universität wien

# DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Bare Plural Noun Phrases in Modern English“

verfasst von / submitted by

Veronika Maria Ableidinger

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra der Philosophie (Mag. phil.)

Wien, 2019 / Vienna, 2019

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet: | A 190 313 344 |
| Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet: | Lehramtsstudium<br>UF Geschichte, Sozialkunde, Polit.Bildg.<br>UF Englisch |
| Betreut von / Supervisor: | Univ.-Prof. Mag. Dr. Nikolaus Ritt |

# Acknowledgments

First of all, I am very grateful to my supervisor, Prof Nikolaus Ritt, for his guidance and encouragement. This thesis has benefited greatly from his comments and suggestions on draft chapters.

I am especially indebted to my family for their endless support and patience during the writing process.

Finally, I thank Learose for proofreading and for all the comments and remarks on the present work.

# Table of Contents

# LIST OF TABLES



# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| BP | Bare Plural |
| D | Determiner |
| DP | Determiner Phrase |
| EModE | Early Modern English (1500-1700) |
| LF | Logical Form |
| LModE | Late Modern English (since 1700) |
| ME | Middle English (1100-1500) |
| ModE | Modern English (since 1500) |
| MP | Minimalist Program |
| N | Noun |
| NF | Non-Fiction |
| NP | Noun Phrase |
| OE | Old English (before 1100) |
| OED | Oxford English Dictionary |
| PdE | Present Day English |
| Pl. | Plural |
| pmw | per million words |
| POS | Part of Speech |
| UG | Universal Grammar |

# 1. Introduction

## 1.1. Research question and hypothesis

This thesis focuses on the diachronic development of bare plural noun phrases within the Modern English period. Achieving an adequate analysis of the development of bare plurals (henceforth BPs) is a challenging task since much of the work done over the past decades in generative linguistics has focused on the syntactic structure and semantic interpretation of BPs. Little attention has been paid to their diachronic development, except in connection with the emergence of the determiner system in the English language.

The core question that this study investigates is whether BP noun phrases (NPs) with indefinite reference have decreased or increased in frequency in Modern English (ModE). In view of the development of the determiner system in English, a trend towards increasing use of overt reference markers can be observed. The use of the article *the* to mark definiteness is obligatory in Present Day English (PdE). Likewise, indefinite singular count nouns are not supposed to occur bare, but are to be preceded by the indefinite article *a(n)*. Plural count nouns seem not to have arrived at this stage yet. To convey indefiniteness, weak *some* is used with plural nouns, but the compulsory use of this plural indefinite marker is restricted to specific NPs only. BPs thus continue to be present in the English language.

The increasing obligatoriness of overt reference marking in English lends to the hypothesis that BPs have decreased in frequency, while overt reference marking in plural indefinite NPs has increased over the past 500 years.

## 1.2. Goals and methods

The aim of this thesis is to trace the development of the BP construction diachronically from Early Modern English (EModE) through to present day. Frequencies of occurrence of BPs and plural NPs headed by *some* will be tracked by using text corpora from EModE and Late Modern English (LModE). It will then be investigated if a tendency towards obligatory indefiniteness marking in plural NPs can observed. In order to deal with this question, it will first be necessary to discuss the role of BPs in the English language and outline the development of the English determiner system in a theoretical part.

The corpora used in this thesis include the 400-million-word Corpus of Historical American English (COHA) and the comparatively small Penn-Helsinki Parsed Corpus of Early Modern

English (PPCEME). These two corpora are described in further detail in Chapter 5 and in Appendices I and III. The data is further supplemented with information derived from the Oxford English Dictionary (OED).

The methodology used for the investigation of COHA and PPCEME can be referred to as diachronic corpus linguistics, which is aimed at examining both the spreading and diminishing use of certain language features. Corpus linguistics in general has significantly influenced the study of languages as the increasing availability of varied corpus resources allows for conducting research on the state of a language and on linguistic change. Corpus linguistics as a sub-discipline of linguistics has existed since the 1960s, but the term 'corpus linguistics' only made its appearance in the 1980s. It can be defined as "the study or analysis of language through the use of (computer) corpora" (Leech *et al.* 2009: 24).

The methodology described in this section consists of several steps, beginning with information deriving from PPCEME and COHA and ending with an abstract discussion of the results:

1) **Use of Antconc and the web-based COHA interface to identify the number of plural NPs with and without overt reference marking in PPCEME and COHA and the corresponding subcorpora.** Since the two corpora used in this study have already been Part-of-Speech (POS) tagged (see Appendices II and IV for the respective tagsets), the first step is to determine the total number of plural NPs. The result will probably point to a general trend towards an increased use of nouns in written English. Furthermore, to compare the occurrences of BPs in EModE and LModE, it is necessary to separate plural nouns preceded by overt determiners from BPs and to compare their ratios in PPCEME and COHA.
2) **Create tables to present quantitative findings.** In order to produce meaningful results, frequencies normalized to occurrences per million words (pmw) are compared between the corpora and subcorpora. The term 'subcorpus' refers, on the one hand, to the division of PPCEME into its three sub-periods and, on the other hand, to the division of COHA into decades. Tables, bar charts and circular charts are given to present and discuss the quantitative findings.
3) **Use of Excel spreadsheets to classify findings according to syntactic and semantic criteria.** Since the global numbers do not form sufficient basis to test the hypothesis that plural nouns are increasingly headed by overt indefinite markers, a more in-depth analysis of BPs in ModE is indispensable. For this purpose, a sample of BPs will be

created and analysed according to specific syntactic and semantic aspects. First, however, all fixed expressions involving BPs will be excluded from the count, so that they do not distort the results. Then the development of both argumental and predicative BPs and of indefinite and kind-referring BPs, respectively, will be subject to investigation. In order to determine a possible trend towards obligatory indefiniteness marking, the share of BPs with indefinite reference marking would need to increase between EModE and LModE.

4) **Visualise findings to facilitate further qualitative analysis.** By examining individual examples in PPCEME and COHA according to the categories identified at stage 3 and illustrating the findings in the form of tables and charts, a more detailed understanding of the use of BPs in both corpora will be gained.
5) **Interpretation of findings in the framework of a generative model of language change.** In order to embed the results within a greater context, the development of the determiner system in the history of English will be reconsidered. How the present analysis of BPs fits into the general trend towards obligatory reference marking in English will be discussed.

## 1.3. Approach

The theoretical framework that has been applied in this thesis is referred to as generative grammar, a research program first developed by Noam Chomsky in the 1950s. It serves as an alternative model to behaviourist and structuralist frameworks since the focus has shifted away from existing structures in the language to the mind of the language learner. Generativists claim that input is not a decisive factor in the language learning process because the existing linguistic knowledge of speakers far exceeds the output that language input can ever yield. Naturally, the question arises where linguistic knowledge comes from. The answer to this question is, according to Chomsky, Universal Grammar (UG), the original state of a mental organ which serves to acquire linguistic knowledge and which is referred to as the language faculty. UG is considered to consist of principles (shared by all human languages) and parameters (the aspects according to which languages vary) (Gelderen 2011: 9; Lightfoot 2006) and, consequently, allows the learner to understand the structure of his/her own language in spite of the exposure to a limited amount of language input (Giorgi & Langobardi 1991: 7).

Essentially, a theory of UG should meet four key criteria. First, universal grammar needs to be applicable to all natural languages, thus it must be universally valid. Next, it needs to explain the universal properties of natural languages, an aspect referred to as explanatory adequacy. Furthermore, a theory of UG must be maximally constrained, that is, a theoretical apparatus must be developed, which is only applicable to natural languages in order to distinguish these languages from artificial languages, for instance. The forth criterion is that of learnability, which implies that grammar must be easily learnable by young children (Radford 2004).

Since the very early days of the generative approach to syntax in the 1950s, a number of models of generative grammar have been developed. The most recent theory within generative syntax is the Minimalist Program (MP) (Radford 2003). As this thesis focuses on the NP, it is particularly important to discuss how phrases are structured within the MP. Merge, the primary structure-building operation in the MP, is thus explained in section 2.1.1. Moreover, it is essential to explain the nature of language change and how linguistic change is triggered since this thesis analyses possible syntactic change in plural NPs in the history of English. Here, the generativist framework offers again some concrete ideas on what this process of language change involves (see chapter 4.1.).

## 1.4. Limitations

This study is subject to some obvious limitations. Since this thesis is a corpus-based study, typical problems linked to the domain of Corpus linguistics, such as representativeness and reliability of data, naturally arise.

To start with, the samples used in the present study are restricted in size, in the selection of texts and in genre. The 1.7-million-word Penn-Helsinki Parsed Corpus of Early Modern English is very small compared to the 400-million-word Corpus of Historical American English. Thus, the results of the search of PPCEME may not be as meaningful as the ones of COHA. The selection of texts differs between PPCEME and COHA and can be found in Appendices I and III. As to the genre, both corpora are restricted to written language. However, analysing changes in spoken material would be interesting as well since spoken language changes before the written language does.

Second, a drawback of COHA is linked to geographical aspects. It is an American corpus and thus limited to American English. Since there are differences between American and British

English not only in terms of lexis but also in terms of syntax, including a British Corpus of LModE would deliver important additional information on the subject.

Finally, the two corpora cover only two specific time periods: The period from 1500 to 1710 is compared to the period from 1810 to 2009. Since change appears to happen gradually and can take hundreds of years, it would also be worthwhile to pay attention to long-term patterns in the change of the BP construction. There are established views that the prenominal structure underwent great changes between OE and ME due to the emergence of the article-system and that, at the same time, the number of determinerless NPs decreased significantly. Taking into account the OE and ME period would definitely yield interesting insights in terms of the long-term development of BPs.

In terms of the theoretical part of this thesis, it is important to mention that the literature consulted focuses exclusively on generative grammar. Insights from other theories of grammar have not been included since this would have gone beyond the scope of the present study.

### 1.5. Structure

This thesis consists of two parts. The first looks at some theoretical issues linked to the NP and the theory of language change. Part I contains three chapters, the first of which deals with the syntactic and semantic analysis of the NP. The second chapter focuses on the BP construction from a syntactic and semantic point of view. The third chapter examines what causes language change in general and presents some well-known views on the development of determiners in the English language.

Part II discusses the results of the corpus study. It has two chapters, one on the structure of the two corpora used in the study, the other on the analysis of the results. Finally, the discussion section is concerned with bringing together theory and empiricism. In the conclusion, the outcome of the study is discussed as well as possible shortcomings and ideas for further research.

Finally, the references used in this thesis are listed and several appendices are included, which contain further information about the corpora and the search queries used for analysis.

# PART I

At the outset of the thesis, it is crucial to provide some theory on the English NP and its role in generative syntactic theory in order to frame the analysis provided in the second part of this work. The syntactic structure of NP/DP, as well as some facts about the grammatical category of the determiner (D), will be presented. This is followed by a discussion of the syntax-semantics interface of the BP construction. Finally, the emergence of the English determiner system will be outlined, which provides the basis for the hypothesis formulated at the end of PART I.

## 2. THE NOUN PHRASE FROM A GENERATIVE PERSPECTIVE

Within the generative tradition, a vast amount of literature on the syntactic properties of the English NP is available. Some key issues will be summarised hereafter.

### 2.1. The syntactic structure of NP/DP

An analysis of the NP typically starts with the following question: What is understood by the term 'noun phrase'? A common answer is that the NP denotes the projection of the noun (N). Since parallelisms between the NP and the clause can be observed, the main features of the syntactic structure of the clause can be adopted for the analysis of the NP (Giorgi & Longobardi 1991: 1f).

#### 2.1.1. Basic phrase structure

To analyse the structure of NPs, it is crucial to first discuss how phrases and clauses are structured. In generative syntax, syntactic structures are assumed to be endocentric. This implies that sentences consist of syntactic constituents, each of which contains a head: "Each head, X, projects a larger syntactic unit (a phrase, XP), and each phrase, XP, must have one head" (Alexiadou, Haegeman & Stavrou 2007: 11). According to this assumption, to give but one example, the head of a verb phrase is a verb. Moreover, this theory allows for the conclusion that every syntactic constituent has the same format, the so-called X-bar format:

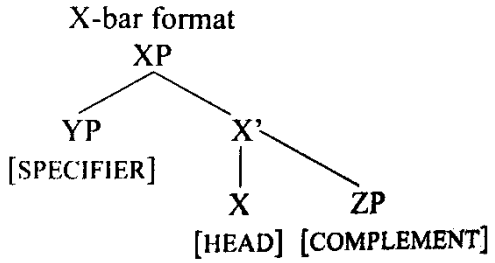


**Figure 1 X-bar format (Alexiadou, Haegeman & Stavrou 2007: 11)**

X constitutes the head of the constituent. Another constituent, ZP, functions as the complement of X. X' is composed of X and its complement and is referred to as the intermediate projection of X. X' combined with a further constituent, YP, called the specifier of X, produces XP, the maximal projection. YP and ZP are also formed according to the structure presented in Figure 1 (Alexiadou, Haegeman & Stavrou 2007: 11).

