

**UNIVERZITA KARLOVA**

**Filozofická fakulta**

Ústav anglického jazyka a didaktiky



Diplomová práce

Bc. Klára Přečková

**Linguistic Markers of Countability in Early and Late Middle English**

Lingvistické ukazatele počitatelnosti v rané a pozdní střední angličtině

Praha, 2025

Mgr. Ondřej Tichý, Ph.D.

## **Prohlášení**

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

Souhlasím se zapůjčením magisterské práce ke studijním účelům.

I declare that the following MA thesis is my own work, for which I used only the secondary literature that is listed in the resources. This thesis was not used as a part of any other university study, nor was it used to gain a different university degree.

I have no objections to the MA thesis being borrowed and used for study purposes.

V Praze dne 14. 1. 2025



Klára Přečková

## **Acknowledgements**

I would like to thank my supervisor Mgr. Ondřej Tichý, Ph.D., for his patient guidance, his unbelievably swift email replies and the incredible amount of time and energy he has dedicated to me throughout this MA thesis. I wish to extend my special thanks to my wonderful family, specifically, my mom and dad, and to my fantastic friends for their remarkable support and life-saving snack breaks during our studies.

## **Abstract**

The main aim of this MA thesis was to explore the linguistic markers of countability in the English language, focusing specifically on their development in the Early and Late Middle English (ME) period (c. 1100-1500). Countability is a key feature influencing noun classification and their grammatical structure. This thesis examines whether and how pluralisation and determiners functioned as countability markers in ME and how these features evolved compared to Present-Day English (PDE). By conducting a detailed analysis, this thesis aimed to trace the development of these linguistic markers across the ME period, providing a clearer understanding of how the concept of countability in English has evolved over time.

The practical part employs a corpus-based analysis using the Penn-Helsinki Parsed Corpus of Middle English, second edition (PPCME2). The analysis combined quantitative and qualitative methods to examine the frequency distribution of plural forms and the usage of determiners.

The findings reveal significant variability in the usage of countability markers during ME. The pluralisation patterns were more flexible, with presumed uncountable nouns forming plural and vice versa. Additionally, the data showed the occasional use of indefinite articles with uncountable nouns like *a water*, a construction now considered ungrammatical in PDE. This tendency suggests that countability rules were not fully standardised in ME yet. These results highlight a transitional phase in the grammaticalisation of countability, reflecting an evolving linguistic system. Future studies could further explore the role of dialectal variation and quantifiers, other essential components of countability distinctions.

**Keywords:** countability, Middle English, number system, definiteness, corpus linguistics, count nouns, mass nouns, diachronic study

## **Abstrakt**

Hlavním cílem této magisterské práce bylo prozkoumat lingvistické ukazatele počitatelnosti v angličtině, konkrétně se zaměřením na jejich vývoj v období střední angličtiny (cca 1100-1500). Počitatelnost představuje klíčový aspekt, který ovlivňuje klasifikaci podstatných jmen a jejich gramatickou strukturu. Tato práce zkoumá, jestli a jak nominální pluralizace a determinátory fungovaly jako ukazatele počitatelnosti ve střední angličtině a jak se tyto prvky vyvíjely ve srovnání se současnou angličtinou. Tato práce měla za cíl za pomoci podrobné analýzy zmapovat vývoj těchto lingvistických ukazatelů napříč daným obdobím a poskytnout tak jasnější představu o tom, jak se koncept počitatelnosti v angličtině historicky vyvíjel.

Praktická část práce je založena na korpusové analýze a jako hlavní zdroj dat využívá Penn-Helsinki Parsed Corpus of Middle English, second edition (PPCME2). Tato práce zkoumá frekvenční distribuce tvarů množného čísla podstatných jmen a užívání determinátorů za pomoci kvantitativních a kvalitativních metod.

Výsledky analýzy nominální pluralizace naznačují, že tvorba množného čísla u počitatelných a nepočitatelných podstatných jmen částečně reflektuje dnešní užití, avšak zahrnuje mnoho příkladů, které se pravidlům současné angličtiny vymykají. Data také poukazují na značnou variabilitu v užívání zkoumaných ukazatelů. Střední angličtina v určitých případech povoluje užití neurčitých členů ve spojení s nepočitatelnými podstatnými jmény, což je konstrukce, která je v současné angličtině považována za negramatickou. Výsledky také naznačují, že pravidla počitatelnosti nebyla ve střední angličtině plně standardizována a poukazují na přechodnou fázi v gramatikalizaci počitatelnosti, která odráží vyvíjející se jazykový systém angličtiny. Budoucí studie by mohly dále zkoumat roli dialektální variace a také kvantifikátorů, které jsou dalším klíčovým prvkem rozlišování počitatelnosti podstatných jmen.

**Klíčová slova:** počitatelnost, střední angličtina, systém čísla, určenost, korpusová lingvistika, počitatelná substantiva, nepočitatelná substantiva, diachronní studie

## List of abbreviations

OE	Old English
ME	Middle English
MED	Middle English Dictionary
PDE	Present-Day English
POS	Parts-of-speech
CQL	Corpus Query Language
C	Countable
U	Uncountable
C/U	Both countable and uncountable meaning
PPCME2	Penn-Helsinki Parsed Corpora of Middle English, second edition

## List of Figures

<i>Figure 1: Number system marking in PDE (Source: Quirk et al.: 1985)</i> .....	13
<i>Figure 2: Data example for U nouns with an indefinite article</i> .....	45
<i>Figure 3: Average plural distribution for Countables, Uncountables, and C/U Nouns</i> .....	46
<i>Figure 4: Comparison of plural forms of Uncountable vs. Substance nouns</i> .....	53
<i>Figure 5: Percentage of plural forms by semantic noun category</i> .....	54
<i>Figure 6: Determinant + 'water' combination visualised in a syntactic tree (ANNIS)</i> .....	56
<i>Figure 7: Percentage of U nouns predetermined by an indefinite article (a/an)</i> .....	57

## List of Tables

<i>Table 1: PDE quantifier classification (Source: Quirk et al. (1985))</i> .....	17
<i>Table 2: Proposed tests for distinguishing countability in PDE</i> .....	23
<i>Table 3: The main syntactic changes (Source: Hogg &amp; Denison., 2006: 111)</i> .....	30
<i>Table 4: Overview of PPCME2 period designations</i> .....	34
<i>Table 5: List of ME word forms (plural), their PDE counterparts and countability category</i> .....	39
<i>Table 6: List of ME word forms (singular), their PDE counterparts and countability category</i> .....	40
<i>Table 7: List of spelling variations for CASTLE and their regional and temporal metadata from PPCME2</i> .....	48
<i>Table 8: List of spelling variations for SPIRIT and their regional and temporal metadata from PPCME2</i> .....	52

## Table of Contents

1. Introduction .....	10
2. Theoretical background .....	11
2.1. Defining countability in the PDE context .....	11
2.1.1. Number .....	12
2.1.2. Definiteness .....	14
2.1.3. Quantifiers .....	16
2.2. Approaches to countability .....	18
2.2.1. Notional account .....	18
2.2.2. Grammatical account .....	19
2.3. The problematic status of countability classification .....	20
2.4. Countability tests .....	22
2.5. Research into countability in languages .....	24
2.5.1. Germanic languages .....	25
2.5.2. Romance languages .....	26
3. Countability in Middle English .....	28
3.1. Shift from Old English .....	28
3.2. Countability markers in Early and Late ME .....	30
3.2.1. Number .....	30
3.2.2. Definiteness and Quantifiers .....	31
4. Method .....	34
4.1. Material .....	34
4.2. Method .....	35
4.3. Analysis .....	36
4.3.1. Pluralisation .....	36
4.3.2. Determiners .....	43
5. Results .....	46
5.1. Pluralisation .....	46
5.1. Determiners .....	56
6. Discussion .....	60

7. Résumé .....	63
References.....	67
Appendix.....	72

## 1. Introduction

The exploration of the countability category constitutes a crucial part of linguistic research, helping us understand how languages categorise and quantify words over time and across different contexts. Many languages employ various degrees and means of this categorisation. The main aim of this MA thesis is to investigate the countability environment of nouns in the development of the English language. Specifically, this text focuses on the linguistic markers of countability in the Middle English period (c. 1100-1500).

The main focus of this work is to explore the category of countability across various levels of Middle English: syntactic, morphological and semantic. To achieve that, the systems of number and definiteness will be examined. In the first, theoretical part, this thesis will review the current approaches to the classification of English nouns into count and mass classes, as well as provide an overview of various means that different world languages employ to express such classes.

The second chapter provides a description of relevant corpora and methods used. This thesis is going to employ a combination of quantitative corpus data analysis and qualitative review, validated by relevant statistical methods. The analysis will focus on the distribution of nominal pluralisation and determiners as signals of the noun countability class. The core will be the second edition of *The Penn-Helsinki Parsed Corpora of Middle English*. Supplementary sources, such as *The Middle English Dictionary* will be incorporated into the analysis as well.

The following part will provide the necessary theoretical background, explain basic terminology and concepts employed in this thesis, and provide information about previous noun countability findings. The research part will describe the material and how we approached the corpus search. Lastly, we will conduct descriptive and statistical data analysis and draw corresponding conclusions from the results.

## 2. Theoretical background

### 2.1. Defining countability in the PDE context

The elementary teaching rule tells us that if an English noun can have a singular or plural form, then it can be classified as countable (C), such as *boat(s)* or *cat(s)*. On the other hand, nouns failing to fulfil this criterion are classified as uncountable (U), e.g. *air* or *peace* (Eastwood, 2005; Quirk et al., 1985; Greenbaum, 1996). In English and other languages, nouns may vary with respect to the property of countability. Traditionally, English nouns tend to be categorised into two main groups based on their countability class: countable (also called 'count') and uncountable (also 'mass' or 'non-count')<sup>1</sup>. Huddleston and Pullum (2002: 334) provide a simple definition: "Count nouns denote entities that can be counted, while non-count nouns denote entities that cannot be counted." While this definition is a helpful onset, the concept of what qualifies as countable is far more complex than it might initially appear, as this work will explore in greater detail. In addition, even though English nouns are frequently categorized within this binary C/U distinction, it is more effective to view countability as a scalar property, where nouns exhibit varying degrees of countability preference (Allan, 1980; Wierzbicka, 1988; Gil, 2013). Many nouns can exhibit both C and U characteristics depending on the semantic context and shift their meaning from one to another. For example, *hair* is uncountable in general reference - *I have long hair*. But in specific reference, it becomes countable, e.g. *I have three hairs in my coffee*.

The first person to make a distinction between countable and uncountable was Otto Jespersen (1924); according to him, if a noun can be associated with numerical quantities such as *one* or *two*, it falls into the category of "countables". These nouns correspond to the syntactic representation of singular, dual, or plural forms. Uncountable nouns he defines as follows:

There is a class of 'things' to which words like one, two are inapplicable; we may call them uncountables, though dictionaries do not recognise this use of the word *uncountable*, which is known to them only in the relative sense 'too numerous to be (easily) counted' (1924: 188).

---

<sup>1</sup> Throughout this thesis, the terms countable/count and uncountable/mass will be used interchangeably, with no difference in their meaning.

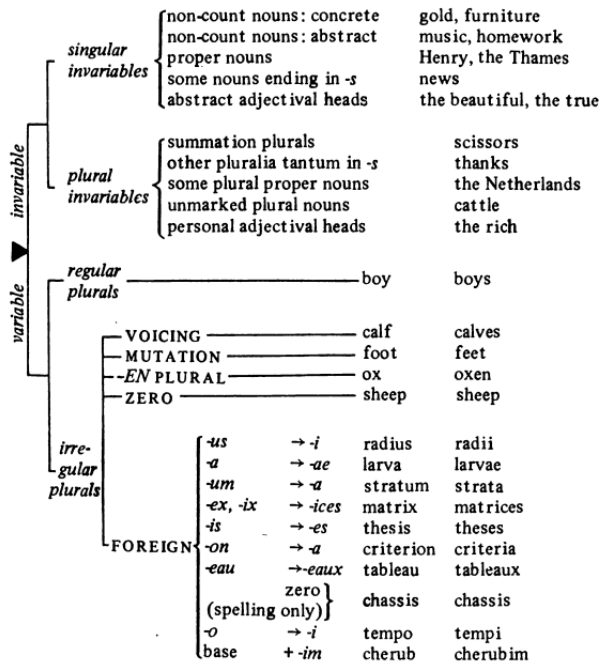
He went on to characterise uncountable nouns, in opposition to thing-words representing countable objects, as “words which do not call up the idea of any definite thing, having a certain shape or precise limits” (Jespersen, 1913: 114). Even though Jespersen further noted syntactic differences between the groups as well, the “reference to countable or uncountable objects seems to have been the defining distinction” (Lasersohn, 2011: 1133). Uncountable nouns have since received more definitions, such as “things that do not naturally divide into separate units” (Eastwood, 2005: 177), but their classification still contains numerous exceptions. Rather than adhering to a strict rule, these definitions often rely on negative definitions of what uncountable nouns are not. What are the fundamental markers of noun countability, then? The general scholarship agrees on the classes of number, definiteness, and a relationship to quantifiers. These categories will be discussed in the following sections.

### **2.1.1. Number**

For nouns, the category of number is tightly intertwined with the concept of countability (Huddleston and Pullum, 2002: 334). In English, the number system is divided into two main noun forms: singular and plural. Singular nouns include both non-count nouns and proper names, but this work will omit proper names in its analysis due to its scope. Countable nouns may either alternate between singular and plural forms, e.g., *boy* and *boys* or stay invariant in the plural form, e.g., *cattle* (Quirk et al., 1985). High countability nouns “denote objects which are conceptualized in terms of highly individuated units” (Gil, 2013). These would be, for example, *a cat* or *a house*. On the other hand, low countability words describe masses that, unlike the previous group, lack a unitary and characteristic shape (Gil, 2013).

Countability has several grammatical implications, particularly regarding how nouns interact with numerals. Nouns that are highly countable can appear directly with numerals, while those that are less countable usually cannot. Instead, low-count nouns require the inclusion of an additional phrase, such as “one glass of water” or “two pounds of sand.” These additional phrases are often called mensural numeral classifiers, as they offer a unit of measurement that allows for counting less countable nouns (Gil, 2013).

Quirk et al. (1985) provide a table (**Figure 1**) summarising possible countability variants distinguishing between invariable and variable forms in terms of countability and plurality:



*Figure 1: Number system marking in PDE (Source: Quirk et al.: 1985)*

Invariable singular nouns encompass non-count nouns, both concrete and abstract (e.g., "gold," "music"), along with certain proper nouns and nouns ending in -s ("news") that resist pluralisation. Invariable plural nouns feature categories where plurality is semantically inherent yet unmarked in typical singular-plural morphology. These include summation plurals (e.g., "scissors"), pluralia tantum nouns (e.g., "thanks"), pluralised proper nouns (e.g., "the Netherlands"), unmarked plural nouns with collective meaning (e.g., "cattle"), and adjectival heads denoting groups or classes of people (e.g., "the rich").

In the variable nouns section, the table outlines regular plural formation (e.g., "boy" to "boys") and irregular morphological processes. These processes include voicing alternation (e.g., "calf" to "calves"), vowel mutation (e.g., "foot" to "feet"), -en suffixation (e.g., "ox" to "oxen"), and zero inflection (no morphological change, as with "sheep"), demonstrating historical morphological patterns.

Lastly, foreign plurals highlight the preservation of Latin, Greek, and French inflectional patterns, where morphological adjustments retain etymological plural markers. The table also acknowledges instances of spelling-only invariance, where pluralisation relies on context rather than

morphological alteration, thus underscoring English's tendency to incorporate foreign inflectional rules alongside native ones.

In sum, the number system in English differentiates singular and plural noun forms and, at the same time, affects their countability behaviour. This connection between countable and uncountable nouns highlights the structural flexibility of English, accommodating regular and irregular pluralization as well as invariable forms that express collective or abstract concepts. Compared to its historical predecessors, which will be discussed in the later sections, PDE has a relatively simple noun number system. Grammatical distinctions between singular and plural nouns are primarily indicated through distinct word forms and the use of determiners. The following section will explore how determiners function in PDE with both countable and uncountable nouns.

### **2.1.2. Definiteness**

In linguistics, definiteness as a category is associated with the use of determiners, mainly the use of the definite article *the* in languages that employ them such as English, Hungarian or Hebrew (Giancollo et al., 2022). “In many languages, a noun phrase may contain an element which seems to have as its sole or principal role to indicate the definiteness or indefiniteness of the noun phrase. This element may be a lexical item like the definite and indefinite articles of English (*the*, *a*), or as an affix of some kind like the Arabic definite prefix *al-* and indefinite suffix *-n*” (Lyons, 1999: 1).

Definite determiners, such as *the*, mark the referent as identifiable to both the speaker and listener, often assuming prior mention or knowledge (Greenbaum, 1996: 173). On the other hand, indefinite determiners, such as *a* and *an*, serve to introduce a new or non-specific referent and are typically used with singular count nouns (Biber et al., 1999: 273). This is particularly significant in noun countability, as articles like *a* cannot be used with non-count nouns (*a water* is ungrammatical) but are freely used with singular count nouns (*a book*, *a car*) (Huddleston & Pullum, 2002, p. 371-72). Huddleston and Pullum (2002: 356) provide a list of the main PDE determiners, where they classify certain quantifiers, such as *many*, *much*, *few*, and *little*, as degree determiners, offering a perspective that overlaps with other classification systems:

i.	<i>the, a</i>	articles
ii.	<i>this, that</i>	demonstrative determinatives
iii.	<i>we, you</i>	personal determinatives
iv.	<i>all, both</i>	universal determinatives
v.	<i>each, every</i>	distributive determinatives
vi.	<i>some, any</i>	existential determinatives
vii.	<i>one, two, three, ...</i>	cardinal numerals
viii.	<i>either, neither</i>	disjunctive determinatives
ix.	<i>no</i>	negative determinative
x.	<i>another</i>	alternative-additive determinative
xi.	<i>a few, a little, several, ...</i>	positive paucal determinatives
xii.	<i>many, much, few, little, ...</i>	degree determinatives
xiii.	<i>enough, sufficient</i>	sufficiency determinatives
xiv.	<i>which, what, whichever, whatever</i>	interrogative and relative determinatives

Quantifiers and determiners both serve to limit or specify the reference of a noun within a noun phrase. Quantifiers like many or few restrict the amount or quantity, which aligns with the role of determiners in narrowing down the scope of a noun's reference. This shared function likely leads to their cross-classification, as both types of words fulfil a similar purpose in guiding the listener's interpretation of the noun.

Additionally, they elaborate on the ways determiners interact with noun countability, noting that some determiners are specifically restricted to count or non-count nouns. For example, *many* and *few* are used with plural count nouns (e.g., *He made many mistakes*), while *much* and *little* are appropriate with non-count nouns (e.g., *Has he got much money?*) (2002: 393). These distinctions align with semantic countability but do not always strictly follow numerical categorization, especially in the case of words that can serve as both count and non-count nouns depending on context (e.g., chicken as an animal vs. chicken as food) (Quirk et al., 1985: 287).

