frequency in the initial decades of the EModE period, and thus the words *not* included

in this data set would have had frequencies too low to begin with. As for the relative (p.p.m.) frequencies, they were calculated with regard to the size of the whole EEBO-TCP corpus.

Each word in the corpus contains a tag from the NUPOS tagset[5] which assigns it to a specific word class. The tags are appended to each word via an underscore and will prove useful in the analysis of the results, especially the tags for foreign words. According to the online documentation, the Trigram tagger was able to identify the correct part of speech tag with an accuracy of 96- 97%[6]. The complete list of tags and their interpretations are provided in Table 1 below.

Tag	Interpretation
_f-la	foreign word – Latin
_f-fr	foreign word – French
_f-mi	foreign word in unspecified other language
_f-it	foreign word – Italian
_f-ge	foreign word – German
_uh	interjection
_av	adverb
_n	noun
_nn	proper noun
_v	verb
_j	adjective
_jn	adjective/noun
_pn	pronoun
_d	determiner
_crd	numeral
_zz	undetermined
_ab	abbreviation
_sy	symbol
_acp	adverb/conjunction/particle/preposition

Table 1: List of tags used in the EEBO-TCP data

---

[5] The full documentation is available here:
http://morphadorner.northwestern.edu/morphadorner/documentation/nupos/
[6] http://morphadorner.northwestern.edu/morphadorner/postagger/

The objective of the EEBO-TCP corpus project is to be as accurate and comprehensive as possible. However, a number of challenges have been encountered thus far and will be described in more detail in the following section. The issues include the disproportionate distribution of tokens over the decades, unbalanced representativeness of the corpus with relation to genre and text type, and incorrectly recognized symbols during the OCR processing.

4.3.1 Distribution of tokens by decade

One of the main disadvantages of this corpus is its uneven distribution of the texts throughout the years, which results in the initial decades being severely under-represented. In fact, the most prominently represented time period 1650-9 contains more tokens than the first ten decades combined. It will therefore be necessary to normalize the data (operating in relative, not absolute, frequencies) in order to yield representative results.

Figure 5: Number of words (in millions) per decade in EEBO-TCP

4.3.2 Representativeness

The representativeness of EEBO-TCP is an essential issue which we must address, because the aim of this research is to provide a comprehensive overview of the developments in Early Modern English word stock. This may not be possible, because although the list of authors is available, there is no information regarding the genres (text types) or regions of origin. Furthermore, all of the corpus data have come from published texts, i.e., we must consider the effect of genre as well as the time lapse between the writing and the publishing of the text.

Another research project based on EEBO-TCP data encountered a significant number of duplicates, as summarized in Figure 6 (Popelíková and Volná 2018). An examination of the sources suggested that the duplicated data are a result of the corpus containing several different editions of the same text, but reprinted years apart, as well as quotations from other publications which themselves are also included in the corpus. The authors were able to search their query results manually in order to find and remove duplicates. However, a more extensive study such as this thesis would undoubtedly benefit from the removal of duplicated sources from the corpus itself.

Figure 6: Duplicates in EEBO data pertaining to reflexive pronouns, ordered by decade

4.3.3   Incorrectly recognized symbols

As the EEBO-TCP corpus contains only original print editions, the text could not always be decoded with ease. In instances where the text was not completely legible or for any other reason could not be properly recognized by the transcriber, the corpus uses a placeholder in the form of the symbol ● (U+25CF in Unicode). At this stage, these words are not being included in the analysis, as it is not clear how they could prove useful. Such words appear in the corpus frequently (a total of 60865 types, see examples 2a-2g).

(2a) wa●_n

(2b) ma●e_n

(2c) ●●pe_n

(2d) af●●r_d

(2e) ●●●●●ing_v

(2f) ●●●●●ly_j

(2g) ●●●●●ment_n

In Table 2 is a list of some incorrectly recognized words which contain a symbol representing several graphemes. There are cases, such as those illustrated in the table, where it would be possible to consolidate them with the correct variants which are also included in the corpus, for example the verbs *ɡclude* (10 hits) could be subsumed under *conclude* (218006 hits) with a total of 218016 hits.

ɡ = <con-> or <com->	**ꝑ = <pro->**
ɡclude_v	ꝑprietas_f-la
ɡmaund_v	ꝑsperitie_n
ɡmaundement_n	ꝑsperous_j
ɡpany_n	ꝑtection_n
ɡsent_v	ꝑtector_n
ɡsequent_j	ꝑtestation_n
ɡsider_v	ꝑuidence_n
ɡtinue_v	ꝑxima_f-la
ɡtra_f-la	ꝑximus_f-la

Table 2: Incorrectly recognized symbols representing groups of graphemes

Furthermore, it is unclear why the long "s," which was frequently used in early modern print, is so often transcribed as "f," for example *fpeak_v,* when there is a number of instances in which the editors actually used the proper symbol *landſ_n.*

4.3.4 Typos and spelling variants

Inevitably, the corpus contains a number of typos or spelling variants, either produced by the transcribers or present in the original texts. Although the data used in this study has been normalized, there are still plenty of cases where one word has several spellings. This may be partly attributable to the process of spelling normalization which was employed, staring with the most frequently used words and was not finished for the entire corpus. Since the words that I am searching for are by definition on the periphery of the language, it can be expected that they will not be fully normalized or lemmatized. On the other hand, even words which we would consider to be relatively frequent by most standards still display a great deal of variation, for example, in the data from Northwestern University, even a word as common as *knowledge* can appear in many forms:

(3a) ●●●●ledge_n

(3b) knouledge_n

(3c) ●…owledge_n

(3d) hnowledge_n

(3e) ●●owledge_n

(3f) konwledge_n

(3g) knwoledge_n

(3h) knewledge_n

It is interesting that although there are several versions with the Unicode symbol placeholder, denoting an unrecognizable grapheme. However, when checking these forms in the EEBO v3, accessible via Lancaster University's CQPweb, there were no types with the placeholder symbol, and the overall variation was also slighter:

(4a) keowledge (EEBO A86477)

(4b) bnowledge (EEBO A13322)

(4c) kgowledge (EEBO A62395)

4.3.5 Other symbols and non-words

At some stage it would also be practical to filter out items which are not considered words, for example various numbers such as *8.3,4_crd*. It remains to be decided how the research should deal with symbols such as the ellipsis "…" used for truncated words, such as *arm●…_n, r●…ason_n* etc.

4.4 Processing the data

A Python 3 script was designed to process the table of words and their respective frequencies by decade in order to retrieve the candidates for potentially obsolete lexical items. The script's primary input is a .csv file containing a table of all the words in the corpus which appeared in a minimum of five documents. Each row represents one word, and the columns represent the temporal axis in which the frequency for each word is recorded by decade. The final two columns contain a total sum of the word's occurrences across the decades, and the total number of documents in which it appeared.

```
1   WORD            1470s   1480s   1490s   1500s   1510s   1520s   1530s   1540s
2   afe_n           0       3       2       1       2       7       8       3
3   clix_crd        0       5       2       0       13      9       9       4
4   Pastorell_nn    0       0       0       0       0       1       0       0
5   ••ged_v         0       0       0       0       0       1       0       0
6   deames_f-la     0       0       0       0       0       2       0       0
7   agad_j          0       1       0       0       0       0       0       0
8   abute_n         0       1       0       0       0       0       0       0
9   espousa_f-la    0       2       0       0       0       0       0       0
10  hunt_j          0       4       6       4       7       9       17      12
11  commercy_uh     0       0       0       0       0       1       0       0
12  laxa_f-la       0       1       3       2       0       0       1       1
13  untrouthe_n     0       1       0       1       0       0       0       4
14  iusticia_f-la   0       4       31      6       3       8       13      19
15  penetrat_f-la   0       2       3       0       0       0       0       0
16  solem_f-la      0       3       2       0       0       1       8       3
17  franco_nn       0       4       0       0       0       0       3       1
18  jerobam_nn      0       1       0       0       0       0       0       0
```

Figure 7: A glimpse of the data prior to processing

For this assignment, the script was tested with minimum and maximum boundaries restrictive enough in order to return a smaller number of candidates for sorting and closer inspection. The aim was to obtain a number in the range from 500 to 1000 tokens. For the purposes of this preliminary probe into the data, the boundaries have been set thus: If a given word occurred more than 50 times in the first 14 decades (1470-1609) and simultaneously had zero hits in the final six decades (1660-1719), it was appended to the list of possible candidates for obsolescence. With these restrictions the script returned a list of 851 candidates.

Naturally, in the future the criteria will be operating with relative frequencies, i.e., instances-per-million words (i.p.m.), however at this moment the exact length of the corpus is not known. At this stage, there are no hard and fast rules dictating what the parameters for the script should be based on. The minimum frequency in the earlier decades and the maximum frequency in the later decades, as well as the delimitation of the actual "earlier" and "later" decades, have so far been arbitrary.

When selecting the minimum frequency in the initial 14 decades, where the required 50 instances (0.25 i.p.m) correspond with Band 5 in OED's frequency groups. Words in this category were still intelligible but restricted to educated discourse. The OED's (2021) classification of frequency bands is described in more detail in chapter 7.1.2.

4.5   *Sorting the obsolescence candidates*

Prior to a qualitative analysis of those 851 candidates which were returned by the script, it was necessary to filter the results and remove any candidates which had unclear readings, or those which we do not consider to be words. This could have also been done automatically at the very beginning, before running the script, but that was unnecessary seeing as there was no upper limit to the number of candidates and it is best done when it can be checked manually with a smaller set of data in order to find all of the relevant cases. For example, the candidates contained incorrectly transcribed characters (5a-5e), words with the placeholder symbol (5f), roman numerals (5g), and incorrectly parsed words (5h-5i).

> (5a) h₃_n, sc₃_n
>
> (5b) de⁹_n, relygyo⁹_n
>
> (5c) pson_n
>
> (5d) amongʃ_n,
>
> (5e) handʃ_n, wordʃ_n, landʃ_n
>
> (5f) 4●_crd, t●e_zz
>
> (5g) xxiij_crd, m.iiii_crd
>
> (5h) the|holy, the|eight
>
> (5i) youre selve_pn

The candidates were sorted according to their suffixes, which showed that not all words were tagged. Although the tags were somewhat helpful in sorting the candidates, the limitation of one tag per word results in a skewed picture of the results. For example, of the 238 nouns which make up the majority of the candidates, there were many which could also be classified as borrowings from either Latin or French, such as *hargabuzier_n.*

Tag	No. of words
f-la	59
f-fr	15
f-mi	11
f-it	1
f-ge	1
n	238
nn	211
v	57
av	28
ab	96
j	40
jn	10
zz	21
crd	16
uh	7
pn	3
acp	3
sy	15
wd	1
---	18

Table 3: Breakdown of the 851 candidates based on their tags

Example 6a could be an older spelling which still shows the word's clear morphemic boundaries, preserves the meaning of the individual components. Candidate 6b is a phrase which has been analysed as a noun. However, it only appears in one single decade, 1570-9, a total of 58 times. It is therefore very likely that all of those 58 hits are to be found in one text, or at least texts from the same author.

(6a) transsubstantiation_n

(6b) for-that-cause_n

There are cases where it is evident that it is not the word itself that has become obsolescent, but an old version of its spelling. Such examples are numerous regardless of the fact that the data has been, to some extent, lemmatized:

(7a) toogither_av

(7b) questyon_n

(7c) kyngdom_n

(7d) souldyour_n

(7e) domynion_n

(7f) Cauntorbury_nn, Cantorburie_nn

(7g) heretyque_n

(7h) pryeste_n, preaste_n

(7i) mischéefe_n

(7j) unbeléever_n

(7k) holieghost_n

(7l) overmuche_av

As there appear to be some issues with capitalization, it cannot be said for certain whether 8a, 8b and 8c differ only in capitalization or whether they denote different referents as their divergent tags _n and _nn suggest (8b and 8c did not make it through the filter for obsolescence candidates and are included here only for comparison). The OED has confirmed that a *simon* was a slang term meaning "a six-pence," whose decline could be explained by the steady rise of the synonymous colloquial term *tanner* following the year 1600 – although it must be said that it has a further two meanings. The situation appears to be similar with the name *jaakob_nn,* whose decline correlates with the rise of the latinized spelling variant *Jacob_nn* in the later centuries of the early modern period.

(8a) simon_n (100 hits)

(8b) SImon_nn (82 hits)

(8c) Simon_nn (33202 hits)

(8d) armenian_nn

(8e) jaakob_nn

The data contains a number of abbreviations, many of which pertain to the books of the Bible, for example 9a-b. The reason for their apparent obsolescence is their replacement with *Rev._ab* and *Eccl._ab*, both of which display a steep increase in frequency in the latter part of the period.

(9a) Eccle._ab

(9b) reve._ab

Following a very extensive weeding out of typos, extinct spelling variants and other false positives, the remaining candidates represent possible cases of true lexical obsolescence. Below is a selection of those which showed the most promise.

Word	Comment
hargabuzier_n	Alternative spelling of *arquebusier*
baptyme_av	Alternative spelling of *baptism*, additionally incorrect tag
travailer_n	Obsolete in the sense of worker, but also alternative spelling of *traveller*
thez_f-fr	Incorrect tag; based on the various EEBO contexts it is most likely the 3rd person pronoun *them* or *they*.
prentice_v	Verb resulting from the rebracketing of the noun *apprentice* (*a prentice*), now archaic and regional
pellette_n	Alternative spelling of *pellet*
peax_zz	Obsolete (latinized) spelling of *peace*
cipriane_jn	Alternative spelling of *cyprian*
pulueris_f-la	In foreign-language (Latin) contexts only
breastlap_n	
sticado_n	No hits in EEBO.
numidy_n	The ancient kingdom of *Numidia*
kneuet_nn	proper noun, e.g., Sir Thomas *Kneuet*
ensignebearer_n	
aegiptiorum_f-la	
consalvo_n	

donqns_f-fr	Suspected error, cause unknown, possibly a typo. No hits in the EEBO.
decius_f-la	
borrow_n	Obsolete noun, cognate of the verb *borrow*, unsure whether all of these instances were correctly identified as nouns.
rosicleer_n	poetic. A name for a type of worthy knight; someone regarded as resembling this knight
cubyte_n	length of measurement
caplm_f-ge	from context looks like it means "chapter"
mengle_v	Spelling - not a verb!
unmeet_v	Excessive in size; immense, huge (not a verb)
spretes_j	Spelling - not an adjective!
thinhabitante_n	Error in parsing? No hits in EEBO
emonge_v	Spelling - not a verb!
fillage_n	French (obsolete). The state of an unmarried woman
pylle_n	Spelling? (Only some meanings are obsolete; v. To beat or strike violently; To hurry, rush; To pillage)
thenemy_n	Incorrectly parsed; the enemy
synnowe_n	Spelling?
conduit_v	To pour forth like a conduit or fountain
thylk_n	(determiner) The very (thing, person, etc.) mentioned or indicated; the same; that; this
commise_v	3 of 7 hits are in French
reappose_v	To repose or place (1567-1601)

Table 4: Selected cases where true lexical obsolescence was suspected

However, not all of these candidates are as obsolete as they might appear at first glance. For example, *hargabuzier_n* is a French borrowing which still exists in present-day English as *arquebusier, n.,* defined by the OED as "a soldier carrying an arquebus" – a type of firearm. Although *baptyme_av* is quite promising at first glance, it soon transpires that the adverb tag must have been incorrectly assigned to a noun, because the OED lists this form only once as an

older spelling of *baptism, n.* Unfortunately, not all of the words could be found in the dictionary, for example *kneuet_nn*, *consalvo_n* and *donqns_f-fr*. A closer reading and analysis of the contexts in which they appear will be necessary in order to determine their meaning and whether they may be counted among cases of lexical obsolescence. Ever more peculiar are the cases where the candidates derived from the Northwestern data are not to be found in the searchable versions, be it via the CNC or Lancaster interfaces. The cause of these erroneous forms is as yet unknown, but it could be due to some sort of parsing error during the disambiguation process. These words include *donqns_f-fr,* and *thinhabitante_n.*

True obsolescence, however, can be found in the form of *travalier_n*, which the OED lists as an obsolete noun meaning a) "one who travails or labours; †one who torments or harasses," or b) "a woman in labour," its last quotation from the year 1611. Other obsolete words returned in the results included *breastlap_n*, found in the OED as *breast-lap,* is an obsolete synonym of *breast-plate,* while the form *ensignebearer_n* is not present in the dictionary at all, but thanks to a clear delimitation of is two morphemes may be parsed as a bearer of an *ensigne*, *n.*, "a sign, token, characteristic mark."

The following words, based on their frequencies in the corpus, could be counted among cases of true obsolescence and will be examined in more detail: *borrow* (n.), *rosicleer* (n.), *cubyte, caplm, unmeet, pylle, conduit, thylk, commise, reappose.* The following chapter features a more detailed analysis of the individual candidates, including an investigation of possible competitors in their respective semantic fields (e.g., *breastlap* vs *breast-plate*).