Forming phrases and sentences as described above is referred to as binary merger operation within the MP. This means that words and phrases are combined to produce a larger category. Apart from projections, heads and phrases, merge includes further principles. The binary character of merge is illustrated in Figure 1 where each node has two downward branches (Radford 2003: 62ff). Furthermore, the specifier c-commands the head and the complement. Finally, recursion is a crucial property of merge (Gelderen 2011: 12).

### 2.1.2. Headedness

The section about phrase structure has shown that every phrase has a head. However, it is not always obvious which word functions as a head. In the nominal domain, there is disagreement amongst linguists as to whether N or D is the head of the NP. Several reasons underlie the dispute as will be discussed later.

Most importantly, lexical heads need to be distinguished from functional heads. Lexical heads contribute to the descriptive content of a sentence and are, thus, referred to as 'content words' in traditional grammar. This group is comprised of nouns, verbs, adjectives and prepositions. Functional heads, on the other hand, refer to words that do not carry descriptive content. Instead, they encode grammatical and structural relationships and can appear as bound morphemes as well as free morphemes (Alexiadou, Haegeman & Stavrou 2007: 12ff).

Alexiadou, Haegeman and Stavrou (2007: 15) identify five properties shared by functional categories:

(i) They constitute closed classes.
(ii) They are generally phonologically and morphologically dependent, and stressless. Often they are clitics or affixes and sometimes they are phonologically null.
(iii) They are usually inseparable from their complement.
(iv) They lack descriptive content.
(v) Functional heads (usually) do not have arguments.

In the nominal domain, D is a so-called functional head. In (1), the article *the* clearly lacks descriptive content, but still contributes to the interpretation of the NP or the determiner phrase (DP) by determining the referential properties of the nouns *cat*, *milk* and *table*.

(1) a. The [$_N$ cat] [$_V$ drinks] the [$_N$ milk].
b. The cat is [$_A$ thirsty].
c. The cat is [$_P$ under] the table.

(Alexiadou, Haegeman & Stavrou 2007: 13)

The definite determiner in (1) reveals that the nouns following it are not newly introduced into the discourse (Alexiadou, Haegeman & Stavrou 2007: 14).

### 2.1.3. The question of DP

The debate on whether N or D is the head of NP was launched in the 1980s. Steven P. Abney (1987), in particular, has successfully influenced the generative research on the syntax of NPs with the introduction of the 'DP-hypothesis'. According to Abney, the NP needs to be reinterpreted as DP, i.e. a projection of D with a nominal complement. The phrase *his three beautiful sisters* could thus be structured as follows:

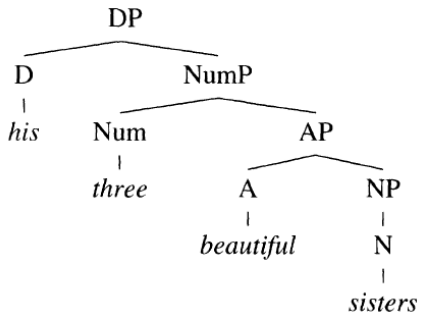


**Figure 2 The DP-Structure (Eynde 2006: 141)**

The introduction of a functional projection that dominates the lexical projection NP can be explained by the fact that parallelisms exist between the nominal domain and the clause, hence the verbal domain. Indeed, while the verb is the semantic nucleus of the clause, the clause is not a mere projection of V. Functional projections such as IP and CP often dominate

the lexical projection VP. This theory has been adapted to the NP as it has been argued that the latter is similarly dominated by functional phrases, one of them being DP. By comparing the V head in the clausal domain to the N head in the nominal domain, it can be concluded that the role of C, the complementizer position, in the clause and can likewise be compared to D, the determiner, in the nominal domain (Alexiadou, Haegeman & Stavrou 2007: XVIII, 3).

In the following sections, morphological, semantic and syntactic/ distributional evidence for postulating functional projections will be elaborated.

### 2.1.4. The functional layering of DP

Since modern generativists assume that parallelisms exist between the nominal and the verbal domain, it is likely that DPs are endowed with functional projections, similar to VPs (TP, AgrP, AspP etc). However, the functional layering of DPs is still a major area of research for linguists. Questions arise as to whether further functional projections exist in the nominal domain, how many such projections there are, how they can be triggered and what kind of properties they possess.

Morphological markers in particular can play a role in postulating functional projections as a similar reasoning applies to the verbal domain. For example, there is some evidence for the existence of an Agreement projection (AgrP) in the functional domain of the clause since verbs are inflected for agreement (Alexiadou, Haegeman & Stavrou 2007: 35).

In the nominal domain, the number ending on N serves as an example of a functional element. Null or plural morphemes are not an inherent part of N. The fact that DPs can either refer to one or to any number of entities serves to prove that a projection to encode Number, called NumP, exists (Alexiadou, Haegeman & Stavrou 2007: 24).

## 2.2. A formal characterization of the grammatical category D

Having touched upon several issues pertaining to the structure of the NP/ DP, the focus now shifts to the grammatical category D. Due to the emergence of the DP hypothesis, the determiner has been assigned a crucial role in the nominal system. At the same time, the reinterpretation of NP has led to a number of questions in terms of the position and interpretation of determiners. In this section, a particular focus is thus placed on the syntactic and semantic interpretation of definite and indefinite determiners as well as on the role of the functional category D.

### 2.2.1. Classification

Determiners typically introduce nominal structures. They determine the reference of the noun, hence the designation determiner (Hope 2003: 25). A number of theories try to explain how determiners are classified, some of which will be presented in this section.

Longobardi (2001: 580f) distinguishes between five classes of determiners: articles, demonstratives, possessives, quantifiers and cardinal numbers. They are all part of a closed class and constitute a functional category (see 2.1.2.). They are mainly assigned to the head or specifier of DP (2.2.2.) and establish the definite/indefinite interpretation of the noun (2.2.3.).

The definite/indefinite distinction was largely abandoned by Milsark (1974, 1977), who prefers to classify determiners in the categories 'strong' and 'weak'. Whereas weak determiners can be used with there-insertion contexts, strong determiners are ungrammatical in the postcopular position of existential sentences. The examples in (2) show that *a, some, a few, many* and numerals are weak, whereas *the, every, all* and *most* are strong:

(2) a. There is/are a/some/a few/many/three fly (flies) in my soup.
b. *There is/are the/every/all/most fly (flies) in my soup."

(Diesing 1992b: 59)

The following table lists Milsark's weak and strong determiners according to their acceptability in existential sentences:

**Table 1 Milsark's strong and weak determiners (Milsark 1977)**

| WEAK | STRONG |
|---|---|
| *a* | definites (*the*, demonstratives, possessives) |
| weak *some* | universals (*all, every, each*) |
| number determiners | *any* when not polarity item of *some* |
| null determiners in nonuniversal reading | null determiner in universal reading |

Milsark's classification of English determiners ties in well with Ioup's (1975: 42) hierarchy of quantifiers, which reflects variation in relative scope preference.

**Ioup's hierarchy**

each > every > all > most > many > several > some ($+NP_{pl}$) > a few

The quantifiers in the left-hand side tend towards wider scope, whereas those in the right-hand side tend towards narrower scope. Interestingly, the leftmost elements correspond to Milsark's strong determiners and the rightmost elements to Milsark's weak determiners, respectively.

### 2.2.2. The syntactic position of determiners

The types of determiners listed in the previous section are not an integral part of all languages. It is true that demonstratives and quantifiers (including numerals) are universal, but articles, for instance, cannot not be found across all languages. Another property that sets demonstratives and quantifiers apart is that they cannot only be used transitively with a nominal complement but also intransitively without a nominal complement and thus assume a more autonomous status in the sentence. The article, in contrast, is morphologically and phonologically dependent on the noun.

In terms of syntactic status, the complementary distribution of determiners may lead to the assumption that determiners occupy the same position, namely the $D^0$ position (Giusti 1997: 99-106). Hence, the basic structure of DP described so far looks as follows:

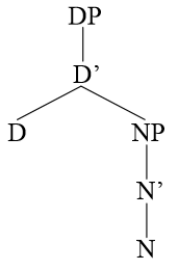


**Figure 3 The basic structure of DP (Alexiadou, Haegeman & Stavrou 2007: 94)**

However, this structure cannot be applied cross-linguistically since the complementary distribution of the different types of determiners is not universal. For instance, there are languages, like Romanian, in which demonstratives and articles can co-occur. Moreover, in a lot of languages, including English, quantifiers co-occur with articles and demonstratives in a number of ways. Cross-linguistic evidence thus clearly suggests that articles, demonstratives and quantifiers occupy different structural slots (Giusti 1997: 100).

In terms of the structural position of demonstratives, Giusti hypothesizes that they are maximal projections which occupy the specifier position of DP. To support the claim, evidence from Romanian and Italian is provided (see 1997: 107-112 for detailed discussion). Hence, the new DP structure looks as follow:

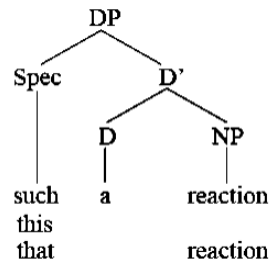


**Figure 4 Articles in D and demonstratives in SpecDP (Alexiadou, Haegeman & Stavrou 2007: 109)**

Alexiadou, Haegeman & Stavrou (108f) add some facts about demonstratives to support the hypothesis that demonstratives occupy SpecDP. Since an interpretative similarity can be observed between the demonstrative *this* and the degree modifier *such,* as well as between the demonstrative *that* and the degree adverb *so*, all these elements occupy the same structural position, namely the specifier of DP. This hypothesis is also consistent with cross-linguistic evidence, which suggests that demonstratives tend to occupy the leftmost position in the DP (Alexiadou, Haegeman & Stavrou 2007: 108f). However, it is necessary to add that the leftmost position of the demonstrative is probably not the 'base position' but a derived position. Linguists often assume that the demonstrative originates lower in the structure since in many languages, e.g. Romanian, Spanish and Greek, the demonstrative does indeed occupy a lower position (Alexiadou, Haegeman & Stavrou 2007: 109ff).

As for the position of quantifiers, Giusti (1991: 444ff) proposes that they neither occupy the $D^0$ nor the SpecDP position. To account for their analysis, Giusti distinguishes two types of quantifiers. Quantifiers that cannot be preceded by an article are external to DP and occupy the head Q position. They select a DP as their complement and project QP, as illustrated in Figure 5. Quantifiers preceded by a determiner, on the other hand, are internal to the DP, are located in a high specifier position and behave as adjectives.

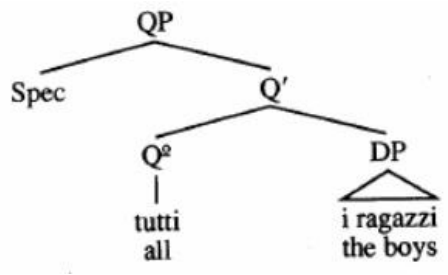


**Figure 5 Quantifiers in the head Q position (Giusti 1997: 114)**

### 2.2.3. The interpretative content of determiners

#### *2.2.3.1. Definite determiners (the, this)*

From a semantic perspective, the demonstrative and the definite article share some central features. They are both definiteness markers and impart referentiality. These parallels are not a coincidence but a result of their diachrony (see 4.2.1.) since processes of semantic and phonological weakening have led to the emergence of the definite article out of the demonstrative pronoun (Alexiadou, Haegeman & Stavrou 2007: 96-99).

And yet, despite such shared characteristics, there are also significant differences in the interpretation of the two elements. Most importantly, it has been argued that demonstratives possess some semantic content as they account for the deictic interpretation of the NP, whereby the abstract feature [+/-DEM] plays an important role in "guiding the hearer's attention to the referent" (Lyons 1999: 21). Consequently, their use and interpretation depends on the context. One important non-deictic usage of the demonstrative is the anaphoric usage. Here, the demonstrative functions like an anaphoric pronoun by referring back to an antecedent linguistic expression which affects the interpretation of that particular demonstrative. Another difference between demonstratives and articles concerns the semantic interpretation of the NP. In contrast to the demonstrative pronoun, definite articles can refer to a kind term. This implies that demonstratives cannot have a generic reading (Alexiadou, Haegeman & Stavrou 2007: 95-105).