Using uncountable nouns in a countable expression can be achieved through the so-called partitive expressions as well. These expressions specify the quantity or amount of substances that do not carry traditional countability features such as plural and thus need these expressions to express their measurement. These include *a piece of, a bit of, some pieces of, two pieces of, a lump of, a bar of...* (Greenbaum, 1996: 98-99). The following section will focus on another type of countability expression within a noun phrase that helps distinguish its countability – quantifiers.

### 2.1.3. Quantifiers

As well as the number system, in PDE, quantifiers function as indicators of noun countability (Quirk et al., 1985; Greenbaum, 1996). Conventionally, a quantifier is a “word or phrase that is used before a noun to show the amount of it that is being considered,” (“Quantifier,” in *Cambridge Learner’s Dictionary*, n.d.). Additionally, quantification can be thus defined as a “category of determiner or pronoun used to express quantity” (McArthur, 1992, p. 747). Quantifiers typically occur in the prenominal position within a noun phrase (NP) but may also function as a pronoun, e.g. *Many have passed*. (Radden, G. and R. D., 2007: 116). They refer to either countable, uncountable or both noun categories. This type of quantity expression is not restricted to languages with obligatory plural marking in the context of numerals (Filip, 2021: 62). Multiple lexical items exhibit quantifier behaviour. Quirk et al. (1985) include as members even adverbs and pronouns, depending on their syntactic and semantic behaviour. Interestingly, among their description of English grammar, a category specifically for quantification is not delineated as a grammatical class. This follows Jespersen’s (1933) categorisation of quantifiers as indefinite pronouns, such as *some* being a pronoun of unspecified quantity and *any* being one of indifference (Vandiver, 2011). Greenbaum (1996: 192) classifies common quantifiers as indefinite determiners and pronouns, dividing them up into sets determined by their quantification of either count or non-count nouns:

1. *many, more, most, a few, fewer, fewest, several, enough* - countable
2. *much, more, most, a little, less, least, enough* – uncountable
3. *few* (count.), *little* (uncount.) – negative (2002: 194)

The categorisation of English quantifiers tends to differ in the number of concrete groups. Sommerer and Hofmann (2021) propose two main categories: universal and existential. Universal quantification, such as *all*, refers to the entirety of a group e.g., “all students”. On the other hand, existential quantification, usually indicated by *some*, signifies a quantity exceeding zero.

Dušková et al. (2009) provide an additional distinction of negative quantifiers e.g. *no, none, neither* + composites, and further include words such as *much, many, (a) few, (a) little, several, enough, one, other* that remain not classified. While Quirk et al. (1985) offer a highly detailed classification of English quantifiers (**Table 1**), it could be argued that his “work has the explicit deficit of concentrating on patterns of syntactic behaviour and aligning the majority of quantifier

descriptions with surface structure phenomena rather than focusing its efforts on semantic behaviour” (Vandiver, 2011).

	NUMBER	FUNCTION	COUNT		NONCOUNT
			personal	nonpersonal	
UNIVERSAL	singular	pronoun	<i>everyone</i> <i>everybody</i>	<i>everything</i>	<i>(it (...)) all</i>
			<i>each</i>		
	plural	determiner	<i>every</i> <i>each</i>		<i>all</i>
			<i>all/both</i>		
ASSERTIVE	singular	pronoun	<i>someone</i> <i>somebody</i>	<i>something</i>	<i>some</i>
		determiner	<i>a(n)</i>		
	plural	pronoun and determiner	<i>some</i>		
NONASSERTIVE	singular	pronoun	<i>anyone</i> <i>anybody</i>	<i>anything</i>	<i>any</i>
		determiner	<i>either</i> <i>any</i>		
	plural	pronoun and determiner	<i>any</i>		
NEGATIVE	singular	pronoun	<i>no one</i> <i>nobody</i>	<i>nothing</i>	<i>none</i>
			<i>none</i>		
	plural	pronoun and determiner	<i>neither</i>		
		pronoun	<i>none</i>		
singular or plural	determiner	<i>no</i>			

*Table 1: PDE quantifier classification (Source: Quirk et al. (1985))*

Both classifications emphasise the main categories of quantifiers: universal, existential, and negative. In each system, universal quantifiers indicate an inclusive scope, covering both count and noncount nouns. Assertive or existential quantifiers are used to indicate the existence of something within a set, aligning closely between the two systems. Likewise, negative quantifiers, appear in both classifications to express the absence of quantity or existence.

Regarding the definitions of individual quantifier words, it is difficult to provide unified information as the semantic aspect of countability expressions is often up for interpretation, even in PDE. For example, according to Jespersen (1949: 49), the quantifier expresses an unspecified quantity, amount or number of persons and things. Biber et al. (1999: 275–8) are more specific, stating that *some* usually specifies a moderate quantity.

One of the questions of this work is whether their function was the same in ME. Bloomfield and Eilbott (1957: 171) commented on the importance of future research on such words already in the 1960s: "Until the exact history of certain indicator words like *any, more, mo, enough, enow, some, such, much, fewer, less, etc.*, is determined, it is difficult to study the development of the concept in the past." They offer examples from the OED and Jespersen, illustrating the classification of words like *mo* and *enow*, historically associated with countable nouns and *more* and *enough* connected with mass nouns. Nonetheless, they also add that this generalisation may not be valid for the ME period and should be examined further.

## **2.2. Approaches to countability**

### **2.2.1. Notional account**

“A major theme of countability research is the relation between contrastive grammatical countability classes and corresponding contrasts, or lack thereof, of notional, or ontological, types of the corresponding referents of the grammatical classes,” (Grimm & Wahlang, 2021: 66). The distinction between count and mass nouns is a linguistic domain closely tied to cognitive concepts like individuation and numerical representation. Several authors have previously explored how languages differentiate between discrete entities, which can be counted, and non-discrete substances, which are measured by mass or volume.

The notional account of noun countability is a theoretical perspective that explains the distinction between count and mass nouns based on the inherent properties of the objects or substances they denote (Joosten, 2003: 219). Some previous authors examining this topic include Parsons (1970), Allan (1980), Link (1983), Bunt (1985), Wierzbicka (1985), Koslicki (1999), and Wisniewski (2010). According to this view, count nouns refer to items that are naturally discrete and individually distinguishable, such as 'apple,' 'book,' or 'car.' These entities can be counted because they exist as separate, distinct units. In contrast, mass nouns refer to substances or entities not

inherently divided into discrete units, such as 'water,' 'sand,' or 'information.' These are measured rather than counted because they are perceived as continuous masses or aggregates. The notional account emphasises that the cognitive basis of countability lies in how we mentally represent and categorise different kinds of objects and substances in the world.

Dušková et al. (2009) comment on countability, stating that the category of countability in language reflects distinctions in extra-linguistic reality between discrete, separable units and continuous entities. Discrete units, with boundaries and separable members, are countable, while continuous entities, lacking boundaries, are uncountable. This distinction influences the grammatical behaviour of linguistic units. However, the degree of grammaticalisation of countability varies across languages, as does the conception of countable and uncountable phenomena in different linguistic contexts. They mention three criteria for distinguishing noun countability: number, quantification, and determination. As these criteria are considered to be decisive for noun countability across various literature, they will be discussed in more detail in their respective sections on Middle English. Additionally, they will be crucial elements for the practical part of the **Analysis** section.

While notional account relies on the inherent cognitive individualisation of numbers, many search for countability cues in the language itself. This grammatical approach will be discussed in the following paragraph.

### **2.2.2. Grammatical account**

In contrast to the notional perspective, the grammatical approach to countability has been examined by scholars such as Bloomfield (1933), Quine (1960), Wierzbicka (1985), and Chierchia (1998). While Jespersen's binary distinction of count and mass nouns was based on a notional account, similarly to Wierzbicka (1985) later on, Bloomfield's distinction followed a grammatical account (Joosten, 2003: 218–219). Bloomfield (1933) proposed formal criteria for count-mass distinction, stating that "Mass nouns never take *a* and have no plural." Additionally, Bloomfield (1933: 205) identifies two key characteristics of count and mass nouns, which he defines as distinguishing properties. Count nouns refer to items that cannot be subdivided or merged without losing their identity (e.g., a pen broken into pieces is no longer a "pen" in function, just pieces of

plastic). Mass nouns, however, can be divided or combined without altering their essence; for instance, splitting or merging portions of water still results in the same substance - water.

The notional account distinguishes the count-mass contrast on a semantic, conceptual level, unlike the grammatical approach, which presumes no connection with extra-linguistic factors (Grimm, 2013). "It encompasses the two dimensions typically associated with grammar: morphology and syntax, and stresses the fact that these dimensions are not related to semantics" (Drożdż, 2017: 16). This view is supported by Allan (1980); he argues against the notion that countability is intrinsic to the noun itself but instead views it as a result of its environment, stating that "countability is a subcategory of the NP, not of nouns, and that nouns have computable countability preferences" (566).

Grammatical countability is driven by the semantic differentiation between object and substance reference, often referred to as the bounded/non-bounded or individuated/non-individuated distinction. Among linguists, there is an ongoing debate about the extent to which grammatical countability is rooted in semantic principles versus how much is arbitrary (Baldwin, 2003 from Wierzbicka, 1988). Baldwin adds that "the prevailing position in the natural language processing community is effectively to treat countability as though it were arbitrary and encode it as a lexical property of nouns."

Gleason (1955: 145) argues that the grammatical view implies that language structures, like grammar rules, often shape how we label and organize concepts, regardless of any inherent qualities those concepts may have. He provides the examples of *rice* (a mass noun) and *beans* (a count noun), despite both consisting of many small particles. Palmer (1983: 34–35, In Drożdż, 2017: 16) supports this view, stating that such distinctions are grammatical rather than tied to inherent differences in meaning.

This thesis will consider both methodologies. The practical section will work on the basis of grammatical differentiation, offering a structure for our examination of language corpora, followed by a semantic, notional analysis of specific cases.

### **2.3. The problematic status of countability classification**

As is the nature of linguistic research, both approaches have their strengths and limitations. Vermote et al. (2016) comment on the problematic aspects of the grammatical approach by noting

its failure to explain the widespread tendency among nouns to prefer either count or mass syntax. For instance, nouns like "car" (count) and "sand" (mass) exhibit a preference that is both intuitively recognisable and shared across individuals. They mention the study of Kulkarni et al. (2013), who conducted a rating task across six languages and implied that this phenomenon suggests a language-specific lexical specification.

While the grammatical approach struggles to explain the inherent preferences of nouns for count or mass syntax, Kiss et al. (2016) in the Bochum English Countability Lexicon (BECL) address three problematic aspects in assessing the count/mass distinction: "First, that the distinction can be applied to the lemma or lexeme. Secondly, the popular view that the distinction is a binary one, and thirdly, that it can be analysed by looking at a small sample of staple nouns only" (2811).

Similarly to Allan (1980), or Quirk et al. (1985), Chierchia does not work with a strict binary distinction between count and mass nouns. "While the mass/count distinction is not altogether indifferent to how things are inherently structured, it appears to be independent of it, which is what makes such a distinction a strictly grammatical one," Chierchia (1998, 57). He presumes that both categories may often have count and mass reading, allowing for a dual class membership. Dual class membership permits a noun to be acceptable in both countable and uncountable forms, with a notable difference in meaning between them. These nouns with dual-class membership can be differentiated by a difference in meaning or by adequate usage of quantifiers, e.g.

1. She's had many difficulties. – countable
2. She's not had much difficulty. – uncountable

Existing research typically categorises these nouns based solely on their form, without clarifying whether a noun inherently possesses both count and mass properties or if its categorisation depends on the specific meaning it conveys in a given context (Kiss et al. 2016: 2811). Additionally, a noun can shift to the opposite countability category, undergoing reclassification. An uncountable noun, *cheese*, may be reclassified as a countable noun with a semantic shift denoting the kind/type/form of the noun, e.g. *What cheese have you got today? We have Cheddar, Gorgonzola, and Danish Blue* (Quirk et al. 1985: 247-48).

While the notion of countability does seem flexible, the majority of English nouns tend to exhibit one default preference. Multiple efforts have been made to define and classify the concept of

countability further. The results of previous linguistic research will be summarised in the section below.

#### 2.4. Countability tests

Many methods and standards have been proposed to define more precise boundaries within the flexible concept of countability. To demonstrate the difference between count and non-count, Greenbaum and Quirk (1990: 71) propose a countability test with five determiner constraints to distinguish whether a noun is *count* or *non-count*: "Can the singular noun occur (a) without a determiner? (b) with the definite article? (c) with the indefinite article? (d) with the partitive *some*, /səm/? Can the plural noun occur (e) without a determiner?"

Similarly, Chierchia (1998), in his work on the universality of countability, discusses five properties important for this distinction:

- a) pluralisation and counting
- b) measure phrases
- c) so-called pseudopartitive constructions with phrases like *quantity of*, *amount of*
- d) so-called fake mass nouns (like *furniture*, *footwear*) and (d) ambiguous nouns (like *beer*, *rope* or *chicken*)
- e) coercion (like *there is apple in the salad*) (3)

Allan (1980) writes that countable items are typically viewed as consisting of one or more distinct entities, while uncountable items are often perceived as a unified whole, though not always. He works with the propositions that: "If the head constituent of the NP falls within the scope of a denominator, it is countable" (541) and "If the head constituent of an NP is plural, it is countable" (542). To divide nouns into respective groups with similar countability properties, he proposed five tests for distinction:

- (1) the EX-PL Test (external plural number registration)
- (2) the A + N Test (a unit denominator, such as *a(n)*, *one* + noun)
- (3) the All + N Test (*all* + noun)
- (4) the F + Ns Test (a fuzzy plural denominator, such as *several*, *about fifty* + noun)
- (5) O-DEN (all other denominators) (562)

**Table 2: Proposed tests for distinguishing countability in PDE**

Countability Tests			
Author	Method	Attribute	Example
Greenbaum & Quirk (1990)	Determiner Constraints	Singular without determiner	<i>Water is essential for life.</i>
		Definite article (the)	<i>The water in this bottle is cold.</i>
		Indefinite article (a/an),	<i>A glass of water quenched my thirst.</i>
		Partitive (some),	<i>Would you like some water?</i>
		Plural without determiner	<i>Trees provide oxygen.</i>
Chierchia (1998)	Properties	Pluralisation/Counting	<i>Three trees were planted in the park.</i>
		Measure Phrases	<i>A gallon of milk is on the counter.</i>
		Pseudopartitive Constructions	<i>A piece of cake is left.</i>
		Fake Mass Nouns & Ambiguous Nouns	<i>Linguistics is a fascinating subject. / Chicken can be bought at the market.</i>
		Coercion	<i>We ordered two coffees.</i>
Allan (1980)	Denominator Tests	EX-PL Test (External Plural)	<i>Those people are waiting outside.</i>
		A + N Test (a/an + Noun)	<i>A cat sat on the windowsill.</i>
		All + N Test (All + Noun)	<i>All children need education.</i>
		F + Ns Test (Several, Fifty + Noun)	<i>Several books were on the table.</i>
		O-DEN (Other Denominators)	<i>A dozen eggs are in the fridge.</i>

Although the literature begins with somewhat similar discussions on how to decide countability classes, there is considerable diversity in the conclusions drawn from them. The countability classes proposed by different scholars vary widely: Bloomfield proposed a binary distinction, Allan suggested eight classes, Wierzbicka (1988) proposed fourteen, Huddleston (1984) six, and

Grimm and Wahlang (2021) proposed fifteen classes. It is apparent that there is a lack of consensus on how many classes we should distinguish.

## **2.5. Research into countability in languages**

This section will look at the previous research on countability systems in other languages, which may provide further insight and possible comparison with the development and findings for English. The main source of influence for Middle English remained both Germanic, with Old Norse; however, borrowings from Latin and Old French played a vital role in the development of the language as well: “Direct contact with Norse was, with a few important exceptions, a geographically specific phenomenon, but French influence could appear in any dialect,” (Corrie, 2006: 97). This text will thus firstly briefly look into world languages, followed by sections on Germanic and Romance languages in more detail.

As mentioned above, in English and numerous other languages, nouns exhibit formal variation to indicate singular or plural number. The addition of an inflectional plural suffix such as *-s*, exemplified by 'dog' changing to 'dogs', the use of an indefinite article to differentiate single entity and the possibility of countability modification with numerals. This characteristic is also observed, for example, in Dutch. These languages are classified as mass-count languages (Lin et al., 2018: 1). Koderá (2011: 42) mentions that certain languages, like Turkish, frequently do not use morphosyntactic markers to indicate plurality, even though they have distinct singular and plural noun forms. Additionally, he mentions that some languages only mark plurality for specific noun categories, such as animate or human nouns. For example, languages like Chinese and Japanese lack grammatical inflections to distinguish between singular and plural forms and maintain the same form regardless of whether the noun denotes a singular entity or multiple entities. They are members of the so-called obligatory classifier languages because numeral expressions have to appear with classifiers when modifying nouns. As a result, these languages have hundreds of classifiers. Filip (2021: 53) further explains this distinction, "In numeral classifier languages, the syntax of nouns with typical count meanings usually resembles the syntax for mass nouns, and numeral classifiers resemble measure words in a language such as English."

Contini-Morava and Kilarski (2013) comment on the difference in the status of nouns between classifier and non-classifier languages, saying:

"The idea is that nouns in classifier languages are like mass nouns in a non-classifier language in that they also need to be unitised in order to be counted. The function of unitising or individuating is usually ascribed to the accompanying classifier, or to the pseudopartitive construction that includes one." (277)

In this regard, it is interesting to cross-linguistically analyse countability systems as they may provide valuable clues to the development of English itself. The subsequent section will focus on Germanic languages, exploring the influence of Dutch and other Germanic sources on English countability and grammatical structures.

### **2.5.1. Germanic languages**

To deepen our understanding of countability systems across languages, it is essential to explore parallels between Germanic and Romance languages, as Bloomfield and Eilbott (1957) emphasize. While countability distinctions manifest differently across linguistic families, Germanic languages like Dutch, Afrikaans, and Standard German exhibit unique mechanisms for distinguishing between countable and uncountable nouns.

In English, adding an indefinite article or plural marking to nouns typically treated as count nouns allows for both kind and unit interpretations, reflecting ambiguity in these constructions. However, in Dutch, this ambiguity is limited when such markers are added to traditionally mass nouns, where only a kind reading is possible (De Belder, 2011: 177). Similarly, in Afrikaans—a language derived from 17th-century Dutch—the lack of number marking in NPs typically signals mass readings. When number marking is applied to a noun usually viewed as a mass noun, it generates a countable, kind reading, indicating a flexible but systematic use of number marking in distinguishing noun interpretations (De Belder, 2011: 183). De Belder also extends this analysis to Standard German, where, as in Dutch and Afrikaans, number marking serves to derive kind readings from mass readings (2011: 185).

As this paragraph has shown, in the Germanic languages mentioned above, the addition of number markers or plural forms systematically shifts the noun's interpretation. How the countability system functions in the Romance languages will be discussed below.

### 2.5.2. Romance languages

The development of countability in Romance languages, particularly French and Latin, had a notable impact on Middle English, shaped largely by the geopolitical changes of the 11th century (See 3.1). As French and Latin integrated into the English lexicon, they introduced grammatical distinctions and nuances in noun countability that would influence English usage.