After sorting through the 600 most frequently occurring words which, at least according to the algorithm, fulfilled the requirements for obsolescence, each word was assigned to a preliminary category as an explanation why it made it through the filter, and the categories are as follows:

- Obsolete (4%): case of suspected obsolescence
- Spelling (44%): obsolete spelling of an extant word
- Error (17%): most often incorrectly assigned word class tag, or OCR error
- Name (18%): Proper name (named entities)
- Foreign (3%): foreign word (used in a foreign language sentence or phrase, e.g., *praeclaram indoletuam **Iesu** benignitas*)
- Abbreviation (13%): truncated forms of standard words, e.g., *novemb., Eccle.*
- Typo (1%): an error presumably already existing in the original text, including phrases written as one word, e.g., *the|eight*)

Figure 8: Categories of 600 most frequent candidates

The most common difficulties encountered were as follows:

1. Deciding whether the candidate is a case of now obsolete spelling or simply a typo or OCR error, for example *thenemy* (the enemy) or *thesayde* (the said)

2. Determining whether candidates with the _f* tags were to be treated as foreign word or as borrowings, for example *imbroccata* or *peax.* The decision was made for each of the problematic candidates separately, based on the immediate context in which they appeared: when they were part of a longer chunk of text written in a foreign language or when they were introduced with "A thyng is saied in latin […]" (A00316), or words to that effect, the word was treated as a foreign word which was not established in the language and therefore not an appropriate candidate for obsolescence.

3. Names with unconventional spellings may be either obsolete variants of the same name, or names of specific foreigners which retained their original spellings, for example *Edmonde* vs. Edmund.

4. Many false positives were found as a result of incorrectly annotated word classes, for example the case of *spretes* (i.e., spirits), which was assigned the verb tag instead of a noun, or for instance the adjective *unmeet* (huge, excessive in size) which was also assigned the verb tag. However, even the supposedly obvious cases or erroneous tagging cannot be overlooked, because then we would be depriving ourselves of the opportunity to find interesting examples of true obsolescence in the form of *borrow (n.), (otherwise, n.)* and undoubtedly many others.

5. There appears to be an unusually large concentration of incorrectly assigned tags among the foreign language group (_f-la, _f-it, _f-fr, _f-ge, _f-mi), for example *kyngis_f-la* (king's)*, ageyn_f-mi* (again), etc.

6. Methodologically, the most difficult task is dealing with polysemy, shifts in meaning, and also drawing a line between the disappearance of a spelling variant and the obsolescence of a form associated with a specific meaning. For example. *hooste, (host)* in the sense "armed company" could be regarded as obsolete, however it is part of the extant lemma *host* and not a separate form, therefore it is not a clear case of obsolescence.

4.6 Summary and further questions

Although there is certainly much to be garnered from the data, it appears as though the normalization process of the EEBO corpus was not as thorough as one might hope, and this has allowed a large number of incorrectly identified candidates slip through the filter. Compared to the version of EEBO accessible through Lancaster University's online interface, this one appears to be halfway between fully normalized spelling and the rich variability that is present in the unlemmatized version of the corpus, accessible via the CNC interface Kontext (ÚČNK 2014). The surprising degree of variability witnessed in the orthography of the Northwestern version is attested in Table 5, which lists all the spellings of *judgement* from the Northwestern EEBO corpus in juxtaposition with those from a fully unlemmatized version of EEBO (such as the one made available by Czech National Corpus).

Northwestern version	CNC version
iudgment	iudgment
iudgmente	iudgmente
iugement	iugement
iugemente	iugemente
iugment	iugment
judgement	judgement
judgemente	judgemente
judgment	judgment
judgmente	judgmente
jugement	jugement
jugment	jugment
-----	jugemente
-----	iudgement
-----	iudgemente
-----	iugmente

Table 5: A comparison of the orthographic variability in the normalized Northwestern corpus versus the raw version available via CNC

Another possible expansion of this project could include n-grams, whose development over the Early Modern period would be an interesting complement to the obsolescence of unigrams as shown here. An issue which must be addressed is the lack of information regarding genres, text types and the number of documents in which each word occurs. It is hoped that these questions will be resolved in the upcoming months.

5 Analysis of obsolete forms

Once the appropriate candidates have been returned by the algorithm, the conditions signaling the decline of particular forms will be selectively analyzed and discussed. Close attention will be paid both to the related changes in language structure and relevant language-external processes.

The key question that needs answering is: Which words can we consider to be obsolete? The methodology is more inclined towards a corpus-driven approach, which in practice means that I will be taking results from the corpus as a point of departure and then confronting these with what is featured in the OED. This can and undoubtedly will lead to discrepancies, such as that a word may appear to be obsolete based solely on the corpus data, but the OED might have some instances from later decades. However unfortunate this may be, I am using the EEBO corpus as a primary source that is considered to be representative of the majority of (if not all) printed texts produced in the given period, and will be treating it as such.

5.1 Qualitative analysis

This section contains a detailed analysis of those candidates whose frequencies indicate the likelihood of obsolescence (as documented below in Table 6). While some candidates were more frequent in the initial decades (before the year 1500), the majority of the candidates has a peak in the respective frequencies between the year 1530 and 1600, which suggests that these forms entered the English language and existed for a century or so (at least as far as we can tell this is true for written texts) before disappearing by the end of the Early Modern period. Unless stated otherwise, all dates and dictionary definitions are taken directly from the online version of the Oxford English Dictionary (OED).

Word	1480-9	1490-9	1500-9	1510-9	1520-9	1530-9	1540-9	1550-9	1560-9	1570-9
travailer_n	0	0	1	1	2	0	1	1	17	69
Caluine_nn	0	0	0	0	0	0	0	4	204	217
debonayr_j	104	13	10	7	0	1	0	1	1	0
meyny_n	53	5	13	4	44	17	5	39	2	5
borrow_n	27	48	20	16	27	78	67	52	72	202
otherwise_n	0	0	0	0	0	1	0	114	2	8
cubyte_n	18	22	15	9	15	192	131	14	1	3
unmeet_v	3	2	0	1	10	55	60	66	72	33
pylle_n	2	4	2	1	30	11	119	32	19	8
conduit_v	61	24	21	11	11	7	11	27	6	6
thylk_n	217	1	0	0	0	0	0	0	0	0

Table 6: A glimpse of selected obsolescence candidates with frequencies by decade (full table available in Appendix A)

5.1.1   travailer, n.

In the case of *travailer,* the situation is somewhat complicated by the homonymy present in the Early Modern English period, where the word exists both as a morphological derivation from the French borrowing *travail, v.* (ME forms *trauaillie, trauaily, trauale, trauaylle…*) whose meaning had already shifted to denote the act of travelling, and as a borrowing from the Old French *travailleor* which evolved into the ME form *travailour* (*1377 --†1611). The meanings which the OED offers for the head word *travailer* are (a) one who travails or labours; (b) one who torments or harasses; (c) a woman in labour. However, *travailer* appeared in English at a time when its synonyms *worker* and *labourer* were apparently already established, and therefore the new word's coexistence with its much more frequent counterparts was relatively short-lived. That *travailer* was doomed from the beginning is evident enough in Figure 9, although in reality the frequency gap is even wider, because the data subsumes even those instances which are likely spelling variants of *traveller*.

Figure 9: Frequencies of *travailer*, *worker* and *labourer* in the EEBO

A close analysis of the individual contexts found in the EEBO results suggests that *travailer* in fact predominantly appears as an earlier spelling variant of *traveller*. Such is the case of "Yet to the fearefull Travailer All wayes were then unſafe" (A11474) thanks to whose clear context the interpretation of *traveller* comes naturally. The most clear-cut case is in the example "the travailer in the Parable Luke 10" (A62137) which refers to the traveller whose life was saved by the Samaritan, a meaning of which we can be certain thanks to the existence of numerous translations.

Similarly Rowlands's text from the 1630s, whose context, albeit ambiguous, is still rather suggestive of the meaning *traveller*, "ſweete is the fountaine to the weary thirſtie travailer" (A11116), while in other, less frequent contexts the word appears in the sense of *worker,* for example in the 1540s translation of Erasmus "the most noble weomen of bloude and estate royall, are no lesse diligent trauaillours then the best, […] ne take any manier skorne or disdeigne in y e labour of drawing this haruest home" (A16036). Both the contextual and orthographic (Table 7 below) overlap makes it difficult to determine exactly how many of these tokens can be counted among the now-obsolete lexeme *travailer,* and which are merely an archaic variant of *traveller.* The most probable causes of *travailer*'s obsolescence are a) the existence of more frequent synonyms, and b) the partial homonymy of *travailer* and *traveller.*

Form	Total in corpus
traveller_n	17288
traveler_n	267
travailer_n	181
travellour_n	89
travailler_n	56
traveiler_n	55
travailour_n	33
travailor_n	25
travellers_n	19
travailler_f-fr	18
travelour_n	18
travelor_n	16
trauailler_f-fr	9
travaler_n	7
travaller_n	7

Table 7: Spelling variants of *traveller* and *travailer* found in the Northwestern EEBO

5.1.2   Caluine, n.

Caluine is apparently an obsolete spelling variant of the proper noun *Calvinist* (adherent of Calvinism). This form is not yet listed in the OED, although it clearly existed as an alternative to variant with the <-ist> suffix. The two lemmas are derived from the name of John Calvin (also written Caluin), whose life and works fall into the Early Modern period (1509-1564), which explains their sudden appearance in the 1550s.

The competition between the forms Caluine ([word="[cC]al[vu][ie]nes?"]) and Calvinist ([word="[cC]al[vu][ie]nists?"]) is recorded in Figure 10 below. Evidently, although it was initially more frequent, the early form *Caluine* soon lost its momentum, was overtaken by *Caluinist* and doomed to extinction.

Figure 10: Competition between *Calvine* and *Calvinist* in the EEBO

5.1.3 debonayr, adj. & n.

Originally a proper phrase *de bonne aire* (11th century Old French), the dictionary definition for the adjective *debonair* (also *debonair, debonnaire* etc.) is "of gentle disposition, mild, meek; gracious, kindly; courteous, affable" (*c1230 – †1847) and although it is marked as obsolete, the dated citations indicate that this usage's disappearance from English took place much later then the EEBO corpus data initially suggested. This was caused by the imperfect normalization of the substantially varied spellings, as seen in Table 8. Nevertheless, the EEBO data suggest a rapid decline of the collective i.p.m frequency for all forms of *debonayr*, so that in Figure 11 we can see that by the year 1530 the word is well on its way to becoming obsolete.

Form	Frequency
debonaire	233
debonayre	146
debonair	134
debonayr	108
debonnaire	37

66

debonnair	12
debonnayre	2
debonnayr	1
debonyr	1

Table 8: The varied spellings of *debonayr* in the EEBO

Figure 11: The frequency of *debonayr* and all of its spelling variants in the EEBO corpus

This citation from Spenser's The Faerie Queene, "Was neuer Prince so meeke and **debonaire**" (A12777), shows the word in its most typical context, a binominal construction, i.e., "a coordinated pair of linguistic units of the same word class which show some semantic relation" (Kopaczyk and Sauer 2017: 2). In fact, of the 673 occurrences in the corpus (results for the query [word="debonn?a[iy]re?"]), 274 contained *debonair* as part of a binominal construction (195 times as a right-hand element, 79 as a left-hand element). The most frequent pairings were with *humble, gentle,* and *meek:*

(10a) "A parfyt mayde **humble and debonayre**" (EEBO A06558)
(10b) "the **gentle and debonair** nature of the inhabitants" (EEBO A06128)
(10c) "by the vertue of the feythe he became **meke and debonayr**" (EEBO A14559)

Although the candidate returned by the script was the adjective *debonayr* which, as it transpires, became obsolete not in the Early Modern period but within the next two centuries, further research revealed the existence of a noun in an identical form. According to the data provided by

the OED, the noun *debonayr* became obsolete over the course of the Early Modern period. Two meanings are differentiated, firstly as a proper noun, "gracious being or person" (*c1366 - †1393), and secondly as the abstract noun "graciousness of manner" (*1697 - †1747), as in "A serious Majesty was attemper'd with such strokes of ***Debonaire***, as won Love, and Reverence." In this period, we may also find two other abstract nouns derived from *debonayr; debonairity* (†1688) and *debonairness* (extant but archaic in present-day usage).

At one point, the three could be found in the same contexts with presumably identical meanings, and we may witness the coexistence of three words derived from the same root by different types of word formation; zero-derivation *debonayr*, *debonairity* with the Romance suffix <-ity> and the Germanic suffix <-ness> in *debonairness*. This type of competition between synonyms is arguably the most unsustainable, as they (very conspicuously) share the same origin and their only difference lies in their word-formation strategy.

5.1.4 imbroccata, n.

The now obsolete *imbroccata*, otherwise known as "a pass or thrust in fencing" (*1595 - †1616), came into English in the 16th century as a direct borrowing from Italian. Although it was a technical term introduced in order to describe a very specific type of action, it was nonetheless employed figuratively in "But then , you haue your passages, and imbroccata's in courtship" (EEBO A04632), which suggests some degree of familiarity with the word allowing it to be used in a context outside of its traditional meaning. On the other hand, it is important to take note of the non-standard formation of the plural using an apostrophe *imbroccata's*, which may be a testament to the word's non-native status in the language. Furthermore, the convention of capitalizing the initial letter presumably fortified its foreignness, as is apparent in this 1650s (apparently posthumous) citation "the Stoccata , Paſſada , Punto , and the Imbroccata, With more Italian poſtures" (EEBO A86166)

5.1.5 meyny, n.

The word *meyny, n.* now spelled *meinie,* is a 14th-century borrowing from the Norman *maigné* which survives to this day but with its meanings confined to Scottish (and Irish in the case of the first two) dialects of English. *Meinie* is still extant in the following senses:

– "a family, a household,"
– "a body of people attending a lord or other powerful person,"
– "a crowd of people; (depreciative) a rabble. Also: the populace, the masses,"
– "a considerable number or collection of items," and
– "a herd, flock, etc., of animals."

The remaining senses of *meyny* are considered to be obsolete based on the dates cited in the OED:

- "collectively: servants" (*1382- †1475)
- "the pieces or men used in the game of chess; a set of chessmen" (*1322- †a1500)
- "a group of people employed together or united by a common purpose; an army, a ship's crew, a congregation, etc." (*1380- †1598)

The word *meyny* (the query form took into account all expected spelling variants including *meiny, meinye* etc.) returns 377 hits in the EEBO corpus. Due to the substantial polysemy of this word, it is sometimes difficult to identify the intended meaning of a given instance, for example in the sentence "the erle of Warwyk had mette with y• (= the) erle of Marche on Cotteſwolde comynge out of Wales [with] a grete **meyny** of walſſhmen" (EEBO A00007) the text could be referring either to an army, servants or even a group of people attending a lord (or, in this case, an Earl).

Another issue is the apparent homonymy with the pronoun *many,* as in the following example: "they were indyted of malice, a great **meyny** of them, which alreadie were in Priſon" (EEBO A03448). For these reasons, it is impossible to determine in exactly how many of the 377 occurrences we are witnessing cases of now obsolete words.

5.1.6   borrow, n. & adj.

Known almost exclusively as a verb in present-day usage, the only sense in which we may encounter *borrow* as a noun in contemporary English is in the language of golf players, where it denotes "the amount which one 'borrows' to allow for the slope of the green", however it is derived from the figurative sense of the verb *borrow*, listed separately in the OED. The following senses of *borrow, n.* are considered obsolete:

- "on/to **borrow**: on security, by way of loan" (*a900 - †1418), allegedly out of use before the beginning of the Early Modern Period,
- "*I dare be borrow*, etc.: 'I'll warrant', 'I'll be bound'" (*c1430 - †a1500), the word's short-lived use in the discourse marker shows a much wider semantic range than that of the verb in present-day usage,
- "A thing deposited as security, a pledge; a guarantee, bail; suretyship; ransom, deliverance. *to borrow*: as a pledge. *to lay to borrow*: to put in pledge, to pawn" (*a975 - †1860),
- "of persons: A surety, hostage; bail, deliverer from prison" (*1000 - †1819)

Of the latter two senses the OED claims that they were "already obsolete or archaic in England in Spenser's time [i.e., the latter half of the 16th century], but [were] retained in Scots Law," which accounts for the 19th-century mentions.

As for the total number of occurrences of *borrow, n.* in the EEBO, there are 692 instances with the _n tag, but it is often difficult to determine whether the word class was correctly assigned, for example "Lat vs to **borgh** our men fra ʒour fals Law"(EEBO A03007) could be interpreted either as a noun or verb.

Further research has led me to uncover the existence of an adjective *borrow*, an antonym to the equally rare and obsolete *transhaw, adj.* (Meaning and origin uncertain: said of the pitch of a wall, perhaps 'exposed' to the blast, opposed to 'borrow' sheltered from the blast.) with only one mention in the EEBO from 1665 by Dudley "by finning or ſetting the finery, **leſſe tranſhaw more borrow** which are terms of art, and by altering and pitching the works"(EEBO A36750). The OED offers the following 1686 citation from Robert Plot's text, "These [walls of blast furnace] according as they may be pitch't **less transhaw, or more borrow**; will mend..or alter the nature of the Iron... The Iron made in a borrow work, is much more tough."

Although these quotes do little to hint at the meaning of either of the adjectives in question, it does suggest an extremely limited usage, probably restricted to a specific trade or craft, whose popularity (or in fact existence) would have been directly linked to the survival (or loss) of the terminology associated with it. Similar cases of obsolescence, where the form disappears in the scope of one word class but remains in another, include the noun *otherwise*.

5.1.7   otherwise, n.

The noun *otherwise* (eOE – †1878) meaning "another way" appeared in the adverbial phrases **in/on other wise** (in another way), **(on) any otherwise** (in any other way), **(in) no also none) otherwise** (in no other way). Although based on the 19th-century citations supplied by the OED we can observe that the noun *otherwise* became obsolete much later than our EEBO data would suggest, it is still evident from the Northwestern data (see Appendix A) that it was already in significant danger as early as in the 17th century.