To sum up, in spite of the fact that definite articles historically developed from demonstratives and share the role of markers of definiteness and referentiality with the later, their lack of semantic content and their use in generic contexts clearly distinguishes them from demonstratives.

#### *2.2.3.2. Indefinite determiners (a, some)*

The indefinite determiner *a* can be interpreted in various ways. It can have a generic reading, as in (3a), a specific existential reading, as in (3b), a non-specific reading, as in (3c) or it can be ambiguous between a specific and a non-specific interpretation, shown in (3d):

(3) a. A blue whale eats about 40 million krill a day
b. I want you to meet a student, her name is Sara
c. I would love to have a unicorn, but they don't exist
d. He will invite a minister to open the conference

(Crisma 2015: 76)

It is important to note that in the examples above, the presence of *a* is always obligatory. Only mass singular nouns may occur bare.

As for the indefinite reading of plural nouns, an overt morpheme again alternates with a bare noun. Plural nouns with a generic or existential interpretation can occur bare (see section 3.2.), but they always yield a non-specific reading. For a specific existential reading, the presence of weak *some* or *sm* is necessary. Thus (4b) sounds odd due to the incompatibility of BPs with a specific reading. Finally, in contrast to (3d), (4c) is not ambiguous due to the presence of the BP (Crisma 2015: 76).

(4) a. I want you to meet sm students: Sara, Giovanni and Priscilla
b. *?I want you to meet students: Sara, Giovanni and Priscilla
c. He will invite ministers to open the conference

(Crisma 2015: 76)

All this is due to the fact that BPs cannot take wide scope, in contrast to singular nouns preceded by *a(n)* and plural nouns preceded by *sm* (Carlson 1977: 417ff). As a consequence, the presence of *an* and *sm* is indispensable for the NP to take wide scope or to be interpreted as specific (Crisma 2015: 80).

### 2.2.4. The role of D

The role of D and its impact on the interpretation of the DP has been a subject of discussion among linguists since the emergence of the DP Hypothesis. In this debate, the article as "the determiner par excellence" (Alexiadou, Haegeman & Stavrou 2007: 55) clearly stands out. To start with, Abney (1987) emphasizes the function of the article as a subordinator which turns its NP complement into an argument. Longobardi (1994) points out that articles impart referentiality, while Lyons (1999: 277f) underlines their role in some languages to grammaticalize the semantic-pragmatic concept of definiteness. Finally, Giusti (1997: 102-107) views the article as a mere grammatical morpheme assigning case to its NP.

The importance of argumenthood in this debate was originally observed by Szabolcsi (1987) and adopted by Stowell (1989) and Longobardi (1994). They argue in favour of a difference in the structural representation of arguments (subject, object) and non-arguments (predicate, vocative). According to this hypothesis, nominal expressions require D to function as arguments, whereas non-arguments are simple NPs. This contrast is illustrated in (5):

(5) a. Bill is a doctor. (predicate)
b. Doctor, come here! (vocative)
c. *Doctor came immediately. (subject)
d. *I saw doctor. (object)

(Delfitto 2005: 214)

At this point, Alexiadou, Haegeman & Stavrou (2007: 131) draw a comparison with the CP, which is turned into an argument by C. The role of the article to turn nominal expressions into arguments can be explained by the fact that D, like the complementizer, specifies the reference of its NP (Alexiadou, Haegeman & Stavrou 2007: 70; 159; Abney 1987: 77). The relation between argumenthood and referentiality is thus a close one. Or, in the words of

Alexiadou, Haegeman and Stavrou, “argumenthood is the syntactic reflex of the concept of referentiality” (2007: 66).

Another important semantic concept is definiteness, represented as a feature [DEF]. Whereas some linguists claim that [DEF] represents the semantic-pragmatic concept of definiteness, which is not realized in a uniform manner across languages, others are convinced that [DEF] is a grammatical feature that has different relations to one or more semantic concepts depending on the respective language (Alexiadou, Haegeman & Stavrou 2007: 57f). A key question, then, is whether definiteness refers to a single, unified phenomenon or whether it represents more than one semantic-pragmatic concept, which seems to have the same morphological realization in a number of languages. Lyons (1999), who cross-linguistically analysed the concept of definiteness, argues that semantic and pragmatic definiteness is universal, but the grammatical realization of definiteness varies greatly across languages. It could therefore be assumed that D is the vehicle of semantic definiteness. Giusti (1997) draws a similar conclusion in saying that the article itself does not carry semantic content. A distinction thus needs to be made between D as a structural position and its lexical realization.

In short, D can be described as a mediator between the description presented by the NP and its application to a specific entity in the real world, as this quote from Cheng and Sybesma (1999: 518) demonstrates:

> The underlying assumption is that there is a division of labor between NP, which describes, and D, which refers. This seems to be a general characteristic of language; for example, the same kind of division of labor exists in the verbal domain between the describing VP and the referring, deictic T, which links the event described in the VP to a particular event associated with a particular point on the time axis. We would like to say, then, that this division of labor is a property of Universal Grammar: some entities describe, whereas other entities perform the deictic discourse function of linking the description to some particular object or event in the real world. In languages with articles/determiners, the deictic function in the nominal phrase is taken care of by the article/determiner. However, this should not lead one to conclude that if a language has no articles/determiners, no element performs the deictic function. If the describing/referring dichotomy is indeed part of Universal Grammar, then if a language has no articles/determiners, some other element in the language must perform the deictic function.

# 3. BARE PLURALS – ASPECTS OF THE SYNTAX-SEMANTICS INTERFACE

In the previous chapter, articles were argued to be obligatory for NPs to function as arguments of verbs. However, this observation appears to be incorrect since a number of languages admit seemingly articleless nouns in argument position. There are both languages that have no articles at all and languages with articles that allow so called 'bare NPs', hence nouns without overt determiners.

English is among the languages that have articles, but that also admit determinerless nominal constituents, which seemingly function as arguments. Generic nouns, proper names and mass nouns in particular do not co-occur with a determiner in English (Carlson 2003: 149). Yet in this section, the focus lies solely on the analysis of the BP construction.

Bare NPs have offered space for debate in the generative theory at least since Carlson initiated the study of BPs in English (Longobardi 2001: 582). The analysis of bare nouns involves a number of unresolved questions concerning the syntax-semantics interface. A central topic is the analysis of bare nominal projections that function as arguments. Are they interpreted as NPs or as DPs despite the lack of an overt determiner? Apart from the role of the functional category D and its lexical realization, the syntactic distribution and semantic interpretation of argument BPs will be discussed (Delfitto 2005: 217).

## 3.1. Syntactic analysis of BPs

### 3.1.1. Syntactic distribution of BPs

English bare NPs have a restricted occurrence, as in many cases the absence of an article causes an ungrammatical phrase. Thus, certain rules provide information about when a determinerless NP is grammatical and when it is not. In many languages, what constitutes a distinguishing feature between grammatical and ungrammatical bare NPs is the count versus mass distinction. In English, the use of bare forms of mass nouns is admitted, whereas (argumental) bare count nouns are only admitted as BPs, as shown in (6) (Delfitto 2005: 215).

(6) a. Water is scarce in this country.
    b. Dogs are intelligent.
    c. *Dog is intelligent.

(Delfitto 2005: 215)

However, not all languages allow the use of bare nouns in argument position, as is shown by the example of Modern French where argument bare nouns are not allowed. Moreover, in

languages in which BPs and bare mass nouns are allowed in argument position, the distribution of those nouns can still be restricted. For instance, in most Romance languages subject BPs are generally excluded from the subject position if they are not modified or coordinated (Delfitto 2005: 215).

Contrary to Romance bare nouns, which have a more restricted syntactic distribution, English BPs can principally occur in all argument positions (Longobardi 2001: 582). Hence, they are accepted both in subject and object position, as shown in (7).

(7) a. Students have occupied the building.
b. I saw students in the building.

(Delfitto & Schroten 1991: 155)

In addition, English bare nominals can also be predicative in sentences containing copular verbs. Interestingly, in non-argumental position (i.e. predicative, vocative position), the use of bare singular nouns is likewise permitted, but this construction is limited to unique roles (Zamparelli 2005: 762f):

(8) a. Anne is head of the department.
b. Bill is employee of the week.

(Swart & Zwarts 2009: 282)

### 3.1.2. DPs and NPs

According to the DP hypothesis, NPs are dominated by DPs. Considering the fact that in the case of bare NPs the functional head D would remain empty, it is natural to ask whether NPs without overt determiners are in fact interpreted the same way as NPs with overt determiners. Is a DP layer needed in this particular case?

Numerous authors are convinced of the universality of D, a category that exists in all languages and that is relevant to the semantic interpretation of the NP, as discussed in 2.2.4. Lyons (1999), however, does not share this opinion and claims that the existence of a functional layer DP depends on whether the language in question encodes semantic definiteness. If this is not the case, D is not projected. This implies that the projection of D correlates with the creation of definiteness marking: “There can be no definite article in languages lacking DP structure, and, to the extent that it is obligatory to have some expression of a projection, languages with DP structure must have a definite article” (Lyons 1999: 323).

In short, the most important question is whether BPs have a null determiner. The alternative is that there is no DP layer for articleless NPs (Alexiadou, Haegeman & Stavrou 2007: 173f).

In the literature two possibilities have been discussed in order to deal with nominal projections without overt determiners that function as referential arguments. Chierchia's solution to the issue is grounded on semantics, whereas Longobardi's approach is more syntactically oriented. Chierchia proposes a set of rules for the interpretation of the noun and claims that in the absence of an article NPs are bare, hence D is not projected. Opposed to this is Longobardi's theory of N to D movement where all arguments are analysed as DPs.

#### *3.1.2.1. Chierchia – a typological parameter*

In 1998, Chierchia proposed the Nominal Mapping Parameter as a semantic parameter to explain the cross-linguistic use and interpretation of bare nominals. This semantic parameter determines whether NPs can be turned into DPs.

As a starting point, Chierchia (1998: 352f) stresses the double role of nouns. On the one hand, they are non-referring in predicate position. On the other hand, they appear in argument position as referring expressions. The immediate question that arises is how these two options are realized in different languages.

One possibility is that the features [±arg], [±pred] control how the noun and the NP are mapped into their interpretations: "For any such feature a, [+a] means that N's can be mapped onto things of type a, [-a] that they cannot" (Chierchia 1998: 353). In a NP[+arg, -pred] language such as Chinese and Japanese, a noun and its phrasal projection can be argumental but not predicative. In these languages, all nouns are mass. In Romance languages, hence in NP[-arg, +pred] languages, nouns and NPs are exclusively mapped onto predicates. Bare arguments are thus not allowed. In Germanic languages like English, which are NP[+arg, +pred] languages, both predicative and argumental NPs are permitted. Moreover, a distinction between mass and count nouns is made. Mass nouns occur as bare arguments, whereas (singular) count nouns do not. BP nouns, however, can also occur as arguments through the possibility of type shift (Chierchia 1998, 353-356).

Summing up this brief survey of Chierchia's account of the cross-linguistic interpretation of bare NPs, one can conclude that argumenthood is not dependent on the category D. Instead, Chierchia proposes that nouns as such tend to be either inherently argumental or predicative.

#### *3.1.2.2. Longobardi – N-to-D Movement*

Contrary to Chierchia (1998), Longobardi (1994) clearly argues in favour of an isomorphism between the status of NPs as arguments and the functional category DP. According to this theory, bare arguments are embedded in a fully developed DP structure that contains an empty D head. This empty D head accounts for the existential reading of BPs and their restricted distribution to lexically governed positions, for empty D heads must be lexically governed and they lead to a default existential interpretation (Longobardi 1994: 640f).

Further interpretations of BPs, including generics and proper names, are, however, neither restricted in their distribution nor introduced by an empty D. For these cases, Longobardi has developed a theory of N-to-D-Movement which takes place either in overt syntax or at the level of Logical Form (LF). According to Longobardi (1994: 659ff), the N-movement is triggered by the referential feature of D, which is uninterpretable and needs to be checked by N. Depending on the quality of the referential feature (weak or strong), N movement takes place in the syntax or at LF. In English and other Germanic languages, the referential feature on D is weak, hence N raises to D at LF. As a consequence of the filling of the D position and the lack of an empty D, the bare noun is no longer interpreted existentially, nor restricted to lexically governed positions.