For example, Vermote et al. (2017, 42) conducted a study analysing the countability in French and Dutch, specifically focusing on nouns for fruit and vegetables. In Dutch, these nouns are more often used in a general mass sense, like "water" (mass noun). In contrast, French conveys the same collective meaning by using the singular definite article, effectively blurring the line between mass and count nouns e.g. *L'or est un metal précieux*. Their study confirms that information about mass and count usage for the analysed languages is stored in the mental lexicon, as evidenced by correlations between their frequency in discourse and acceptability in rating tasks. More frequent and attested mass usages were found to be more acceptable, with acceptability also influenced by semantic-pragmatic context and the identity of the lexical items and speakers. Overall, both frequency and context play key roles in how mass noun usage is perceived (2017,47-18).

Other Romance dialects also reveal unique mechanisms for distinguishing between countable and uncountable nouns, sometimes even through gender. Contini-Morava and Kilarski (2016: 267, in Haase, 2000: 226) provide an interesting example from Central Italian dialects, where masculine gender is used for countable nouns and neuter for uncountable forms. For instance, *lu pane* (masculine) means "the piece/loaf of bread," while *lo pane* (neuter) refers to "the bread" in a general sense, such as "the bread on the table." Similarly, *lu pesce* denotes a live fish (masculine), whereas *lo pesce* signifies "the fish" intended for eating (neuter).

Stark (2008) researched how noun countability in relation to specificity was shaped throughout the development of Old Italian. Specificity marked the distinction between referents intended as particular entities and those viewed as any member of a broader category, with specific referents more likely to be explicitly marked by determiners. Previous literature research introduced different notions of specificity - scopal, epistemic, partitive, and relative - that described whether a noun phrase pointed to a known or undefined entity in a text's context (133). Her results show that specificity indicators, such as certain indefinite determiners (e.g., *certo* and *uno*), often

accompanied noun phrases with perfective verb forms or subject positions, thus marking their referents as clearly specific, while other markers (e.g., *alcuno* and *qualche*) appeared in non-specific contexts (136–137). Stark discusses the grammaticalization process of Italian determiners where they shifted from specificity markers to countability classifiers. Initially, determiners classified only specific noun phrases, but they were increasingly used with non-specific referents, indicating a shift towards a broader role in organizing noun phrases (144).

The development of the countability system and countability markers may draw useful parallels to the evolution of countability in the history of English, particularly the shift from a highly inflectional system to an analytical one.

“The loss of this rich inflectional ‘classification system’ in the Romance languages might then have provoked the appearance of new ‘classification devices’ and the evolution of a new Romance classification system by indefinite determiners” (Stark, 2008: 135). As Stark suggests for Italian, the gradual loss of complex inflectional systems may have required clearer syntactic cues for noun classification. Just as Old English, with its case and gender markers, began to simplify its structure, Middle English compensated with new syntactic strategies, increasingly relying on determiners to clarify noun referents and countability. Similarly, Old Italian, influenced by the loss of Latin’s rich inflectional system, began to mark specificity with determiners like *uno* and *certo*, which paralleled the English indefinite article *a/an*—itself derived from the Old English *ān* (“one”).

Initially, both languages used these markers specifically to denote countable entities with distinct referents, but as grammaticalization progressed, the markers expanded to less specific contexts (133-137). Furthermore, both languages developed systems to differentiate between mass and count nouns, with Old Italian introducing the partitive article for diffuse substances (*del vino* for “some wine”) while English adopted expressions like *some* or partitive constructions (e.g., “a piece of bread”). In essence, the shift in Romance languages from relying on implicit, context-based cues for countability to using explicit, syntactically grounded determiners reflects a broader trend of grammaticalization. This change redefined how specificity and countability are expressed within their noun phrase structures.

### **3. Countability in Middle English**

#### **3.1. Shift from Old English**

The Norman Conquest of 1066 initiated a gradual yet profound transformation in the English language. After the conquest, French largely displaced English as the language of the government and of administration, while Latin remained the main language of the church, pushing English into the peripheries. This both physical and linguistic invasion of French-speaking Normans led to a significant shift as Old English, the language of the Anglo-Saxons, fell out of use in written contexts (Nevalainen, T., & van Ostade, I. T.-B., 2006: 271). The dominance of French and Latin resulted in more writing in these languages, with English writing becoming rare. The surviving work from the 13th century combines Old English tradition, French and Latin elements, and a considerable amount of invention (Hogg, R. & Denison, D., 2006: 59).

As has been mentioned above, the language environment at the end of twelfth-century England was rather diverse and complex. In regard to the linguistic situation, most people, particularly from lower social ranks, spoke English in local dialects without a standardised form. Even though Norman French became the language of the ruling class in England, “approximately four million people, according to some estimates, most densely clustered in the southern half of England - continued to speak English, and written in OE,” (Horobin, 2002: 27). For much of the Middle English period (circa 1100 to the late 1400s), communication in English was essentially local, with first French and then Latin used for government and law. As a result, there was no shared national Middle English dialect, but rather great regional diversity in both speech and writing (Corrie, 2006).

In the former Danelaw region, Scandinavian-English bilingualism persisted into the early twelfth century but eventually declined, resulting in the extinction of spoken Scandinavian and the incorporation of Scandinavian vocabulary into English. Meanwhile, the nobility and royal family primarily spoke Anglo-Norman French. However, by the late twelfth century, many upper-class individuals began learning English, leading to bilingualism, especially in towns where the middle class also adopted French. Latin remained prominent as the language of the church, scholarship, and official records, maintaining its status as a spoken and written medium in religious and scholarly settings. (Townend, 2006: 66-68).

If we focus on the expansion of English vocabulary, the data (Durkin, 2014) reveals a significant increase in the number of French loanwords during the 1150-1299 period, likely due to the Norman Conquest. This trend continues into the 1300-1499 period, although at a slightly reduced rate. Latin influence, while present throughout the periods, experiences a notable surge in the 1300-1499 period, reflecting increased contact with Latin texts and scholarship. On the other hand, Scandinavian loanwords are less prominent but remain a consistent presence.

As French words were adopted, they partially took over the functions once served by derivational morphology, leading to their decline. In comparison to Old English, characterized by its highly inflected and complex grammatical structures, the transition into Middle English brought a shift to a more analytical framework. This resulted in reduced reliance on inflectional markers and, subsequently, higher emphasis on word order and auxiliary constructions. This transition led to significant changes, including the loss of grammatical gender, simplification of pronoun number distinctions, elimination of case markings, and reduction in verb conjugation patterns, among others (Brinton & Bergs 2017: 94-95). However, the changes were not a wholesale adoption of French grammar. Instead, English grammar evolved into a hybrid system, blending Old English structures with French influences.

In the early thirteenth century, English began to re-establish itself as the dominant language, particularly after King John's loss of Normandy to France, which forced nobles to choose between English or French allegiances. Despite the continuing prestige of French culture, by the mid-thirteenth century, English had become the primary language of daily life and business, even among the upper classes. Thus, the combined effects of political shifts, social changes, and the influence of the Norman Conquest led to the resurgence of English and eventual dominance in England (Nevalainen & van Ostade, 2006: 274: Allen, 2017).

The table below outlines significant syntactic changes that occurred during the transition from Old English to Middle English (Hogg & Denison., 2006: 111):

Changes in:	Old English	Middle English
<b><i>case form and function:</i></b>		
<i>genitive</i>	genitive case only, various functions	genitive case for subjective/poss. <i>of-phrase elsewhere</i>
<b><i>determiners:</i></b>		
<i>system</i>	articles present in embryo-form, system developing	articles used for presentational and referential functions
<i>double det.</i>	present	rare
<b><i>quantifiers:</i></b>		
<i>position of</i>	relatively free	more restricted

**Table 3:** *The main syntactic changes (Source: Hogg & Denison., 2006: 111)*

In OE, the genitive case served various grammatical functions, but in ME, its use became more specialized, primarily indicating subjective or possessive relationships. For other functions, it was often replaced by the “of” phrase construction. In terms of determiners, OE had an underdeveloped system, with articles in the early stages of evolution. In ME, however, articles had taken on specific roles, functioning both to introduce new information (presentational use) and to reference known information (referential use). The use of double determiners was present in OE but became rare in ME. Additionally, the placement of quantifiers, such as “some” and “all,” was relatively free in OE, allowing flexibility in sentence structure. In ME, however, this freedom was reduced. These changes may illustrate a shift from a more inflectional system in OE to a syntactic structure that relies more on word order. How nouns indicate their countability preference—such as numbers, definiteness, and quantifiers in PDE—will be thoroughly examined in the context of Middle English in the following section.

### **3.2. Countability markers in Early and Late ME**

#### **3.2.1. Number**

In terms of morphological developments, number marking continued to signal countability in ME. In OE, genitive and dative plural forms were distinct and easily recognizable, serving as clear syntactic markers for plural nouns, with the nominative plural typically ending in *-as* and the genitive singular in *-es*. However, this distinctiveness was gradually lost in ME, as the *-es* ending was increasingly applied across most nouns, regardless of case. As a result, both the nominative

plural and genitive singular of ‘stone’ in ME would appear as *stānas* (Markwardt, 2019: 51; Baber, 2000: 158). This development led to a merging of forms, rendering the genitive formally identical to other plural forms and making partitive constructions syntactically ambiguous.

Furthermore, the OE ending *-an*, which marked both the nominative plural and genitive singular, transformed in ME to *-en*, as seen in *eage* (‘eye’) and *eagan* (meaning both ‘eyes’ and ‘of an eye’). By the fifteenth century, however, the *-es* plural had largely replaced other plural forms, becoming the dominant inflection. Concerning other plural forms, some words followed the so-called front vowel mutation pattern e.g. *foot – feet*, *goose – geese*, *mouse – mice* and do not have plural inflection (Baber, 2000: 158).

Burrow and Turville-Petre (1992: 38) comment on the position of collective nouns in ME, stating that variations exist in how collective nouns like *court* or *folk* are treated. On one hand, it can be considered as one entity. On the other hand, these words sometimes do not have plural endings after numerals. Some remained from Old English with unchanged plurals, while others used to have a genitive plural after numerals.

It is evident that while the process of linguistic standardization was beginning to take shape, and increasing evidence points to some syntactic regularization, it would still take considerable time before established grammatical rules could be consistently applied. As the ME period is characterized by strong dialectal variation, it is difficult to make generalizations about plural form paradigms; rather, ME dictionaries work with possible form lists. This is also partly due to the scribes, who often continued according to the OE writing tradition or applied new conventions of Norman French, thus creating a plethora of spellings and norms (Baber, 2000: 151).

### **3.2.2. Definiteness and Quantifiers**

Tichý (2022) points out that despite its prominent position in PDE grammars, countability is not present in Old and Middle English grammar books. He claims that the PDE count-mass distinction developed, in part, due to the development and distribution of the indefinite article. In Old English, the word *ān* (one) was used to indicate an indefinite singular noun. Similarly to PDE, it did not extend its use to plural nouns. In ME, Sommerer and Hofmann (2021: 2) believe that the word *sum(e)* began to take on a grammatical role before plural and mass nouns to fill the gaps left

by *ān*. At the same time, *sum(e)* lost its original function of individualising singular count nouns, leaving that role to *ān*.

Regarding the development of article usage, in OE the definite article showed three genders (*se* for masculine, *sao* for feminine, *þæt* for neuter), and was declined through all four cases, singular and plural. The singular had a fifth case, the instrumental, *þy* or *pon*. The form *the* arose as Late OE *þe*, which replaced *se* and *seo*. In the course of ME, the other forms disappeared, and *the* became used for all of them and the sole form used for definite article. Meanwhile, *that*, which originally represented the nominative/accusative singular neuter form of the definite article, has become a contrasting demonstrative with its own distinct meaning (Baber, 2000: 160).

A notable shift in ME was the formal separation of the indefinite presentative *an* from the numeral *ān*. The OE numeral *ān* evolved into ME *oon*, while the article's vowel was reduced to a short [a], and its final nasal was dropped before consonants. This phonetic reduction allowed *an* to encroach on the territory previously held by the zero article and the other indefinite marker *sum*, thus becoming more common and semantically bleached—a typical feature of grammaticalization. As a result, *sum* itself became more specialized in ME, occurring primarily with plural nouns and those used generically, a usage that has continued into PDE (Hogg & Denison, 2006). In their analysis of texts from PPCME, Sommerer, L., and Hofmann, K. (2021) showed that in Old English, both *sum(e)* and *an* were used to indicate indefiniteness before singular nouns. However, over time, *an* became the preferred choice, pushing *sum(e)* out of this role. This change was due to competition between the two forms, with *an* eventually becoming the standard way to express indefiniteness with singular nouns.

Additionally, they explain that in later ME, *sum(e)* increasingly functioned like the modern word *some* to indicate an unspecified quantity of plural nouns (e.g., *some books*). This usage became particularly common in informal writing. Preference for complex noun phrases: From early on, *sum(e)* was favoured in complex noun phrases with modifiers both before and after the noun (e.g., *sum(e) of the olde bookes*). These complex phrases seem to have facilitated the shift of *sum(e)* towards its modern indefinite meaning

Other determiners commonly used throughout English history include possessive and genitive phrases, interrogative, indefinite, and relative pronouns, and quantifiers (Fisher & Van Der Wurff, 2017: 118).

Marckwardt (2019: 51 – 54) further focuses on the development of the quantity word *fela* (to a great extent, much)<sup>2</sup>, noting its prevalence in partitive constructions and its use with count nouns. This word had cognates in other Germanic languages and was often followed by the genitive plural of the substantive. Over time, the distinctiveness of the genitive plural forms diminished as ME adopted a generalized plural *-es* ending, making the partitive construction unrecognizable. Despite this, *fela* continued in use throughout the ME period, but with declining frequency, eventually becoming obsolete by the end of the sixteenth century.

Lastly, let us mention some additional quantifier examples. The terms indicating indefinite quantity or unit such as *any*, *some*, *each*, and *every* evolved distinct roles with subtle variations in meaning. *Any* (from OE *ænig*) denotes an indefinite unit among several, akin to 'any' in modern usage. In affirmative contexts, *any* could imply 'every,' representing all members of a group, as in "for any lord to leggen in his bedde" (countable context). The word *some* (OE *sum*) referred to 'a certain one' or 'something' selected from a group, often used substantively in lists, as in "He moot be deed, some in his bed, som in the depe see, som in the large feeld." Meanwhile, *each* and *every* (OE *æghwyle* and *ælc*) were used interchangeably to imply totality. Though *each* and *every* could function alone, adding *one* (e.g., "each one") or *thing* (e.g., "everything") clarified specific references to persons or objects (Brunner, 1965, pp. 68-69).

---

<sup>2</sup> Oxford English Dictionary, s.v. "fele (adv. & adj.2)," August 2024, <https://doi.org/10.1093/OED/7996137369>.

## 4. Method

### 4.1. Material

The core data for this study comes from the Penn-Helsinki Parsed Corpus of Middle English, second edition (PPCME2) (Kroch et al., 2000). This corpus provides a rich resource for analysing ME language use, encompassing approximately 1.35 million words across 55 distinct texts. The PPCME2 specifically targets the ME period, with texts dating from 1150 to 1500 AD. This timeframe allows for the exploration of the development and change within ME over nearly four centuries. It is divided into periods with corresponding annotations. The main periods are M1 to M4, texts originally composed in one period but surviving only in manuscripts from a later period are assigned two-digit period designations:

PPCME2 periods		
Period designation	Composition date	Manuscript date
<b>MX1</b>	unknown	1150-1250
<b>M1</b>	1150-1250	1150-1250
<b>M2</b>	1250-1350	1250-1350
<b>M23</b>	1250-1350	1350-1420
<b>M24</b>	1250-1350	1420-1500
<b>M3</b>	1350-1420	1350-1420
<b>M34</b>	1350-1420	1420-1500
<b>MX4</b>	unknown	1420-1500
<b>M4</b>	1420-1500	1420-1500

*Table 4: Overview of PPCME2 period designations*

The corpus includes both syntactic and parts-of-speech (POS) tags. Although no corpus is free from annotation errors, the PPCME2 is a widely recognized and well-established resource that ensures a high degree of reliability. The corpus has also been widely used in previous linguistic research and thus was selected as a credible source for this thesis. This corpus was accessed via the Czech National Corpus, KonText interface (Machálek, 2014).

## 4.2. Method

The methodology employed in this thesis involves a combination of quantitative and qualitative analysis, with a focus on identifying patterns and rules governing the usage of determiners and nominal pluralisation with nouns of various countability preferences. These methods were employed to explore the following research questions:

- Q1. Did pluralisation function as a distinguishing feature of countability in ME nouns?
- Q2. To what extent do determiners in ME correlate with the countability status of nouns?
- Q3. How do patterns observed in ME compare to PDE English in terms of countability markers?

As mentioned in the theoretical part, ME is a period strongly characterised by linguistic variation, specifically in the earlier parts of the period. The hypothesis is that countability markers in ME differ from those in PDE. To deal with the variability, it is necessary to incorporate as many dialectal variances as possible. This corpus, while comprehensive, does not provide tags for word lemma. Thus, KonText evaluates different spelling variants of the same word as separate words because of the lack of lemmatisation. This was one of the methodological problems that needed to be accounted for, as ME is a period of vast orthographic variation. All queries were done using the Corpus Query Language (CQL). In preparation for my analysis, OpenAI ChatGPT (November 2024 version) was used to generate the CQL queries in order to make the search for all word forms more efficient. It was prompted with regular expressions employed by KonText and relevant orthographic variations that needed to be included. After using this tool, I have further reviewed and corrected the queries as needed and take full responsibility for its content.

Two possible approaches to conducting linguistic research in corpora exist. Either they are used as a source of data from which linguistic theories are deduced, or they serve as a collection of empirical examples on which already established theories are confirmed or disproved. This study follows the latter approach, the corpus-based method.

### 4.3. Analysis

#### 4.3.1. Pluralisation

The first analysis examined the features of noun countability by investigating their plural forms. Nominal pluralisation had already been established as a feature of English nouns in OE and remained a productive feature in ME. Nonetheless, the morphological features of these plural forms were highly variable and had yet to be standardized. Our presumptions arise from PDE observations and grammar rules, and this analysis aimed to determine whether nominal pluralisation functions as a distinguishing feature of countability in ME lexis. This would presume that presumed C nouns would occur in a plural form more frequently, while presumed U nouns would stay in their singular form. Nouns that may have both meanings, putative C/U nouns, would presumably occur in both or fluctuate.

First, the frequency list of nouns was extracted from PPCME2, giving us 200 of the most frequent noun word forms. Nouns with multiple forms due to this variation lead to homonymy and require additional attention. For example, the plural ending *-en* can signify both the genitive singular noun form and the nominative plural form, e.g. ME *stones* instead of OE *stānas* ‘stones’ and *stānes* ‘of a stone’ (Barber, 2000). The PPCME2 authors further comment on their decisions in the POS tagging process of earlier periods: “In early texts, before the universalization of plural *-s*, it can be quite difficult to distinguish reliably in all cases between singular and plural. Therefore, for the period M1, we have tried to follow the translation accompanying the edition used when one is available, or else a separate translation,” (Kroch et al., 2000).