It is only right to be suspicious of the sudden surge in occurrences of the noun *otherwise* in the single decade 1550-9, and indeed an examination of the concordances proves that the ambiguity often makes it difficult even for a human annotator, let alone an automatic tagger) to determine whether the given example is a noun or adverb. Although I am inclined to interpret the following examples as nouns, there is evidently still a considerable degree of ambiguity there:

(11a) "None **otherwise** do I restrain the vowes of the olde lawe" (EEBO A02621)

(11b) "many man hadde ben hangyd & drawen and slayne in **otherwise**" (EEBO A08936)

(11c) "they cannot hook him in **otherwise**, than by streining hard" (EEBO A31089)

Originally written as two separate words, the OED's first recorded example of *otherwise* as a compound dates to c1430. Although the corpus data suggests that this became the norm certainly

by the 1550s (see Figure 12), there are still singular instances of it written as two separate words in later centuries, take this example from Playfere's 1623 text: "but in Gods Dictionarie it hath no such name. In the holy Scripture , it ***in other wise*** called . It is called Adams goodly apple" (EEBO A09744)

The EEBO corpus contains 730 instances of the phrase *other wise,* while the compound *otherwise* 88427 hits (of course, this includes the adverbs proper, which cannot be filtered out when working with an untagged corpus). In the 1470s the two-word variant was still prevalent, but after the year 1530 it was overtaken by the single word to such an extent that including the decades 1560 and onwards in the graph would have made it impossible to discern the movement in the earlier centuries, where we can observe something like a competition between the two orthographic norms.

Figure 12: *Other wise* vs. *otherwise* in the early decades of the EEBO corpus

In analogy to the structure *other + wise,* we may also observe the appearance of *otherways* (*a1225 - extant) and *elsewise* (*1548-†1888)*, neither of which became particularly widespread in the language.

5.1.8   ordnance, v.

The verb first appears formed by conversion of the noun *ordnance* in the 16th century, meaning "to provide with ordnance or military equipment, esp. artillery" (*1531 - †1548) *Ordnance* existed as a variant of the French borrowing *ordinance,* from whose original meaning it diverged significantly, nevertheless both of these spellings are found in the EEBO data under the relevant

military-related meaning. The verb is transitive and, according to the OED, usually in passive voice. This is confirmed by its occurrences in the EEBO corpus, for example:

> (12a) "This was a strong toune well walled, dyched and **ordinaunced** but not manned" (EEBO A02595)

> (12b) "a woonderfull strong Bulwarke , well **ordinanced** and well manned" (EEBO A67926) – although the OED's last citation is from 1548, this example comes from a text written in 1583.

Interestingly, in its passive usages, the verb is often accompanied by other verbs of similar origin (i.e., derived from nouns), such as *walled, ditched* and *manned,* which we can observe in the above examples. Other than *equip* (although it covers the intended meaning in a much more general sense)*,* there is no obvious synonym for the verb *ordnance,* leaving only its restricted usage (restricted both semantically and grammatically, since it almost exclusively appeared in the passive participle) as the most likely cause of its disappearance from English.

5.1.9   rosicleer, n.

Rosicleer is recorded in the OED as a borrowing from Spanish (where it was originally used as a word for "red silver, ruby silver ore"), a rare and poetic term meaning "a type of worthy knight; someone regarded as resembling this knight" (*1578 - †1631)

Strangely, the EEBO data does not reflect this meaning, as it only contains instances where *Rosicleer* is a given name - with just one mention of *rosicleer* (again, the query was constructed so as to include all conceivable spellings [word="[rR]oss?[ei]cli?e?er"]), and even here it appears to be a proper name, in this text translated from Spanish in 1599: "went to speake with **rosicleer** at the Monasterie" (A08545) The conclusion that this, too, is a given name, is supported by the absence of determiner; and although it could have been accidentally omitted (either by the translator or at any time during the digitization process), given the evidence (i.e., all other instances of *Rosicleer* which are present in the same text), we can safely assume that this is also the case:

> (13a) "behold the mighty Trebatio, & his sonne **Rosicleer**" (EEBO A08545)
> (13b) "This valyaunt **Rosicleer** trauayled with so great ioy and contentment" (EEBO A08545)

In fact, of the total 718 hits, 714 were from one and the same text (which recount the adventures of, among others, a knight named *Rosicleer*), and 2 of the remaining 4 occurrences are obvious allusions to the original text. This leaves only two instances where we may consider the possibility of *Rosicleer* having become an generic term as suggested by the definition available in the OED, although the link to the eponym is still strong:

(13c) "He had no Arthurs , nor no **Rosicleer's** , No Knights o ' the Sunne" (EEBO A04658) from a 1631 text by Jonson, also note the use of apostrophe in the plural form (in juxtaposition to *Arthurs* this is a possible indication of the name's non-native status)

(13d) "Which of you is the valiant **Rosicleer**, Dares breake his Launce on me" (EEBO A01779) from a 1640 text by Glapthorne.

An additional two definitions exist in the OED which attest that *rosicleer* also existed firstly in the field of mineralogy as "any of the varieties of ruby silver ore, as proustite and pyrargyrite" (†1855) and secondly in poetic registers for "a rose-coloured light associated with the dawn" (†1883). Neither of these meanings was found in the EEBO corpus. In the case of *rosicleer*, the form's obsolescence is hardly surprising, given the fact that its existence was so wholly dependent on one single work of literature.

5.1.10 cubyte, n.

A *cubyte* is a historical term that signifies "the part of the arm from the elbow downward; the forearm." and by extension also a unit of measurement. The former meaning is marked as obsolete (†1882), but the latter (measure of length) is not, although the last citation in the OED is from the year 1875, "He is four cubits high".

This raises the question of what makes a word obsolete – in this case, it has not been used as a real-life unit of measurement in over a hundred years, but it must have been evaluated as more familiar, possibly due to mentions in historical texts. Nevertheless, it did not become obsolete within the Early Modern period, and the script only evaluated it as a potential candidate due to the spelling variation (*cubit, cubite, cubyt,* and *cubite*).

5.1.11 caplm, n.

Was not cited in the OED but based on the contexts found in the corpus, it is an abbreviated form of the Latin *capitulum* (chapter). Moreover, it was used by only 7 authors (see Figure 13).

	Filter	doc.author	Freq	i.p.m.	
1	p / n	Hilton, Walter, d. 1396.	404	1793.87	
2	p / n		325	5.96	
3	p / n	Alexander, ab Alexandro, d. 1523.	4	17.93	
4	p / n	Alcock, John, 1430-1500.	3	68.3	
5	p / n	Boethius, d. 524.	1	6.16	
6	p / n	John, of Garland, ca. 1195-ca. 1272.	1	13.62	
7	p / n	Peter, of Blois, ca. 1135-ca. 1212.	1	77.44	

Figure 13: Authors using the *caplm* abbreviation

5.1.12 unmeet, adj.

The word *unmeet* was a false positive due to an incorrectly assigned verb POS tag; The verb unmeet does not exist in the OED and moreover I was unable to find any concordances in the EEBO which would indicate otherwise. However, this led me to the identically formed adjective for which the OED provides 5 definitions, the first 4 of which are listed as obsolete:

- "Excessive in size; immense, huge; Immoderate or excessive in amount" (OE – †1475)
- "Unequal; unevenly matched" (*a1393 – †a1793)
- "Not closely connected; distinct; remote" (*a1393 - †1598)
- "Horrible to look at, ugly" (*a1425 - †a1450)
- "Not having or showing the necessary qualities or skills for something; incompetent; unworthy. Now archaic and somewhat rare" (*1425 –archaic and rare)

Of the 900 instances of *unmeet,* it seems that they are all examples of the still extant meaning, i.e., unworthy (in fact they often appear in binominal constructions with *unworthy* or *unfit*), and it may be noted that these texts all appear to be religious treatises, where the concept of unworthiness is quite frequently mentioned:

(14a) "how unmeet our shallow Wit is to judge of the things of infinite Wisdom" (EEBO A26879)

(14b) "it is unmeet to be Loved , but it is meet to be Used" (EEBO A26905)

(14c) "we shall be wholly unqualified and unmeet for God ' s Kingdom" (EEBO A67108)

The OED also lists obsolete (and rare) the word *unmeet* in adverbial uses (†c1600)*,* with the meanings *immensely*, *unsuitably*, *unkindly* and *unequally,* although these uses were not backed by the data in the EEBO. The form *unmeet* also led me to the rare (and very much obsolete) form *unmeetly, adj.* (†1534) although the EEBO data (results for [word="[uU]nmeetl[iy]e?"]) suggests later usage, as in the 1641 text "was full unmeetly to bee matched with his" (EEBO A51324) Although we cannot yet speak of obsolescence in connection with the adjective *unmeet,* the related forms *unmeetly, adj*. and *unmeet, adv.* have apparently well and truly disappeared from use by the end of the Early Modern period.

5.1.13 conduit, v.

A rare verb derived from the noun *conduit,* meaning "An artificial channel or pipe for the conveyance of water or other liquids; an aqueduct, a canal," from the Old French *conduit* or Latin *conductus,* and in a more specific, now obsolete sense of "a structure from which water is distributed or made to issue; a fountain" (*1430 - †1552)

The meanings attributed to the verb are the literal "to pour forth like a conduit or fountain" (†1591) and the figurative "to transmit or convey as through a conduit" (†1628), the latter of which can be found among the examples from the EEBO: "whom sir Edwarde Pounynges conduited to the toune gate" (EEBO A02595)

Due to the variety of spellings in the Early Modern period, there was a polysemy with the closely related verb *conduct,* take for example "Sodaynly apperid by me a ship conduited by one man only" (EEBO A68341). The most likely cause of the obsolescence of the verb *conduit* is the availability of synonyms (the OED thesaurus lists *conduct*, *derive, channel, carry, convey,* and *transmit*). Of these, likely the most threatening was the verb *conduct* due to the similarity in form and identity of origin.

5.1.14 thylk, adj. & pron.

As defined in the OED, *thylk* exists as an adjective used as a determiner meaning "the very (thing, person, etc.) mentioned or indicated; the same; that; this", and pronoun "That (or this) person or thing" , however it can be found only in archaic or dialectal contexts. However, its plural use is marked as obsolete: either as a determiner with a plural noun (†1490) or as a plural pronoun (†1450). The 420 occurrences (returned by the query [word="[tT]h[yi]lke?"]) in which *thylke* appears as a determiner with plural nouns, for example:

(15a) "it were noyfulle to charge this place with all ***thylke reaſons***" (EEBO A00005)

(15b) "closed within one walle alle ***thylke cytees*** aboute and made one grete cyte"(EEBO A03319)

The most likely reason for the disappearance of the plural usage is the synonymy with the incomparably more widespread and established (yet still equally native) adjective/pronoun *those,* which appears 299,243 times as a determiner with plural nouns, as opposed to 420 for *thylk.*

5.1.15 commise, v.

A borrowing from French, the verb *commise* has the meaning "To commission or appoint (a person) to carry out a task" (*a1470 - †1518) or "To place (a thing, matter, person, etc.) in the care, custody, or charge of another; to entrust" (*1474 – †1591).

In the EEBO corpus, whenever a variation of *commise* appears an English-language context, it is predominantly in the sense of "to perform, carry out (an act); esp. (in a negative sense) to commit, perpetrate (a crime, offence, etc.)"(*1475 – †1547), for example in the texts: "I shal shewe to you that ye ***comyse*** two synnes" (EEBO A14559) and "had hardynesse to ***commyse*** and doo this cryme soo moche deffamed" (EEBO A14476). More ambiguous examples which are open to interpretation include "vnder the couerture of this gracious courtosie I haue not ntcion to ***commise*** one so grete an euyll" (EEBO: A68341). It may also be relevant to mention the fact that

of the 25 results returned by the query [word="[cC]omm?[yi]se"], 5 instances were in French texts, for example "Erreur qu ' elle a commise en vous persécutant" (EEBO A44775).

Among the causes of this form's obsolescence we may count its perceived foreignness (as suggested by the presence of French language contexts), competition with its related synonym *commission, v.* (formed by conversion from the noun *commission*) , as well as near homonymy with the verb *commit* as a result of the variable spelling.

5.1.16 reappose, v.

Defined as "to repose or place" (*1567 - †1601), the form *reappose* is believed to be an alteration of the more widespread *repose*. There are only 3 occurrences in the EEBO corpus, including "on whose iudgement a ma(n) may safely reappose his whole fortunes" (EEBO A04577), again surprisingly given the number of instances (216) recorded in the Northwestern data.

However, there is a second meaning for *reappose* listed in the OED which appears in present-day usage in the field of medicine, as "to bring (the edges of a wound, parts of a fractured bone, etc.) back into apposition". Interestingly, this is first recorded in a 1918 text in the form *re-apposed.* This suggests that the verb became obsolete around the year 1601, most likely due to the prevalence of its closely related synonym *repose,* but was resurrected a little over three centuries later in 1918, presumably when the specialized field of medicine needed to name a (supposedly new) concept, and the not-so-new verb was formed by derivation from the noun *apposition* ("action of putting or placing one thing to another")*.*

5.1.17 putcase, n.

The compound *putcase,* obsolete since the 18th century, is listed in the OED with the meanings "a supposition, a hypothesis"(*1565 - †1577) and "A person skilled in putting or arguing hypothetical cases" (*1590 - †1742). Throughout the EEBO corpus, there is an inconsistency to how it is written. At times, *putcase* appears in the corpus as one word or hyphenated as *put-case,* but in the majority of examples it is recorded as a phrase *put case*, its discrete verb + object components clearly identifiable. The corpus returns only 2 results for the full compound and 4 results for the hyphenated version. On the other hand, there are 380 instances of *put case* appearing as a phrase, which goes to show that the compound was not firmly rooted in the language and its two components were not as closely linked as the OED entry would indicate.

(16a) *"**Putcase** thou hadst a good cause […]"* (EEBO A02783)

(16b) "VVhat a fooliſhe ***putcaſe***, and what a fond what if is that, to ſaie, VVhat if a pirate inuade the Arke of Noe?" (EEBO A04468)

(16c) "which ***put-case*** with that intent is worthy to be put into a cap-case" (EEBO A20733)

The most notable discovery was that not all usages are fully covered by the definition in the OED, namely that not all of the phrases function as a noun, and a surprisingly large number of concordances had the phrase *put case* in the role of a discourse marker similar to what one would expect from an imperative clause such as *imagine* or *take, for instance:*

(16d) "and **put case** it be as true as it may be" (EEBO A02129)

(16e) "For, **put case** that this had not beene his owne natiue Country" (EEBO A01020)

In fact, the following examples clearly show that *put case* was also commonly used (for want of a more appropriate term) as a sort of phrasal verb. The meaning which I have elicited from the context would be something like *propose* or *stipulate:*

(16f) "But I **put case** they haue none" (EEBO A04479)

(16g) "The similitude is this I **put case** thou haue a servaunte who thou put test" (EEBO A01273)

5.1.18 pylle, v.

The obsolete verb *pylle,* also spelled *pilyie,* comes from the French "to pillage, plunder" (either a region or a person) and was extant in 16th and early 17th century English, with its last citation from 1626. In the Early Modern period, it coexisted with at least two homonyms; first, the closely related *pylle* of Latin origin, meaning to "beat or strike violently" (spelled *pell* in later centuries); and second, the native verb *pylle* (Old English, "strip, pluck"), still common in modern-day usage as *peel*, though its older spelling is now obsolete. The context of "pilliage (a person)" and "beat (a person)" are so similar that there are many instances in the corpus where the meaning is uncertain, such as in the final two of these examples:

(17a) "Cutte brede Wasshe the mortier And the pestel Make vs somme gharlyk […] I **pylle** the gharlyk" (EEBO A14548)

(17b) "it is couted a lyght offence to **pylle** to defraude to oppresse wydowes and other poore folke" (EEBO A00391)

(17c) "**pyllars** and robbers were comen in to the felde To **pylle** and to robbe many a ful noble knyghte of brochys" (EEBO A21703)

The *pyllars* (pillagers) in the latter example lead us to the noun derived from the verb, spelled *pillar* or *pillour* in later centuries, which survived the verb by several centuries but is now also obsolete, its last citation in the OED dating back to 1823. We could speculate as to why this is, and what comes to mind is that there was significantly less homonymy, ergo less competition, in

the case of the noun, though confirming this hypothesis would require a thorough examination of all alternatives available in the Early Modern period.

5.1.19 thesayde, v.

Thesayde is tagged as a verb, but that is the result of an error in the annotation process, as the contexts of this word unanimously point towards an adjectival function in the sentence. Incorrect tag notwithstanding, this candidate may offer some insight into the orthographic norms in the Early Modern period. In the case of the form *thesayde,* which has 154 occurrences attested in the EEBO (as many as 537 hits when expanding the query to [word="thesa[yi]de?"]) it would appear that we are not dealing with a case of lexical obsolescence per se, but with the idiosyncratic style of a small handful of authors (in this case 16, of which two displayed an uncommon affinity for this form, as seen in Figure 14 below).

From all of the contexts it is clear that the intended meaning is in fact identical to *the said,* for example "he and his shoulde enioye the possessyon of **thesayde** castell quyetely and peaceably durynge thesayde truce and amitie" (EEBO A02595), however this form is marginal to say the least, with the 537 occurrences accounting for only 0.2% of the larger group containing all variants of *the said* and *thesayde.*

I decided to search for the alternative *the said* in the texts by the author with the highest frequency of *thesayde,* Edward Hall, and the search resulted in 325 instances of the phrase *the said* ([word="the"] [word="sa[yi]de?"]) written separately. This may have been caused by errors in the OCR digitization of the original texts, but it could also indicate that the orthographic norms were not as constricting at the time of writing. The explanation may even be as pragmatic as typographic concerns, so we may imagine that the printer needed to remove one character in order to fit the words on one line, and linking these two grammatical words into one would have been the easiest way to condense the text.