This theory, however, broaches the question as to why existential BPs can occasionally occur in subject position, as in (9), despite the lexical government requirement.

(9) Students have occupied the building.

(Delfitto & Schroten 1991: 160)

Since this case only exists in English but not in the Continental Germanic languages, Longobardi (1994: 645) considers the acceptability of (9) as a marked phenomenon: “Such predicted markedness of the existential interpretation of bare noun subjects appears to be empirically reflected by the influence that different lexical choices in the predicate have on it” (1994: 645).

### 3.2. Semantic interpretation of BPs

Having analysed the BP construction from a syntactic point of view, the focus now shifts to the interpretation that bare NPs receive in argument position. Bearing in mind that the study of English BPs spans forty years of semantic research, it is hardly surprising that a large number of concepts and theories have been developed. A natural starting point for discussing

BPs is Carlson's influential work *A Unified Analysis of the English Bare Plural* in which the three primary readings of bare NPs – kind, generic and existential – are discussed. In subsequent years, Carlson's original theory has been extended for cross-linguistic application (Dayal 2011: 1088f).

### 3.2.1. Carlson's theory

Carlson's (1977) analysis of BPs is the most influential approach to the topic within the framework of generative grammar. In his theory, BPs are assigned a uniform analysis, according to which English bare plurals are proper names of kinds of things. Nevertheless, this theory allows for a large range of different uses of BPs.

Carlson's (1977) study concentrates in particular on the distinction between the 'generic' use of BPs (e.g. *Dogs bark*) and their existential or 'indefinite plural' use (e.g. *Dogs were sitting on my lawn*). Crucially, Carlson argues that the decisive factor for this distinction is not the NP itself but the context of the sentence, which acts on the BP and thus selects the universal or existential reading. Carlson therefore claims that the BP has no inherent quantificational force and "never, in and of itself, gives rise to an ambiguity" (1980: 24).

Particularly noticeable with the analysis of BPs is the diversity of their possible readings. The generic uses of BPs in particular require further subdivisions. To account for the variety of generic readings, Carlson provides examples for different generic interpretations that arise. To start with, he claims that it is natural to regard the generic as a type of universal quantifier. In sentences like (10a), the expression *all horses* seems to give an accurate paraphrase as the meaning of the sentences *horses are mammals/creatures/material objects* is synonymous with *all horses are mammals/creatures/material objects*. However, there are exceptions to this 'universal'. In the examples of (10b), the null determiner seems to have the force of *most*, but the sentences are still true.

(10) a. Horses are mammals/creatures/material objects.
b. Horses are smart/larger than mules/good pets.
c. Horses are widespread/extinct/indigenous to eastern Chile.

(Carlson 1977: 413f)

It may not be concluded, though, that generic BPs are ambiguous between the abovementioned quantificational forms, i.e. between the determiners *all* and *most*, since in the examples of (10c), the use of those quantifiers would be inappropriate. There is in fact no

quantifier paraphrase available in these cases as the predicates *be widespread*, *be extinct* and *be indigenous to* cannot be assigned to particular individuals but only to kinds, i.e. to species, as shown in the examples of (11). These so-called 'kind-referring' NPs are in opposition to 'object-referring' NPs found in (10a) and (10b). This finding does not affect Carlson's opinion, though, as to the relationship between particular and kind-level individuals, which he assumes to be a close one (Carlson 1977: 444).

(11)

| | | |
|---|---|---|
| *Fred<br>*All goats<br>Goats<br>This kind of animal | are<br>is | widespread/numerous/extinct<br>rare/common/indigenous to… |

(Carlson 1977: 444)

In terms of the existential use of the BP (see examples in (12)), Carlson notes that it has often been referred to as the 'indefinite plural' as this use appears to be the semantic plural of the singular NP preceded by the indefinite singular article *a(n)* having the force of *some*.

(12) a. *Doctors* tried to save the dying boy.
b. Knute threw *rotten peaches* at the library.
c. *Mice* will come out of that wall if you pound on it.

(Carlson 1977: 414)

In his study, though, Carlson rejects the idea that the zero determiner in plural NPs serves as the counterpart of *a(n)* since the two do not share all relevant semantic properties. Indeed, Carlson (1977: 415-429; 1980: 7-21) successfully illustrates that they behave differently with respect to opaque contexts, relative scope, the differentiated scope phenomenon and pronominalization.

To account for the different interpretations of bare NPs, Carlson (1977: 448ff; 1980: 66ff) elaborates his idea of unifying the generic and the existential use of the BP. For this purpose, he assumes a tripartite ontology: kinds, objects and stages. Crucially, kinds and objects are referred to as 'individuals' while stages are spatially and temporally bound manifestations of individuals. To distinguish between objects and kinds, Carlson points out that the former occupies only one place at a time, the latter, in contrast, may occupy many places at a time. Another difference is that kinds may tie together objects as well as stages. Objects, on the other hand, may tie together stages only, as shown in Figure 6 (Carlson 1980: 69).

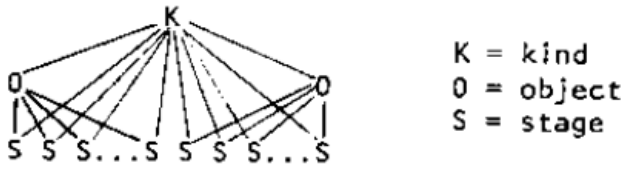


**Figure 6 Kinds, objects and stages (Carlson 1980: 69)**

As stated above, Carlson assumes that the meaning of the predicate is responsible for the different readings of BPs. In this respect, he essentially follows Milsark (1974: 210-216) who distinguishes between 'state-descriptive predicates' and 'property predicates':

> Properties are those facts about entities which are assumed to be, even if they are not in fact, permanent, unalterable, and in some sense possessed by the entity, while states are conditions which are, at least in principle, transitory, not possessed by the entity of which they are predicated, and the removal of which causes no change in the essential qualities of the entity. (Milsark 1974: 212)

Milsark observes that state-descriptive predicates (ex. 13a&14a) – in contrast to property predicates (ex. 13b&14b) – allow for there-constructions. This is true for both verbal and adjectival predicates.

(13) a. There are students smoking in the classroom

b. *There are students knowing French.

(14) a. There are firefighters available.

b. *There are firefighters intelligent.

(Krifka *et al.* 1995: 23)

Milsark's state-descriptive predicates correspond to Carlson's 'stage-level predicates' (S-level predicates) and make up the majority of predicates in English. These include many verbs, progressives, passives, certain adjectives such as *available* or *drunk* and most prepositional phrases that apply only to stages of individuals and speak of 'happenings'. The small group of property predicates correspond to Carlson's 'individual-level predicates' (I-level predicates). Included here are the 'characteristic' verbs, passives, the verb *can* in the sense of *be able to*, most verbs lacking progressive form such as *resemble*, adjectives such as *intelligent* or *large*, certain prepositional phrases and all predicate nominals. They apply only to individuals (objects and kinds) and speak of 'characteristics'. This fundamental difference in meaning between S-level and I-level predicates is primarily responsible for the different interpretations of BPs as Carlson clearly demonstrates: "Thus, *being intelligent* is a characteristic, but *being drunk* is a happening; *being a linguist* is a characteristic, but *being on the corner* is a happening. A VP like *is running* is a happening, but VP's like *be eaten alive* and *runs* are ambiguous between characteristics and happenings" [original emphasis] (Carlson 1980: 75).

Milsark (1974) and Carlson (1980) both assume that S-level predicates select the existential reading of the BPs as well as the indefinite article *a* and the unstressed variant of *some* (*sm*), as shown with the adjective *available* in (15).

(15) a. Doctors are available.
b. A doctor is available.
c. Sm doctors are available.

(Carlson 1980: 74)

With I-level predicates, though, only the universal reading of the BP is acceptable, as illustrated in (16). The use of *a* and *some* is not appropriate in this case.

(16) a. Doctors are intelligent.
b. (*)A doctor is intelligent.
c. *Sm doctors are intelligent.

(Carlson 1980: 74)

Example (17) illustrates that the prepositional phrases as well as the progressive and passive participles may select the existential readings. The passive is ambiguous, though (Carlson 1980: 74).

(17) a. { Sm cats were / Cats were / A cat is } on the corner.

b. { Sm cats were / Cats were / A cat was } running.

c. { Sm cats were / Cats were / A cat was } attacked by Jules.

(Carlson 1980: 74)

Predicate nominals always apply to individuals and, thus, select the universal reading of the subject. Hence, in the examples of (18) the use of *sm* and *a* are unacceptable.

(18) a. *Sm cats are mammals.
b. (*)A cat is a mammal.
c. Cats are mammals.

(Carlson 1980: 74)

The following sentence is an example of a PP which only selects the universal reading of the BP:

(19) Tires are in short supply.

(Carlson 1980: 74)

As far as direct object position is concerned, in most cases verbs select the existential reading of BPs. One exception is a class of verbs that does not appear in the progressive form and which includes *fear, hate, love, respect, loathe* and *admire*. They are all part of Levin's (1993: 191f) class of verbs of psychological state with experiencer subjects. The full list of verbs (taken from Levin 1993: 191) can be found below.

- POSITIVE VERBS: admire, adore, appreciate, cherish, enjoy, esteem, exalt, fancy, favor, idolize, like, love, miss, prize, respect, relish, revere, savor, stand, support, tolerate, treasure, trust, value, venerate, worship
- NEGATIVE VERBS: abhor, deplore, despise, detest, disdain, dislike, distrust, dread, envy, execrate, fear, hate, lament, loathe, mourn, pity, regret, resent, ?rue

According to Carlson, these verbs select the universal reading of the BP object. Thus (20) can only be used to express that Betty generally hates dogs.

(20) Betty hates dogs.

(Carlson 1980: 113)

Moreover, the use of *a* and *sm* is strange with NPs in the direct object position of these verbs:

(21) a. ?Bill fears a duck.

b. ?John admires sm Congressmen.

(Carlson 1980: 113)

As it has become evident over time that Carlson's theory faces some empirical difficulties, numerous influential ideas in terms of the syntax-semantics interface of BPs have since been proposed. Two major theories will be presented hereafter.

### 3.2.2. The Ambiguity Approach

Carlson's analysis of the semantics of English BPs clearly represents the basis for the further development of the theory. Advocates of the Ambiguity Approach (Wilkinson 1991; Diesing 1992b; Gerstner-Link & Krifka 1993; Kratzer 1995) agree with Carlson that kinds are included in the ontology. Furthermore, they adopt Carlson's view that the quantificational force of BPs is external to the NP.

However, the Ambiguity Approach differs strongly from Carlson's original theory in that the idea that BPs always refer to kinds is rejected. Instead, it is proposed that English BPs are ambiguous, i.e. they are either proper names of kinds in connection with kind-level predicates or they behave as weak[1] indefinites (i.e. the plural counterparts of indefinite singular NPs) in connection with I-level and S-level predicates.

Proponents of the A-Approach believe that there exists a mapping between syntactic structures and tripartite logical representations (consisting of a quantifier, the restrictive clause and the nuclear scope) and draw on the Kamp-Heim approach to the semantics of NPs, which suggests that singular indefinites have either universal or existential force depending on whether they are mapped onto the nuclear scope or the restrictor of a tripartite structure. This proposal has subsequently been applied to BPs.

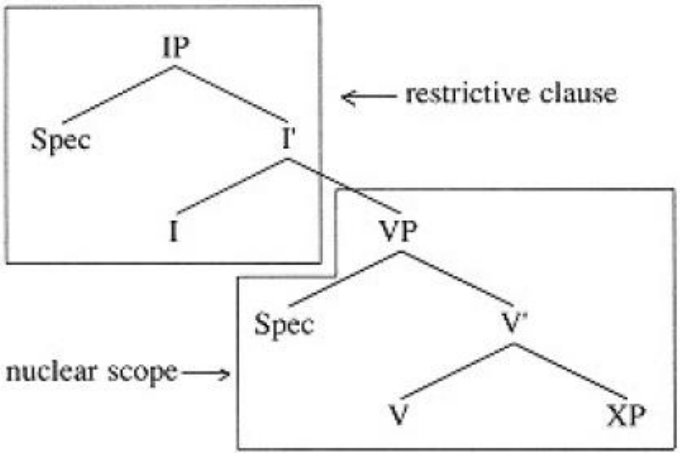


**Figure 7 Mapping Hypothesis (Diesing 1992b: 9)**

Diesing (1992b) first explains that the mapping procedure, which creates a close link between syntactic structures and semantic representations, splits the syntactic tree into two parts (Figure 7). These two parts are subsequently mapped into the restrictive clause and the nuclear scope of the logical representation. Crucially, Diesing's approach builds on a two-subject model of phrase structure, also known as the VP-Internal Subject Hypothesis (see Radford 2003: 151ff for detailed discussion), according to which the subject appears either in the specifier of IP or in the specifier of VP. The Mapping Hypothesis suggested by Diesing involves the mapping of different portions of a sentence into the restrictive clause and the nuclear scope depending on their syntactic position: "Material from VP is mapped into the nuclear scope. Material from IP is mapped into a restrictive clause" (Diesing 1992b: 10).