The problem of ambiguous ME noun endings, that could signify both singular and plural forms, was addressed by checking the POS tags included in the corpora. Those forms with ambiguous endings were thus separated either into singular or plural tables according to their tags. Word forms in possessive/genitive form were not included in the analysis as they were too scarce in our dataset:

- N            Noun, singular and collective
- N\$          Noun, possessive/genitive
- NS          Noun, plural
- NS\$        Noun, plural, possessive/genitive

Additionally, proper nouns were excluded from the analysis as well. They often behave grammatically differently from common nouns, and more importantly, proper nouns are inherently outside the category of countability. Focusing on common nouns, thus, ensures that the analysis captures general grammatical patterns and avoids inconsistencies introduced by the unique behaviour of proper names.

To further ensure accuracy, a manual correction process was conducted. For example, the ME word *sonne*, which was one of the most frequent ones in PPCME2, could both mean ‘son’ and ‘sun’. This correction process involved consolidating spelling variants and distinguishing between homonymous nouns to produce a refined list that accurately represented distinct lexical items in their ME forms. For this, the online *Middle English Dictionary* (MED) run by the University of Michigan was employed. It provided a non-exhaustive list of spelling and inflectional variations of selected nouns, including accounts of erroneous spellings, inflectional forms and some plural forms. Another example would be the word *man*, which has all the following spelling variations, and thus was excluded from the analysis:

**man** *n.* Also *mane, manne, mon(e, monne, (?error) ma.* Forms: sg. gen. *man(e)s, mannes, etc. & man, (error) men; dat. man(ne, etc. & (early) men, manna, monnen; pl. men, menne, (errors) meyn, man, me, mennes & (early) mæn, monne, mannen; gen. men(e)s, mennes, (error) mannes & men & manne, monne, menne, mennene, (early) manna, -æ, monna, -æ, mone, mannen(e, monnen(e, -an, mennen; dat. men, menne & manne, monne, (early) mannen, -a(n, -um, monnen, -an, -um, -am, monen, mennen & (error) mannes.*

Additionally, it proposes a problematic task of distinguishing between the indefinite pronoun and noun, which could only be resolved by laborious manual checking of every corpus account. Due to homonymy and a vast number of orthographic forms, words such as *man* were not taken into account for the scope of this analysis and to further eliminate human error during manual analysis. Lastly, the fact that the MED list of word forms is non-exhaustive, the query aimed to utilize as many wildcards (regular expressions with special meaning) as possible, e.g. `word="h[ae]rm(e|es|en)?"` would include all of these words: *harm, herm, harme, harmes, harmen, herme, hermes, hermen*. It is vital to note that although this thesis strives to encompass the majority of variations for a single word, it recognizes that the extensive range of spelling differences prevents it from claiming to offer a comprehensive list.

Following the creation of this preliminary noun list, the nouns were subsequently cross-referenced with their PDE equivalents. For this purpose, we employed the *Oxford Advanced Learner's Dictionary*, a reliable resource for determining noun countability. Each noun was categorized based on its PDE countability status: 'C', 'U', and 'C/U' for nouns that can be used in both contexts e.g. *punishment* (C: *I was sent to bed as a punishment, U: He could take any amount of punishment*). This selection differentiated selected nouns based on their expected countability preference in ME. As has been mentioned at the beginning of this thesis, even though countability is often structured in a binary framework, it should still be seen as a fluid, scalar value. Thus, even though some words are categorized in one group, it does not imply that they may not be used in different countability contexts. Instead, it suggests that the chosen category represents the dominant meaning and that the word most frequently appears in this context.

As a result, a list of 60 nouns was created, representing both the prototypically PDE countable nouns, e.g. *king* and prototypically uncountable nouns, e.g. *reason*. The following tables provide a summary list of selected ME words, **Table 5** for plural and **Table 6** for singular forms. The number of instances found in PPCME2 is indicated in brackets:

PDE COUNTABILITY CATEGORY					
countable		uncountable		both meanings	
PDE word	ME pl. variants	PDE word	ME pl. variants	PDE word	ME pl. variants
<i>day</i>	dayes (317), daies (136), daȝes (26), dages (12)	<i>grace</i>	graces (14), gracis (4)	<i>time</i>	times (8), timis (5), timys (1)
<i>heart</i>	heortan (6), heorte (21), hertes (53), heorten (5), herte (2)	<i>folk</i>	folkes (37), folk (20)	<i>land</i>	lande (1), londe (1), londes (63), landes (30), laundes (1)
<i>castle</i>	castelles (21), castles (13), castels (7), castelle (1), casteles (12), castellys (12)	<i>bliss</i>	blissen (12), blisses (3)	<i>manner</i>	manieres (1), maneres (88), maner (2)
<i>horse</i>	horse (1), horses (1), hosen (4), hors (1), hoses (3), horses (11)	<i>dread</i>	dredes (3)	<i>word</i>	wored (3), weord (2), worden (8), words (2), wordis (158), wurdes (1), wordus (5), word (37), wodes (10)

<i>father</i>	fadres (24), faderes (8)	<i>age</i>	ages (9)	<i>church</i>	0
<i>place</i>	places (77)	<i>pride</i>	0	<i>honour</i>	honours (6)
<i>end</i>	endes (10)	<i>gold</i>	0	<i>passion</i>	passyons (3)
<i>pope</i>	papes (1), popes (16)	<i>charity</i>	charites (1)	<i>reason</i>	resones (2), reses (3), resounes (1)
<i>world</i>	worldes (18), worldis (1)	<i>mind</i>	0	<i>stone</i>	stanen (1), stane (2), stones (76)
<i>earl</i>	eorles (7), eorless (1), eorle (1)	<i>flesh</i>	0	<i>home</i>	hames (1)
<i>book</i>	bokes (42)	<i>mercy</i>	messys (1), mercies (13)	<i>desire</i>	desires (13)
<i>god</i>	goden (1), godes (78), guode (1)	<i>love</i>	louves (1), lufes (2)	<i>country</i>	contrees (41), cuntrees (3), countrees (3), contreies (1)
<i>kingdom</i>	kyngdomys (1), kyngdoms (7), kyngdammes (3), kyngdomes (12)	<i>joy</i>	joyes (8), yoyes (1)	<i>life</i>	liften (1), lifes (3), leves (7)
<i>mother</i>	modren (1), modres (4)	<i>blood</i>	0	<i>gospel</i>	gospelles (5)
<i>child</i>	childre (6), children (218)	<i>help</i>	helpes (1)	<i>power</i>	0
<i>house</i>	hous (1), houses (22), huses (4)	<i>earth</i>	erthes (1)	<i>judgement</i>	0
<i>name</i>	name (1), nomen (2), names (66)	<i>glory</i>	0	<i>glory</i>	spiritis (8), spiritys (2), spirites (12)
<i>temple</i>	temples (4)	<i>hope</i>	0	<i>penance</i>	penaunce (1), penances (1), penonce (1)
<i>emperor</i>	emperoures (9)	<i>wisdom</i>	wisdoma (1), widsomes (2)	<i>bed</i>	beden (1), beddes (2), bede (8)
<i>woman</i>	wemmes (1), wimmon (3), wimmen (24), wummen (10), wommen (71), wimon (1), women (51), wemmen (3), wemen (3), wummon (2)	<i>water</i>	waterse (1), watteres (1), wattres (2), waters (28), weteres (1), wateres (6)	<i>punishment</i>	pinen (14), pine (6), peynes (37), peines (2), paynes (26), pinen (26)

*Table 5: List of ME word forms (plural), their PDE counterparts and countability category*

**Table 6:** List of ME word forms (singular), their PDE counterparts and countability category

PDE COUNTABILITY CATEGORY					
countable		uncountable		both meanings	
PDE word	ME sg. variants	PDE word	ME sg. variants	PDE word	ME sg. variants
<i>day</i>	daies (1), dai (164), day (1341)	<i>grace</i>	grace (849)	<i>time</i>	time (216), teme (5), tima (1)
<i>heart</i>	hertes (1), heart (1), heorte (276), herte (707)	<i>folk</i>	fólkes (6), folk (371)	<i>church</i>	chirche (194), churche (8), cherche (22)
<i>castle</i>	castelle (49), casteel (1), castel (149)	<i>bliss</i>	blise (1), blisce (11), blissen (1), blisse (452), blis (9)	<i>manner</i>	manire (1), manere (384), maner (618)
<i>horse</i>	horse (179), horsse (1), hors (163), hor (1), hosen (1)	<i>mercy</i>	mercer (4), merceye (13), Mercy (10), mercy (520), merci (66)	<i>word</i>	woerdes (1), wode (52), vord (1), word (508), wort (3), wurð (6), weord (1), wod (3)
<i>father</i>	fader (275), vader (3), fadere (7), feder (17)	<i>love</i>	luve (1), laf (8), lufen (1), lofe (126), leve (36), love (112), lufe (382), lowe (3), lave (1), lof (4), luffe (5)	<i>land</i>	londe (157), lande (349), land (150), lote (2), lond (498)
<i>place</i>	plase (8), place (489)	<i>pride</i>	priede (1), prudu (1), prede (28), prude (27), pride (114)	<i>honour</i>	honor (1), onour (5), honur (8), honour (77)
<i>end</i>	end (9), inde (3), ende (756), ynde (1)	<i>gold</i>	golt (5), gold (232), golde (62)	<i>passion</i>	passyoun (2), passyon (32), passion (41), passioun (30)
<i>pope</i>	pope (384), pape (2)	<i>joy</i>	joie (4), joy (130), joye (63)	<i>life</i>	lijf (75), lefe (9), life (41), liffe (24), live (1), leve (36)
<i>world</i>	world (649), world (88), wurld (1)	<i>mind</i>	munde (2), minde (9), mynd (15), mynde (174), mende (62)	<i>stone</i>	stane (5), stone (59), stones (2)
<i>earl</i>	eorl (37), eorldom (3), eorle (3)	<i>flesh</i>	flesshe (29), fleche (2), flesh (7), flesse (1), fleish (1)	<i>home</i>	hom (114), home (77), hame (12)
<i>book</i>	buce (1), bok (8), boke (108), book (243), boc (109)	<i>dread</i>	dret (1), drede (261)	<i>desire</i>	desier (2), desire (99), desir (21)
<i>god</i>	godes (5), guode (26), god (221), goð (2)	<i>age</i>	age (134), ayge (10)	<i>country</i>	contree (155), countree (2), cuntree (1)

<i>kingdom</i>	kyngedome (5), kingdome (1), kingdam (2), kyngdam (21), kyngdome (8), kingdom (2), kyngedom (4), kyngdom (140)	<i>charity</i>	cariteð (1), carite (1), kariteð (4), charite (219), chearite (1), cherite (19), cherte (2), cariteþ (1), carited (1), charitee (8)	<i>reason</i>	reson (61), resun (22), reson (2), reisun (21), resoun (172), resoune (1), resone (7), reson (1), reason (11)
<i>mother</i>	modir (63), modur (10), moder (215)	<i>blood</i>	blod (204), blood (130), blode (75), bloode (25), blude (15)	<i>gospel</i>	gospelle (4), godspel (16), gospel (106)
<i>child</i>	cild (1), chil (1), child (279), childe (152), childre (1)	<i>help</i>	helpes (1), help (161), helpe (169)	<i>power</i>	pouwer (10), power (307), pouer (7)
<i>house</i>	howe (1), hus (97), house (53), hous (273), husa (1)	<i>earth</i>	þerthe (2), eurthe (2), erthe (155)	<i>judgement</i>	jugement (8), jugge (4)
<i>name</i>	nam (9), name (734), nome (89), noma (1), nomen (1)	<i>glory</i>	glorie (98), gloire (2), glory (12)	<i>glory</i>	glorie (98), gloire (2)
<i>temple</i>	tempul (1), tempulle (1), tempil (1), teple (202)	<i>hope</i>	hope (169)	<i>penance</i>	penans (2), penance (78), penonce (10), penaunce (87)
<i>emperor</i>	emparour (13), emperour (165)	<i>wisdom</i>	wysdome (2), wysdom (79), wisdom (114), wysdome (25), wisdom (20), wisdomes (1), wysdam (3), wisdome (2), wisdom (16), wysedom (3), wisdom (6)	<i>bed</i>	bed (78), bede (26), bedde (64)
<i>woman</i>	wimmon (21), wummon (47), wimmen (2), womman (202), wumman (2), women (2), wummone (1), womon (2), wimman (12), woman (247), wommen (1)	<i>water</i>	wæter (1), weattre (1), watre (6), water (274), wetere (4), waterr (11), watir (73), watire (2), veter (11), watere (23), watur (26), weater (5)	<i>punishment</i>	payne (103), pinen (2), peyne (139), paine (8), peynes (2), peine (2), payn (2)

As previously mentioned, PPCME2 does not include word lemmatization. Thus, a CQL query that would include all possible plural variations had to be created. The query was structured in a broader manner so as not to omit any possible plural form and include plural forms that were shared with singular e.g. *herte* (heart), signifying both singular and plural form (in PPCME2 distinguished by POS tag). An example of a CQL query for the word *castle* (see the full list of CQL queries in **Appendix**):

**castle** - [word="cast[e,ē]l\*[e]." | word="castl.\*" | word="kestel.\*" | word="chastel.\*" | word="castel.\*"]

As a result, this query gave us the following variations (manually limiting the search only for plural and singular non-inflected forms), singular - *casteel*, *castel*, *castelle*, and plural – *castels*, *casteles*, *castelles*, *castles*, *castelle*, *castellys*. Finally, the absolute and relative frequency (i.p.m.) for each word and its spelling variations were extracted. Similarly, each noun was labelled according to its semantic countability category: ‘concrete,’ ‘abstract,’ and ‘substance,’ as this classification also affects its expected countability preference as suggested by the notional account. Below is an example entry for the word *day* in plural forms:

count.	cat.	PDE word	tag	ME word	Freq	i.p.m.	Freq. 2	i.p.m. 2
C	concrete	day	NS	dayes	317	233,41	491	361,53
			NS	daies	136	100,14		
			NS	dazes	26	19,14		
			NS	dages	12	8,84		

To calculate the percentage of plurals derived from all noun forms, the same methodological steps were applied as those used for singular forms:

tag	ME sg.	Freq	i.p.m.	Freq. sum	i.p.m. sum
N	daies	1	0,74	1506	1108,89
N	dai	164	120,76		
N	day	1341	987,4		

Subsequently, this study closely analysed the behaviour of nouns in relation to their semantic category mentioned above, grouping them into three distinct categories: ‘abstract,’ ‘concrete,’ and ‘substance’ nouns. The classification of nouns into these categories is based on the different ways

in which they represent concepts in the real world. Abstract nouns refer to ideas, qualities, or states that are intangible and cannot be physically perceived. These nouns typically describe emotions, conditions, or intellectual concepts, such as *hope*, *love*, or *freedom*. Concrete nouns, on the other hand, denote physical objects or things that can be perceived by one's senses, such as *day*, *horse*, or *emperor* in our dataset. These nouns represent entities that exist in the physical world and can be observed or touched. Lastly, substance nouns are those that refer to materials or substances that do not have a specific shape or boundary, such as *water*, *gold*, or *fire*. These nouns typically refer to substances or qualities that are continuous and do not easily divide into distinct, countable units.

The dataset used in this study contains a relatively limited number of substance nouns, with only five instances categorized as such: *gold*, *water*, *blood*, *flesh*, and *fire*. In comparison to abstract and concrete nouns, substance nouns tend to be more specialized in use and, in our data, showed less frequently. In future research, focusing on specialized corpora or expanding beyond PPCME2 to include cookbooks, medical texts, or trade records could provide richer data.

The following words were used for the semantic analysis:

- **Abstract:** *age, bliss, charity, desire, dread, end, glory, god, grace, help, honour, hope, judgement, joy, kind, life, love, mercy, mind, pain, passion, penance, power, pride, reason, spirit, time, wisdom.*
- **Concrete:** *bed, book, castle, child, church, country, day, earl, emperor, father, folk, gospel, heart, home, horse, house, kingdom, land, mother, name, place, pope, stone, temple, woman, world.*
- **Substance:** *blood, fire, flesh, gold, water.*

The results of this analysis can be found in the subsequent **Chapter 5.1**.

#### **4.3.2. Determiners**

Following the analysis of nominal pluralisation, this work examines the use of determiners in relation to the selected dataset. This analysis aimed to address the question: To what extent do determiners in ME correlate with the countability status of nouns?

The authors of PPCME2 POS annotation guidelines list the following words as determiners in the corpus: A(N), THAT, THE, THESE, THIS, THOSE, YON, YONDER. Due to the lack of

lemmatisation, the corpus includes 294 forms of them in total. The authors add that “demonstratives are always tagged D, regardless of whether they precede a noun,” (Santorini, B., 2016). For our analysis, we have selected the most frequent determiners - the indefinite articles (*a/an*) and definite (*the*), in all their spelling variations.

Our assumptions were grounded in the usage of articles in PDE. English nouns must be predetermined, typically by indefinite articles (*a/an*) or their definite counterpart (*the*). The indefinite articles are used with a singular countable noun when the noun referred to is nonspecific or generic e.g. *My mum wants to eat an apple*. Only countable nouns can be used with the indefinite articles. On the other hand, the definite article is used to show specific reference, e.g. *My mum wants the apple*. This sentence refers to an object that has been previously mentioned or described through possibly spatial deixis, so it has a specific referent. It can be used with both singular and plural nouns and both C and U nouns. A singular C noun always requires an article, which can be either definite or indefinite. In contrast, when a C noun is plural, it takes the definite article if it refers to a specific group or entity. However, if the plural noun is used in a general sense, no article is needed e.g. *Lions eat meat*. On the other hand, in PDE, uncountable nouns never take the indefinite article. Our analysis wanted to explore whether the last statement stands true for ME data.

Regarding the material, the same set of words (**Table 4**) was used for the analysis. Firstly, we took the grammatical rule that uncountable singular nouns should never occur with an indefinite article and would result in an ungrammatical construction, e.g. *\*Give me a water*. See an example CQL query for uncountable noun *grace*:

```
[tag = "D"] + [word="[gc]ra[c,s,z][ie].?"]
```

This search returned 161 results, including other determinants. **Figure 2** represents an example of our data after manual extraction. The percentage value of indefinite determination was calculated from the total number of all instances with determination preceding the U noun. It is essential to address the arising methodological problem that this approach could miss instances where an element is positioned between the determiner and the noun. Additionally, it assumes a fixed word order which may not always apply to ME nouns. To address the issue of elements between the determiner and the noun, ANNIS (Krause & Zeldes, 2016) tool was used, which works with syntactically tagged corpora such as PPCME2. ANNIS allows to create complex queries to find

patterns where elements appear within noun phrases. This approach will help move past the limits of assuming fixed word order and ignoring syntactic variation, giving a clearer view of the data (see **Figure 6**).

The results of this analysis can be found in the following **Chapter 5**.