	Filter	doc.author	Freq	i.p.m.	
1	p / n	Hall, Edward, d. 1547.	462	611.45	
2	p / n	Numan, Philippe, d. 1617.	33	377.17	
3	p / n	Parsons, Robert, 1546-1610.	19	7.11	
4	p / n	Harvey, Gabriel, 1550?-1631.	5	51.08	
5	p / n	Fitzherbert, Thomas, 1552-1640.	4	14.84	
6	p / n	Carew, George, Esq.	2	9.05	
7	p / n	Niclaes, Hendrik, 1502?-1580?	2	7.17	
8	p / n		2	0.04	
9	p / n	Scotland. Court of Session.	1	1.79	
10	p / n	Hume, David, 1560?-1630?	1	1.52	
11	p / n	Howell, James, 1594?-1666.	1	0.53	
12	p / n	Soranzo, Lazzaro.	1	14.58	
13	p / n	Painter, William, 1540?-1594.	1	1.9	
14	p / n	More, Thomas, Sir, Saint, 1478-1535.	1	0.77	
15	p / n	Estienne, Charles, 1504-ca. 1564.	1	1.54	
16	p / n	Erasmus, Desiderius, d. 1536.	1	0.39	

Figure 14: Complete list of authors with at least one occurrence of *thesayde*

5.2   Observed factors in lexical obsolescence

This section offers a brief summary of the observed factors which were evaluated as being the most pertinent to the eventual obsolescence of the lexical items surveyed in the preceding chapter. These observations will then serve as a basis for the proposed method of classification in chapter 6.

Firstly, there are cases of obsolescence which go hand in hand with the disappearance of a real-word concept, for example *borrow* (adj.) and *transhaw* (adj.). In such instances when a word is no longer needed in reference to an obsolete object or concept, it might organically shift in meaning and find its calling elsewhere, but there is no data supporting that this was the case of the adjectives *borrow* and *transhaw* and so they promptly disappeared from the language in their entirety.

Another aspect factoring into the death of a word is the existence of several synonyms with little or no specialization of meaning, which is to say that they can be used interchangeably in all of their contexts. It is not unusual for one or more of these synonyms to gradually disappear from use, and presumably such was the fate of the noun *travailer*, which coexisted in Early Modern English with the nouns *worker* and *labourer*. In a similar vein we have death by extreme synonymy, which is what I like to call obsolescence within a pair or group of effectively identical words which differ only in their affixes, such as the three abstract nouns *debonairity*,

debonairness, and *debonair*, and the pair of nouns denoting adherents of Calvinism, *Caluine* and *Caluinist.*

Furthermore, I observed several cases of partial obsolescence, where loss occurs in the main dialect, but the word survives as a regional variant. These words occupy the indeterminate space between full obsolescence and present-day (albeit archaic) usage and are worthy of our attention – they may have survived in one or more dialects of English, but as far as standard English usage is concerned, these words are effectively obsolete. Examples from our data include *meyny,* which is limited to the Scottish and Irish varieties of present-day English, and *thylk,* which is present in dialects around Cornwall and the West Midlands, according to the OED, and appears in the forms *thick*, *thicky*, *thickee*, or *thicka*.

Not infrequently was obsolescence limited to a single word class in a derivational family, such as the case of *borrow* (n.)*, survived by its verbal counterpart*, *otherwise* (n.) which remains in use as an adverb and adjective, and *ordnance (v.),* where the noun is still in existence, limited to a more or less specific domain of military words. The data also showed the disappearance of words which had been used exclusively by one author or group of authors (possibly linked to one printing house), such as *thesayde* (v.)*, thenemy* (n.) and *thimperial* (j.) and these I am hesitant to include among cases of true obsolescence due to their idiosyncratic usage.

The data also features the unlikely phenomenon of resurrected words, exemplified by *reappose* (v.), which disappeared in the early 17th century and cropped up again some three hundred years later in a medical text in 1918, though it is unclear whether those responsible for the reinstitution of *reappose* were aware of the existence of the closely related obsolete word.

6 Classification of obsolete forms

Based on the obsolete lexical items which have been examined in the previous chapter, and in extension to previous research (Visser 1949, Görlach 1991), I will present a twofold classification of obsolescence. The proposed classification takes into account both the conditions and circumstances of a word's decline, as well as the actual realization of obsolescence, which may range from the obsolescence of a spelling variant, the word's confinement to a dialect, all the way to the complete disappearance of a form and the meaning associated with it. With the latter in mind, I suggest that each realization has its place on a scale featuring varying degrees of obsolescence, based on a number of factors including the lexical item's frequency and distribution in historical and synchronic corpora.

Of all the 851 candidates that were sorted, there were 134 errors or typos, 100 abbreviations, 26 foreign-language words, 195 proper nouns, 357 obsolete spellings, and finally 39 lexical items which were identified as instances of true obsolescence (as defined at the beginning of this thesis). In the process of analyzing the results, an additional 9 obsolete words were found in connection to the candidates that I already had extracted from the EEBO corpus. For example, when researching the obsolete noun *borrow,* I consequently came upon the existence of the corresponding adjective *borrow,* and by extension its own antonym, the adjective *transhaw.* In summation, the 9 obsolete words found outside the EEBO are *brandon (n.), prentice (n.)* 2 counts due to polysemy, *debonairity (n.), pylle (v.), unmeetly (j.), borrow (j.)* and *transhaw (j.).* This leaves a total of 48 lexical items which allegedly became obsolete before the year 1660, since the methodology specified that the words had to have 0 hits after this year, otherwise they would not have made the cut.

The previous chapters contain a description of the methodology used for the extraction of the results, and a detailed analysis of the obsolete forms themselves, with information regarding the form's frequencies in the corpus, the years in which it was last sighted (in fact, cited) according to the OED, as well as the specific conditions and processes which led to each form's ultimate decline. In addition to the methodology and extraction of obsolete forms, one of the objectives of this thesis was to propose a classification of obsolete forms which would be clear, concise, and systematic. In this chapter, I will attempt to do just that and, wherever possible, illustrate each category by providing relevant examples from my results.

This thesis operates on the basis of a corpus-driven methodology, which is why the point of departure were the observed cases (from the EEBO and also OED), based on which I modelled my classification schemes but while at the same time allowing for the possibility of categories which were not in our observed data, e.g., some of the examples from previous research (Visser 1949; Görlach 1991). An alternative strategy would be to first construct a more general framework and then see how the observed cases fit into it, and make the necessary adjustments to the framework ex post.

Observed data are the best means of assessing the merits of a classification system but relying solely on them as a source for such a scheme would not be wise, since we are always limited by the size and extent of the corpus/language data and can never be sure that all potential cases are accounted for. For this reason, a robust theoretical classification that can potentially address any new cases as they may appear, would probably be our safest bet, in the case of this thesis it would be the form/function dichotomy. The following section contains an in-depth survey of the existing classifications or, rather, previously defined categories of obsolescence (Visser 1949, Görlach 1991) in relation to the cause of the form's disappearance.

6.1   Earlier classifications

When dealing with the history of a language's word stock, emphasis is invariably placed on word-formation processes, borrowings resulting from language contact, semantic changes, etc. In short, we tend to take special interest in how a language gains new words and how they evolve and survive but, if previous literature can be taken as a reliable indicator, we are considerably less curious about how a language can lose some of its words (or even kill them off). Although previous literature dealing with lexical obsolescence is sparse and usually tends to be quite brief, there are two main sources which provide some valuable insights into this phenomenon, and together define 13 main categories into which we may divide all cases of lexical obsolescence.

In the space of an approximately three-page section in his *Introduction to Early Modern English*, Görlach (1991) provides a succinct and fairly comprehensive summary of the main causes of word loss, where he cites Visser (1949) as the main source on this topic. The causes of lexical obsolescence mentioned by Görlach are as follows:

1) disappearance of the word's original referent
2) the word is restricted to a dialect
3) weakening of emphasis through overuse
4) political correctness and use of euphemisms
5) homonymic or homophonic conflict
6) polysemy
7) phonic inadequacy (I have included this in a more inclusive category *formal inadequacy*)
8) word formation patterns are no longer productive
9) levelling of endings

In his 1949 lecture *Some causes of verbal obsolescence,* Visser presents a very thorough examination of word loss, with a rich selection of examples. Although the examples all serve to illustrate obsolescence in the Old English word stock, the processes described therein are universal and may be extended to the cases of lexical obsolescence in Early Modern English,

which is the focal point of this thesis. Visser's lecture includes four more causes which, although they may have been mentioned by Görlach (1991) in passing, were not defined as separate categories. They are:

10) language contact
11) synonymy
12) social levelling
13) fashionable language or slang words

Although these causes are defined as 13 separate categories, some are more closely related than others, for example *social levelling* and *fashionable language/slang words* both stem from the desire to belong to a certain speech community. Additionally, these categories are not completely independent of each other, and in some cases we may see them as one cause leading to another. For example, *synonymy* is a very frequent cause of lexical obsolescence, but in many cases it may be said that the root cause is actually *language contact*, resulting in the synonymy which proved fateful for the given form. Although the links between some of these causes are strong, they are not unidirectional; For example, the *levelling of endings* invariably led to *homonymic conflict*, but *homonymic conflict* was not always the result of *levelled endings*.

Of these categories, almost all relate to a cause of obsolescence, save for one; *word is restricted to a dialect.* The fact that a word is forced out of the standard and into strictly dialect usage (regional or social) is not a cause in itself, but rather the concrete manifestation of this phenomenon in the language. A word may be driven to the periphery of the language due to its competition with a synonym or as a result of the stigmatization of a social group associated with it, just to name a few possible causes. In this sense, *restriction to dialect* does not belong among the list of causes but is one of the possible realizations of lexical obsolescence. I attempt to resolve this discord by introducing a dual classification as seen below in section 6.2, and it is within the scope of the *realization* classification that I will be dealing with the category of *dialect* in section 6.2.1.11.

Below, each of the remaining 12 categories is described in more detail, with reference to the examples provided in Table 9, as well as additional examples from the results section or, wherever necessary, from earlier literature.

6.1.1   Disappearance of the word's original referent

Perhaps the most obvious cause of obsolescence is the disappearance of the object, custom, institution, etc. which the word denotes, as it is reasonable to assume that once an object no longer exists, it will be referred to less and less until eventually, the word itself disappears. In Table 9 I have included the example of *cubyte,* a unit of measurement which was gradually replaced by other units and became restricted to historical usages. Visser (1949: 7) provides Old

English examples of this type of obsolescence reflecting the cultural history of that period, most of which fall into the domain of weaponry, magic and witchcraft, garments and law. These include *malswyrd* (a sword decorated with an inlaid ornament), *dægredoffrung* (morning sacrifice), *sciccels* (a type of cloak), and, perhaps the most charming of them all, *ealugafol* (a tax paid in ale).

This category also includes instances, especially numerous among cases of obsolete Old English words, where the original referent does not cease to exist, but is simply no longer relevant within the context of the current society. A popular example of this is the loss of the two Old English nouns denoting the paternal (*fædera*) and maternal (*ēam*) uncle once this this distinction was no longer of social or legal relevance. In the case of Early Modern English, we witness a similar situation with the nouns *transhaw* and *borrow,* which were used to describe parts of a wall (although these concepts were most likely never as central to society as the words for *uncle*).

Finally, a concept can disappear as a result of political correctness, taboo or censorship, for example following the prohibition of religious rituals and institutions such as those rejected by the Anglican Church (Visser 1949: 13), although this is very rare, since such usually survive in closed circles and never become obsolete per se.

6.1.2   Weakening of emphasis through overuse

The perceived weakening of a word's impact due to its overuse is one of the psychological factors mentioned by Görlach in the list of possible causes of words loss. Although this process is not limited to the category of intensifiers, they do seem to be the most frequently cited victims in previous literature. Initially, when a word is used as a new intensifier, there is a strong association with its meaning, for instance *wondrous* was used as an intensifier in the sense of "to a wonderful degree," but with frequent use its impact became less and less powerful, and its link to the original meaning of "wonderful" may have also weakened. At that point, if the language contained a number of intensifiers whose meanings were vaguely somewhere in the "wonderful" field, it would be faced with the question of the sustainability of its synonyms, and over the course of time a preference for a portion of these synonyms would inevitably emerge, leading to the marginalization and, ultimately, loss of the remaining intensifiers. Aside from *wondrous,* Görlach includes the ME and EModE intensifiers *al, ful, right* and *wel* as examples of this phenomenon (1991: 142).

However, in the results of this thesis, I did not find any cases of obsolescence which could be attributed to the weakening of emphasis through overuse; for one, all cases of word loss found in this thesis had a more plausible or even obvious cause, and furthermore the obsolete forms were never particularly frequent in the EEBO corpus, so the cause of overuse is very unlikely.

6.1.3   Political correctness and use of euphemisms

As Visser argues (1949: 21), political correctness and "tabus (sic.) of indecency" are in many ways the driving forces behind changes in the vocabulary. The need to avoid a word which is considered offensive or untasteful leads to that word being replaced by a euphemism which, through frequent use, in turn comes to be perceived as equally undesirable, is replaced by another euphemism, and so the process continues, leading to the obsolescence or dialectalization of many such words. The following selection of words denoting what Visser euphemistically refers to as "the apartment" (ibid.), spanning from Anglo-Saxon to 20th-century English: *gangern, gangstol, privy, stool (of ease), jaques, water closet, convenience.*

A closely related phenomenon is the avoidance of previously innocuous words due to the fact that they are evocative of offensive subjects, widespread especially in the Victorian period, where *trousers* were referred to as *inexpressibles, unwhisperables, sit upons,* etc., (ibid.), although in this case it was the euphemistic forms which became obsolete once they no longer served their purpose while the original word survived.

The motivation to avoid words which are considered offensive or stigmatizing primarily stems from the cultural and social norms, but history has witnessed cases where more factors come into play, such as the interaction of political correctness and homophonic conflict, which can be illustrated using the example of *queen* and *quean* (prostitute), whose homophony created an unacceptable situation, leading to further tabooization and the ultimate demise of *quean* (Görlach 1991: 142).

6.1.4   Homonymic or homophonic conflict

When two words have an identical pronunciation and, in the case of homonymy, this is accompanied by identical spelling, the result is an ambiguity of meaning that language generally does not favour. *Homonyms* are two or more words which share the same spelling *(homographs),* as well as the same pronunciation *(homophones),* but, unlike *polysemes* (covered in the following section)*,* they are not related to each other in meaning*.*

The victims of obsolescence due to homonymic or homophonic conflict are most likely to be words which "[belong] to the same class and sharing syntactical and semantic features […] since they are easily confused in many contexts, which leads to misunderstandings and, in consequence, reduced frequency" (Görlach 1991: 142). This claim is confirmed by the cases cited in earlier literature as well as the results of this thesis; Homophony is represented by the already mentioned fate of *queen* and *quean*, and as for homonymy, it was most likely homonymic conflict that caused the obsolescence of *travailer* (meaning "worker") in Early Modern English, which for a while coexisted with its homophonic counterpart *travailer* ("traveller"). Both pairs fit Görlach's description of endangered homophones/homonyms, since they belonged to the same word class and were likely to cause confusion in a number of similar

contexts – as mentioned previously, the meaning of *travailer* in most of the EEBO hits was difficult to determine with confidence due to the ambiguous contexts in which it appeared.

Causes of *homonymy* and *homophony* include borrowings from other languages, sound changes, and, less commonly, shifts in meaning. Morphological levelling also resulted in abundant homonymy and homophony most notably in the Old English and Middle English lexicon and can be illustrated using the example of *lettan* and *lēten* which concerns the levelling of endings specifically (more on that in 6.1.8). As for borrowings, there is a great deal of homonymy attributable to this process, wherein the foreign word was introduced into the language in the same form as an already established native form. It is not at all difficult to find examples such as *band* (OE "that with or by which a person or thing is bound") and *band* (from French "an organized company; a troop"). Aside from *travailer,* the EEBO results also include the obsolete verbs *pylle* (from French "to pillage, plunder" and probably Latin "beat or strike violently"), in both cases homonyms of the native verb *pylle* (OE, "to peel, strip, pluck").

6.1.5   Polysemy

A similar motivation exists in polysemous pairs or groups of words, which result from semantic changes due to which "words acquired new senses while at the same time retaining their earlier ones" (Nevalainen 2006: 70). Since the multiple senses of one word are too close together, that word is semantically overburdened and often incapable of conveying the intended meaning with sufficient clarity; OE examples provided by Visser include *winn* meaning "toil; labour; trouble; hardship" or *ræd* meaning "advice; deliberation; plan; conspiracy; intelligence; profit; remedy […]" or *hæmed* "marriage; adultery" (1949: 17). The potentially embarrassing ambiguity stemming from this inevitably leads to obsolescence, but rather than losing the word entirely, polysemy usually leads to the loss of one or more sememes. For instance, when we look at the noun *cheere*, only a handful of its original sememes survive to this day; "the expression on or appearance of a person's face, as indicating emotion or character," "habitual behaviour; bearing, manners" and "insincere show of affection" have all been obsolete since the 16th century.

From the results in this thesis, I am mentioning the verb *pylle*, which, in addition to having lost several of its homonyms, also lost the sememes "to seize forcefully, snatch", "to become bald", and "by picking out foreign matter" to obsolescence, while the sememe "to peel, strip, pluck" survived.

6.1.6   Formal inadequacy

This category subsumes under the name *formal inadequacy* the question of word length, grammatical function and phonic inadequacy, since all of these causes are related to the word's formal characteristics and salience. There is no consensus on the topic of word length and its role in the loss or survival of words, because although shorter words are generally thought to be preferred over long ones, there is an opposing view that "shortness, or insignificance of sound, is

the bane of words" (Visser 1949: 14). In fact, there are arguments to support both sides: on the one hand, there is a strong tendency to shorten words such as *laboratory: lab, telephone: phone,* but on the other hand, there do seem to be limits to this, as attested by the fact that the Latin *apis* "bee" cannot be found in any modern dialect of French, since its pronunciation would be the single phoneme [e] (ibid.).