[1] The weak/strong distinction refers to Milsark's work, discussed in section 2.2.1. Most importantly, strong indefinites presuppose, in contrast to weak indefinites, the existence of individuals satisfying their restriction (Chierchia 1998: 341).

This procedure makes some predictions concerning the interpretation of BP subjects. Diesing (1992b: 16) assumes that BPs are not inherently quantificational, but instead introduce variables into the logical representation. The variables of BPs that are realized VP-internally and mapped onto the nuclear scope end up bound by existential closure. This is the source of the existential reading. The variables of BPs that are realized VP-externally and mapped onto the restrictive clause, on the other hand, are bound by an abstract generic operator *Gen*, which yields a generic reading.

Referring to Carlson's distinction between I-level and S-level predicates, Diesing (1992a: 356f) shows that these two types of predicates differ in terms of where their BP subjects appear in the logical representation. In short, Diesing proposes that subjects of S-level predicates can be bound by existential closure or the generic operator *Gen* and appear in the nuclear scope or the restrictive clause, respectively, whereas subjects of I-level predicates can only be bound by *Gen* and are mapped onto the restrictive clause. This approach implies that subjects of S-level predicates, in contrast to subjects of I-level predicates, are ambiguous between existential and generic readings depending on the context. The multiple readings that S-level predicates allow are illustrated in the following example:

(22) Firemen are available.

(Diesing 1992a: 356)

The S-level predicate *available* in (22) allows both an existential and a generic reading of the BP subject. On the existential reading, there are firemen available at a specific point in time. The generic reading expresses that it is a general property of firemen that they are available. Diesing also points out a third reading, the so-called "existential generic" interpretation, which will not be discussed here further.

Similarly, Kratzer (1995) states that BP subjects of S-level predicates can have two readings and uses the following example.

(23) Typhoons arise in this part of the Pacific.

(Kratzer 1995: 139)

Since (23) is a generic sentence, Carlson's analysis predicts a generic reading of *typhoons* (24a). However, Kratzer (1995: 139f) claims that both the generic (24a) and the existential readings (24b) are possible.

(24) Typhoons arise in this part of the Pacific.

a. Typhoons have the general property of arising in this part of the Pacific.
b. This part of the Pacific has the general property that there are typhoons arising there.

So, the BP subject can either be mapped into the restrictive clause or the nuclear scope of the quantifier construction. The former yields a generic reading, the latter an existential reading.

As for BP objects**,** Diesing (1992b: 29f) states that no generic readings should be allowed from a syntactic point of view. English is a language in which objects always appear within VP and VPs are mapped onto nuclear scopes. Diesing agrees with Carlson, though, that objects of experiencer predicates are indeed interpreted generically. It appears that they are mapped onto the restrictor and bound by the generic operator *Gen*. In addition, in some habitual contexts also other verb types allow generic readings for BP objects.

(25) Esther reads novels.

(Diesing 1992b: 30)

In (25) the BP *novels* can be assigned a generic reading in order to express that Esther reads a novel whenever she comes across one.

The ambiguity approach has the advantage that the quantification is no longer dependent on the properties of the verb, as is the case in Carlson's original theory. It is thus possible to map the BP more freely in case of ambiguity. In addition, a unified explanation for the behaviour of singular indefinites and BPs has been made possible (Dayal 2003: 71). However, Carlson (2003: 154) stresses that these benefits are gained at the expense of a unified analysis of BPs.

### 3.2.3. The Neocarlsonian Approach

The discussion of the kind-reference analysis of bare NPs and the Ambiguity Approach showed that there are two ways to interpret BPs. While Carlson's theory states that BPs uniformly refer to kinds, the second approach holds that BPs are either kind-referring or indefinites (Krifka 2003: 184). Advocates of the Neocarlsonian approach (Carlson 1989; Chierchia 1998; Dayal 2004) combine elements of both methods and are particularly concerned with a cross-linguistic investigation.

Carlson (1989) revises his original theory in the late 1980s since his unified analysis of the English BPs faces some problems. Most importantly, it fails to account for the relationship between, on the one hand, the existential and universal readings of BPs, and, on the other hand, the generic and non-generic nature of the sentence. Moreover, Carlson is concerned with the syntactic/semantic (and pragmatic) context of bare NPs, which determine the interpretation of BPs.

A key variant of the Neocarlsonian approach has been developed in Chierchia (1998) in a study dealing with the syntax-semantics mapping across languages (see 3.1.2.1.). Chierchia follows Carlson's original theory by assuming that BPs uniformly denote kinds. However, they can be turned into indefinites whenever they serve as arguments of object-level predicates. This process requires a type shifting operator called 'Derived Kind Predication' (DKP) which applies on demand (Chierchia 1998: 364f).

### 3.2.4. Tests for NP genericity

For the empirical part, it will be essential to distinguish generic sentences of any type from nongeneric sentences. Since characterizing sentences and kind-referring NPs might not always be recognized as such, some linguistic tests developed over the last decades will assist in classifying the corpus data. The key is to differentiate, on the one hand, between characterizing and particular sentences and, on the other hand, between kind-level and object-level predicates.

#### *3.2.4.1. Characterizing vs. particular sentences*

To distinguish between characterizing and particular sentences Chierchia (1995: 177ff) summarizes six key properties that are crucial for the characterization of I-level predicates.

**Stable states**

First, I-level predicates express 'stable' states. Statives that express 'transient' or 'episodic' states (e.g. being drunk, sick, etc.), as well as locatives (e.g. being on the roof), are thus classified as S-level predicates. The distinction between 'transient' and 'stable' states, however, is not always clear-cut. One possibility of distinguishing between them is to look at their behaviour with temporal adverbs. The example of (26a), which involves an S-level predicate, clearly allows the use of *yesterday, last month* and *one year ago*, in contrast with the I-level sentence in (26b), which does not seem compatible with temporal modifiers. While *being drunk* expresses a single state, *being tall* is regarded as a permanent state.

(26) a. John was drunk yesterday/last month/a year ago.
b. ??John was tall yesterday/last month/ a year ago.

(Chierchia 1995: 177)

It is important to add, though, that some adjectives seem to belong to both classes. The adjective *being sick*, for example, can be classified as stable with the meaning of 'chronically sick' or as transient in case of reference to a temporary illness.

Apart from temporal modifiers, the use of progressive aspect in combination with I-level predicates results in ungrammaticality, or rather, characterizing sentences transformed into the progressive lose their generic character. Thus, the characterizing reading of the examples in (27) is clearly lost in (28) (Krifka *et al.* 1995: 12).

(27) a. The Italian drinks wine with his dinner.
b. An Italian drinks wine with his dinner.
c. Italians drink wine with their dinner.
d. Luigi drinks wine with his dinner.
(28) a. The Italian is drinking wine with his dinner.
b. An Italian is drinking wine with his dinner.
c. Italians are drinking wine with their dinner.
d. Luigi is drinking wine with his dinner.

(Krifka *et al.* 1995: 12)

**Locative modifiers**

Secondly, the insertion of a locative modifier helps to differentiate between I-level and S-level predicates. In general, I-level predicates cannot be modified by a locative since the location of I-level predicates is unrestricted. This fact accounts for the oddity of the sentences in (29). The compatibility of S-level predicates with locative modifiers in (30), on the other hand, can be attributed to the fact that they are located in space (Chierchia 1995: 178).

(29) a. ??John is a linguist in his car.
b. ??John is intelligent in France.
c. ??John knows Latin in his office.
(30) a. John is always sick in France.
b. John works in his office.

(Chierchia 1995: 178)

**Perception sentences**

Thirdly, the use of I-level predicates as complements of perception verbs is impossible:

(31) a. *I saw John a linguist.
b. *I saw John tall.
c. *I heard John like Mary.
(32) a. I saw John drunk.
b. I heard Mary beat John.

(Chierchia 1995: 178)

**Existential sentences**

Similarly, the existential construction with *there* seems to single out I-level predicates. The examples of (33) prove that the coda position of *there*-sentences allows only S-level adjectives.

(33) a. There are two men drunk/ sick/ available…
b. ??There are two men intelligent/white/altruistic…

(Chierchia 1995: 178)

**Bare plurals**

Another property of I-level predicates is that they select the universal interpretation of BPs. However, this rule does not always apply. According to Chierchia (1995: 180), BP subjects of I-level predicates can, in some cases, be interpreted existentially due to certain grammatical relations involved. Two rules summarize the distribution of universal readings of BPs (taken from Chierchia 1995: 180).

(i) The bare plural subject of non-unaccusative i-level predicates must be interpreted universally.
(ii) The bare plural subjects of i-level unaccusatives and passives, as well as other bare plural arguments of i-level predicates, can be interpreted existentially.

The second rule accounts for the existential interpretation of the BP subjects in (34). The reading of (34a) seems to be that there are ponds that belong to this property. The same applies to (34b).

(34) a. Ponds belong to this property.
b. Counterexamples to this claim are known to me.

(Chierchia 1995: 180)

**Adverbs of quantification**

The last property of I-level predicates concerns their interaction with adverbs of quantification. Chierchia illustrates the different behaviour of I-level and S-level predicates:

(35) a. ??When John knows Latin, he always knows it well.
b. ??When John is intelligent, he is always pleasant.
(36) a. When John speaks Latin, he always speaks it well.
b. When John is drunk, he is always obnoxious.

(Chierchia 1995: 180)

While the I-level sentences in (35) sound odd, the S-level sentences in (36) seem to be well-formed. However, the I-level sentences become grammatical if the NPs in the when-clauses are replaced by indefinites or BPs, as the examples in (37) show.

(37) a. When a Moroccan knows French, she knows it well.
b. When a student is intelligent, it is a pleasure to work with him or her.

(Chierchia 1995: 180)

In the absence of a when-clause, a similar pattern can be observed:

(38) a. John always speaks French
b. ??John always knows French
c. A Moroccan always knows French.
d. Moroccans always know French.

(Chierchia 1995: 181)

Thus, the examples in (38) clearly show that sentences involving an I-level predicate and an adverb of quantification require an indefinite or a BP as argument.

The test which Krifka *et al.* propose involving adverbs of quantification goes in a slightly different direction. Here, characterizing and particular sentences are combined with adverbs such as *usually* and *typically*. If the original sentence is characterizing as in (39), then the inserted adverb merely emphasises the fact that there may be exceptions to the rule which the sentence contains. If the original sentence is particular, however, the change in meaning is more drastic. While the original sentence generally reports a specific event or a fact, the new sentence expresses a general rule, as shown in (40).

(39) a. A lion has a bushy tail.
b. A lion usually has a bushy tail.
(40) a. A lion stood in front of my tent.
b. A lion usually stood in front of my tent.

(Krifka *et al.* 1995: 9)

#### *3.2.4.2. Kind-level predicates vs. object-level predicates*

Having discussed the six key properties of I-level predicates, the focus now shifts to the distinction between kind-level and object-level predicates. Zamparelli (2002: 309ff) and Krifka *et al.* (1995: 9ff) summarize a variety of tests which have been developed in the literature.

To start with, kind-level predicates cannot refer to ordinary individuals. This explains the contrast between (41a) and (42). Moreover, they may be episodic (Zamparelli 2002: 309).

(41) a. Domestic dogs evolved from jackals/ appeared 100000 years ago/ will become extinct/ become bigger and bigger as you look back into the history of civilization/ have increased by 4% throughout the country

b. Light-bulbs were {created/ invented/ perfected/ analysed} by Edison.

(42) #Fido evolved from jackals/ appeared 100000 years ago/ will become extinct/ becomes bigger and bigger as you look back into the history of civilization/ has increased by 4% throughout the country.

(Zamparelli 2002: 309)

Next, the BP subject of kind-level predicates may be replaced by a definite singular NP but not by an indefinite NP, as the sentences in (43) show (Krifka *et al.* 1995: 10).

(43) a. The lion will become extinct soon.

b. Lions will become extinct soon.

c. *A lion will become extinct soon.