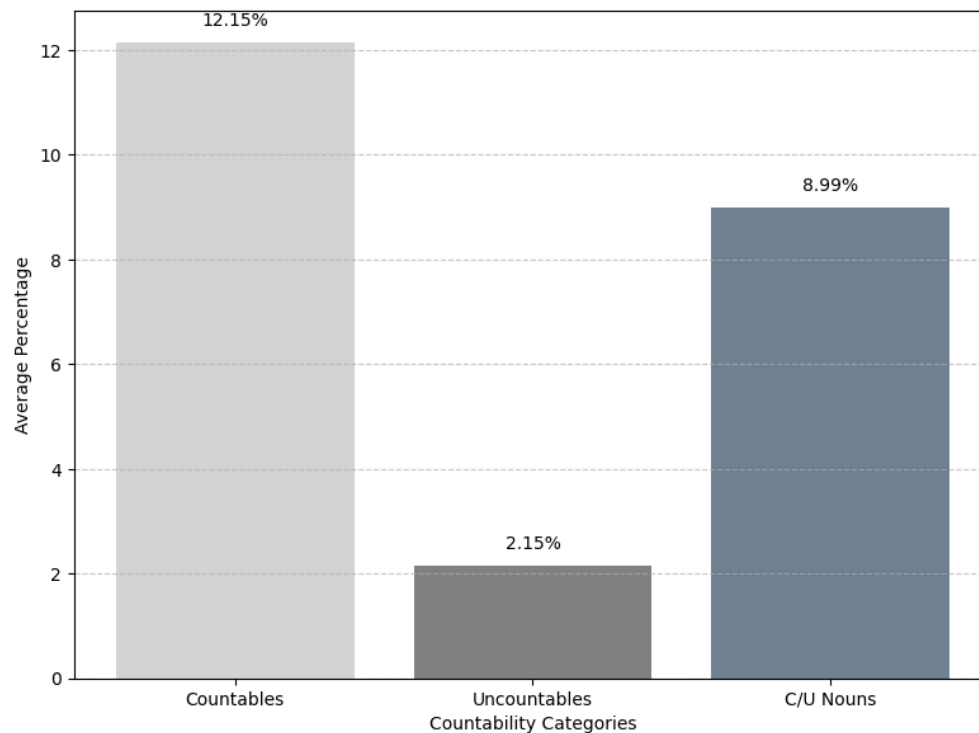
PDE word_U	Det_type_U	Det_ex_U	ME form_U	tag_U	Freq_U	i.p.m._U	Freq_sum_U	Freq_total_D	%_IND_U
grace	D	be	grace	N	77	56,7	115	154	4,55%
	D	the	grace	N	36	26,51			
	D	be	grace	N	1	0,74			
	D	The	grace	N	1	0,74			
I	a	grace	N	5	3,68	7			
	e	grace	N	1	0,74				
	ane	grace	N	1	0,74				

*Figure 2: Data example for U nouns with an indefinite article*

## 5. Results

### 5.1. Pluralisation

The main task of this analysis was to measure and compare the percentage of plural forms across three distinct countability categories – ‘C’, ‘U’, and ‘C/U’. **Figure 3** illustrates the results of the comparison of the average percentage distribution of plural forms across three distinct countability categories:



*Figure 3: Average plural distribution for Countables, Uncountables, and C/U Nouns*

The findings on nominal pluralisation as an indicator of countability preference are as follows: the category of countable nouns shows the highest average percentage of plural forms at 12.15%. That means that out of all noted word forms in PPCME2, 12.15% of them occurred in the plural. This result suggests the predicted assumption that countable nouns are likely to appear in the plural, as they can be directly counted and categorized in both singular and plural contexts. This is consistent with the PDE countability principles. It is important to recognize that the 12.15% occurrence of plural forms for C nouns in ME indicates that the majority of instances (88.85%) were still in singular form.

To check the statistical significance of our findings, we have conducted relevant statistical tests, in our case for non-parametric, small sample data. The Mann-Whitney U test gives us a p-value of 0.0563, indicating that there is no statistically significant difference between the percentages of C nouns and C/U at the conventional significance level of 0.05. This does not mean there is absolutely no difference; it just means the data do not provide enough evidence to confidently claim one. Nonetheless, the difference between C and U nouns did prove statistically significant with  $p = 0.000016$  ( $\alpha = 0,05$ ). The test strongly suggests that countable and uncountable nouns have significantly different patterns in how their percentages are distributed and the difference is not a result of chance. Similarly, there is a statically significant difference between U and C/U nouns with  $p = 0.00594$  ( $\alpha = 0,05$ ).

The following words occurred with above 10% of instances in plural form (a threshold selected by the author for this analysis):

<b>PDE word</b>	<b>i.p.m. plural</b>	<b>i.p.m. singular</b>	<b>sum</b>	<b>perc.</b>
<i>temple</i>	16.94	150.94	167.88	10.1%
<i>kingdom</i>	16.94	134.75	151.68	11.2%
<i>place</i>	56.70	365.95	422.64	13.4%
<i>earl</i>	6.63	31.66	38.29	17.3%
<i>woman</i>	124.44	399.08	523.52	23.8%
<i>god</i>	58.91	187.02	245.93	24.0%
<i>day</i>	361.52	1108.89	1470.42	24.6%
<i>castle</i>	48.60	146.53	195.12	24.9%
<i>child</i>	164.93	319.56	484.49	34.0%

The word *child* was shown to be the most commonly used in both number forms. With 224 examples in plural and 434 in singular form, it offers the most balanced distribution. Nonetheless, the noun *castle* was selected as it fulfils the criteria for C nouns while remaining manageable for qualitative analysis within the scope of this thesis.

The results generally align with the PDE countability status, as distinct plural forms are recorded consistently across periods and regions with examples from 21 unique sources. Abbreviations in the third column represent the main dialectal regions in 12<sup>th</sup> to 15<sup>th</sup> century England: East Midlands (EM), Southern dialect (S), West Midlands (WM), Northern dialect (N), and Kentish dialect (K). The date refers to the presumed date of origin of the document. The variability in plural forms reflects regional orthographic conventions rather than different concepts of countability:

form	ME form	dialect	text name	date	instances	
pl	castelle	WM	<i>The Brut or The Chronicles of England</i>	c1400	1	
		WM	<i>The Brut or The Chronicles of England</i>	c1401	11	
	castelles	EM	<i>Capgrave's Chronicle</i>	a1464	2	
		EM	<i>The Tale of Melibee</i>	c1390	1	
		EM	<i>Mandeville's Travels</i>	?a1425 (c1400)	3	
		WM	<i>Mirk's Festial</i>	a1500 (a1415)	1	
		S	<i>John of Trevisa's Polychronicon</i>	a1387	3	
		casteles	K	<i>Ayenbite of Inwyt</i>	1340	4
			EM	<i>Capgrave's Chronicle</i>	a1464	3
			EM	<i>The Peterborough Chronicle</i>	no MED date	3
			EM	<i>Vices and Virtues</i>	a1225 (c1200)	2
		castellys	S	<i>Gregory's Chronicle</i>	c1475	11
	WM		<i>Mirk's Festial</i>	a1500 (a1415)	1	
	castels	EM	<i>The Earl. Complete Engl. Prose Psalter</i>	x	2	
		WM	<i>Malory's Morte Darthur</i>	a1470	4	
		S	<i>Purvey's General Prologue to the Bible</i>	a1450	1	
	castles	WM	<i>Ancrene Riwe</i>	no MED date	2	
WM		<i>The Lambeth Homilies</i>	a1225	1		
EM		<i>The Peterborough Chronicle</i>	no MED date	10		
sg	castelle	S	<i>Gregory's Chronicle</i>	c1475	47	
		WM	<i>Malory's Morte Darthur</i>	a1470	2	
	casteel	EM	<i>English Wycliffite Sermons</i>	c1400	1	
		castel	WM	<i>Ancrene Riwe</i>	no MED date	16
	K		<i>Ayenbite of Inwyt</i>	1340	4	
	WM		<i>The Brut or The Chronicles of England</i>	c1400	39	
	EM		<i>Capgrave's Chronicle</i>	a1464	14	
	WM		<i>The Lambeth Homilies</i>	a1225	8	
	WM		<i>Malory's Morte Darthur</i>	a1470	24	
	EM		<i>Mandeville's Travels</i>	?a1425 (c1400)	2	
	WM		<i>St. Margaret</i>	c1225 (?c1200)	1	
	WM		<i>The New Testament (Wycliffite)</i>	a1425	3	
	EM		<i>The Peterborough Chronicle</i>	no MED date	10	
	S		<i>John of Trevisa's Polychronicon</i>	a1387	17	
	EM		<i>Caxton's History of Reynard the Fox</i>	1481	3	
	N		<i>Richard Rolle, Epistles</i>	a1450 (?1348)	1	
	WM		<i>Sawles Warde</i>	c1225 (?c1200)	2	
	WM		<i>English Wycliffite Sermons</i>	c1400	5	

**Table 7:** List of spelling variations for CASTLE and their regional and temporal metadata from PPCME2

From a semantic point of view, all of the analysed countables are concrete nouns, as they refer to specific, tangible entities that exist in the real world. These nouns describe things that can be physically experienced, and they represent objects or concepts that can be counted individually. Whether referring to a physical structure like a *temple* or *castle*, a political entity such as a *kingdom*, or people like a *woman*, *emperor* or *earl*, each term denotes a distinct, countable unit and they all point to things with real-world references that can be quantified. Although these words denote concrete concepts with a clear, perceivable existence, their varying percentage values reveal that the group of countable nouns is not homogeneous when it comes to nominal pluralization.

While this thesis includes metadata of dialectal distribution for two words, a full study of dialectal and time-based influences on countability markers was not carried out. The focus of this research was to map the general patterns of the countability system in ME. A broader analysis of dialect and temporal changes would have required a larger dataset, which would have expanded the project significantly. Future research could look into these areas to better understand how dialects and time periods affected countability markers in ME.

Some nouns do not have many instances in plural potentially not due to their low-countability preference, but possibly due to the real-world circumstances. For instance, the term *pope* appears in the plural form only 4.2% of the time, representing an institution that, aside from certain historical crises, has typically functioned as a singular entity. Thus plural forms may have been possible but not recorded as they were not needed. Similarly, *mother* (1.7%) may have referred to the religious meaning of ancestress or spiritual mother and even in more secular meaning it may not have been common to refer to multiple female parents as its singular form is frequently recorded in the corpus (288 entries).

In contrast, the category of uncountable nouns shows a significantly lower percentage at 2.15% which aligns with the grammatical characteristics of PDE uncountable nouns. These nouns, such as *water* or *gold*, generally do not take plural forms because they refer to substances, qualities, or abstract concepts that cannot be easily counted. The low average percentage suggests that pluralization of uncountable nouns does occur in ME, but is rare, typically occurring in specific contexts or exceptions. Only one word occurred with above 10% in the plural form, *folk*:

PDE word	i.p.m. plural	i.p.m. singular	sum	perc.
<i>folk</i>	41.97	277.59	319.56	13.1%

*Folk* is a collective noun. While collectives semantically denote a plural meaning, they grammatically have a singular form. They may become plural, especially in complex sentences or when distinguishing between multiple groups, e.g. *The folk is gathering at the hall.* X *The folks are coming from different villages.* The plural form is used when the focus is on the individual members of the group rather than the group as a single entity:

- a) **singular:** And God woot that in alle these langages and in many moo han these conclusions ben suffisantly lerned and taught, and yit be diverse reules ; right as diverse pathes leden diverse folk the righte way to Rome. (*A Treatise on the Astrolabe*)
- b) **plural:** [...] and on a stede þeron munegeð some of þe wordes þe weren spoken bitwine two folkes. þat on was of ierusalem. and þat oðer of babilonie. (*Trinity Homilies*)

The following words do not have any records of nominal pluralization in PPCME2:

PDE word	i.p.m. plural	i.p.m. singular	sum	perc.
<i>hope</i>	0	124.44	124.44	0.0%
<i>pride</i>	0	125.91	125.91	0.0%
<i>mind</i>	0	181.87	181.87	0.0%
<i>gold</i>	0	49.33	49.33	0.0%
<i>blood</i>	0	330.60	330.60	0.0%
<i>flesh</i>	0	24.30	24.30	0.0%
<i>glory</i>	0	73.63	73.63	0.0%

As will be further evident in the semantic analysis, many of the words above represent substance nouns and these almost never distinguish between singular and plural numbers in PDE. Our data suggests that this pattern may stand even for ME lexis.

The last category analysed words that can occur in both countable and uncountable contexts, the C/U nouns. Their average percentage of 8.99% suggests a moderate frequency of nominal pluralization. This category reflects the variability of certain nouns that can shift between C and U classifications, resulting in a more balanced occurrence of plural forms. The following words occurred with above 10% of instances in the plural form:

PDE word	i.p.m. plural	i.p.m. singular	sum	perc.
<i>spirit</i>	16.20	103.08	119.28	13.6%
<i>country</i>	35.34	116.38	151.72	23.3%
<i>word</i>	164.93	423.38	588.31	28.0%
<i>stone</i>	58.17	48.59	106.77	54.5%

A closer look at an exemplary word from the group above, *spirit*, shows that the ME word exhibits variability in its usage, functioning as both a C and a U noun depending on context. The following list provides a summary of sources and their origin:

number	ME form	dialect	text name	date	instances
pl	<i>spiritis</i>	EM	The Old Testament (Wycliffite)	a1425 (c1395)	1
		S	Purvey's General Prologue to the Bible	a1450 (a1397)	2
		EM	English Wycliffite Sermons	c1400	5
	<i>spiritys</i>	EM	The Book of Margery Kempe	c1450	2
	<i>spirites</i>	WM	Aelred of Rievaulx's De Institutione Inlusarum (Ms. Vernon)	c1400	2
		EM	Aelred of Rievaulx's De Institutione Inlusarum (Bodley 423)	a1450	1
		EM	Capgrave's Chronicle	a1464	1
		EM	The Parson's Tale	c1390	1
		EM	Mandeville's Travels	?a1425 (c1400)	2
		S	John of Trevisa's Polychronicon	a1387	3
		N	Richard Rolle, Epistles	a1450 (?1348)	1
	EM	English Wycliffite Sermons	c1400	1	
	sg	<i>spirith</i>	EM	English Wycliffite Sermons	c1400
<i>spiryte</i>		EM	The Cloud of Unknowing	a1425 (?a1400)	1
		N	Richard Rolle, Epistles	a1450 (?1348)	1
		EM	Capgrave's Chronicle	a1464	1
<i>spirite</i>		EM	The Cloud of Unknowing	a1425 (?a1400)	8
		EM	The Book of Margery Kempe	c1450	1
		N	Richard Rolle, Epistles	a1450 (?1348)	1
		N	The Northern Prose Rule of St. Benet	a1425	1
<i>spirit</i>		WM	Aelred of Rievaulx's De Institutione Inlusarum (Ms. V.)	c1400	6
<i>spirit</i>		WM	Aelred of Rievaulx's De Institutione Inlusarum (Ms. V.)	c1400	6
		EM	Aelred of Rievaulx's De Institutione Inlusarum (Bodley 423)	a1450	4
		EM	A Treatise on the Astrolabe	a1450 (c1391)	1
		K	Ayenbite of Inwyt	1340	1
		EM	Capgrave's Chronicle	a1464	2
EM		The Cloud of Unknowing	a1425 (?a1400)	34	

EM	The Parson's Tale	c1390	5
WM	The Mirror of St. Edmund (Vernon Ms.)	c1390	15
EM	Hilton's Eight Chapters on Perfection	a1450 (a1396)	9
EM	The Book of Margery Kempe	c1450 (probably before)	4
WM	The New Testament (Wycliffite)	a1425	11
EM	The Old Testament (Wycliffite)	a1425 (c1395)	4
S	John of Trevisa's Polychronicon	a1387	2
S	Purvey's General Prologue to the Bible	a1450 (a1397)	10
N	Richard Rolle, Prose Treatises from the Thornton Ms.	c1400	1
EM	The Book of Vices and Virtues	c1450	8
EM	English Wycliffite Sermons	c1400	6

**Table 8:** List of spelling variations for SPIRIT and their regional and temporal metadata from PPCME2

As a countable noun, it appears in plural forms such as *spiritis*, *spiritys* and *spirits*, recorded in mostly religious texts like *The Old Testament of Wycliffite Bible*, *Purvey's General Prologue to the Bible*, and *English Wycliffite Sermons*. These instances reflect the noun being treated as individual, countable entities, with distinct forms indicating numerical quantity:

- i. The thridde shal be in heuen with blessed spirites and aungels, if we wel do.  
(*Aelred of Rievaulx's De Institutione Inclusarum*)

On the other hand, in its uncountable sense, *spirit* appears in singular forms like *spirith*, *spiryte*, *spirte*, *spirit* and *spirite*, found with most instances in works such as *The Cloud of Unknowing* and again the *English Wycliffite Sermons*. In this usage, the word refers to a more abstract, unquantifiable concept, such as an essence or force, and does not take the plural form:

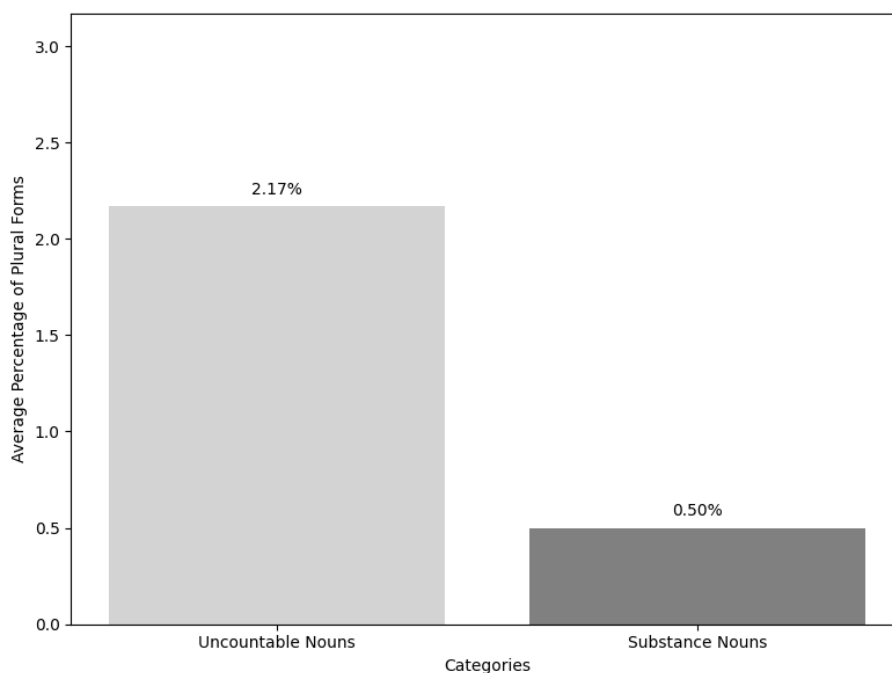
- ii. Bot I bid þee do þat in þee is to hide it. And whi bid I þus? Sekirly for I wolde þat þou castedest it into depnes of spirite, fer fro any rude medelyng of any bodelines.  
(*The Cloud of Unknowing*)

In the second part of the analysis on nominal pluralisation, the data was analysed based on its differences in word meaning. The analysis of the semantic influence on countability included the comparison of uncountable nouns and their semantic subgroup of substances. These groups overlap in **Figure 3**. However, for **Figure 4** the words were divided, and those classified as

substances appearing in both groups were included exclusively in the substance category. See the word list included in the analysis below, along with their corresponding graph comparison:

<b>PDE word</b>	<b>i.p.m. plural</b>	<b>i.p.m. singular</b>	<b>sum</b>	<b>perc.</b>
<i>water</i>	8.11	321.77	329.87	2.5%
<i>gold</i>	0	220.16	220.16	0.0%
<i>blood</i>	0	330.60	330.60	0.0%
<i>fire</i>	0	36.08	36.08	0.0%
<i>flesh</i>	0	29.45	29.45	0.0%

**Figure 4** displays a comparative analysis of the average percentage distribution of plural forms for two categories of uncountable nouns: Uncountables and Substances. The data reveals a great contrast in the frequency of plural forms between these two groups. Uncountable nouns show a higher tendency to form plurals (2.17%) compared to substance nouns (0.50%).