From the results of this thesis, consider the transitive verb *ne* (from French "to deny"), whose length and inadequate salience I believe to be among the causes of its demise. Another type of formal inadequacy is the inability of a form to express the necessary syntactic functions, for example there is insufficient information on tense in the preterite/participle *put,* which led to the form's replacement in sentences with a "past time-sphere" by the unambiguous *putted* in selected dialects (ibid. 15).

Phonic inadequacy, although named by Görlach as one of the causes of lexical obsolescence, was mentioned very briefly with reference to phonaesthemes such as *pl* and *sl.* The issue is that if, due to phonological changes, a cluster such as *pl* changed to *fl,* the originally desired onomatopoeic effect would be lost, leading to the word's replacement by a more appropriate one (Görlach 1991: 143; Visser 1949: 15). However, neither of the sources provided any examples from English, nor was I able to find any in the EEBO corpus. Related to this is the ousting of "difficult" or unfamiliar pronunciation, and in this context Visser (1949: 12) mentions the loss of all Anglo-Saxon words beginning with the sound clusters *fn, wl* and *hw,* which the French speakers assumedly perceived as foreign and certainly difficult to pronounce.

6.1.7   Word formation patterns

Throughout the history of the English language, the number and type of word formation strategies currently in use has been evolving due to a number of factors. For example, with the decline of the strong verb system in Middle English, the use of internal modification became a much rarer word formation strategy, while stress modification became one of the central morphological processes in the shaping of English (Nist 1966: 194). Likewise, with less emphasis on the native strategy of compounding, an increasing interest was taken in affixation.

The status and productivity of word formation patterns in Early Modern English was exceptionally influenced by language contact, and most markedly in the borrowing of affixes, which Görlach considers to have been "moderately productive" and used predominantly with foreign bases. The foreign affixes coexisted with native ones, for example the type Adj -> N included the suffixes *-th, -head, -ness, -ment, -esse, -ion, -ity,* etc. (1991: 175). Asa result of all this, the word formation strategies were much more flexible than they had ever been in Old and Middle English. Importantly, there was a great deal of freedom regarding the number of synonymous derivational forms that could coexist, in blatant disregard for the economy principle, for example the derivations of *throne* in the sense of "to remove a ruler from his position of power" included *disthrone, dethrone, unthrone,* and *dethronize* (Nevalainen 2006:

60). Naturally, most of these creations were short-lived, and the result was generally the survival of one and the obsolescence of the rest. The obsolete forms found in the EEBO included the noun *debonayr* (presumably a product of zero-derivation of the adjective) as well as its synonyms *debonairity* (†1688) and *debonairness* (rare usage), formed using the affixes *-ity* and *-ness*. Of these three, *debonairness* survived the longest, which is not unexpected since the suffix *-ness* was the most productive of its group (Görlach 1991: 175). Sound change is another factor with the potential to influence the viability of a word formation pattern, and can, for example, lead to the levelling of endings as seen in the following section.

6.1.8 Levelling of endings

The levelling of endings is a process that took place predominantly in the centuries between Old English and Middle English and resulted in the obsolescence of myriad forms. The process goes hand in hand with the language shifting from a synthetic (inflecting) system to an analytic one, where the latter (for example, present-day English) relies on auxiliary words and word order rather than inflectional endings to express relations between parts of speech. Levelling of endings involves two or more affixal forms from one linguistic paradigm converging, through sound change or analogy, and resulting in less variation within that paradigm.

Consider the example OE *þing,* where the levelling of endings would have resulted in the loss of forms in plural genitive (*þinga*) and dative (*þingum*) forms. In addition to the obvious obsolescence of forms, the convergence of paradigms resulted in cases of homonymy which in turn likely played a role in the complete loss of one of the homonyms. For example, the Old English words *lēten* ("to let") and *lettan* ("to hinder") became homonymous due to the phonological reduction of their respective endings, resulting in the interchangeably used forms *lete*, *lett* which put the sense of "hinder, prevent" in grave danger (it is labelled *archaic* in the OED, although the last citation dates back to 1885, making obsolescence a real possibility).

Compared to Early Modern English, this cause of word loss has a considerably higher representation in the Old and Middle English periods – in fact, I did not find a single instance of obsolescence that could be attributed to the levelling of endings in the EEBO data, whereas instances dating back to earlier centuries are numerous.

6.1.9 Language contact

Cited as the chief cause of enormous word loss following the Norman conquest (Visser 1949: 7), language contact is certainly one of the key catalysts for language change in general but vocabulary in particular, arguably the least stable component of a language. The conditions under which language contact takes place vary greatly, and it has been so throughout the history of English, whether it be warfare, trading, religious discourse, foreign invasions, contact through literature, or even academic/scientific writings. Furthermore, a number of factors can influence the degree and effectivity of borrowing.

One of these factors is the reception of the foreign language by native speakers, which in the case of the Norman conquest meant the mutual sympathy between the conquerors and the conquered, and it is thought that the respectful and peaceful coexistence between the Anglo-Saxon and Norman population is the reason why one language was not entirely replaced by the other, but rather the result was the amalgamation that we know from later centuries. The conditions following the Norman conquest determined which native words would be replaced by the French equivalent and vice versa; Visser argues that as a result of bilingual families, where the mother spoke English and the father spoke French, most of the household words remained English, while the non-domestic words were frequently replaced by their French translations, for instance *beorg* (mountain), *cneatung* (inquisition) and *hror* (calamity).

Long-term, everyday contact with French is also thought to have changed the speaker's speech habits and their perceptions of certain groups of words, specifically there was an avoidance (and ultimate obsolescence) of compounds, a typically Germanic structural feature (ibid.: 10). This lead to the loss of many Old English compounds such as *hwitleac* (onion), *wanderweorpe* (mole) and *handfæstan* (to betroth).

The phonetic and morphological similarities between the languages also played an important role in determining the extent to which borrowings could take place. A great example is the borrowing of Old Norse personal pronouns, which was possible a) thanks to the close contact between the language speakers, but also importantly b) the closeness of the two languages themselves.

The language contact that took place between English and other European languages in the Early Modern period was of a different nature; the majority of words which were borrowed into English from (predominantly) Latin, Italian and French at this point were of specialized domains, most often the sciences, the arts, architecture and fashion, and the motivation for borrowing was twofold; firstly, these languages were overwhelmingly perceived as superior to English and there was a certain prestige associated with the use of this vocabulary and, secondly, a sizeable portion of the borrowed words did not have a native equivalent. Moreover, many such lexical items would have been borrowed via written text and not through spoken contact between speakers.

Language contact is the language-external cause which subsequently sparks some of the more direct causes of lexical obsolescence; whether it be the closely related language-external causes of *social levelling* and *fashion/slang,* or language-internal causes such as *synonymy*, *polysemy*, or *changes* in word-formation patterns as a result of borrowings from foreign languages.

6.1.10 Synonymy

Synonymy proved to be one of the chief causes of word loss in the EEBO data, and its origins stem from language contact. A substantial number of Norman French words entered the English language after 1066, and several centuries later their Latin equivalents were added to the mix in a

new wave of borrowing (I say Latin, but in fact some of these were borrowed indirectly through French). This "simultaneous borrowing" resulted in an extensive collection of synonyms, realized by groups of three words with essentially the same meaning, differing only in their connotations and stylistic effect (Crystal, 194). Take for example the synonyms *ask, question, interrogate*; the Old English word is generally the most colloquial (often also the shortest), the Norman French is associated with a higher register or even literary genres, and the Latin is usually considered to be the most academic of the three, although of course in the case of *interrogate* we would speak of a semantic narrowing which has resulted in highly specialized usage. Presumably, it is owing to this specialized usage that all three synonyms have so far withstood the test of time.

However, there are countless examples of words which have not been so fortunate, since language does not take kindly to redundancy. From the EEBO corpus, I will mention *forepoint* ("to determine beforehand"), *putcase* ("hypothesis, supposition") *and unmeet* ("immense, huge"). Such was the fate of *nim,* which was replaced by its synonym *take* in general usage before the end of the 13th century, and was confined to cant, the secret language used by criminals, where it gained the specialized meaning of "to steal". Regardless, it too became obsolete by the end of the 17th century (Görlach 1991: 143).

6.1.11 Social levelling

Early Modern England was a classist society with an indisputable sense of hierarchy, but gradually, the opportunities for social mobility increased. The era following the Restoration allowed for the emergence of urban class (Everitt 1966: 72), attributable in part to the ease of travel, broadening of social networks and increased literacy within the population. Those who aspired to higher classes had a chance, possibly greater than ever before, to better themselves and move up in the world economically, socially, and intellectually. Language was a major class marker, and so there was a tendency to adopt the speech and mannerisms of higher classes. This led to the stigmatization of words which were primarily associated with regional dialects or language of the lower classes, leading to the loss of those forms or their confinement to dialects.

6.1.12 Fashion and slang words

Every new generation of language users replaces familiar words with innovative and fashionable slang. This category shows a great deal of variation and new words spread quickly, which Visser (1949: 23) demonstrates with the succession of words for "dandy" throughout history; *coxcomb* (1573), *puppy* (1598), *jack-a-napes* (1592), *bright* (1600), *flash* (1603), *trig* (1610) and many others. Most of these slang words are short-lived, soon coming to be replaced by a more fashionable term, so their effect on the language's vocabulary is negligible (Visser 1949: 23), but many of them can be found among obsolete forms. However, the results of this thesis do not contain any slang words, most likely due to their genre-specific and fleeting nature.

6.2   Proposed classifications

This chapter contains an outline of potential classification systems. The proposed methods of classification take into account both the causes of lexical obsolescence as defined by previous sources (Visser 1949, Görlach 1991) as well as the concrete ways in which obsolescence takes place, ranging from obsolescence of spelling to the complete obsolescence of form and function. An evaluation of their respective merits and drawbacks, especially from the perspective of needs specific to lexicographers and dictionary users, is provided in the discussion that follows.

6.2.1   Cause and realization

Earlier literature has a strong tendency to focus on the causes of lexical obsolescence, and these causes have been examined and covered in sufficient detail in the previous section. However, I propose that the manner in which this phenomenon manifests is equally as fascinating, and certainly deserving of attention in my proposed classification. For this reason, alongside the *causes* of lexical obsolescence, I have included the additional dimension of *realization,* which shows concrete ways in which the loss of a lexical item may manifest in the language.

In Table 9 below, I have encapsulated the causes of obsolescence as identified in previous research (Visser 1949, Görlach 1991), and further divided them into the sub-categories of *external* and *internal* – and how serendipitous that there is an equal number of language-external and language-internal causes! The second component of my proposed classification is that of *realization*, provided in Table 10, in which obsolescence is understood in the widest possible sense. That is, lexical obsolescence does not only constitute the disappearance of both form and meaning, but a whole spectrum of phenomena associated with lexical loss. Depending on how strictly one wishes to define lexical obsolescence, the different types of realization may be understood as degrees of obsolescence or, in some cases, steps leading to full obsolescence.

Language-external (Cx)	Language-internal (Ci)
real-world concept disappears *(cubyte)*	synonymy *(conduit, nim)*
language contact *(athel)*	homonymy *(travailer)*
censorship / taboo *(quean)*	polysemy *(cheere)*
weakening through overuse *(wondrous* as intensifier)	productivity/restrictions of word formation pattern *(debonairity)*
social levelling	levelling of endings *(lettan/lēten)*
fashion/slang	formal inadequacy* *(ne, v.)*

Table 9: A comprehensive overview of the causes of lexical obsolescence, both internal and external.

Realization (R)
obsolescence of form and meaning
disappearance in one word class *(borrow, n.)*
replacement by synonym *(travailer)*
disappearance of sememe *(host)*
dialectal/historical usage *(meyny)*
rebracketing/reanalysis (*gydre*)
spelling obsolescence *(throwghe)*

Table 10: A comprehensive overview of how the causes of lexical obsolescence are realized.

When applying this classification to the 48 suspected cases of obsolescence in the data extracted from the EEBO (not counting the 357 obsolete spellings), most often the form's loss could not be attributed to one single cause; on the contrary, it was frequently and interplay of several language-external *as well as* a related language-internal one. For example, the noun *meyny* ("a household, a family"), whose obsolescence was most likely caused by language contact (Cx) leading to synonymy (Ci), as well as subsequent stigmatization and social levelling (Cx), which ultimately led to the confinement of *meyny* to dialectal usage (R). If a form's obsolescence had a number of potential causes, I attempted to assign that form to the most prominent category whenever possible. The following subchapters contain a more detailed overview of the individual causes and realizations in terms of their applicability to the obsolete forms found by the methodology in the present study.

6.2.1.1 *Language-external causes of obsolescence*
The language-external causes were adopted from earlier classifications (Görlach 1991, Visser 1949), all of which have been described in detail in chapter 6.1. The cases of lexical obsolescence in the EEBO results which were linked to language-external causes can be seen in Table 11 below. The most numerous categories were a) the disappearance of the original referent, and b) language contact. There were two categories in the language-external group for which my methodology (keeping in mind that there were certain specifications that had to be met, cf. chapter 4.1) did not return any examples from the EEBO data, and those were a) censorship/taboo (*quean*), and b) fashion/slang (*nim*).

real-world concept disappears	*borrow_j, transhaw_j, rosicleer_n, cubyte_n, paynim_n, imbroccata_n, hackbutter_n, ensignebearer_n, levyte_n, breastlap_n, prentice_n, brandon_n, redcrosse_j*
language contact	*unmeet_j, unmeetly_j, unmeet_av, unmeetly_av, ne_v, meyny_n, paynim_n, debonayr_j*
weakening through overuse	*capitane_j*
social levelling	*putcase_n*

Table 11: Language-external causes of obsolescence and examples found in the EEBO data

When speaking of language-external causes, it is necessary to establish whether this signifies causes from outside the language or causes from another language (which would essentially limit the category to *language contact*). In this case, what is meant is the effect of social and historical realities on changes in the language.

6.2.1.2 *Language-internal causes of obsolescence*

Also following from earlier classifications were the categories for language-internal causes of obsolescence (see chapter 3.1). Given the abundant language contact that was taking place throughout the Early Modern period, it is not surprising that the category of synonymy was by far the most frequent. The individual categories and their examples can be viewed in Table 12 below. Cases of obsolescence attributable to the language-internal causes a) levelling of endings (*lettan/leten*) and b) polysemy (*cheere*) were not found among the obsolete forms in the EEBO data.

synonymy	*borrow_n, unmeet_j, unmeetly_j, conduit_v, reappose_v, meyny_n, paynim_n, debonayr_j, heerewithall_av, capitane_j, forepoint_v, breastlap_n, prentice_n, putcase_n, barnis_f-la, certess_n, for-that-cause_n*
homonymy	*pylle_v, travailer_n*
phonic inadequacy	ne_v
productivity/restrictions of word formation pattern	*commise_v, caluines_nn, debonairity_n, civil_v, myrte_zz, conjoint_n, constraint_j*

Table 12: Language-internal causes of obsolescence and examples found in the EEBO data

6.2.1.3 *Realizations of obsolescence on a cline*

In efforts to acknowledge that obsolescence is not simply binary, but that there are many nuances in what constitutes word loss, I have proposed a cline which shows various types – or levels – of obsolescence (Figure 15). The purpose of this visualization is to demonstrate the difficulty of determining where to draw the line between a word that is obsolete and one that is not. How much of a lexical item must be stripped away before it can be considered obsolete?

obsolescence of form and meaning

replacement by a synonym

disappearance of one sememe (polysemy)

disappearance from one word class

restriction to historical usages

restriction to regiolect/sociolect

rebracketing

reanalysis

obsolescence of spelling

Figure 15: Proposed cline covering types of lexical loss, from spelling obsolescence to full obsolescence of form and meaning.

The realizations on the cline may be roughly divided into groups based on the degree of word loss. In some cases, these could possibly be viewed as stages leading to full obsolescence, for example should a word be restricted to a dialect, where only some facets of its original meaning remain, until it ultimately disappears from the language altogether. The following four levels were defined in the context of formal obsolescence, since that is the primary concern of this thesis:

1. **Full obsolescence**: disappearance of form and meaning, or replacement by synonym
2. **Selective obsolescence**: disappearance of sememe or word class
3. **Disappearance from standard usage**: restriction to historical or regional usages
4. **Orthographic**: rebracketing, reanalysis and obsolescence of spelling

6.2.1.4   Realization categories for the EEBO results

The 48 obsolete forms were divided into categories based on the type of obsolescence, as well as the 357 cases of obsolete spellings, as listed below in Table 13. Using the same data, Figure 16 shows a graph of the realizations of obsolescence, but excluding the category of spelling, which is significantly more numerous and would overshadow the remaining categories.