(Krifka *et al.* 1995: 10)

Moreover, Zamparelli (2002: 310) underlines the taxonomic properties typical of kinds. Thus, only kinds can be divided into more specific subkinds, as can be demonstrated by the example of *dogs*, which are *mammals*, which are in turn *animals*.

(44) a. Dogs {have several breeds/ come in many sizes/ have very diverse subkinds}

b. *Fido and Spotty {have several breeds/ come in many sizes/ have very diverse subkinds}

(Zamparelli 2002: 310)

Another popular test for kindhood hypothesizes that quantitative predicates cannot be combined with overt determiners, as shown in (45) (Zamparelli 2002: 323).

(45) a. *{All/most/those/the} dogs are {common/rare/abundant/widespread/scarce}.

b. *{Twenty/ many/ some/ one} dog(s) is/ are {common/ rare/ abundant/ widespread/ scarce}.

(Zamparelli 2002: 323)

Last but not least, certain predicates require kind-referring terms in object position. For example, the object arguments of *invent* and *exterminate* must be kinds as only kinds (and not objects) can be invented or exterminated (Krifka *et al.* 1995: 10f).

# 4. THE DEVELOPMENT OF THE DETERMINER SYSTEM

## 4.1. Grammaticalization, generative grammar, and the linguistic cycle

Up to this point, this thesis has mainly focused on the analysis of NPs on a synchronic level. Since the empirical part is concerned with the diachronic development of bare nouns, it seems all the more important to outline the generative approach to language change, which differs significantly from other approaches. As mentioned in the introduction, Noam Chomsky has conducted ground-breaking work in the field of generative grammar over the last decades. The application of the generative approach to language change is usually attributed to David Lightfoot (1979).

In recent years, a focus in historical linguistics has been placed on investigating the relation between morphological and syntactic changes in the English language. In this context, researchers have often dealt with grammaticalization as a root cause of language change (Kemenade & Vincent 1997: 1). Grammaticalization, in general, describes a phenomenon through which an originally independent word transforms into a grammatical element (Meillet 1921). It is important to recognize, though, that grammaticalization theory does not provide a full explanation of diachronic processes: "Grammaticalization does not address the question of what is occurring in the underlying grammar of individuals who use new patterns and in that of older speakers who do not" (Ackles 1997: 41). These questions are key issues in the generativist approach to language change.

Both the generativist and the grammaticalization approach address interesting questions, but the priorities of the two types of study differ greatly from one another. To summarize:

> If language is the genetically determined language faculty of the brain as it is developed in an individual in response to (early) life experiences (Chomsky's I-language), study of communication, discourse, and socio-historical change is of little importance. If instead, language is defined as an unplanned but structured system which arises as humans use their biological endowment to interact with one another (an invisible hand definition), study which tries to identify the nature of the biological endowment cannot be central (Ackles 1997: 43).

Despite the different approaches, there have been efforts in the past to integrate the two theories. Ackles (1997: 44) introduces one possibility to profit from the insights gained through the two theories. Whereas historical research, thus also grammaticalization research, supplies evidence of parametric differences in languages, generative linguistics supplies a model for analysing the way in which linguistic changes become encoded in the language faculty.

### 4.1.1. Reanalysis

For the development of the determiner system 'reanalysis' plays a crucial role, which is also considered as the "key mechanism of change" in the generative tradition (Kemenade & Vincent 1997: 2). However, the meaning of this term is rather ambiguous as numerous scholars have used it in different senses.

A look at the history of the term 'reanalysis' reveals the existence of a variety of definitions. In the early 20th century, the French linguist Antoine Meillet (1921) seems to have used the term as a synonym for 'grammaticalization'. In recent years, however, Lightfoot (1979) and other scholars apply the concept of reanalysis very differently from the concept of grammaticalization as it is used by modern grammaticalization theorists such as Hopper & Traugott (1993).

The generative interpretation of reanalysis is reflected in Langacker (1977) who defines reanalysis

> as change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation. Reanalysis may lead to changes at the surface level, (…) but these surface changes can be viewed as the natural and expected result of functionally prior modifications in rules and underlying representations (1977: 58).

Since existing structures of utterances are reorganised in a way that the change is not visible on the surface, reanalysis is referred to as "a hidden change" (Kemenade & Vincent 1997: 4). An example of reanalysis is the recategorization of English modals shifting from full lexical verbs in Old English (OE) and Middle English (ME) to a separate category of modals in ModE (Lightfoot 2006: 27-32).

### 4.1.2. Problems with the generative approach to change

A problem with generative work on change is that emphasis is placed on abrupt change. On the surface, however, language change is gradual. Generative syntacticians have thus developed different solutions to account for long-term change from a generative perspective. Anthony Kroch (2005), for instance, claims that during a period of change, speakers acquire two grammars of their native language. Hence, for a limited time the old form and the new form co-exist. Eventually, the old form is not used anymore, leaving the innovative form as the only option. This illustrates that syntactic change goes through various phases: "Syntactic change, indeed change in general, follows an S-shaped curve, taking off slowly, booming

after it has passed a certain threshold level, and petering out slowly when nearing completion" (Kemenade & Vincent 1997: 5). Even though this approach provides a conclusive argumentation in terms of describing the time course of change, it disregards decisive factors such as the causality of change or the crucial trigger for the emergence of innovative forms (Kemenade & Vincent 1997: 5).

In terms of the causation of change, different explanations are offered. Lightfoot (1991) relates the output of the older generation to the grammar of the new generation and claims that language change is triggered by robust linguistic data that comes from the language environment. By robustness Lightfoot means simple and unembedded material. Clark and Roberts (1993), on the other hand, relate language change to preferences inherent in the system. In a similar manner, Henry (1997) and Kiparsky (1997) view language change as the consequence of preferences. According to Henry, the language learner can ignore available evidence and select from a set of alternatives offered by UG. Kiparsky considers change as a push-pull-mechanism where (strong) preferences may replace (weak) evidence.

### 4.1.3. Cyclical change and grammaticalization

Elly van Gelderen (2011: 3ff) provides a somewhat unusual generative perspective on linguistic change. Contrary to Lightfoot, who claims that external factors such as input are the source of change, Gelderen is convinced that change comes above all from the inside or, more precisely, "internal principles (…) bias the learner toward certain structures" (2011: 4).

Moreover, Gelderen claims that language change is a cyclical change. A cycle is in general a period of time in which a certain sequence of events or phenomena is observed. At the end of a cycle, similar phenomena emerge again, proceeding at a different pace, though. From this point of view, linguistic change is unidirectional. Crucially, cyclical change provides a new perspective on the principles dominating the language faculty. In detail, this means that Economy Principles or, more precisely, Feature Economy, present in the UG of the child, is responsible for linguistic cycles: "semantic features become grammatical features, which in turn need semantic ones again" (Gelderen 2011: 4). Feature Economy does also account for grammaticalization according to Gelderen (2011: 19). The changes that grammaticalization involves are illustrated below.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| a. phrase | > | word/head | > | clitic | > | affix | > | O |
| b. adjunct | > | argument | > | (argument) | > | agreement | > | O |

While (a.) shows the morphosyntactic changes, (b.) represents the changes in argument status. As soon as the right side of these lines is reached, new words and phrases are generated through renewal and borrowing and the grammaticalization process starts again. This means that cyclical change takes place (Gelderen 2011: 6).

Corpus-based studies have shown that the effects of grammaticalization can be observed far earlier in spoken than in written English. Specifically, this means that older forms are used much longer in formal writing than in spoken language while, at the same time, newly grammaticalized forms are accepted more slowly in written than in spoken English (Leech *et al.* 2009: 239). Anthony Kroch (2005: 3) goes even further by describing written language as conservative and resistant to change.

Having discussed some theoretical issues associated with language change, the focus now shifts to a concrete example of language change, namely article development in English.

## 4.2. English determiners – From OE to PdE

The English language has changed significantly since the OE Period. Due to the emergence of new words, the disappearance of old ones and the change in meaning of words, but also due to morphological and syntactic changes, OE (from the first Anglo-Saxon settlements in England to about 1100) or ME (1100 to about 1500) appears to be a different language from the English spoken today (Barber 1993: 38f).

And yet, no great changes can be observed in terms of the structure of the NP. With few exceptions, the basic pattern Determiner-Adjective-Noun has remained the same since the OE Period. Exceptions are that OE expressions like 'all' (*eall*), 'both' (*bēgen*) and the adjective ending in *–weard* precede the determiner, plus, adjectives and determiners may follow the noun. Moreover, demonstratives and possessives may co-occur in OE and even beyond, until the 17$^{th}$ or 18$^{th}$ century (Barber 1993: 120; Wood 2007: 339ff).

Even though the structure of the NP has hardly changed from OE to PdE, significant differences in terms of the system of determiners can be observed. A closer examination of the history of English reveals that definite and indefinite articles have not always been an integral part of the English language (Philippi 1997: 62; Gelderen 2011: 210). In this section, established views on the development of definite (*the, this*) and indefinite (*a(n), some*) determiners will thus be discussed.

### 4.2.1. Definite determiners

There is no general consensus among linguists on the question of when the definite article first appeared in the English language. Despite claims that OE does not have a definite article, scholars such as Crisma (2011: 175f) and Sommerer (2011) argue that in OE the definite article exists already, albeit not formally distinct from the demonstrative morpheme *se,* which is either used as a demonstrative or in an article-like manner. Hence, the debate over the emergence of the article appears to be a matter of terminology.

OE demonstratives play a decisive role in case-distinction as OE nouns do not necessarily show distinctive endings. The two OE demonstratives *sē* ('the, that') and *þes* ('this') are declined according to three genders, five cases and two numbers. Hence, there is a large variety of forms compared to the five remaining demonstratives in PdE, (*the, this, these, that, those*) (Barber 1993: 120). In addition, OE demonstratives function as definiteness markers, but their use is restricted to empathic contexts. The distinction between definite and indefinite NPs in OE is thus generally indicated by structural case markings.

In the ME Period, the masculine nominative form *sē* is reanalysed as the definite article *the* and assumes the function of definiteness marking as a result of the reduction of the inflectional system and in particular the loss of the verbal genitive. Consequently, the restrictions on the use of the lexical definiteness marker are lost. In ME, it must be used in indirectly anaphoric contexts and is optionally used with abstract and generic NPs (Philippe 1997). In addition, the neuter nominative/accusative *þæt* is reanalysed as the demonstrative *that* with a meaning differing from the definite article. The plural demonstrative *those* develops from the nominative/accusative plural form *þa* (Gelderen 2011: 210f).

In EModE, the definite article *the* remains invariable. In contrast to PdE, however, the article is occasionally spelt *th* or *th'* and pronounced [ð] before vowels. In the 16th century, this form of the definite article appears across all literary genres, whereas in the 17th century, its use is more restricted. Due to metrical purposes, it is mainly found in dramatic dialogue and poetry. In addition, forms such as *ith* 'in the', *ath* 'of the', 'on the' are placed before vowels and consonants in dramatic dialogue.

As for the use of the definite article, some minor changes can be observed over the past 500 years. On the one hand, in EModE, it is sometimes found in positions in which it is omitted in PdE, for example before titles, vocative expressions and the names of branches of learning, arts, crafts, games and pursuits. Moreover, the definite article is inserted before the names of diseases and, occasionally, before the names of parts of the body when the possessive is used

in PdE. Finally, expressions such as *the death, the life* and *the heaven* and phrases such as *at the last* or *at the least* are common in EModE. On the other hand, the definite article in EModE is sometimes omitted in positions where it is used today, for instance, before river names and in numerous preposition phrases such as *at door, by help of* and *in presence of* (Barber 1997: 159f).

The demonstratives *this/these* and *that/those* were already used in their present meaning in the beginning of the EModE period. In addition, the form *tho*, an alternative to *those*, existed at that time, but it was seldom used. In one aspect, the system of demonstratives has nevertheless changed. Whereas in PdE a binary system exists (*this/these* vs. *that/those*), EModE has a ternary system consisting of *this/these*, *that/those* and *yon*/*yond(er)*. While *this* means 'near the speaker' and *that* 'remote from the speaker', *yon* is used in the sense of 'remote from both speaker and hearer'. Even though *yon/yond(er)* slowly disappeared from most text genres and from the speech of the upper classes by the end of the 17th century, it was still used in poetry and by the lower classes as well as in Scots and in a number of rural English dialects (Barber 1997: 161-163).