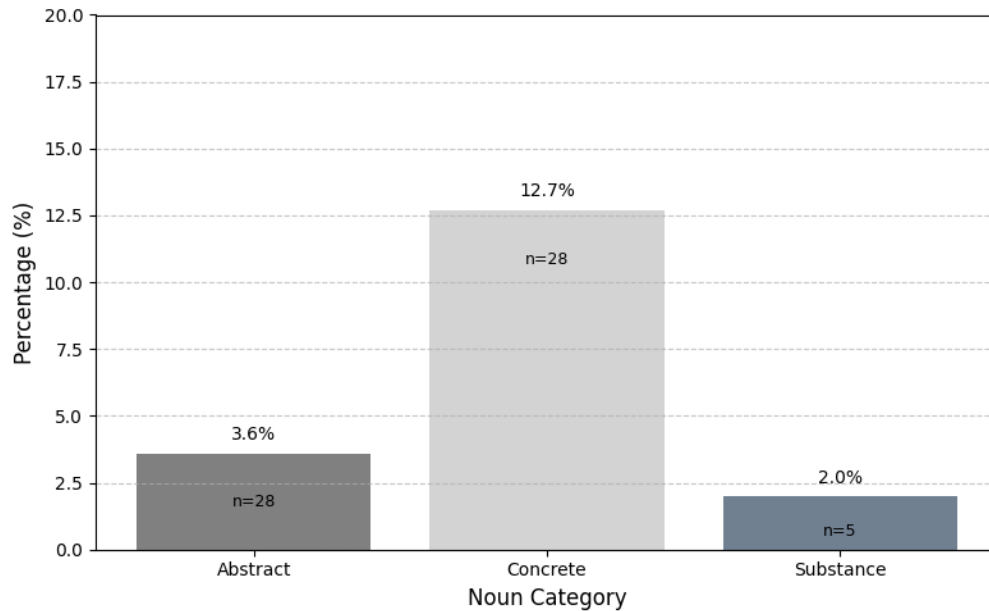


**Figure 4:** Comparison of plural forms of Uncountable vs. Substance nouns

This difference can be attributed to semantic differentiation and grammatical flexibility. When talking about countability being a scalar attribute, substance nouns represent its near end since they represent entities that are seen as a cohesive whole or an undifferentiated mass. For example, *gold* refers to the substance in general, not individual units of blood. Since they describe an unbounded

quantity, pluralizing them would contradict their core meaning. Our ME data supports this statement as the pluralisation within substance nouns nears zero value.

Using the non-parametric Mann-Whitney U test, which is better suited for our small sample size and its non-normal distribution, the difference between U nouns and substance nouns does not provide statistically significant results with  $p = 0.2634$  ( $\alpha = 0,05$ ).



*Figure 5: Percentage of plural forms by semantic noun category*

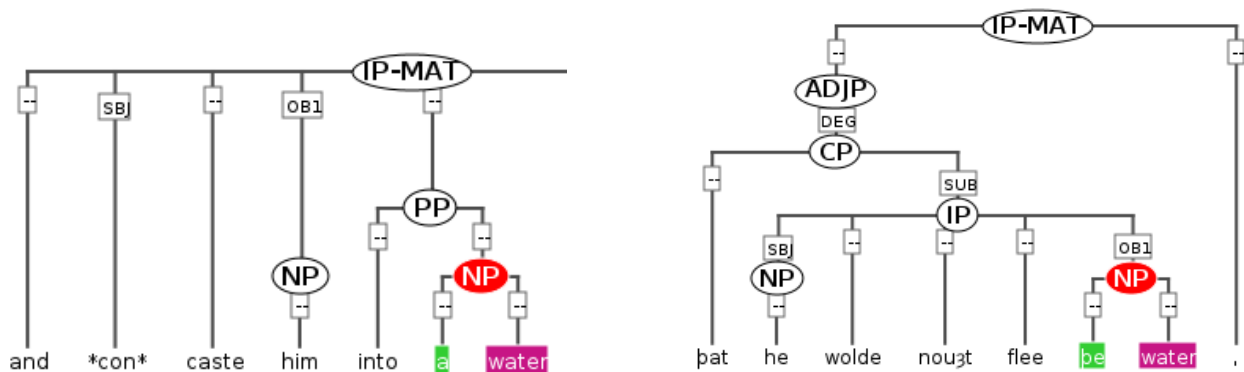
In contrast, **Figure 5** illustrates the differences between percentage values of plural forms between abstract, concrete and substance nouns. Concrete nouns show the highest tendency to form plurals (12.7%), followed by Abstract nouns (3.6%), and Substance nouns (2.0%). This trend can be attributed to several factors. Concrete nouns, representing tangible objects, are inherently countable and more likely to occur in plural forms. Abstract nouns, often referring to intangible concepts, have a limited potential for pluralization. Substance nouns, typically treated as indivisible masses, are least likely to be pluralized.

However, the graph has limitations. The smaller sample size (5) for Substance nouns limits the generalizability of the findings for this category. The Mann-Whitney U test implies that the observed differences in the proportion of plural forms between these two noun categories could be due to random variation rather than an underlying systematic difference with  $p = 0.107$  ( $\alpha = 0.05$ ).

Nonetheless, the p-value is relatively close to 0.1 which may imply a weak significance or a trend worth further investigation. Thus, while the data does not provide convincing evidence of a difference, it may still suggest a potential area for further exploration, particularly with a larger dataset or refined analysis methods.

## 5.1. Determiners

Another possible countability marker analysed in this thesis was noun determination. This part aimed to answer our second research question: To what extent do determiners in ME correlate with the countability status of nouns? Specifically, whether uncountable nouns in ME allow predetermination by an indefinite article, which in PDE is considered an ungrammatical combination. The main aim of this part was to confirm whether this rule has already been established in ME.



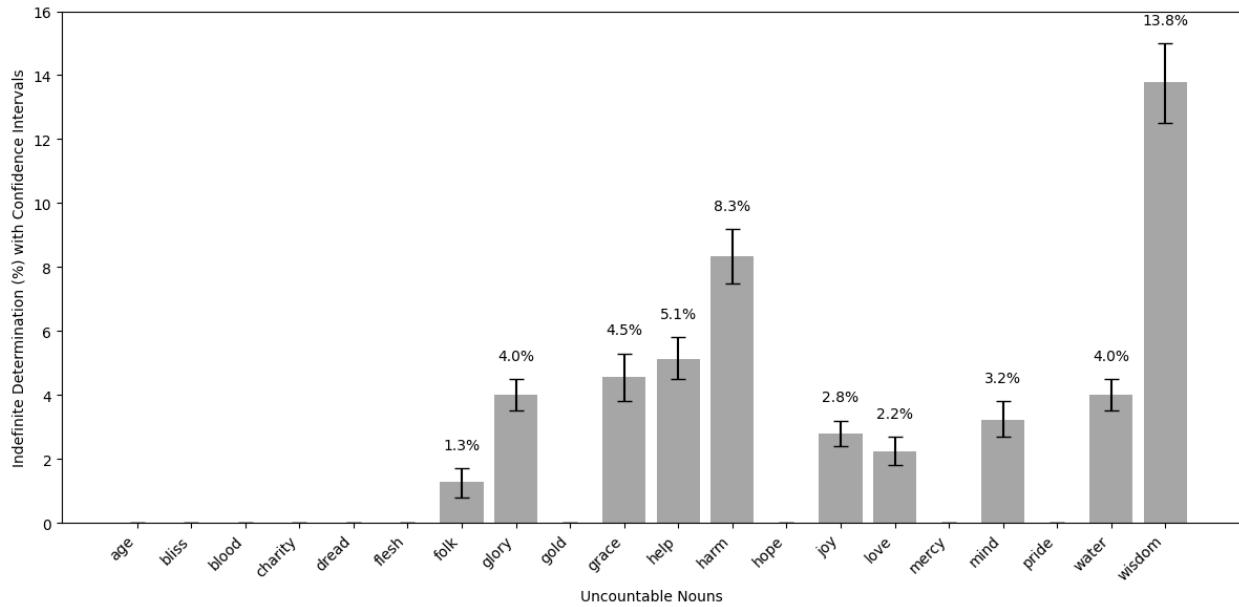
**Figure 6:** Determinant + 'water' combination visualised in a syntactic tree (ANNIS)

**Figure 6** illustrates the syntactical structure of a noun phrase where a U noun, in this case, *water*, is predetermined both by indefinite *a* and definite *the* articles:

- a) and caste him into a water
- b) þat he wolde nouȝt flee þe water

Both examples pertain to a singular, uncountable substance; however, the first example allows the use of an indefinite article, which is something we would not expect in PDE based on contemporary grammar.

**Figure 7** illustrates the percentage occurrence of U nouns in ME that appeared with an indefinite article. Each bar represents a specific U noun, with its respective percentage labelled above the bar. As can be seen from the bar chart, a substantial number of words do not occur with an indefinite determinant at all. Those are *age*, *bliss*, *blood*, *charity*, *dread*, *flesh*, *gold*, *hope*, *mercy*, and *pride*. On the other hand, *wisdom*, with 13.8%, was the most common occurrence of a U noun with indefinite determination by *a/an* article.



**Figure 7:** Percentage of U nouns predetermined by an indefinite article (a/an)

It can be observed that while the uncountable noun *folk* allowed pluralization in 13.1% of cases on average, it did not permit predetermination by an indefinite article. This may suggest that the countability preferences in ME had not yet been fully standardised and fluctuated in terms of regularity. This behaviour would align with the development of the English language during the 12<sup>th</sup> to 15<sup>th</sup> century when standardisation and grammatical rules were slowly being employed.

Nouns such as *grace* (4.55%) and *water* (4.02%) occur frequently in the corpus but exhibit low indefinite article usage, reinforcing the rarity of this construction for U nouns. Similarly, some nouns like *bliss*, *flesh*, *charity*, and *blood* show no instances of predetermination by an indefinite article, aligning with PDE norms.

As previously mentioned, *wisdom* most commonly appeared with an indefinite article, precisely four times. The recorded instances in PPCME2 are:

- i. Prudence after him is a wisdom to lede man that he erre not [...]  
(Richard Fitzjames' *Sermo die Lune*)
- ii. y e wysdome of the worlde myghte neuer lede the hauer [...]  
(Richard Fitzjames' *Sermo die Lune*)
- iii. But bi þe þridde is helid and ʒoten in-to þe soule sich a wisdom and sich a deep knowynge [...] (Hilton's *Eight Chapters on Perfection*)

- iv. so that sapience is a wysdome to lede man y t he erre not from his ende naturally after Aristotle [...] (Richard Fitzjames' *Sermo die Lune*)

A closer look at the examples shows that despite the word being the most frequent example in our data, it may also be an idiosyncratic feature of Richard Fitzjames's writing. Three out of four instances come from his work *Sermo die lune in ebdomada Pasche* from 1495. This concentration within a single author's text suggests that the occurrence of the word with an indefinite article might not reflect a broader linguistic trend. Indeed, the rest of the data agree that uncountable nouns in Middle English were not traditionally predetermined by an indefinite article. Thus, the frequency of this specific example should be interpreted with caution, as it likely represents an individual stylistic choice or scribal practices rather than a disruption to the overall findings and grammatical norms.

The data below show the determinant distribution between indefinite (I) and definite (D) articles for the word *wisdom*. The percentage value in the last column was calculated for indefinite articles from the total numbers of recorded determiners in PPCME2, not limited to articles:

PDE word	det. type	article	ME word	freq.	i.p.m.	sum	Det. sum <sup>3</sup>	perc.
<i>wisdom</i>	D	þe	wisdam	1	0,74		29	13,79%
	D	the	wisdom	1	0,74			
	D	þe	wysdam	1	0,74			
	D	þe	wysedome	1	0,74			
	D	þe	wysdom	3	2,21	18		
	D	þe	wysdome	1	0,74			
	D	The	wysdome	1	0,74			
	D	þe	wisdom	7	5,15			
	D	the	wysdome	1	0,74			
	D	þene	wisdom	1	0,74			
	I	a	wysdome	1	0,74			
	I	a	wisdom	2	1,47	4		
	I	e	wysdome	1	0,74			

<sup>3</sup> This value refers to the total number of recorded determiners in PPCME2, not limited to articles.

Similarly, the word *joy* offers an example where *a* precedes *joy* and it comes from *Mirk's Festial* originated around 1420-1500, again highlighting structure considered scarce in PDE:

- i. Then, when our lady herd þys word, anon þer come such a spirituall swetness and a joy yn hor hert, þat non erþly mon couþ tell hit.

The rest of the recorded instances occurred with definite article:

PDE word	det. type	article	ME word	freq.	i.p.m.	sum	Det. sum	perc.
<i>joy</i>	I	a	joy	1	0,74	1	36	2,78%
	D	the	joye	3	2,21	35		
	D	The	joye	1	0,74			
	D	þe	joy	23	16,94			
	D	þe	joye	8	5,89			

For the majority of uncountable nouns, ME usage aligns with PDE norms, as they are rarely if ever, preceded by an indefinite article. This supports the hypothesis that the grammatical rules governing countability were largely established in Middle English.

The subsequent statistical analysis found important patterns in how definite and indefinite articles were used with presumed U nouns in ME. The Chi-Square Test revealed a statistically significant association between article type and noun ( $\chi^2 = 44.39$ ,  $p = 0.0008$ ), indicating that certain nouns are significantly more likely to occur with either definite or indefinite articles, and this is not due to random variation. Fisher's Exact Test, applied individually to each noun, provided more detailed insights. For most nouns, the p-values were above the set threshold of significance ( $p > 0.05$ ), suggesting no strong preference for either article type. However, two nouns, *charity* ( $p = 0.004$ ) and *wisdom* ( $p = 0.017$ ), showed a statistically significant preference for definite articles.

Unfortunately, sparse data was a major limitation for some nouns as many of them had zero occurrences of one article type. This led to infinite Fisher statistics which strongly prevents the interpretation. Overall, we may assume that in ME, nouns do exhibit a preference for a particular article type.

## 6. Discussion

This MA thesis aimed to map the linguistic markers of the countability category in ME. Although there has been an extensive discussion about the semantic and psycholinguistic aspects of countability, its historical development remains under-researched. This thesis focused on tracing how this linguistic category developed during ME. This was done through a corpus analysis of 60 nouns ranging across 3 different countability categories: countable, uncountable and nouns that may occur in both contexts. These nouns appeared among the 200 most common nouns in the PPCME2, our chosen ME corpora. The analysis explored two main countability markers: nominal pluralisation and the use of determiners, more specifically indefinite determiner. Our working hypothesis presumed that the previously mentioned markers would differ from those in PDE. The question of how they may differ was further examined in our research questions. Overall, this thesis set out to investigate whether countability in ME operated in the same way as it does nowadays.

The first part of this work aimed to define countability as a category in PDE and its historical development. From the 1920s, when Jespersen coined the distinction ‘countable’ and ‘uncountable’, the category itself grew from a binary distinction to a group with up to 15 classes (Grimm & Wahlang, 2021). This section consulted multiple PDE grammar books (Quirk, 1985; Greenbaum, 1996; Biber, 1999; Huddleston & Pullum, 2002), to sum up the characteristic linguistic traits of count and mass nouns, focusing on the categories of number, definiteness and quantifiers. Subsequently, the theoretical part discussed the two dominant methods of differentiating countability: the notional and the grammatical approach. While the first one emphasizes the cognitive basis, the latter turns to grammatical cues and finds countability encoding arbitrary. Numerous tests were later presented to define more precise boundaries of such a flexible concept, and their summaries can be found in Table 2.

According to our assumptions, the findings suggest that ME nouns did exhibit different results than would their PDE counterparts. The analysis of pluralization revealed that nouns presumed to belong to the countable category according to their PDE counterpart exhibited a higher average percentage of plural forms (12.15%) compared to assumed uncountable nouns (2.15%) and nouns that can be used in both contexts (8.99%). While each group displayed the expected behaviour of their PDE classification category, the results showed considerable variation, suggesting that the

countable-uncountable distinction had not yet fully taken shape. However, as the corpus-based behaviour of these nouns in PDE was not examined in this study, these findings should be approached cautiously and would benefit from further verification. The observed patterns of ME countability were not consistent enough to confirm that these nouns adhered to established grammatical rules. Instead, the behaviour of countable and uncountable nouns may have been more closely influenced by the real-world properties of the objects they referred to rather than by a well-defined set of grammatical principles governing their classification. Additionally, conducted statistical tests proved many differences between the groups to be significant and thus worthy of further research.

The semantic analysis further revealed that concrete nouns (12.7%) had a higher tendency to form plurals than abstract (3.6%) and substance nouns (2.0%). This finding is in accordance with the literature review, which states that substance nouns are among the least likely noun category to be counted or pluralized. Thus, the lower rate of pluralization for abstract and substance nouns in ME highlights their uncountable nature, a feature widely recognized in studies of noun countability in PDE.

The second part of the analysis explored the extent to which determiners in ME correlate with the countability status of nouns. The chosen methodology examined the use of indefinite articles *a/an* with uncountable nouns, a construction which is considered ungrammatical in PDE. The data revealed that a small percentage of uncountable nouns in ME did occur with indefinite articles, suggesting that the strict grammatical rule of PDE was not fully established in ME. The analysis also showed that certain nouns, such as *wisdom*, exhibited a higher percentage of indefinite determination (13.8%), while others, such as *blood* or *gold*, did not occur with indefinite determiners at all. The statistical tests showed a significant relationship between the type of article used and the noun in ME, with certain nouns more likely to appear with either definite or indefinite articles ( $p = 0.0008$ ), indicating that the patterns were not due to chance.

The overarching hypothesis of this thesis was that the linguistic markers of countability in ME would differ from those in PDE. The analysis of nominal pluralization and the use of indefinite article provided some evidence supporting this hypothesis. Nonetheless, our ME data showed a higher degree of variability in the use of plural forms and a less rigid correlation between determiners and countability status. These findings suggest that the countability system in ME was

undergoing a process of grammaticalization, and the strict rules of PDE were not yet fully established.