Type	Count
spelling	357
rebracketing	3
dialectal	4
historical	7
word class	5
sememe	5
synonym	18
full obsolescence	6
Total	405

Table 13: The classification of obsolete words in the EEBO in terms of realization

Figure 16: The classification of obsolete forms (excluding 357 cases of spelling) in the EEBO in terms of realization ordered with respect to their degrees of obsolescence.
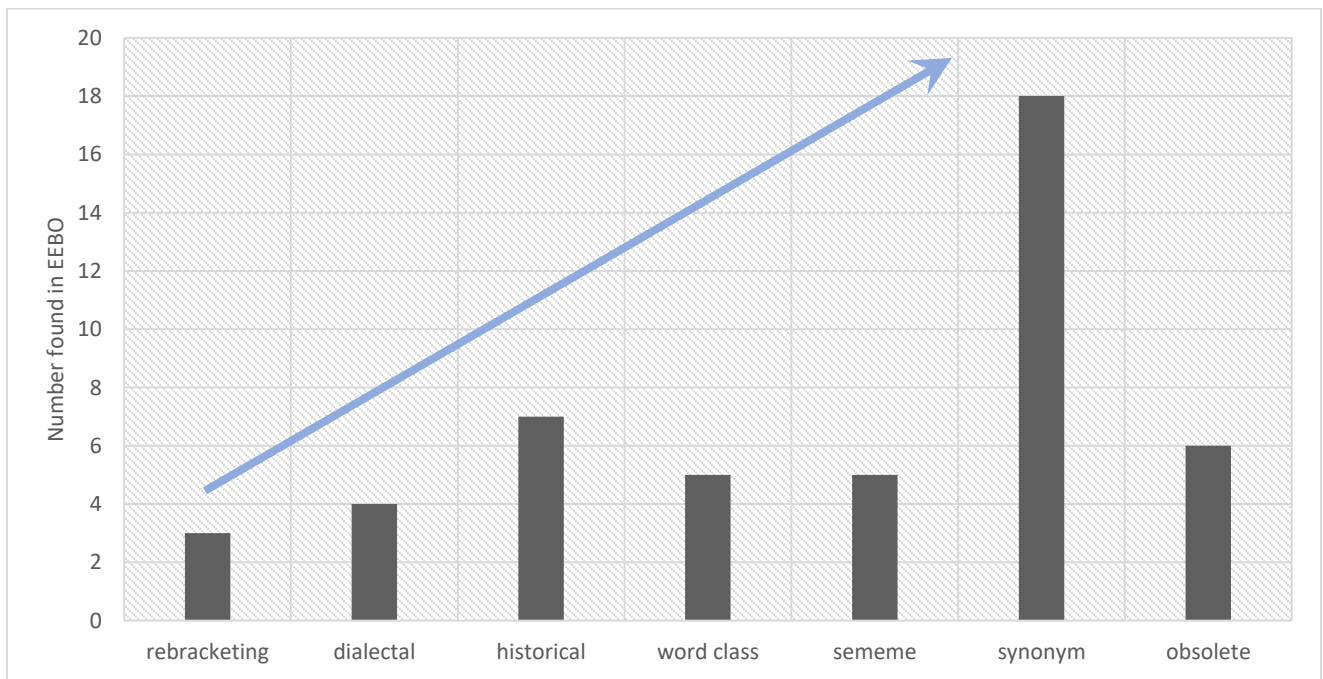
6.2.1.5 Consolidating cause and realization

Based on the results available from this study, we can observe how the two components of the dual classification work together and represent two ways of looking at the same issue. Although there is no single exclusive link between a cause (C) and realization (R), it stands to reason that there are certain logical links and tendencies. In my analysis of the results, I have found that it is often the case that one word's decline may be attributed to an interplay of several causes, for example the cause of language contact more frequently than not goes hand in hand with the language-internal cause of synonymy. Likewise, there is frequently a direct link between cause and realization, for instance we may take the above case of synonymy, which virtually always manifests as the replacement of the word in question by one of its synonyms, rather than, say, the restriction of one of the synonyms to peripheral (dialectal) usages.

- language contact (Cx) frequently leads to synonymy (Ci) and the replacement of one word by its synonym (R)
- the dispreference for a word formation pattern (C) often leads to the replacement of the undesired word by its synonym (R)
- the disappearance of a concept (Cx) inevitably leads to full obsolescence (R)

6.2.1.6 Obsolescence of form and meaning

Perhaps the most prototypical realization of lexical obsolescence is the disappearance of the original referent (concept, institution or physical object), which in turn renders the lexical item pointless. The loss of words denoting things that disappeared was not uncommon in earlier periods, namely the loss of Old English words following "the introduction of French substitutes or innovations" (Visser 1949: 7), however, in the EEBO data I found only 6 instances of this full obsolescence, including the noun *rosicleer* ("a type of worthy knight") and the adjective *transhaw* ("said of the pitch of a wall, perhaps 'exposed' to the blast"). This category is without a doubt an example of *full obsolescence*.

6.2.1.7 Replacement by a synonym

Synonymy is perhaps the most common cause of redundant items in the lexicon (chapter 6.1.10) and subsequently one of the chief causes of word loss. The disappearance of a form and its replacement by an already existing synonym is a realization of lexical obsolescence even if we were to apply the strictest definition, and I would therefore consider it *full obsolescence*. Examples from the EEBO include the verbs *conduit* ("conduct, pour forth") and *forepoint* ("forebode"), and the noun *putcase* ("hypothesis).

6.2.1.8 Disappearance of one sememe

The disappearance of one or more sememes is the most straightforward realization of obsolescence brought about by the language-internal cause polysemy (discussed in detail in

chapter 6.1.5), wherein the word survives along with at least one component of its meaning. For example, the EEBO corpus bears witness to the obsolescence of selected sememes of the noun *levyte* (now Levite), which retained its primary meaning "descendant of Levi" but lost the secondary senses "deacon" and "clergyman".

6.2.1.9   Obsolescence limited to word class

This type of obsolescence involves pairs or groups of words which are identical in form, related in meaning, but belong to different word classes. The most obvious origin of such groups is zero derivation, where the derivate is identical to the original word, such as the adjective *capitane* (Obs. "chief, main") derived from the noun *capitane* ("captain"). Another source are groups of words which existed in Old and Middle English as distinguishable forms such as the OE verb *wyrcan* and noun *weorc*, but phonological and morphological changes led to their convergence into the identical forms *work* (verb and noun).

However, unlike the case of polysemes, it is highly unlikely that the identity of forms from different word classes (e.g., adjective and noun) would lead to unwanted ambiguity and establish sufficient motivation for the obsolescence of one or the other. Rather, the loss of a word would be caused by a process unrelated to that word's shape, and the fact that an identical form survived in another word class was only a happy coincidence. For instance, the synonymy which led to the decline of *capitane* affected only the adjective, because the noun *capitane* did not have any synonyms which would be in competition.

Examples of obsolescence limited to one word class from the EEBO corpus include the nouns *borrow* ("by way of loan")*, otherwise* ("another way") and the verb *ordnance* ("to provide with military equipment")*.

6.2.1.10 Restriction to historical usages

This category contains words which the OED labels as *Historical usage.* Although these words are effectively obsolete in the sense that they are not actively used in contemporary language, they are still encountered and understood in specific contexts in relation to literature or past events. The categories used by the OED to label its entries include *archaic, historical* and *obsolete,* and there is no documentation providing any kind of definition that could help us differentiate between these three. Indeed, in many cases it is difficult to draw the line between *historical usage* and *obsolescence,* and I expect that the decision was left to the discretion of the editors.

6.2.1.11 Restriction to dialect

Obsolescence in the scope of the main dialect may often involve the word's continued usage in regiolect or sociolects. One of the key causes of this is language contact, which was often localized to the standard dialect and did not always reach the more peripheral dialects. This serves as a viable explanation for why the Old English *ēam,* although replaced by the French

equivalent *uncle* through language contact*,* still survives to this day in Scottish English in the form of *eam*. Social levelling and the stigmatization of local varieties

Another possible cause for the disappearance of a word from standard but not dialectal usage is the fact that sound laws operated differently across dialects, and so a word which became a homophone in – and subsequently disappeared from – the standard dialect very often survived in other varieties of the language (Visser 1949: 20).

The results from the EEBO data include the following examples of forms that are obsolete in the standard variant but survive in regional dialects: the nouns *barnis* ("children") and *meyny* ("family, household," now exists as a Scottish and Irish variant)*,* possibly also the determiner *thylk* ("the very [thing/person/etc.]")*.* Admittedly, this is a group of words which occupies the no man's land between obsolescence and present-day regional (also possibly archaic) usage. An argument for obsolescence is that if a word exists only in the scope of a regiolect or sociolect – although it is not obsolete in the sense of having disappeared from the language entirely – it can be considered obsolete from the perspective of the standard.

6.2.1.12 Rebracketing and reanalysis

This category is concerned with formal obsolescence due to a shift in the perceived morphemic boundary between two words, and the subsequent creation of a new form which replaces the old. The most notorious examples of rebracketing are the OE *nædre* and ME *naperonn*, where the initial <n> was taken for an article and resulted in the forms *an adder* and *an apron*. The examples of possible reanalysis that were found in the EEBO corpus are not as straightforward; the noun *gydre* used in the phrase "to gydre", which would later become *together,* might better be more accurately called compounding. Compounding is also evident in the adjective *thesayde* ("the aforesaid").

For cases of rebracketing which result in obsolescence, the word loss is invariably driven by a change in orthography, one might be inclined to equate it with spelling obsolescence. However, given the fact that the process involves a shift in morphemic boundaries, I would argue that it is a level above spelling obsolescence.

6.2.1.13 Spelling obsolescence

Spelling is arguably the least stable facet of a language and therefore formal obsolescence due to spelling changes is not at all uncommon in the history of English. The list of causes in Table 9 does not include *phonological change*, which almost exclusively associated with spelling obsolescence. The language-external (Cx) and language-internal (Ci) causes in relation to spelling are as follows:

- **Language contact**: Following the Norman conquest, French spelling conventions were applied to English (Nevalainen 2006: 31). Uncontrolled simultaneous borrowing led to

multiple spellings such as the 17th-century borrowing *juvenile/iuuenile*, of which only one survived.

- **Social levelling**: Standardization, most notably through the dissemination of printed material, "the first national models for spelling, the royal writing offices and early printing presses, were based in the capital and referred to the southern rather than the northern dialects for spelling norms" (Nevalainen 2006: 32). Related to social levelling was the perceived prestige of French and Latin, which encouraged the introduction of etymological spellings such as *debt* for ME *dette,* which was a more accurate reflection of the Latin etymon *dēbitum*, (ibid.)

- **Levelling of endings**: when inflectional endings were reduced to <-e>, this development was reflected in the spelling of those words, and later the final <-e> was lost completely (ibid.)

- **Phonological change**: Spelling may change in order to accommodate the changed pronunciation of words. In Middle English, the spelling was phonemic and reflected the different pronunciations of a word across dialects, resulting in a great deal of variation which can be witnessed well into Early Modern English.

6.2.2   Form and function

Aside from the cause and realization approach, another possible way of classifying types of obsolescence is from the perspective of form and function, namely which of these two components is affected. On the most fundamental level, the nature of the form or function leading to obsolescence can be characterized by *Unsustainability* and *Duality* (Table 14).

	Form	**Function**
Unsustainability	*the form disappears:* • spelling obsolescence • levelling of endings • phonic inadequacy	*the concept is no longer denoted:* • disappearance of referent • taboo • fashion
Duality	*form causes confusion and loses one of its functions:* • homonymy • polysemy	*several forms share one function*: • synonymy

Table 14: Characteristics of form and function which may trigger lexical obsolescence

	Form	Function
Inadequacy	• phonic inadequacy *(ne, v.)* • spelling obsolescence *(throwghe)*	• real-world concept disappears *(cubyte)*
Modification	• levelling of endings *(lettan/leten)* • productivity/restrictions of WF pattern *(debonairity)*	• weakening through overuse *(wondrous* as intensifier) • disappearance in one word class *(borrow, n.)*
Perception	• social levelling, fashion/slang *(nim)*	• taboo *(quean)*
Competition	• homonymy *(travailer)* • polysemy *(cheere)*	• synonymy *(conduit)*

Table 15: Processes affecting form and function which may lead to lexical obsolescence, with examples from Early Modern English

This idea is further developed in Table 15 above, but with a greater degree of granularity, wherein *Duality* and *Competition* comprise the same types of obsolescence, but *Unsustainability* can be further split into *Inadequacy* (the form or the concept is no longer viable), *Perception* (shift in the perception of language users leading to a – to some degree – conscious decision) and *Modification* (a shift on the paradigmatic or structural level).

7   Discussion

Every methodology has its own merits and pitfalls, and the one used in this thesis is no exception. Below, I will attempt to summarize the most important stages of the process and assess their possible effect on the results of this study. For a timeframe as broad as the Early Modern period (1470-1710 is the exact timespan covered in this research) and a corpus as extensive as the EEBO, it may come as a surprise that the methodology has uncovered a mere 48 confirmed cases of lexical obsolescence. However, given the limitations and characteristics of the EEBO corpus, as well as the strict criteria used for filtering the candidates, a number in this range was to be expected. The process is described in more detail in the following chapters.

7.1   Limitations of the methodology

In the methodology for the extraction of potentially obsolete words, the final specifications in the Python3 script were the result of great degree of freedom that was afforded to me in this process. This involves the question of setting a word's minimum frequency of occurrence in the earlier decades, determining that potential obsolescence would be signaled by the word's complete absence in the last decades of the EEBO, and likewise the decision of where to draw the line between the earlier, middle and later decades.

7.1.1   Demarcation of decades

The minimum frequency for a potential candidate was 50 occurrences in the first 14 decades (1470-1609), which falls into the OED's Frequency Band number 5, i.e., words which are restricted to educated discourse but still understandable. Of all the words in the EEBO which met these criteria, the candidates for manual analysis were the ones that simultaneously met the condition of zero hits in the final six decades (1660-1719), where the pseudo-arbitrary boundary was set for the 1660s, which corresponds with the restoration of monarchy in Britain and the onset of a more utilitarian approach to language. However, it is a fact that I could have just as well chosen to draw the line at an earlier or later decade, since it is doubtful that, for instance, the first ten years of Charles II's reign would have made for a decisive delimitation between a word's living and obsolete status, and in this respect the specifications were somewhat arbitrary.

For the EEBO corpus, the time periods are predefined, in 10-year increments. Language change, on the other hand, is not so straightforwardly linear, and so there are likely to be periods where the given phenomenon (e.g., word frequency) does not exhibit variation, whereas in other periods greater granularity would be conducive to more accurate results with clear trends. Although for the purposes of this research the 10-year time periods returned a sufficient number of obsolete forms needed for the subsequent classification, an alternative route might be data-driven periodization. It is an approach that could produce more meaningful time periods and obsolescence trends based on probability distributions which are compared using the Kullback-

Leibler Divergence (or relative entropy) method designed for "detecting features involved in diachronic linguistic change and discerning periods of change without pre-selection of features and periods" (Degaetano-Ortlieb and Teich 2018: 30). Similarly, Hilpert and Gries propose a variability-based neighbour clustering (VNC) principle of data-driven periodization, which better allows for periods of stagnation and ensures that "(1) parts of the data that exhibit similar characteristics [...] form part of the same corpus period and (2) breaks between different periods [are] inserted at points in time where there are measurable shifts in the characteristics of the data" (2020: 48).

Tichý (2018b: 34) addresses the problem of disparate sample size by assigning confidence intervals to the individual time periods, making it possible to draw more meaningful conclusions regarding the orthographic variability in the data. For future research, Tichý (ibid.: 38) suggests a sliding window sampling method as a way of achieving balanced time intervals with the same number of tokens. Such an approach promises to more accurately determine periods in the corpus where obsolescence was most pervasive as well as providing an alternative method for identifying obsolescent words or constructions.

7.1.2 Frequency thresholds

In Tichý's (2018a) study on obsolescence in Late Modern English, the first round of identifying candidates involved taking those n-grams that were in the top 40,000 most frequent in at least one of the Google Books decades, but with frequencies below 0.03 i.p.m. in the last decade. An important component of Tichý methodology is the Obsolescence Index which helps identify the most interesting candidates. The Obsolescence Index does this by comparing the maximum relative frequency of a given word and its relative frequency in the final decade, making it possible to catalogue and filter suitable candidates for subsequent analysis.

In the present study it did not prove necessary to calculate the Obsolescence Index, since the frequency limits were designed to extract only those candidates where a) the cases of obsolescence were as clear-cut and unambiguous as possible with a frequency of zero in the final decades, and b) the number of candidates returned by the script would be approximately between 500 to 1000 words. To address the need for clear examples, we needed to find words which had not been conspicuously rare at the onset of the Early Modern period, and yet regardless of this had completely disappeared from the (written) language by the end of the 17th century. Therefore, the minimal frequency limit in the initial decades was set so that all obsolescence candidates would be words with frequencies corresponding to the OED's (2021) frequency bands 5 to 8, i.e., common words in everyday use. This is in line with the frequency constraints by Tichý, who notes that based on the frequency bands it is "possible to postulate that common words are those with a frequency over 1 ppm [i.p.m.] or, in other terms, common words are the 40,000 most frequent words in a given decade." The premise is that we can estimate a given

word's likely status in the language based on where it appears in these frequency bands, which are defined as follows:

Band 8 consists of words with a frequency of over 1000 i.p.m., the most common English words including determiners, pronouns, prepositions, modals, auxiliaries, and conjunctions. Examples: *a, the, it, they, who, of, from, can, may, do, have, and, but.*

Band 7 includes words between 100 and 1000 i.p.m., the rudimentary vocabulary in modern everyday usage such as basic terms denoting and describing people, time, measurements and objects. Examples: *man, woman, hour, day, animal, house, big, small, good, bad.*

Band 6 contains words with frequencies between 10 and 100 i.p.m., the vocabulary covering almost all instances of everyday communication, including nouns and adjectives denoting "specific objects, entities, processes, and ideas [and] describing the qualities of particular situations, states of affairs, etc., or people's actions in particular contexts" (OED 2021). Examples: *machine, explosion, traditional, successful, red, green, Scottish, Irish, democracy, capitalism.*

Band 5 consists of words which fall between 1 and 10 i.p.m., these words are generally associated with a more learned vocabulary but will still appear easily recognizable to most native speakers. These words are not overly technical and in certain demographics can even be used in everyday conversation. Examples: *surveillance, authorized, jeopardize, gravitate, empirically, disproportionately, Platonic, Darwinian*.