In order to put the above observations into the generative framework of this thesis, it can be supposed that the definite article develops from a demonstrative and eventually becomes a non-generic marker as Gelderen's DP cycle below predicts.

| | | |
|---|---|---|
| a. demonstrative | > definite article | > Case/non-generic |
| b. specifier | > head | > affix |
| c. i-F | > u-F | > (u-F) |

(Gelderen 2011: 201)

Based on the assumption that the demonstrative occupies the specifier position and the article the head of DP (see 2.2.2.), the development of the article involves the reanalysis of the demonstrative as head whereby the deictic character of the demonstrative is lost during this process. Additionally, interpretable features (i-F) reanalyse as uninterpretable features (u-F) (Gelderen 2007: 278f; Gelderen 2011: 197ff). The first two stages of this process are visualized in Figure 8.

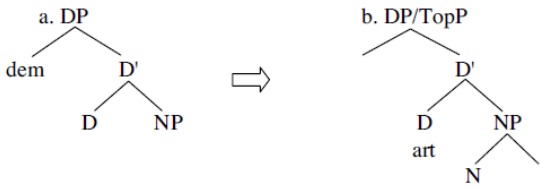


**Figure 8 Reanalysis of the demonstrative as head (Gelderen 2007: 288)**

### 4.2.2. Indefinite determiners

The emergence of the indefinite article is less controversial than the evolution of *the*. It is generally agreed that *a(n)* developed from the unstressed form of the numeral *one* and becomes obligatory as indefiniteness marker in EModE (Philippi 1997: 63).

Crisma (2015: 84) describes the development of *a(n)* in three different stages. In early OE, *a(n)* functions as cardinal corresponding to PdE *one*; a usage which exists up until the EModE period (Barber 1997: 160). Later in the OE period, *a(n)* used as an existential operator can be observed. Here, it can mark specificity and acts as the singular counterpart of PdE *sm*. Its use is obligatory when a noun takes wide scope or receives a specific interpretation. Finally, in ModE *a(n)* functions as an indefinite article, comparable to an expletive. Bare singular count nouns are no longer grammatical. Besides this, *a(n)* can be combined with generics.

Breban (2012: 273ff) focuses in particular on those functions of OE *a(n)* that are not inherent in the modern article. Crucially, in OE *a(n)* does not occur in all indefinite NPs as it generally acts as a presentative maker. It is thus restricted to NPs denoting referents which are introduced into the discourse and will have an important role therein. Those referents are 'specific' (as opposed to 'arbitrary') and 'persistent', i.e. they are repeatedly mentioned in the discourse. The restricted occurrence of *a(n)* implies that NPs with an unidentifiable referent appear without a determiner in OE. The development from presentative marker *a(n)* to unidentifiability marker is due to the loss of its original features. Around the 13$^{th}$ century, the occurrence of *a(n)* is extended to any NPs which denote a discourse-new referent. At the same time, the indefinite article *a(n)* and the numeral *one* clearly appear as two different elements.

The development of the quantifying determiner *some* is, in contrast to the indefinite article *a(n)*, more disputed. A central point of debate is whether quantifiers developed from adjectives, as argued by A. Carlson (1976: 14f) and Lightfoot (1979: 168f) who assume that OE adjectival quantifiers were reanalysed as a separate category of quantifiers in the late 16$^{th}$ century. According to Fischer and Leek (1981: 311ff), though, the inflection and distribution of these elements are evidence of their status as a separate category in OE.

Interestingly, *sum* – the OE form of PdE *some* or *sm* – competes against *a(n)* as presentative marker with singular count nouns in the OE period. Only from the 10$^{th}$ century onwards, the occurrence of *sum* is restricted to NPs with plural and generic referents. The function of *sum* as a presentative marker gets lost in the transition from ME to EModE. Instead, it functions as a marker of specific unidentifiable referents (Crisma 2015: 86: Breban 2012: 285f).

In EModE, *some* is used particularly frequently to mark specific reference when the referent is associated with special characteristics that set it apart from others of the same type, as in (46):

(46) hit shold, as by the Quenes Lettre appereth, have bene th'occasion of *some great and good effecte* [original emphasis].

(PPCEME 2004, *Letter from Sir Thomas More to Cardinal Wolsey*, cited in Breban 2012: 286)

(47) *Some instinct, some strange warning*, sent the sleeper on the bed flying from it, dazed as she was [original emphasis].

(CLMETEV 2006, *The Happy Foreigner*, cited in Breban 2012: 287)

In (46) it is not just any effect but one with the specific properties 'great' and 'good'. Also, in (47) *some* is used to express certain features linked to this instinct. A similar usage of *some* can be observed in PdE.

In LModE, a new function of *some* is observable. In (48) the presence of *some* means that the addressee needs to pick one referent from a set, no matter which one. Thus, in (48) any mountain-top can be chosen. In sharp contrast to the examples in (46) and (47), *some* in (48) conveys emphatic non-specificity as to the set of referents (Breban 2012: 287).

(48) when the character of the country is scanned by a stranger *from some mountain-top*, the very act of traversing it appears impossible [original emphasis].

(CLMETEV 2006, *Eight Years' Wanderings in Ceylon*, cited in Breban 2012: 287)

Despite the fact that *some* can mark both specificity and non-specificity, Breban (2012: 287) refuses the idea of a semantic equivalence to *a(n)*. One difference between the two determiners is that a NP preceded by *a(n)* is ambiguous between a specific and a non-specific reading. A NP with *some*, on the other hand, can yield an emphatically specific or non-specific interpretation. According to Breban (2012: 287), *a(n)* conveys 'unidentifiability' in singular NPs, whereas *some* conveys 'specificity'. Another function of *some* is to mark choice between referents from "*a set* as non-specific 'any of these'", non-specific *a(n)*, on the other hand, means "any referent of *the type*" [original emphasis] (2012: 287).

As with the definite determiner, Gelderen (2011: 202) describes the reanalysis of the numeral as indefinite marker in terms of a cyclical change whereby the syntactic change depends on the question of which structural slots numerals and indefinite articles occupy. In terms of Feature Economy, Gelderen states that interpretable features (in this case number) are lost.

| | | |
|---|---|---|
| a. numeral | > indefinite pronoun | > indefinite article |
| b. Q/A | > D | > zero |
| c. i-F | > u-F | > zero |

(Gelderen 2011: 202)

## 4.3. The determiner phrase in the history of English

So far, it has been assumed in this thesis that despite the lack of dedicated definite and indefinite articles, there exists a DP structure in OE with the demonstrative occupying the specifier position. However, this hypothesis is disputed. There has been considerable disagreement among linguists regarding the question of whether DP already existed in OE, or whether a new functional structure was developed in the course of the history of English.

The key argument against the interpretation of OE nominals as DPs is the apparent lack of prenominal structure in OE. According to Yamato (1989), prenominal elements in OE were adjectives, hence lexical words, and their word order differs significantly to PdE by allowing double or even multiple determination on a single head noun in NP structure. In the course of the ME period when the articles as a category emerged, the number of determinerless NPs decreased rapidly. Yamato (1989: 14) suggests that these changes triggered the establishment of the functional category D. In sum, he postulates a parameter shift from [-D] to [+D] which accounts for contrasts between OE and PdE.

In a similar way, Osawa (2000) argues that the English NP developed into DP via the emergence of the D-system. The decisive factor is the R(eferential)-role of N which is bound by morphological case in OE in order to turn predicate nominals into arguments. The demise of morphological case in English triggered the introduction of the functional category D to bind the R-role, thus changing the status of nominal projections from NP to DP.

Despite the word order evidence in favour of the interpretation of OE nominals as NPs, Wood (2005) and Allen (2006) refute the theory of the non-existence of DP in OE. They both reject the idea that the various positions in which OE demonstratives and possessives occur indicate that they are freely ordered adjectives. Instead, it is proposed to treat them as different syntactic structures. Crucially, Wood (2005) argues that different combinations of possessive and demonstrative pronouns are possible in OE because the possessive is not yet definite. The possessive thus functions either as an adjective or moves from below the D to the specifier position of DP. The latter implies that the demonstrative is already reanalysed as a D head. The reanalysis of the possessive from indefinite to definite features only occurred in the transition from OE to ME and causes the disappearance of the possessive-demonstrative construction as both elements occupy the head D position (Gelderen 2011: 212f).

In short, the hypothesis that OE nominals are interpreted as DPs implies the existence of a prenominal structure. Crisma (2011: 177f) points out that this theory only applies to arguments. Vocatives and predicates are considered to be NPs (see section 2.2.4.).

## 4.4. Summary and outlook

In the first part of this thesis, some theoretical considerations on the English NP have been presented, which will provide a framework for the empirical part. First, it has been argued that there is a functional projection DP which dominates the lexical projection NP. However, there seems to be no consensus as to whether BPs, in the absence of a determiner, need a DP layer. According to Chierchia, whose argumentation is grounded on semantics, BPs are interpreted as NPs. Longobardi's syntactic approach to the issue proposes, though, that all arguments are DPs. Additionally, differing views on the semantic interpretation of BPs have been presented. Carlson, who initiated the study of BPs, claims that they unambiguously refer to kinds, whereas proponents of the ambiguity approach suggest that BPs are ambiguous between reference to kind and an indefinite reading.

Furthermore, determinants of syntactic change have been discussed in order to gain the knowledge necessary to interpret the historical development of determiners. According to generativists, reanalysis is the main cause of change. It refers to the creation of new patterns, hence to new associations of form and content. In the generative tradition, emphasis is placed on abrupt change. It is only on the surface that languages appear to change gradually, but the grammar of the individual does not gradually change. In terms of trigger and causation, Lightfoot claims that robust evidence from the language environment is a decisive factor for language change, whereas others believe that language change comes in fact from the inside. According to Gelderen, language change is a cyclical change for which Feature Economy is responsible.

Finally, the development of definite and indefinite determiners has been discussed. There seems to be a consensus that the indefinite article developed from numeral *one*, whereas the definite article developed from the demonstrative pronoun *se*. Both determiners already appear in their present form in the ME period. What should be kept in mind is that the use of the articles has greatly changed. In the process of becoming an indefinite article, OE *a(n)* lost its restricted occurrence and discourse-related functions. These functional shifts can likewise be observed with the development of the definite article.

Closely linked with the emergence of the determiner category in English is the functional structure of NPs. Whereas some linguists claim that the lack of the determiner category in OE necessitates the analysis of OE nominals as NPs that only develop into DPs with the evolution of the D-system, others argue in favour of the existence of DP in OE in showing that OE has a prenominal structure. The debate as to whether the development of determiners in the history

of English involves new functional structure will doubtlessly continue. However, this question is not determinative in the present study as the focus in the 2nd part of this thesis is placed on the ModE period.

The theoretical considerations presented in the first part of this thesis have produced important insights that raise new questions as to the development of BPs in ModE. Bearing in mind the development of definite and indefinite determiners in the history of English (see 4.2.), the following hypothesis can be formulated.

In the history of English, a trend towards obligatory marking of definiteness and indefiniteness is observable. In OE, the demonstrative pronoun *se* as marker of definiteness alternates with bare nominals, but from the ME period onwards the use of the definite article has been obligatory as a result of a concomitant process of reanalysis. A similar pattern can be noticed for marking of indefiniteness. The obligatoriness of *a(n)* with bare singular count nouns traces back to the EModE period in which the former numeral *one* attained the status of a true indefinite article.

In the light of these developments, it seems worthwhile to examine the BP construction in ModE. Along with mass nouns and proper names, plural count nouns are part of those nouns that still appear bare in PdE, or else, they are modified by weak *some* or *sm* to mark indefiniteness. However, the observed trend towards obligatory reference marking in English allows for the hypothesis that the frequency of BPs has decreased between EModE and LModE, whereas in the same time period the number of plural nouns headed by *some* has risen.

In order to test this central hypothesis, it is necessary to investigate how frequent the cases of those BPs with an indefinite reading are in EModE and LModE, as only a decreasing number of indefinite BPs is an indication of a trend towards obligatory marking of indefiniteness. The development of non-referential BPs such as generics and non-argumental BPs is thus not relevant for this study, even though they cannot be completely disregarded in the empirical part.

The aim of the following chapters is to verify the above-mentioned hypothesis by examining and analysing the development of BPs in ModE quantitatively and qualitatively. Afterwards, possible explanations for an increasing or decreasing occurrence of BPs in PdE will be presented.

# PART II

# 5. THE CORPORA

Before the results of the empirical investigation are presented, some information on the corpora used in this study will be provided. As discussed in the introductory chapter, this study focuses exclusively on the Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME) and on the Corpus of Historical American English (COHA).