It is important to acknowledge the factors that could have influenced the results of the data analysis and the limitations of our research. One significant challenge was the vast linguistic variation in the English language during this period. Dialectal differences and the lack of standardisation rules made it difficult to identify consistent patterns and infer generalizable results. To ensure accurate results, including all word forms in the search query is essential and omitting them may distort the subsequent analysis. Presuming we obtained practically all orthographic variations, another challenge was to ensure we omitted any cases of homonymy. The solution to this problem involved manual correction of data and deleting any words that caused semantic ambiguity. While this part proposes a potential for human error, problematic words were carefully reviewed multiple times, consulting their quotations and cross-referencing other sources to ensure accuracy and fully understand the intended meaning.

Regarding future research and recommendations for implementation, it would be beneficial to undergo this analysis on a more substantial number of examples. For the intentions and scope of this MA thesis, we employed 60 examples. Nonetheless, it would be highly useful to find more examples of substance nouns and potentially extend the number of examples in each countability group. This thesis suggests that certain characteristics of countability may hold true, though further research would be needed to confirm or challenge these findings. Additionally, while this thesis mentions dialectal metadata for some exemplary words, it did not analyse how dialects might affect the presumed ME countability categories. The same is true for time periods; the influence of time on countability markers was not explored. This choice was made to focus on the main patterns in ME first, but future research could look into the potential influence of dialects and the specific time periods of ME, as outlined in PPCME2, on countability preferences. Additionally, the research could include comparisons between ME and PDE as well as analyses of the same nouns in OE. In the future, the analysis of countability markers and overall behaviour should be extended to cover the entire history of the English language, providing much-sought insights into its origins.

## 7. Resumé

Tato magisterská práce měla za cíl zmapovat lingvistické ukazatele počitatelnosti ve střední angličtině. Přestože se o sémantických a psycholingvistických aspektech počitatelnosti vedou rozsáhlé diskuse, její historický vývoj zůstává stále nedostatečně prozkoumán. Tato práce se zaměřila na sledování vývoje této kategorie prostřednictvím korpusové analýzy 60 podstatných jmen. Ta byla rozdělena do 3 kategoriích podle jejich postavení v současné angličtině: počitatelná, nepočitatelná a podstatná jména, která se mohou vyskytovat v obou kontextech. Tato podstatná jména byla vybrána, jelikož se objevila mezi 200 nejčastějšími podstatnými jmény v námi zvoleném korpusu PPCME2. Následná analýza zkoumala dva hlavní jevy: nominální pluralizaci a použití determinátorů, konkrétně neurčitého členu. Naše obecná hypotéza předpokládala, že výše zmíněné ukazatele se budou lišit oproti jejich užití v současné angličtině. Následnou otázku, jak se mohou lišit, jsme dále zkoumali skrze výzkumné otázky. Celkově si tato práce kladla za cíl prozkoumat, zda počitatelnost ve střední angličtině fungovala stejným způsobem jako v současnosti.

Teoretická část této práce se snaží vymezit počitatelnost jako kategorii v současné angličtině a také její historický vývoj. Od 20. let 20. století, kdy Otto Jespersen zavedl rozlišení „počitatelné“ a „nepočitatelné“ podstatné jméno, se samotná kategorie rozrostla z binárního rozlišení na skupinu s až 15 třídami (Grimm & Wahlang, 2021). V této části byly konzultovány četné gramatické příručky moderní angličtiny (Quirk, 1985; Greenbaum, 1996; Biber, 1999; Huddleston & Pullum, 2002) a byly shrnuty charakteristické jazykové rysy počitatelných a nepočitatelných substantiv se zaměřením na kategorie čísla, určitosti a také kvantifikátorů. Později jsou v Kapitole 2 rozebrány testy substantivní počitatelnosti, které mají za snahu vymezit přesnější hranice této flexibilní kategorie.

Kapitola 3 popisuje, jak se během přechodu od staré ke střední angličtině měnily tři hlavní složky počitatelnosti: determinátory, kvantifikátory a kategorie čísla. Po obsazení Normany v roce 1066 začala angličtina méně využívat pádové koncovky a více se spoléhala na slovosled a pomocná slovesa. Mezi hlavní změny patřilo zjednodušení množných tvarů, kdy většina slov začala používat koncovku *-es*, a zánik gramatického rodu. Určitý člen *the* se sjednotil do jedné podoby a neurčitý člen *an* se stal běžným pro označení počitatelných podstatných jmen.

Kvantifikátory jako *some* a *any* získaly jasnější funkci při označování počitatelnosti. Následující kapitoly se věnovaly praktické části této diplomové práce a její metodologii.

V tomto výzkumu byla zvolena metodologie kombinující kvantitativní a kvalitativní analýzu zaměřenou na vzorce a pravidla užívání pluralizace a determinátorů u podstatných jmen s různou preferencí počitatelnosti. Analýza vycházela z dat korpusu Penn-Helsinki Parsed Corpus of Middle English, druhé edice (PPCME2). Kvůli výrazné ortografické variabilitě ve střední angličtině bylo nutné využít regulární výrazy pro zachycení všech relevantních tvarů slov. Vybraná podstatná jména byla rozdělena podle jejich počitatelnosti v současné angličtině a dále zkoumána z hlediska frekvence výskytu v množném čísle a následně také použití neurčitých členů. Statistické testy pak ověřily, zda se pozorované rozdíly mezi jednotlivými kategoriemi podstatných jmen ukázaly jako významné.

V souladu s naší hypotézou výsledky naznačují, že středoanglická podstatná jména skutečně vykazovala jiné výsledky než jejich současné ekvivalenty. Analýza nominální pluralizace odhalila, že podstatná jména, u nichž se předpokládá, že podle svého současného protějšku patří do kategorie počitatelných, vykazují vyšší průměrné procento plurálových tvarů (12,15 %) ve srovnání s předpokládanými nepočitatelnými substantivy (2,15 %) a substantivy, která lze použít v obou kontextech (8,99 %). Přestože každá skupina vykazovala chování odpovídající své klasifikační kategorii v současné angličtině, výsledky ukázaly výrazné odchylky, což naznačuje, že rozlišení mezi počitatelností a nepočitatelností nebylo ve střední angličtině plně ustálené. Pozorované vzorce počitatelnosti nebyly dostatečně konzistentní, aby potvrdily, že tato podstatná jména dodržují určitá zavedená gramatická pravidla. Provedené statistické testy navíc prokázaly, že mnohé rozdíly mezi skupinami jsou statisticky signifikantní, a tudíž hodné dalšího zkoumání. Sémantická analýza dále ukázala, že konkrétní podstatná jména (12,7 %) měla vyšší tendenci tvořit plurál než abstraktní jména (3,6 %) a substance (2,0 %). Toto zjištění je v souladu s literaturou v teoretické části, která často zmiňuje látková podstatná jména jakožto příklad silně nepočitatelných jmen.

Druhá část analýzy zkoumala, do jaké míry determinátory ve střední angličtině korelují s kategorií počitatelnosti. Analýza zkoumala užití neurčitých členů *a/an* s nepočitatelnými podstatnými jmény, což je konstrukce, která je v současné angličtině považována za

negramatickou. Výsledky ukázaly, že malé procento nepočitatelných podstatných jmen se ve střední angličtině skutečně vyskytuje s neurčitými členy, což naznačuje, že ve střední angličtině ještě pravděpodobně nebylo ustáleno toto striktní gramatické pravidlo. Analýza také ukázala, že některá podstatná jména jako například *wisdom* měla vyšší procento predeterminace neurčitým členem (13,8 %), zatímco jiná, například *blood* nebo *gold*, se s nimi nevyskytovala vůbec. Statistické testy ukázaly signifikantní vztah mezi typem použitého členu a podstatným jménem ( $p = 0,0008$ ), což naznačuje, že tyto vzorce nebyly způsobeny náhodou.

Tato práce vycházela z hypotézy, že lingvistické ukazatele počitatelnosti ve střední angličtině se budou lišit od těch v současné angličtině. Analýza nominální pluralizace a použití neurčitého členu poskytla určité důkazy na podporu této hypotézy. Data z PPCME2 korpusu však ukázala větší variabilitu ve využití plurálních forem a slabší vazbu mezi determinátory a statutem počitatelnosti. Zjištění naznačují, že systém počitatelnosti ve střední angličtině byl stále v procesu gramatikalizace a striktní pravidla známá ze současné angličtiny ještě nebyla ustálená.

Je důležité zmínit faktory, které mohly ovlivnit výsledky analyzovaných dat, a následná omezení této práce. Jedním z významných problémů byla obrovská ortografická variabilita angličtiny v tomto období. Dialektální rozdíly a absence standardizačních pravidel komplikovaly identifikaci jednotných vzorců a zobecnění výsledků. Pro dosažení přesných závěrů bylo nutné zahrnout do vyhledávacích dotazů všechny možné slovní tvary, protože jejich opomenutí by mohlo ovlivnit přesnost analýzy. Po snaze zahrnout co nejvíce ortografických variant se objevila další výzva – odstranění případů homonymie. Tento problém jsme řešili ruční úpravou dat a vyřazením všech slov s potenciální sémantickou nejednoznačností. Přestože tento krok zahrnoval riziko lidské chyby, problematická slova byla opakovaně pečlivě kontrolována, porovnávána s citovanými zdroji a ověřována v dalších referenčních materiálech, aby byla zajištěna maximální přesnost a správné pochopení významu.

Co se týče doporučení pro budoucí výzkum, bylo by přínosné tuto analýzu podstoupit s výrazně vyšším počtem příkladů. Pro záměry a rozsah této magisterské práce bylo použito 60 substantiv. Výsledky naznačují, že určité charakteristiky počitatelnosti v současné angličtině mohou platit i v té střední, ačkoli pro jakoukoliv generalizaci by bylo potřeba získání dalších příkladů, které by případná zjištění potvrdily nebo zpochybnily. V neposlední řadě tato práce sice uvádí dialektální

metadata u dvou příkladových slov, ale detailněji nezkoumala, jak tyto dialekty mohly ovlivnit předpokládané kategorie počitatelnosti ve střední angličtině. Stejně tak se nezabývala vlivem časového období. Nicméně budoucí výzkum by mohl zkoumat potenciální vliv dialektů a konkrétních časových období střední angličtiny, jak je uvedeno v PPCME2, na preference v počitatelnosti. Dále by výzkum mohl zahrnovat srovnání mezi střední a současnou angličtinou, stejně jako srovnání stejných podstatných jmen ve staré angličtině. Následné rozšíření analýzy lingvistických ukazatelů počitatelnosti na celou historii vývoje anglického jazyka by mohlo přinést cenné poznatky o jejích počátcích.

## References

### Primary sources

Krause, T., & Zeldes, A. (2016). ANNIS3: A new architecture for generic corpus query and visualization. *Digital Scholarship in the Humanities*, 31(1), 118-127.

<https://doi.org/10.1093/llc/fqv054>

Kroch, A. and A. Taylor. (2000). *Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*. Department of Linguistics, University of Pennsylvania.

<https://www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4>.

Machálek, T. (2014): *KonText – aplikace pro práci s jazykovými korpusy*. FF UK, Praha.

Available online from: <http://kontext.korpus.cz>

McSparran, Frances et al. (2000-2018). *Middle English Compendium*, Ann Arbor: University of Michigan Library. Available online from <http://quod.lib.umich.edu/m/middle-english-dictionary>

*Oxford English Dictionary Online*. (2023). Oxford: Oxford University Press. Available online from <http://www.oed.com>.

### Secondary sources

Allan, K. (1980). Nouns and Countability. *Language*, 56(3), 541–567.

Allen, C. (2017). Middle English. In *Oxford Research Encyclopedia of Linguistics*.

<https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-263>.

Barber, C. (1993). *The English Language: A Historical Introduction*. Cambridge, Cambridge University Press.

Baber, C. (2000). *An introduction to Middle English*. Cambridge: Cambridge University Press.

Berg, T., & Neubauer, M. (2014). From unit-and-ten to ten-before-unit order in the history of English numerals. *Language Variation and Change*, 26(1), 21–43.

doi:10.1017/S0954394513000203

Biber, D., K. M. Broussard, S. Johansson, G. Leech, S. Conrad, and E. Finegan. (2000).

*Longman Grammar of Spoken and Written English*. London: Longman.

Blake, N. (Ed.) (1992). *The Cambridge History of the English Language* (Vol. 2). Cambridge:

- Cambridge University Press.
- Bloomfield, L. (1933) *Language*. New York: H. Holt and Company.
- Bloomfield, M. & Eilbott, B. A. (1957). "A Diachronic Approach to Lexical Number: Middle and Modern English". *American Speech* 32. 170–5.
- Brinton, L., and Bergs, A., eds. (2017). *Middle English*. Berlin/Boston: De Gruyter.
- Brunner, K. (1965). *An Outline of Middle English Grammar*. Oxford: Basil Blackwell.
- Burrow, J. A., and Turville-Petre, T., A. (1992). *Book of Middle English*, Oxford: Blackwell.
- Contini-Morava, E., & Kilarski, M. (2013). Functions of nominal classification. *Language Sciences*, 40, 263–299. <https://doi.org/10.1016/j.langsci.2013.03.002>
- Corrie, M. (2006). Middle English – dialect and diversity. In Mugglestone, L. (Ed.), *The Oxford History of English*, (pp. 86-119). Oxford: Oxford University Press.
- De Belder, M. (2011). A morphosyntactic decomposition of countability in Germanic. *Journal of Comparative Germanic Linguistics*, 14(3), 173–202. <https://doi.org/10.1007/s10828-011-9045-0>
- Denison, D. (1998). Syntax. In: Romaine, S. (ed), *The Cambridge history of the English language*. Vol. 4, 1776–1997. Cambridge: Cambridge University Press.
- Doetjes, J. (2012). Count/mass distinctions across languages. In: Maienborn, C., K. von Heusinger and P. Portner (eds) *Semantics: International Handbook of Natural Language Meaning* Vol. 3, 2559–2580. Berlin: De Gruyter Mouton
- Dryer, M. S. (2013). Indefinite Articles. In: Dryer, M. S. & Haspelmath, M. (eds.) *WALS Online*
- Drożdż, G. (2017). *The puzzle of (un)countability in English: A study in cognitive grammar*. University of Silesia in Katowice. Retrieved from <https://www.researchgate.net/publication/329340274>
- Dušková, Libuše et al. (2009) *Mluvnice současné angličtiny na pozadí češtiny*. Electronic edition. Ústav anglického jazyka a didaktiky, FF UK. Available at <http://www.mluvniceanglictiny.cz>
- Eastwood, J. (2005). *Oxford learner's grammar*. Oxford: Oxford University Press.
- Filip, H. (2021). *Countability in Natural Language*. Cambridge: Cambridge University Press,
- Fischer, O. (1992) Syntax. In: Blake, N. (ed), *The Cambridge History of the English Language*. Volume II. 1066–1476, 207–408. Cambridge: Cambridge University Press
- Gianollo, C., von Heusinger, K., & Napoli, M. (Eds.). (2021). Determiners and

- Quantifiers. Leiden, The Netherlands: Brill. <https://doi.org/10.1163/9789004473324>
- Gil, D. (2013). Numeral Classifiers. In: Dryer, M. S. & Haspelmath, M. (eds.) *WALS Online*
- Gleason, H.A., Jr. (1955). *An introduction to descriptive linguistics* (Rev. ed.). Holt, Rinehart and Winston: New York.
- Greenbaum, S. (1996) *Oxford English Grammar*. Oxford University Press.
- Grimm, S. and Wahlang, A. (2021). Determining Countability Classes. In: Kiss, T., F. J. Pelletier and Husic (eds) *Things and Stuff: The Semantics of the Count-Mass Distinction*, 357-376. Cambridge: Cambridge University Press.
- Hogg, R., and Denison, D. (2006). *A History of the English Language*. Cambridge: Cambridge University Press.
- Huddleston, R. (1984). *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Horobin, S. (2002). *An Introduction to Middle English*. Edinburgh: Edinburgh University Press.
- Jespersen, O. (1924). *The Philosophy Of Grammar*. London: George Allen & Unwin Ltd.
- Joosten, F. (2003). Accounts of the count-mass distinction: A critical survey. *Nordlyd*, 31(1), 216–229. Proceedings of the 19th Scandinavian Conference of Linguistics. Edited by A. Dahl, K. Bentzen, & P. Svenonius.
- Kiss, T., Pelletier, F. J., Husic, H., Simunic, R. N., & Poppek, J. M. (2016). A sense-based lexicon of count and mass expressions: The Bochum English Countability Lexicon. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2810–2814). Portorož, Slovenia: European Language Resources Association (ELRA).
- Kodera, M. (2020). An overview of the count/noncount distinction of English nouns. *阪南論集 人文・自然科学編*, 55(2), 1–41.
- Lass, R. (1992). "PHONOLOGY AND MORPHOLOGY." Chapter. In *The Cambridge History of the English Language*, edited by Norman Blake, 2:23–155. Cambridge: Cambridge University Press. doi:10.1017/CHOL9780521264754.003.
- Lasersohn, P. (2011). “Mass Nouns and Plurals,” in Claudia Maienborn, Klaus von Heusinger

- and Paul Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning*, Volume 2, Berlin: de Gruyter, 1131-1153.
- Ledgeway, A., & Roberts, I. (2017). Introduction. In A. Ledgeway & I. Roberts (Eds.), *The Cambridge Handbook of Historical Syntax*, *Cambridge Handbooks in Language and Linguistics* (pp. 1–4). chapter, Cambridge: Cambridge University Press.
- Lin, J., Hacoen, A., & Schaeffer, J. (2018). The interpretation of the mass-count distinction across languages and populations: Introduction. *Glossa: A Journal of General Linguistics*, 3(1), 70. <https://doi.org/10.5334/gjgl.638>
- Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511605789>
- Marckwardt, A. H. (2019). *Much and many: The historical development of a Modern English distributional pattern*. In Rosier, J.L. (Ed.), *Philological Essays: Studies in Old and Middle English Language and Literature in Honour of Herbert Dean Meritt*. 50–54. De Gruyter Mouton.
- McArthur, T. (Eds.). (1992). *The Oxford Companion to the English Language*. Oxford: Oxford University Press.
- Nevalainen, T., & van Ostade, I. T.-B. (2006). Standardisation. In R. Hogg & D. Denison (Eds.), *A History of the English Language* (pp. 271–311). chapter, Cambridge: Cambridge University Press.
- Quirk, R., S. Greenbaum, G. Leech and J. A. Svartvik (1985). *Comprehensive Grammar of the English Language*. London: Longman.
- Radden, G., & Dirven, R. (2007). *Cognitive English Grammar*. Amsterdam/Philadelphia: John Benjamins.
- Santorini, Beatrice. 2005. Annotation manual for the PPCME2, PPCEME, and PCEEC. <http://www.ling.upenn.edu/hist-corpora/annotation/index.htm>.
- Sommerer, L. and Hofmann, K. (2021). "Constructional Competition and Network Reconfiguration: Investigating Sum(e) in Old, Middle and Early Modern English." *English Language and Linguistics* 25, no. 1, 1–33.
- Srinivasan, M., & Barner, D. (2020). Lexical, syntactic, and pragmatic sources of countability: An experimental exploration of the mass-count distinction. In F. Moltmann & L. Tovena (Eds.), *Mass and Count in Linguistics, Philosophy and Cognitive Science*.