Band 4 contains words with frequencies between 0.1 and 1.0 i.p.m., this vocabulary is "marked by much greater specificity and a wider range of register, regionality, and subject domain" (OED 2021). Nevertheless, the majority of these words would be easily understood by native speakers if they were to appear in fiction or journalism. Examples: *overhang, embouchure, astrological, egregious, insolent, galvanize, skyrocket, befuddle, methodically, pleasurably, surreptitiously*.

Band 3 includes words between 0.01 and 0.1 i.p.m., words in this category only rarely appear in written texts such as newspaper articles, but they are not entirely incomprehensible to the English-speaking population. The adjectives, adverbs and verbs in this category are either very technical or colloquial. Examples: *Ebullition, merengue, amortizable, agglutinative, cutesy, teensy, badass, emote, mosey.*

Band 2 contains words under 0.01 i.p.m., these are highly technical terms virtually unintelligible to most language speakers. Examples: *ennead, scintillometer, geogenic, abactinal absterge, haver-cake, hidlings, unwhigged, supersubtilized, gummose, cloit, stoothe, lawnly, whethersoever.*

Band 1 consists of only "extremely rare words" which are usually either particularly technical, archaic or restricted to historical use, and "unlikely ever to appear in modern text"

(OED 2021). However, they cannot yet be considered obsolete. Examples: *abaptiston, abaxile, grithbreach, gurhofite, zarnich, zeagonite.*

The bands were defined using contemporary English data (1970 to present), which begs the question whether the distribution of lexis in an older variety of English will be roughly the same, making these frequency categories applicable to Early Modern English also. In an experiment aiming to determine the size of a language's core vocabulary, Cvrček (2014: 4) was able to compare the core vocabulary size of three typologically distinct languages (English, Czech and Italian) by observing changes in the hapax-type ratio, which increased as new words were gradually added to the corpus, until eventually reaching a turning point; for English, the number of types at minimal point was 80,753 for word forms and 31,218 lemmas (Cvrček 2014: 13). The same method, if applied to a lemmatized version of an Early Modern English corpus and compared to the results for present-day English, could provide an answer to the question of whether the OED frequency bands are in fact applicable to Early Modern English.

7.1.3 Degree of obsolescence permitted

At the onset, the intention was to focus solely on the complete loss of lexical items, i.e., the final stage of obsolescence, but ultimately the data showed many cases which were not in the final stage, but were for example words restricted to dialectal usage, i.e., this obsolescence is a process but also (if we look at it as a state) several different realizations on a cline. Rudnická (2019) provides this definition of (grammatical) obsolescence:

> "Grammatical obsolescence describes a situation in which a previously popular and productive construction is, often gradually, losing its productivity and popularity over time until the construction disappears or there are only residues or fossilised forms left. The function of the obsolescent construction may discontinue or continue to be (fully or partially) expressed by alternative means."

Being obsolete and disappearing from the active and productive language usage is one of the potential final stages of the process of obsolescence. Furthermore, the area of interest as defined in chapter 1, *formal obsolescence*, is only concerned with the disappearance of a word's form, and not its meaning. The main reason for this decision was the fact that disappearance of meaning does not easily lend itself to a quantitative corpus analysis, such as the one employed in this methodology. However, owing to the presence of the part-of-speech tags provided in the Northwestern University version of EEBO corpus, some of the cases which I managed to identify in the analysis are in fact not examples of formal obsolescence, but rather loss of one function/sememe, such as the noun *borrow.*

7.2   Limitations of the corpus

7.2.1   Representativeness

Although the EEBO corpus is supposed to be a representation of all Early Modern English printed texts, it by no means covers all texts produced at the time. It is also important to bear in mind that, owing to its nature, the EEBO does not hold record of personal correspondence nor of the vernacular as it was spoken at the time, unless of course it appeared as a quotation in one of the contemporary publications. That is to say that although it provides as comprehensive a picture of the language as we could hope for, it does not guarantee that it contains all of the words that became obsolete in this period, much less in the frequencies that this methodology required (i.e., at least 50 occurrences in the first 14 decades).

It cannot be ruled out, therefore, that the EEBO corpus contains words which did become obsolete in the later decades but, due to frequencies lower than 50 in the earlier decades, flew under the radar. Conversely, the fact that a word's frequencies indicate that it became obsolete, it must be noted that it became obsolete within the scope of the corpus, and it is by no means certain that the words disappeared from the language at the same time. Even so, this should not be considered a major drawback in the case of this particular research endeavor, given that the principal aim was to produce a potential system for classifying obsolete forms, rather than producing a definitive list of all obsolete forms in Early Modern English.

7.2.2   Spelling normalization

An important feature of the EEBO corpus is that the spelling has been normalized only to an extent. The normalization of spelling in a body of text of this size (and with spelling as varied as that in Early Modern English) is a very challenging and time-consuming undertaking, and so the editors began with the most frequent words, and only covered a portion of the total corpus. For the purposes of this thesis, I have decided to use rather strict criteria for determining whether a lexical item is to be considered obsolete. This means that I made the choice to disregard any manifestations of obsolescence limited to the orthographic level, which would include cases of rebracketing, reanalysis and, most notably, obsolete spellings of extant words, which my results contained in great quantities. Of the 851 candidates returned by the script, if spelling changes, i.e., the lowest tier of obsolescence, were to be counted, we might add another 357 words, and boast a total of 405 (see Table 13) cases of obsolescence.

Seeing as the primary interest of this research lay outside of orthographic variation, the fact that the spelling in the EEBO corpus was only partially normalized has resulted in two issues:

a) Since over the course of the Early Modern period spelling gradually became more standardized, the later decades saw a decrease in the number of spelling variants per word, for example in the years 1470-1599, the EEBO corpus contains 13 different spelling variants of *judgement,* with a total of 35,596 hits including *iudgement, judgment, iudgemente,* etc. In

contrast, out of 33,442 hits there are only 6 distinct forms of *judgement* in the later periods (1660-1710). The number of instances in which *judgement* appears (in one form or other) in both these periods is comparable, but the later period suggests a significantly higher level of standardization. Owing to the fact that the script's specifications were designed to return forms which appeared in the years 1470-1599 and were lost by 1660, the list of potentially obsolete items contained a high number of outdated spelling variants of words which, however, are extant and therefore would not count as cases of lexical obsolescence in the narrowest sense of the word. As a result of this, the manual sorting and analysis of the results was a rather lengthy process, since there was a substantial amount of filtering and checking to see if the form at hand in fact was just an outdated spelling and not another separate lexical item.

b) When comparing the normalized EEBO data from Northwestern University and the EEBO version available at the Czech National Corpus, I encountered a minor issue regarding the frequencies of the individual items, which was that the frequencies provided in the data from Northwestern did not always perfectly correspond to the frequencies I found for the same word in the Czech National Corpus. Since the Northwestern data were the first point of contact with the corpus and they were seen as a tool to help quickly filter out the most likely candidates for obsolescence, the exact numbers were not seen as an issue. In the section where the individual instances of obsolete words are reviewed, I use the data from the EEBO version accessed via the Czech National Corpus interface, since it is the version where I also have access to the specific examples. However, this raises the question of how insufficiently normalized spelling might have skewed the numbers, and due to the high numbers of orthographic representations of one lexical item we can only speculate how many cases of obsolescence went undiscovered. Take, for example, the varied spellings of *debonayr* (n.) which were found in the EEBO corpus using the following query: [word="[d]ebb?onn?[iy]rr?e?"]. The query returned a total 674 instances with 9 spelling variants (see Table 8 in chapter 5.1.3), whereas the Northwestern data returned a mere 148 hits for the word debonayr, and the remaining spelling variants were analyzed separately.

7.3   Further questions regarding the proposed classification

Although the proposed systems of classification aim to strike a balance between comprehensiveness and practicality, they are not without their shortcomings when dealing with complex cases of obsolescence. As mentioned in chapter 6, for a number of examples there has proved to be some uncertainty as to the possible cause of obsolescence, which raises the question of whether we should not simply choose one primary cause which we deem the most salient. If we do decide to take this route, determining the primary, most salient cause is surely bound to be subjective. Perhaps it should be the first identifiable factor in a chain of causes leading to a form's obsolescence or, conversely, the last one. Görlach likewise acknowledges that "the number of concurrent factors involved [in lexical obsolescence] often makes it difficult or

impossible to reconstruct the specific causes that have led to the loss of an individual word," (1991: 140)

There is a category of words which on the surface so appear to have been the victims of obsolescence but have been excluded from the classification. Ghost words are those which come into existence by some oversight, perhaps due to an error in spelling at the printing house or confusion of pronunciation which was then duly recorded in a dictionary. Such words were never actually part of the lexicon, in fact they hardly ever existed in the first place, which is why they are not counted as obsolete in the scope of the present study.

The results derived from the EEBO contain several examples of what we might call marginal obsolescence. These are words which existed in the core vocabulary but became limited to dialectal use. The question here is whether disappearance from the main dialect is reason enough for us to speak of obsolescence. There is a sense that a line should be drawn, but I would venture that is not necessary to strictly divide what is and what is not obsolescence if we concede that obsolescence might be perceived as a cline (shown in Figure 15), ranging from spelling obsolescence to the aforementioned marginal obsolescence (limited to dialect or word class) and finally to full obsolescence of form and meaning. Lastly, there is the singular example of a word that becomes obsolete at one point in time, such as the case of the verb *reappose* (chapter 5.1.16) but due to productive word formation patterns it is reconstructed centuries later (possibly with a shift in meaning, Is the word then resurrected, no longer obsolete, or do we treat the new emergence as a coinage in its own right?

7.3.1 A system of classification for dictionaries

One of the main objectives of this thesis was to propose a lexicographical classification of obsolescence, in the hope that it may provide a stable framework for a more systematic practice of labelling obsolete forms, which could then be utilized in contemporary lexicography and English dictionaries specifically. In order for this system to have the potential for practical application in lexicography, it should be clearly structured, unambiguous and prepared to address even the most complicated cases of obsolescence without being confusing or cumbersome. Given the frequent interplay of a number of processes involving formal obsolescence, I have found it useful to distinguish between cause and realization, and furthermore divide the causes into language-external and -internal (chapter 6.2.1), since it is often the case that one leads to the other (such as language contact (Cx) and synonymy (Ci)).

In light of all the difficulties encountered when attempting to classify the results of this research, it is obvious that the greatest challenge has been to achieve a balance between complexity and clarity. A good point of departure would be to consider the needs of lexicographers and dictionary users for whom this classification is designed; in the OED, it is possible to search based on the categories *Origin, Subject, Region* and *Usage.* In the latter category, the types of usage potentially relating to obsolescence are *Obsolete, Historical,*

Archaic, Disused, Irregular, Rare and Regional. Although we might intuitively guess what some of these labels mean, there are no definitions of these available and it is difficult to imagine what they mean in relation to the degree of obsolescence. The metadata that is completely lacking in the OED concerns the cause of obsolescence, indicating that this category is either a) difficult to determine, or b) of secondary importance.

7.4   Further suggestions

One of the difficulties encountered in this research was the disparity between the frequencies found in the Northwestern data and the unlemmatized version provided by the Czech National Corpus. In what was essentially a trade-off between convenience and accuracy, convenience won out. However, it is likely that the original methodology applied to an unlemmatized version of the same corpus would result in a more exhaustive list of obsolescence candidates, even after discarding the cases of spelling obsolescence, though such an endeavor would certainly require extensive manual sorting.

Another possible improvement to the results might be a different candidate mining strategy. In further research, an alternative approach to the corpus should include data-driven periodization, firstly for the purpose of directly mining the candidates and secondly for a better delimitation of the time periods in which significant language change took place. The methodology would allow us to dispense with any preconceived notions of what the minimum frequency should be. Further steps might be the inclusion of n-grams, making it possible to find not only obsolete word forms but entire collocations or grammatical constructions (*to have compassion on sb.*, *on the morrow*, *to fall among sb.*), which fell beyond the scope of this thesis.

The key positive outcome of the present research is the proposed classification of lexical obsolescence, modelled on authentic examples from Early Modern English. An invaluable next step would be to conduct a round of user testing, which would provide us with an evaluation of the viability of the proposed systems of classification and how successful they would be when applied to obsolescence in different periods, or even different languages. In any case, it is vital that the practicality of the classification in lexicography and dictionaries should be tested and evaluated by those who would ultimately be using it.

8   Conclusion

Lexical obsolescence is an under-researched topic in general, and so it is not particularly surprising that the obsolescence of the Early Modern English word stock is covered by only a handful of publications, in varying degrees of detail. In existing literature describing the English lexicon, far greater attention has been paid to the question of how new words enter the language, whether it be by way of interlanguage borrowing or native word formation strategies. The types of word formation have been thoroughly documented in countless handbooks of morphology (Plag 2018; Bauer L., R. Lieber and I. Plag 2015; Lipka 1992) and so we may consequently categorize any nascent form with reasonable ease, whereas the problem of classifying lexical obsolescence and loss is addressed in two key sources (Visser 1949; Görlach 1991: 139-143) and little else is available that might help us sort obsolete words into coherent categories.

This thesis set out to map the key factors which are thought to have contributed to lexical obsolescence and loss in Early Modern English whilst proposing a classification framework for this phenomenon using authentic examples from contemporaneous printed media. The first step was corpus-driven research in which I analyzed the frequencies of Early Modern English words as they appear in the individual decades of the Early English Books Online corpus, compiling a preliminary list of 851 words whose frequencies indicated that they likely became obsolete over the course of the Early Modern period. The candidates were then manually sorted and the false positives were assigned provisional categories such as "incorrectly parsed word" or "abbreviation", leaving a total of 48 items which were evaluated as instances of true obsolescence.

Factors thought to have contributed to the obsolescence of each individual candidate were recorded, and it was found that of the 48 words, the most frequent cases of formal obsolescence constituted the disappearance of a real-world object, process or concept (such as *trenshaw)*, obsolescence within the main dialect (such as *meyny*, currently limited Scottish and Irish dialects of English), obsolescence of a word in one word class but not another (the noun *otherwise*, which still survives as an adverb), and synonymy caused either by borrowing a word for the one and the same concept from several languages (the nouns *travailer*, *worker,* and *labourer*) or inconsistent word formation (like the synonymous nouns *debonairity*, *debonairness*, and *debonair*).

Over the course of the research, it became increasingly clear that aside from the *causes* of lexical obsolescence as defined by Visser (1949) and Görlach (1991), it would be worthwhile to view the data from the perspective of *realization,* i.e., how a given form's obsolescence or loss manifests in the language. This might be anything from obsolescence of spelling to the complete obsolescence of form and function. Additionally, the recorded cases of true obsolescence confirmed the view put forth by Fischer (2007: 32) and Smith (2006: 120), among others, that the binary distinction of external versus internal causes is a flawed concept, because it is the interaction of these two types of causes that results in language change and it would be

imprudent to treat them as completely detached from one another. For example, in Early Modern English it was often the case that language contact and the resulting borrowings (Cx) led to an unsustainable synonymy (Ci) which then led to the loss of one or more forms, such as the case of the words *travailer, labourer* and *worker.*

An alternative method of classification is concerned with the form and function of a given word, and how those may be characterized in terms of obsolescence. In relation to form and function, this classification method distinguishes between the effects of *duality/competition* (e.g., synonymy) and *unsustainability* (e.g., disappearance of the original referent). Zooming in on the *unsustainability* component we may further divide it into *inadequacy* (e.g., the phonological representation of the form becomes too weak and can no longer exist as a word carrying meaning), *perception* (e.g., language users associate undesirable social status or connotation with the word) and *modification* (e.g., a structural shift in the language results in the loss of endings).

Moving from theory to practice, a crucial next step in this research will be testing the viability of the above-mentioned frameworks for the classification of lexical obsolescence. As proposed in chapter 6, the classification systems are designed to be as comprehensive as possible, considering the corpus data at our disposal. Refinement following user testing by lexicographers and dictionary users would ensure that we create a truly viable system, since the proposed classification (especially given the interplay of external causes, internal causes and realization of obsolescence) could be cumbersome if used in its entirety.

The data-driven approach used in the initial stages of this research proved essential to the gathering of sufficient examples of obsolescence and the compilation of a comprehensive classification framework. Although a qualitative approach is not without its merits and was, in fact, adopted in the latter part of the research, it was the sheer volume of data from the Early English Books Online corpus and the frequency- based method that allowed for the obsolete forms to be identified with relative ease, providing a solid foundation for the subsequent classification development. The applicability of the proposed system of classification remains to be user tested in further research.

9 Bibliography

Aarts, B., A. McMahon and L. Hinrichs (2021). *The Handbook of English Linguistics*. West Sussex: Wiley Blackwell.

Bailey, R. W. (2006). "English Among the Languages." In Mugglestone, L. (Ed.), *The Oxford History of English,* 334-359. Oxford: Oxford University Press.

Barber, C. (1997). *Early Modern English*. Edinburgh University Press.

Baron, A. and A. Hardie (2012). Prerequisites to a corpus-based analysis of EEBO-TCP Corpus Linguistics with EEBO-TCP. *Conference paper at EEBO-TCP 2012.* University of Oxford, United Kingdom.

Bauer, L. (2002). "Inferring Variation and Change from Public Corpora." In Chambers, J., P. Trudgill, and N. Schilling-Estes (Eds), *The Handbook of Language Variation and Change,* 97-114. Oxford: Blackwell Publishing.