- Stark, E. (2008). "The Role of the Plural System in Romance." In *The Paradox of Grammatical Change. Perspectives from Romance*, 57–84. Amsterdam, Philadelphia: John Benjamins.
- Tichý, O. (2022). Countability in the history of English: Evidence from grammars and dictionaries. *Linguistica Pragensia*, 32, 39–58.  
<https://doi.org/10.14712/18059635.2022.1.3>
- Townend, M. O. (2006). Contacts and Conflicts: Latin, Norse, and French. In L. Mugglestone (Ed.), *The Oxford History of English* (pp. 61-85). Oxford University Press.
- Vandiver, W. R. (2011). *A semantic analysis of quantification in English*. Lafayette, IN: Purdue University doctoral dissertation.
- Vermote, T. , P. Lauwers and L. De Cuypere. (2017). Transcending the Lexical vs. Grammatical Divide regarding the Mass/Count Distinction. Evidence from Corpus Studies and Acceptability Surveys in French and Dutch. *Language Sciences* 62: 37–51.  
<https://doi.org/10.1016/j.langsci.2017.02.002> [Google Scholar]
- Wierzbicka, A. (1988). *The semantics of grammar*. Amsterdam: John Benjamins.

## Appendix

### CQL queries

#### PDE countable nouns:

**book** - [word="b[ōo][ck](e|es)?" | word="b[ua]c(e|es)" | word="bowk(e|es)"|word="b[ēo]c" | word="bōken"|word="book"]

**castle** - [word="cast[e,ē]l\*[e]." | word="castl.\*" | word="kestel.\*" | word="chastel." | word="castel.\*"]

**child** - [word="child" |word="childe" | word="cild" | word="chil" | word="shild[e]" | word="sheld" | word="ched" | word="childre" | word="childrer?" | word="childr[ei]n" | word="childr[ei]n(g|e|o)s?" | word="chelder" | word="chulder" | word="childres"]

**day** - [word="da[i,y,ʒ,g]es"|word="dau"|word="daʒhes"]

**earl** - [word="ē?r[l].\*" | word="eerl.\*" | word="erel.\*" | word="eorl.\*" | word="[ou]rl.\*" | word="yerl.\*" | word="ʒorl.\*"]

**emperor** - [word="[e,i,au,a]mperour" | word="[e,i,au,a]mperour.." | word="emparour.." | word="emprour.."]

**end** - [word="[ēeiʒ]nd(e|es)?" | word="[ēeiʒ]in(de|des)"]

**father** - [word="f[āa]der(e|es|en)?" | word="fadres" | word="[vf][ea]der" | word="father"]

**god** - [word="g[ōo]d(e)" |word="god(de|es)?" | word="goid" | word="guode" | word="cod" | word="got" | word="goð" | word="god[en]his|edes|æs"]

**heart** - [word="hert(e|es)"|word="heort(e|æ|an|ne|en)"]

**horse** - [word="hor[cs]\*(es|en)?" | word="horis(es|en)?" | word="ors(es|en)?" | word="ho[rs](es|en)?" | word="heors(e|es|en)?" | word="horses?"]

**house** - [word="hou[cs](es|en)?"|word="housse(s|en)?" | word="ouese(s|en)?" | word="h[wu]s(es|en)?" | word="husse(s|en)?" | word="huis(es|en)?" | word="ouse(s|en)?" | word="hou(s|ses|sen)?" | word="husa(s|es|en)?" | word="how(e|es|en)?" | word="houses?"|word="hūs[as|um|an]" | word="hūsæ(n)?"]

**kingdom** - [word="[k,c][i,y][un]gdom.\*" |word="[k,c][i,y]ng[is|h]e]dom.\*" |word="[k,c][i,y]ngd[am|an|on].\*" |word="[k,c][i,y]ngham.\*" | word="[k,c][i,y]ngdams.\*" |word="[k,c][i,y]ngdome[ns].\*" |word="[k,c][i,y]ngdans.\*" | word="[k,c][i,y]ngghams.\*" |word="[k,c][i,y]engdams.\*" | word="[k,c][i,y]ugd[oa]ms.\*"]

**mother** - [word="mōd[ra|er]"|word="mod(e|i|u|the)r?"|word="m[au]der" |word="mother(e|ir)?"|word="muther"|word="modern"|word="mōderes"|word="modre[ns]"]

**name** - [word="n[aā]m(e|a)?"|word="n[aā]m[eæ][sn]"|word="n[oō]m(e|a)?"  
|word="n[oō]m[eæ]n"|word="n[oō]m[eæ]s"]

**place** - [word="pla[cs]h[e]s[en]"|word="plais[es|en]"| word="plecis"  
word="plase"|word="place[ns]"]

**pope** - [word="p[ōoa]p(e|es)?" | word="pwop(e|es)" | word="pāp(e|es)?" | word="pap[ae](n|es)?"  
| word="popa"]

**temple** - [word="temple(s|en)?" | word="temppl(e|s|en)?" |word="temp[ieu]l(e|es|en)?"  
|word="tempille(s|en)?" | word="tempul(le|les|len)?" |word="templ[en|ee|id|us]"  
word="temples?"]

**woman** - [word="w[ooiuea]m{1,2}[en|ene|es|enne]"| word="w[ooiuea]m{1,2}en(ne)?"  
word="w[ooiuea]fmen" | word="w[ooiuea]fmon(en|ne)?" |word="w[ooiuea]fman(ne)?" |  
word="w[ooiuea]mm[ooiuea]ne?" |word="w[hvw][oiuea]m{1,2}[en|enne]"  
|word="w[hvw][oiuea]m{1,2}e(s|n)?" | word="w[ooiuea]m{1,2}on(ne)?"|word="woman"]

**world** - [word="w[o,u]rld"|word="w[o,u]rld.."|word="wor[e,u]ld[e,æ][s,n]"|word="wor[e,u]ld"]

#### PDE uncountable nouns:

**age** - [word="age" |word="a[ay]ge" | word="age[ns]"|word="aage[ns]" | word="ayge[sn]"  
|word="aygen"]

**bliss** -[word="blis(es|en)?" |word="blis[cs]e(s|en)?" | word="blisce(s|en)?" | word="blise(s|en)?"  
|word="blesse(s|en)?" | word="blisses?" |word="blissen"]

**blood** - [word="bl[ōo][d]e?" | word="blo[oi]d"| word="bloud"| word="blud[e]?"  
word="blood[e]?"| word="bloude"]

**charity** - [word="charit[ēe].\*"|word="chearite.\*" | word="cherite.\*" |word="carite.\*" |  
word="kariteð.\*" |word="ch[ae]rte.\*" | word="chierte.\*"]

**dread** - [word="dred(e|es)" |word="dred(e)" | word="dret" | word="dr[ia]de" | word="dridde"  
|word="dredes" |word="dreddes"]

**earth** - [word="[ei3y]?[eo]rthe(s|n)?" | word="[ei3y]?[eo]rte(s|n)?" | word="[ei3y]?[uo]rthe(s|n)?"  
| word="[ei3y]?[uo]rte(s|n)?" ]

**flesh** - [word="fl[ee,eæ,a]sh[e,es,en]" | word="flæs[e,es,en]" | word="flees[e,es,en]"  
|word="flesh[e,es,en]" | word="fless[e,es,en]" |word="vlesh[e,es,en]" | word="vles[e,es,en]" |  
word="fleis(h|s)" | word="vleis[e,es,en]" | word="flish[e,es,en]" |word="flech[e,es,en]" |  
word="vlech[e,es,en]" |word="flexs[e,es,en]"|word="flesh"]

**folk** - [word="[v,f][o,ö,oo]lk" | word="folkes"]

**glory** - [word="glorie" | word="glor[iy]" | word="gloire" | word="gleyre" | word="glor(ie|y|ie)s" | word="gloires" | word="glories" | word="gloren" | word="gloiresen"]

**gold** - [word="gold.\*" | word="glod.." | word="gol.." | word="goud.." | word="gould.." | word="guld.." | word="golden" | word="zolde" | word="golt.\*"]

**grace** - [word="[gc]ra[c,s,z][ie].?"]

**harm** - [word="h[ae]rm(e|es|en)?" | word="h[ae]rm[ei](s|n)?" | word="h[aeæ]rmene" | word="h[aeæ]rmen" | word="harem(e|es|en)?" | word="h[ae][ri]m(e|es|en)?" | word="h[ae]rmome?" | word="herm(e|es|en)?"]

**help** - [word="h[aei]lp(e|es|en)?" | word="helpe(s|n)?" | word="healpe(s|n)?" | word="holp(e|es|en)?"]

**hope** - [word="h[o,ō]pe" | word="hop(pe|e)" | word="hoipe" | word="ope" | word="h[oō]pen" | word="hoppen" | word="hoipen" | word="open" | word="h[oō]pes" | word="hoppe[s]" | word="hoipes" | word="opes"]

**joy** - [word="j[oe][iy](e|æ|e)?" | word="jo[iɜ](e|es|en)?" | word="y[oe][iy](e|es|en)?" | word="ǵ[oiy](e|en|es)?" | word="g[ei](e|es)?" | word="jai" | word="joies" | word="joien"]

**love** - [word="l[uo][fv](e|æ|ve)" | word="l[uo][fv](es|en)" | word="l[uo][fv](e|æ|ve)(s|n)?" | word="luf[fv](e|æ|en|es)" | word="l[ea][fv](e|ve|s|n)?" | word="l[o][fn](e|ve|n)?" | word="low(e|es)" | word="lou[vf](e|es)" | word="lune" | word="lone"]

**mercy** - [word="merc[ie].?" | word="mercy." | word="mercy" | word="mercie." | word="mercize." | word="mersi." | word="marci.\*" | word="marsi." | word="messy." | word="merce." | word="marce." | word="marse." | word="mearci." | word="vercy." | word="versy."]

**mind** - [word="m[iy]nd[e,es,en]" | word="m[iy]nd" | word="miende[s,n]" | word="m[y,e,u]nd[e,es,en]" | word="m[e,a,u]ind[e,es,en]"]

**pride** - [word="prid(es|en)?" | word="priɜd(es|en)?" | word="pr[ui]ede(s|en)?" | word="pritte(s|en)?" | word="pru[dt]e(s|en)?" | word="pr[eu]d[eu](s|en)?" | word="pruten(s|en)?" | word="pruid(es|en)?" | word="pruite(s|en)?" | word="prides?" | word="prudene"]

**water** - [word="w[a]t[e]r(e|es)?" | word="water(r|e|es)?" | word="wat[ui]r(e|es)?" | word="watier(e|es)?" | word="watre" | word="watter(e|es)?" | word="whater" | word="vater" | word="weter(e|es)?" | word="wæter(e|es)?" | word="weater" | word="weattre" | word="wettre" | word="water" | word="wateran" | word="wætera" | word="wattera" | word="wa[ru]ter" | word="watered" | word="water" | word="wartre" | word="wter" | word="wade" | word="water[en|s|es]" | word="watirris" | word="wattres" | word="waterum"]

**wisdom** - [word="w[iy]sd[ouaei]m(e|es|en)?" | word="w[iy]sd[ouaei]m{2}(is)?" | word="w[iy][sz]edom(e|es|en)?" | word="wisdoma" | word="w[hi]sd[ouaei]m(e|es|en)?" | word="v[iy]sd[ouaei]m(e|es|en)?" | word="w[iy]dom(e|es|en)?" ]

PDE both meanings:

**bed** - [word="bed" | word="bedd(e)" | word="bæd(de)" | word="bydde" | word="bede" | word="beds" | word="bedd(es)" | word="bæd(des)" | word="bydden" | word="beden" ]

**church** - [word="ch[u,i,e]rch[en,es]" ]

**country** - [word="c[o]u[ntr[e]æ](e|o)[s|n]" | word="cuntre(e)[s|n]" | word="countre(e)[s|n]" | word="co(u)ntreo[s|n]" | word="contrei(e)[s|n]" | word="contra(i)e[s|n]" | word="co(u)ntreth[s|n]" ]

**desire** - [word="d(e|i|is|es)s(ir|ier|ire|er)(e|es|en)?" | word="des(er|ir|ier|ire)(e|es|en)?" | word="dis(ir|ire|er)(e|es|en)?" | word="diss(ir|ire|er)(e|es|en)?" ]

**gospel** - [word="(gospel|gospelle|gospille|gospel|cospel|godspel|goddspel|godespel|gotspel)" | word="gospel[s|es]" | word="gospelle[s|es]" | word="gospille[s|es]" | word="gospel[s|es]" | word="cospel[s|es]" | word="godspel[s|es]" | word="goddspel[s|es]" | word="godespel[s|es]" | word="gotspel[s|es]" ]

**home** - [word="h[ao]m(e|es|en)?" | word="h[oō]me(s|n)?" | word="ham(e|es|en)?" | word="whom(e|es|en)?" | word="[oa]m(e|es|en)?" | word="hem" | word="heom" | word="h[ae][oi]m(e|es|en)?" | word="hāmen" | word="heomen" ]

**honour** - [word="honour" | word="hon[uo]r" | word="honnour" | word="honowur(es|en)" | word="onour" | word="honours" | word="hon[uo]rs" | word="honnours" | word="honoures" | word="honores" | word="honnoures" ]

**judgement** - [word="[y]ugement(s|es|en)?" | word="jugg[ie](s|es|en)?" | word="juje(s|es|en)?" | word="jeuge(s|es|en)?" | word="jouge(s|es|en)?" | word="goge(s|es|en)?" | word="judgement" | word="judmens" | word="jud(s|es|en)?" | word="juggemet(s|es|en)?" | word="judgemet(s|es|en)?" ]

**kind** - [word="man[ēaei]r(e|es)?" | word="man[ēaei]r[aeiou](e|es)?" | word="man[io]er(e|es)?" | word="moner(e|es)" ]

**land** - [word="l[oō][nlt][d(e|es|en|um)?" | word="land(e|es|en)?" | word="laund(e|es)?" | word="laind" | word="lend" | word="loand" | word="leond" | word="lænde" | word="lon" | word="longde" | word="londen" | word="londum" | word="l[oō]nd(e|es|um)?" | word="-[loun][d](e|es)?" ]

**life** - [word="lif(s|es|en)?" | word="lijf(s|es|en)?" | word="life(s|es|en)?" | word="li[wv]e(s|es|en)?" | word="liffe(s|es|en)?" | word="lifve(s|es|en)?" | word="livf(s|es|en)?" | word="lief(s|es|en)?" ]

|word="lefe(s|es|en)?" | word="leve(s|es|en)?" | word="leif(s|es|en)?" | word="leive(s|es|en)?"  
|word="lifæs" | word="lifæ" | word="lifen" | word="lives" | word="liven"]

**pain** - [word="p[ae]i[nm](e|i|g|en)?" | word="p[ae]in(e|es|en)?" | word="p[ae][ig]ne(e|s|n)?" |  
word="pa[iy]n(e|es|en)?" | word="p[ae]i[eai][gn](e|s|n)?" | word="pe[yi]ne(e|s|n)?" |  
word="pene" | word="p[aoi]n(e|es|en)?" | word="beyn" | word="peimis" | word="peines" |  
word="pey[mi]s" | word="painnens" | word="poynt"]

**passion** - [word="pass(y|i)oun" | word="pas\*(y|i)on" | word="pass(y|i)o" |  
word="pas(y|i)c(y|i)oun" | word="pa(y|i)ss(y|i)oun" | word="pa(y|i)sson" |  
word="pa(y|i)sso(y|i)ne" | word="poss(y|i)one" | word="pass(y|i)um" | word="pass(y|i)c(y|i)on" |  
word="pass(y|i)ons" | word="pass(y|i)ounes" | word="pas(y|i)ons" | word="pass(y|i)os" |  
word="pas(y|i)c(y|i)ons" | word="pa(y|i)ss(y|i)ouns" | word="pa(y|i)ssons" |  
word="pa(y|i)sso(y|i)nes" | word="pass(y|i)c(y|i)ouns"]

**penance** - [word="penaunce" | word="penauns(e)" | word="pen[oa]nce" | word="penans" |  
word="pen[oa]n[sz]" | word="penawun[sc]e" | word="penawince" | word="penanche" |  
word="peinaunce" | word="painance" | word="panans" | word="penaunces" |  
word="penanc[en|es]"]

**power** - [word="pouer(e(s|es|en)?)?" | word="pou[oa]ir(e(s|es|en)?)?" | word="pouare(s|es|en)?"  
|word="pour(e(s|es|en)?)?" | word="puere(s|es|en)?" | word="pouwer(e(s|es|en)?)?" |  
word="p[ou]wer(s|es|en)?" | word="pwower(s|es|en)?" | word="pōer(e(s|es|en)?)?" |  
word="po[aē]r(e(s|es|en)?)?" | word="po[ei]ia[ai]r(s|es|en)?" | word="pore(s|es|en)?" |  
word="poier(s|es|en)?" ]

**reason** - [word="r[eai][s][aoei]\*(n|nn|m|s)?(e|es|en)?" | word="r[ai][s][ouei][n|nn|m|s]?(e|es|en)?"  
| word="resoun(n)?(e|es|en)?" | word="ressone" | word="resun(n)?(e|es|en)?" |  
word="resan(e|es|en)?" | word="resen(e|es|en)?" | word="reasoun(e|es|en)?" |  
word="reason(e|es|en)?" | word="reisoun(e|es|en)?" | word="reissoun" | word="reison(e|es|en)?" |  
word="raisoun(e|es|en)?" | word="raison(e|es|en)?" | word="reisun" | word="reaisun" |  
word="rosoun"]

**spirit** - [word="spir[iy][dt]?e(s|n)?" | word="spir[iy][dt]?h(s|n)?" | word="spir[iy]tus" |  
word="spir[iy]t[iy]s" | word="sp[eiy]r[dt]e(s|n)?" | word="sp[eiy][rdt]h(s|n)?" | word="spirit" |  
word="spriryte" | word="sp[iy]rett(e|es)?" ]

**stone** - [word="st(o,y)ne" | word="st(o,y)nne" | word="st(o,y)ine" | word="stan(e|n)" |  
word="st(a,y)(i,y)n(e|n)" | word="st(o,y)an" | word="stau(n|m)" | word="st(o,y)ou" | word="sto-"  
| word="sta-" | word="s(i,y)tne" | word="st(o,y)wyn" | word="st(o,y)nd" | word="s(o,y)n" |  
word="st(o,y)r" | word="ston(e|es|n|us)" | word="stan(e|en)" | word="st(o,y)anne(s|n)" |  
word="st(a,y)n(a,s)" | word="stanum" | word="stunnes"]

**time** - [word="time(s|en)?" | word="tim(is|es|en)?" | word="teime(s|en)?" | word="teme(s|en)?" | word="timæ(s|en)?" | word="tima(s|en)?" | word="timon(s|en)?" | word="tyne(s|en)?" | word="tome(s|en)?" | word="timeus" | word="timus" | word="timene"]

**word** - [word="w[oō][rðd](e|es|um)?" | word="wort" | word="wourd(e|es|us)?" | word="wurd(e|es|us)?" | word="wrđ" | word="vord" | word="weord" | word="worth" | word="wored" | word="woerdes" | word="word(d|is|us|a|æ|um|en|s)?" | word="wur(d|n|ð|dyes)" | word="wirdes" | word="wurpes" | word="owrdes"]