Bauer L., R. Lieber and I. Plag. (2015). *The Oxford Reference Guide to English Morphology.* Oxford: Oxford University Press.

Beal, J. C. (2010). "Prescriptivism and the suppression of variation," In Hickey, R. (Ed.), *Eighteenth-Century English: Ideology and Change*, 1-20. Cambridge: Cambridge University Press.

Blank, P. (2006). "The Babel of Renaissance English." In Mugglestone, L. (Ed.), *The Oxford History of English,* 212-239. Oxford: Oxford University Press.

Bybee, J. (2015). *Language Change*. Cambridge: Cambridge University Press.

Campbell, G. (2010). *Bible: The Story of the King James Version 1611–2011*. Oxford: Oxford University Press.

Chaira, F., and Hamada, H. (2018). "On Lexical Obsolescence in Tacawit: The Case of Six Berber Fauna Terms." *Journal of Translation and Languages*, 17(02), 51-75.

Coleman, J. (2012). *The Life of Slang.* Oxford: Oxford University Press.

Considine, J. (2008). *Dictionaries in Early Modern Europe*. Cambridge: Cambridge University Press.

Corrie, M. (2006). "Middle English—Dialects and Diversity." In Mugglestone, L. (Ed.), *The Oxford History of English,* 86-119. Oxford: Oxford University Press.

Crawforth, H. C. L. (2013). *Etymology and the Invention of English in Early Modern Literature*. Cambridge: Cambridge University Press.

Crystal, D. (2002). *The English Language.* London: Penguin books.

Cvrček, V. (2014). "Quantitative delimitation of the core of a language." *Časopis pro moderní filologii* 96(1), 9-26.

Dawson, H. (2007). *Locke, Language, and Early-Modern Philosophy*. Cambridge: Cambridge University Press.

Degaetano-Ortlieb, S., and E. Teich (2018). "Using relative entropy for detection and analysis of periods of diachronic linguistic change." In *Proceedings of the second joint SIGHUM workshop on computational linguistics for cultural heritage, social sciences, humanities and literature*, 22-33.

Durkin, P. (2009). *The Oxford Guide to Etymology*. Oxford: Oxford University Press.

Dworkin, S. (1995). "Two Studies in Old Spanish Homonymics." *Hispanic Review*, 63(4), 527-542. Doi:10.2307/474740

EEBO – Early English Books Online. Ústav Českého národního korpusu FF UK, Prague 2014. Available at: http://www.korpus.cz

Everitt, A. (1966). "Social Mobility in Early Modern England." *Past & Present*, (33), 56-73. Retrieved from http://www.jstor.org/stable/649802 (Accessed 02/2021)

Fischer, O. (2007). *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: Oxford University Press.

Fisiak, J. (1993). *An Outline History of English. Volume One: External History*. Poznań: Kantor Wydawniczy Saww.

Gelderen, Elly van. (2006). *A History of the English Language.* Amsterdam, Philadelphia: John Benjamins Publishing Company.

Gilman, C. (1979). "Lexical Loss in Cameroon Pidgin." *Anthropos*, 74(1/2), 174-180. Retrieved August 28, 2021, from http://www.jstor.org/stable/40459957

Görlach, M. (1990). *Studies in the History of the English Language*. Heidelberg: Winter.

Görlach, M. (1991). *Introduction to Early Modern English*. Cambridge: Cambridge University Press.

Haugen, E. (1950). "The Analysis of Linguistic Borrowing." *Language*, 26(2), 210–231.

Hickey, R. (2012). "Internally and externally motivated language change." In Hernández-Compoy J. M. and J. C. Conde-Silvestre (Eds), *The Handbook of Historical Sociolinguistics*, 401-421. Malden. MA: Wiley-Blackwell.

Hickey, R. (2010). Attitudes and concerns in eighteenth-century English. In Hickey, R. (Ed.), *Eighteenth-Century English: Ideology and Change,* 1-20. Cambridge: Cambridge University Press.

Hilpert, M and S. T. Gries (2020). Quantitative approaches to diachronic corpus linguistics." In Kytö, M. and P. Pahta (Eds), *The Cambridge Handbook of English Historical Linguistics*, 36-53. Cambridge: Cambridge University Press.

Hogg, R. and D. Denison (Eds) (2006). *A History of the English Language.* Cambridge: Cambridge University Press.

Joby, C. (2015). *The Dutch Language in Britain (1550-1702): A Social History of the Use of Dutch in Early Modern Britain.* Leiden: Brill.

Klein, L. (1993). "'Politeness" as linguistic ideology in late seventeenth- and eighteenth-century England." In D. Stein and I. Tieken-Boon van Ostade (Eds), *Towards a Standard English: 1600 – 1800*, 31-50. Berlin, Boston: De Gruyter Mouton.

Kleinberger, A. (2011). Ottoman lexical obsolescence in the Arabic dialects of the Galilee region. Şarkiyat Mecmuası, (15), 49-88.

Knowles, G. (1997). *A Cultural History of the English Language.* London: Arnold.

Kopaczyk, J. and H. Sauer (Eds) (2017). *Binomials in the History of English*. Cambridge: Cambridge University Press.

Labov, W. (1999). *Principles of linguistic change. Volume 1: Internal factors.* Oxford: Blackwell.

Lamb, J. (2014). *Rules of Use: Language and Instruction in Early Modern England*. London: Bloomsbury.

Lancashire, I. and J. Damianopoulos (2014). "The Early Modern English Law Lexicon." In Mac Aodha, M. (Ed.), *Legal lexicography : a comparative perspective,* 31-43*.* Surrey: Ashgate Publishing.

Lass, R. (2006). "Phonology and morphology." In Hogg, R. M and D. Denison (Eds), *A History of the English Language,* 43-108. Cambridge: Cambridge University Press.

Lass, R. (1993). "Proliferation and option-cutting: The strong verb in the fifteenth to eighteenth centuries" In D. Stein and I. Tieken-Boon van Ostade (Eds), *Towards a Standard English: 1600 – 1800*, 81-113. Berlin, Boston: De Gruyter Mouton.

Lavagnino, J. (2012). *Scholarship in the EEBO-TCP Age.* Retrieved from: http://www.slideshare.net/jlavagnino/scholarship-in-the-eebotcp-age (Accessed 07/2019).

Leith, D. (1996). "English – colonial to postcolonial." In Graddol, D., D. Leith and J. Swann (Eds). *English history, diversity and change,* 180-212*.* London: Routledge.

Linn, A. (2021). "English Grammar Writing." In Aarts, B., A. McMahon and L. Hinrichs (Eds). *The Handbook of English Linguistics,* 63-80. West Sussex: Wiley Blackwell.

Lipka L. (1992). *An Outline of English Lexicology: Lexical Structure Word Semantics and Word-Formation*. Tübingen: Max Niemeyer Verlag.

Locke, J. (1689). *An Essay Concerning Human Understanding.* Edited by Kenneth P. Winkler (1996). Cambridge: Hackett Publishing Company.

Locke, J. (1975). *An Essay Concerning Human Understanding.* Edited by Peter H. Nidditch. Oxford: Oxford University Press.

Lowe, E. J. (2005). *Locke.* London: Routledge.

Martin, S. (2006). Collaboration in electronic scholarly communication: New possibilities for old books. *Journal of the Association for History and Computing*, *9*(2).

Martin, S. (2007). "Digital scholarship and cyberinfrastructure in the humanities: Lessons from the text creation partnership." *Journal of Electronic Publishing*, *10*(1), 1–5. Retrieved from: https://doi.org/10.3998/3336451.0010.105 (Accessed 08/2019).

McMahon, A. (1994). *Understanding Language Change.* Cambridge: Cambridge University Press.

McMahon, A. (2010). "Perspectives on Weakness from English /r/." In Minkova, D. (Ed.), *Phonological Weakness in English: From Old to Present-Day English*, 99-115. London: Palgrave Macmillan.

McInnes, A. (1969). "The Revolution and the People." In Holmes, G. (Ed.), *Britain after the Glorious Revolution, 1689-1714,* 80-95. London: Macmillan.

Miller, D. G. (2012). *External Influences on English. From its Beginnings to the Renaissance.* Oxford: Oxford University Press.

Millward, C. M. and M. Hayes (2012). *A Biography of the English Language.* Boston: Wadsworth.

Mueller, M. (2012). *Towards a book of English: A linguistically annotated corpus of the EEBO-TCP texts*. 1–23. In Bodleian Libraries, University of Oxford, "Revolutionizing Early Modern Studies"? The Early English Books Online Text Creation Partnership in 2012: EEBO-TCP 2012. Retrieved from: http://ora.ox.ac.uk/objects/uuid%3Aa2b1af11-5278-4e5a-8aa0-ee9c16550c44 (Accessed 08/2019).

Mueller, M. (2009). "NUPOS: A part of speech tag set for written English from Chaucer to the present." Northwestern University. *Abgerufen am, 20*, 2013.

Mugglestone, L. (Ed.) (2006). *The Oxford History of English*. Oxford: Oxford University Press.

Nevalainen, T. (2006). *An Introduction to Early Modern English.* Oxford: Oxford University Press.

Nevalainen, T. and I. Tieken-Boon van Ostade (2006) "Standardisation." In Hogg, R. and D. Denison (Eds), *A History of the English Language,* 271-311. Cambridge: Cambridge University Press.

Nevalainen, T. and H. Raumolin-Brunberg (2003). *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. London: Longman.

Nist, J. (1966). *A Structural History of English.* New York: St. Martin's Press.

Oxford University Press. (2014). *OED Online.* Retrieved from http://www.oed.com/

Oxford English Dictionary. (2020). *Period Coverage 1700-1799*. Retrieved from https://oed.hertford.ox.ac.uk/quotations/period-coverage/1700-1799 (Accessed 02/2023)

Oxford University Press. (2021). "Key to Frequency." *OED Online.* Retrieved from https://public.oed.com/how-to-use-the-oed/key-to-frequency/

Petré, P. (2010). "The functions of weorðan and its loss in the past tense in Old and Middle English." *English Language and Linguistics* 14, 457-84.

Plag I. (2018). Word-Formation in English. Cambridge: Cambridge University Press.

Popelíková, J. and V. Volná (2018). *"I feare me that...": Reflexive Verbs in Early Modern English.* Paper presented at ICEHL XX, University of Edinburgh, 27-31 August.

Pulcini, V. (2008). "Lexical obsolescence among Italian Anglicisms." In Pezzini, D., Iamartino, G., Maggioni, M. L., and Facchinetti, R. (Eds), *Thou sittest at another boke... Studies in Honour of Domenico Pezzini.* Milano: Polimetrica, 471-488.

Read, S. (2008). "Thy Righteousness is but a menstrual clout": Sanitary Practices and Prejudice in Early Modern England. *Early Modern Women*, 3, 1–25.

Read, A. W. (1978). The Sources of Ghost Words in English, *Word*, 29:2, 95-104. Retrieved from: https://doi.org/10.1080/00437956.1978.11435650

Rini, J. (1990). Excessive Analogical Change As An Impetus For Lexical Loss: Old Spanish Connusco, Convusco. *Romanische Forschungen,* 102 (1), 58-64. Retrieved August 28, 2021, from http://www.jstor.org/stable/27940047

Rudnicka, K. (2019). *The Statistics of Obsolescence: Purpose Subordinators in Late Modern English.* 10.6094/978-3-928969-75-8.

Rudnicka, K. (2015). "The Statistics of Obsolescence: Case Studies from Late Modern English Grammar." Presentation at Changing English: Integrating cognitive, social and typological perspectives (ChangE 2015), University of Helsinki, Finland, 8–10 June.

Sandler, M. (2003). The Early English Books Online - Text Creation Partnership. *The Charleston Advisor* 4*,* 47-9.

Smith, J. J. (2007). *Sound Change and the History of English*. Oxford: Oxford University Press.

Smith, J. J. (2006). "From Middle to Early Modern English." In Mugglestone, L. (Ed.), *The Oxford History of English,* 120-146. Oxford: Oxford University Press.

Sprat, T. (1667). *History of the Royal Society.* London: J. Martyn and J. Allestry.

Southcombe G. and G. Tapsell (2010). *Restoration politics religion and culture: Britain and Ireland 1660-1714.* Palgrave Macmillan.

Stockwell, R. and D. Minkova (2001). *English Words: History and Structure*. Cambridge: Cambridge University Press.

Tichý, O. (2021). Corpus driven identification of lexical bundle obsolescence in Late Modern English. 10.1075/slcs.218.04tic.

Tichý, O. (2018a). "Lexical obsolescence 1700-2000." In Kopaczyk, J and Tyrkkö, J (Eds), *Patterns in text: Corpus-driven methods and applications.* Studies in Corpus Linguistics, John Benjamins.

Tichý, O. (2018b). "Kvantitativní analýza ortografické variability v Korpusu rané anglické korespondence (A quantitative analysis of spelling variation in the Corpus of Early English Correspondence (CEEC))." *Studies in Applied Linguistics* (Special Issue), 27-41.

Tichý, O. and J. Čermák (2015, March 25-27). *Lexical obsolescence and loss: two methodological probes.* International Symposium on Approaches to English Historical Lexicography and Lexicology, Oxford, UK.

Tieken-Boon van Ostade, I. (2006). "English at the Onset of the Normative Tradition." In Mugglestone, L. (Ed.), *The Oxford History of English,* 240-274. Oxford: Oxford University Press.

Tiersma, P. M. (1999). *Legal Language.* Chicago: University of Chicago Press.

Visser, F. T. (1949). *Some Causes of Verbal Obsolescence: Rede Uitgesprocken Bij de Aanvaarding Van Het Ambt Van Hoogleraar in de Engelse Taalkunde Aan de RK Universiteit Te Nijmegen Op 11 November 1949.* Utrecht: Dekker & van de Vegt.

Welzenbach, R. (2012). *Transcribed by hand, owned by libraries, made for everyone: EEBO-TCP in 2012*. Paper presented at 'Revolutionizing Early Modern Studies?', Oxford, 17-18 September. Retrieved from: https://deepblue.lib.umich.edu/handle/2027.42/94307 (Accessed 10/2018).

Appendix A

Word	1470-9	1480-9	1490-9	1500-9	1510-9	1520-9	1530-9	1540-9	1550-9	1560-9	1570-9	1580-9	1590-9
travailer_n	0	0	0	1	1	2	0	1	1	17	69	29	8
caluine_n	0	0	0	0	0	0	0	0	0	3	0	60	0
Caluine_nn	0	0	0	0	0	0	0	0	4	204	217	413	106
debonayr_j	9	104	13	10	7	0	1	0	1	1	0	0	0
imbroccata_n	0	0	0	0	0	0	0	0	0	0	0	1	126
meyny_n	0	53	5	13	4	44	17	5	39	2	5	3	1
borrow_n	11	27	48	20	16	27	78	67	52	72	202	52	8
otherwise_n	0	0	0	0	0	0	1	0	114	2	8	4	12
ordnance_v	0	0	0	0	0	0	0	0	0	2	1	141	1
rosicleer_n	0	0	0	0	0	0	0	0	0	0	0	491	108
cubyte_n	0	18	22	15	9	15	192	131	14	1	3	4	0
caplm_f-ge	0	0	260	105	0	0	0	0	0	0	0	0	0
unmeet_v	0	3	2	0	1	10	55	60	66	72	33	8	0
pylle_n	0	2	4	2	1	30	11	119	32	19	8	7	2
conduit_v	37	61	24	21	11	11	7	11	27	6	6	0	0
thylk_n	0	217	1	0	0	0	0	0	0	0	0	1	0
commise_v	8	84	71	12	16	6	3	3	2	0	7	0	7
reappose_v	0	0	0	0	0	0	0	0	0	45	165	2	0
thesayd_j	0	0	0	0	0	0	0	56	4	0	0	0	3

Table A-1: Obsolescent forms and their frequencies in the Early English Books Online corpus

(contd.)	1600-9	1610-9	1620-9	1630-9	1640-9	1650-9	1660-9	1670-9	1680-9	1690-9	1700-9	1710-9	Total
travailer_n	0	12	6	18	3	3	0	0	0	0	0	0	**171**
caluine_n	0	1	20	1	0	0	0	0	0	0	0	0	**85**
Caluine_nn	0	269	232	5	0	2	0	0	0	0	0	0	**1452**
debonayr_j	0	0	0	0	1	1	0	0	0	0	0	0	**148**
imbroccata_n	0	1	0	0	0	1	0	0	0	0	0	0	**129**
meyny_n	0	0	1	1	0	0	0	0	0	0	0	0	**193**
borrow_n	0	4	1	0	0	7	0	0	0	0	0	0	**692**
otherwise_n	0	1	10	5	0	1	0	0	0	0	0	0	**158**
ordnance_v	0	7	2	0	0	0	0	0	0	0	0	0	**154**
rosicleer_n	0	0	0	0	0	0	0	0	0	0	0	0	**599**
cubyte_n	0	1	0	0	0	0	0	0	0	0	0	0	**425**
caplm_f-ge	0	0	0	0	0	0	0	0	0	0	0	0	**365**
unmeet_v	0	0	0	0	0	0	0	0	0	0	0	0	**310**
pylle_n	0	7	0	0	1	0	0	0	0	0	0	0	**245**
conduit_v	0	1	0	0	0	0	0	0	0	0	0	0	**223**
thylk_n	0	0	0	0	0	3	0	0	0	0	0	0	**222**
commise_v	0	0	0	0	0	2	0	0	0	0	0	0	**221**
reappose_v	0	2	2	0	0	0	0	0	0	0	0	0	**216**
thesayd_j	0	0	0	0	0	0	0	0	0	0	0	0	**63**

Table A-2: Obsolescent forms and their frequencies in the Early English Books Online corpus (contd.)