## 5.1. The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)

The Penn Helsinki Parsed Corpus of Early Modern English is part of a project of the University of Pennsylvania and the University of York which aims at producing syntactically annotated corpora for all periods in the development of English. The corpus comprises over 1.7 million words, spread over 448 text samples, each of which can be accessed in three forms: parsed, POS tagged and unannotated text. Due to its syntactic annotation, it is possible to search not only for words and word sequences but likewise for syntactic structures. The texts are split into three successive periods of 70 years each, starting from 1500. Table 2 presents the distribution of word counts over the periods E1, E2 and E3 and across the three subcorpora. Even though it seems as if the texts included in the PPCEME were all published before 1710, the Helsinki Corpus and its supplements contain at least one text from the 1710s (Kroch 2011a; Kroch, Santorini & Delfs 2016).

**Table 2 Word count summary by time period and subcorpus (Kroch 2011c)**

| **Period** | **Helsinki** | **Penn 1** | **Penn 2** | **Total** |
|---|---|---|---|---|
| *E1 1500-1569* | 196,754 | 194,018 | 185,423 | **576,195** |
| *E2 1570-1639* | 196,742 | 223,064 | 232,993 | **652,799** |
| *E3 1640-1710* | 179,477 | 197,908 | 187,631 | **565,016** |
| **Total** | **572,973** | **614,990** | **606,047** | **1,794,010** |

The corpus was originally constructed from the Helsinki Corpus samples of the EModE period. Texts and text types were subsequently extended with the aim of ensuring that each genre, in each time period, was represented, wherever possible, by a sample of roughly 50,000 words (Kroch 2011b). The PPCEME covers a wide range of genres, including literate, religious and administrative texts. The distribution of word counts as related to genres is shown in Appendix I.

The corpus consists of three subcorpora. The Helsinki directories comprise approximately 573,000 words and contain the first supplement to the Helsinki corpus. The Penn1 directories, containing a first supplement to the Helsinki Corpus, consist of roughly 615,000 words. The Penn2 directories, another supplement to the Helsinki Corpus, consist of approximately 606,000 words. As for the supplements, the authors endeavoured to use, where possible, texts written by the same authors and from the same editions as in the Helsinki Corpus. Penn2, though, contains more additional material than Penn1 (Kroch 2011c).

Compared to the Corpus of Historical American English (5.2.), the Penn-Helsinki Parsed Corpus is a rather small corpus. However, small corpora can be particularly useful for examining high-frequency constructions.

## 5.2. The Corpus of Historical American English (COHA)

The Corpus of Historical American English (COHA) is a historical database of roughly 400 million words released in late-2010. It contains approximately 100,000 texts published in the period from 1810 to 2009 and drawn in equal measure from popular magazines, newspapers, fiction and non-fiction (NF) books across all 20 decades. The composition of the corpus can be found in Appendix III. The corpus is annotated for lemma and part-of-speech (Davies 2012: 121f).

Compared to other English historical corpora (the Brown family of corpora, the ARCHER corpus, etc.), COHA stands out in numerous ways. It is 100 times larger than other structured corpora of historical English. In addition, it is well balanced by genre and sub-genre across the decades. As a consequence, the Corpus of Historical American English allows for research on a wide range of topics including changes in lexis, morphology, syntax and semantics (Davies 2012: 121f).

The 100,000 texts in COHA originate from a variety of sources. As Table 3 shows, some texts are part of text archives such as Project Gutenberg and Making of America, others have been

converted from PDF files to text and a number of texts have been scanned from printed works (Davies 2012: 124f).

**Table 3 Sources of COHA (Davies 2012: 125)**

| **Genre** | **Sources** |
|---|---|
| **Fiction** | Project Gutenberg (1810-1930), Making of America (1810-1900), scanned books (1930-1990), movie and play scripts, COCA (1990-2010). |
| **Magazines** | Making of America (1810-1900), scanned and PDF (1900-1990), COCA (1990-2010) |
| **Newspaper** | PDF > TXT of at least five newspapers (1850-1980), COCA, etc. (1990-2010) |
| **NF** | Project Gutenberg (1810-1900), www.archive.org (1810-1900), scanned books (1900-1990), COCA (1990-2010) |

# 6. CORPUS-BASED ANALYSIS

In this chapter, the results of the corpus-based analysis will be presented. They will provide new insights and a better understanding of the development of bare plural NPs during the ModE period. The main aim of this chapter is to investigate whether the trend towards obligatory reference marking, which is particularly evident in the transition period from OE to ME, continues to play an important role in ModE with regard to the BP construction. In order to tackle this question, several studies will be conducted.

First, the frequency of NPs with and without overt determiners will be determined to collect some basic information about the development of the BP construction over the last five centuries (6.1.). The next step is to provide in-depth analysis of the factors leading to these results. For this purpose, syntactic and semantic evidence will be analysed. The large number of BPs in PPCEME and COHA, however, necessitates the creation of a database first (6.2.). A sample of manageable size from both corpora will thus be selected to determine, on the one hand, the role of fixed expressions containing bare NPs (6.2.3.) and, on the other hand, differences in the development of argumental and non-argumental, as well as of kind-referring and indefinite, BPs (6.2.4 & 6.2.5.). The tests for NP genericity summarized in section 3.2.4. will be particularly helpful assisting in distinguishing between generic and non-generic subject BPs. Additionally, the use of bare NPs with existential *there* and in object position will be subject to investigation (6.3. & 6.4.).

To obtain reliable comparative figures, the raw frequency and the frequency per million words (pmw) will be given if required. For the sake of simplicity, the results will be rounded off to two decimal places.

## 6.1. Plural nouns in PPCEME and COHA

First, the frequency of plural nouns in PPCEME and COHA was identified. The results of the search and the rate of change were recorded in Table 4 which shows a significant increase in frequency of plural nouns (+14.44%) between EModE and LModE.

**Table 4 Historical use of plural nouns**

| | PPCEME | | COHA | | Change % |
|---|---|---|---|---|---|
| | raw freq. | pmw | raw freq. | pmw | |
| TOTAL | 68,930 | 38,422.31 | 17,861,469 | 43,968.64 | **+14.44%** |

A possible explanation for the increase in frequency of plural NPs is that nominal structures in writing have increased in general. Biber and Gray (2011: 228) state that, "nouns have increased in use in academic research writing and in newspaper prose over the past three centuries, while their use has remained relatively constant in drama and fiction". Also, Leech et al. (2009: 245ff) identify an increase in frequency of nouns as an ongoing trend in the English language. The scope of this thesis, though, does not allow expanding on this issue.

As a next step, the ratio of NPs with and without overt reference marking in PPCEME and COHA was looked at in more detail. The searches were restricted, so that they would only retrieve nouns in sentence-initial and post-verb position. This way, it could be ensured that mainly argumental NPs in subject and object position as well as predicates would be retrieved from the corpus. Including other NPs in the count, such as objects of prepositions, would have complicated the analysis. All the search queries for this section are given in Appendices V & VI.

Against all expectations, the proportion of plural NPs preceded by overt D has declined strongly from 72.22% to 60.81%, while the proportion of BPs has risen significantly from 27.78% to 39.19% between EModE and LModE (see Table 5). Moreover, Figure 9 provides a visual impression of the differences in the ratio of BPs and NPs with overt D between the two corpora.

**Table 5 The ratio of plural nouns preceded by overt or null D in PPCEME and COHA**

| | PPCEME | | | COHA | | |
|---|---|---|---|---|---|---|
| | raw freq. | pmw | % | raw freq. | pmw | % |
| overt det. | 6,373 | 3,552.38 | 72.22% | 1,460,940 | 3,596.32 | 60.81% |
| BPs | 2,452 | 1,366.77 | 27.78% | 941,461 | 2,317.55 | 39.19% |

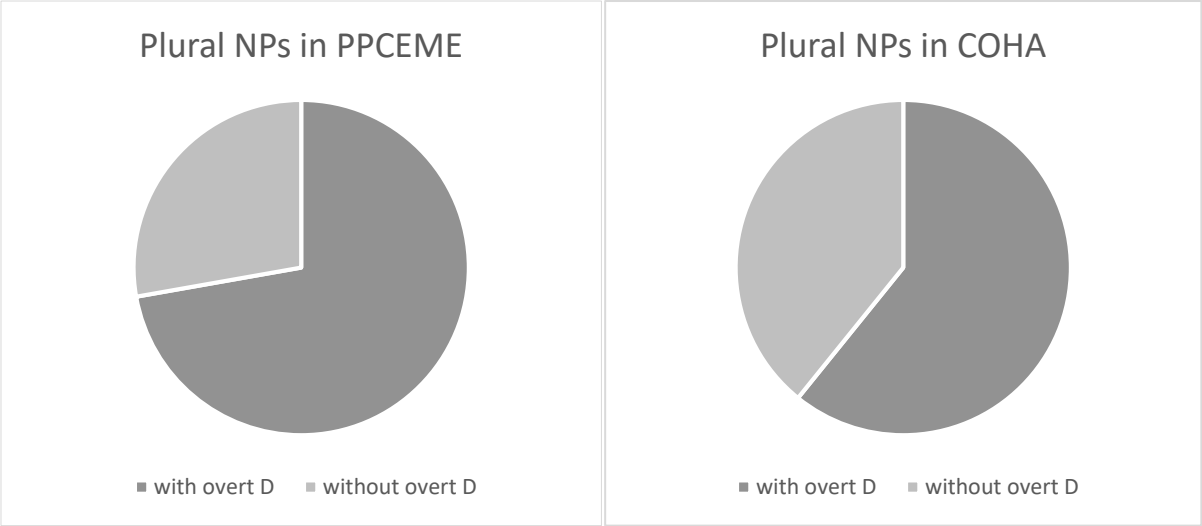


**Figure 9 The ratio of BPs and NPs with overt reference marking in PPCEME and COHA**

To obtain a more accurate picture of the development of BPs and NPs with overt reference marking, the gradualness of the change was traced using further quantitative data. Hence, the frequencies of NPs with and without overt reference marking for the three subperiods E1 (1500-1569), E2 (1570-1639) and E3 (1640-1710) of PPCEME as well as for each decade in COHA were included in Table 6.

**Table 6 The ratio of BPs and NPs with overt D in the subperiods of PPCEME and COHA**

| | overt det. | | bare plurals | |
|---|---|---|---|---|
| | raw freq. | % | raw freq. | % |
| 1500-1569 | 1,876 | 72.54% | 710 | 27.46% |
| 1570-1639 | 2,594 | 72.60% | 979 | 27.40% |
| 1640-1710 | 1,903 | 71.38% | 763 | 28.62% |
| 1810 | 4,858 | 65.00% | 2,616 | 35.00% |
| 1820 | 27,674 | 72.67% | 10,407 | 27.33% |
| 1830 | 53,396 | 68.90% | 24,107 | 31.10% |
| 1840 | 60,612 | 69.92% | 26,080 | 30.08% |
| 1850 | 59,912 | 70.17% | 25,475 | 29.83% |
| 1860 | 59,998 | 66.55% | 30,161 | 33.45% |
| 1870 | 62,673 | 67.06% | 30,782 | 32.94% |
| 1880 | 68,504 | 65.50% | 36,090 | 34.50% |
| 1890 | 72,546 | 67.30% | 35,243 | 32.70% |
| 1900 | 75,454 | 64.96% | 40,703 | 35.04% |

| | | | | |
|---|---|---|---|---|
| 1910 | 78,333 | 64.20% | 43,681 | 35.80% |
| 1920 | 88,021 | 60.72% | 56,951 | 39.28% |
| 1930 | 86,045 | 59.21% | 59,269 | 40.79% |
| 1940 | 88,978 | 58.73% | 62,516 | 41.27% |
| 1950 | 87,982 | 58.30% | 62,925 | 41.70% |
| 1960 | 86,083 | 57.20% | 64,402 | 42.80% |
| 1970 | 85,767 | 56.30% | 66,577 | 43.70% |
| 1980 | 96,441 | 56.05% | 75,614 | 43.95% |
| 1990 | 106,204 | 53.73% | 91,448 | 46.27% |
| 2000 | 111,459 | 53.62% | 96,414 | 46.38% |

Figure 10 demonstrates that the percentage of BPs within all Pl. NPs in PPCEME and COHA increases almost steadily: from the first subperiod, where the share of BPs amounts to 27.46%, to the last period with 46.38%. In the EModE period as well as in the 19th century, the percentage of BPs fluctuates between 27% and 35%. The spike in the period from 1810 to 1819 does not necessarily present a significant change as only a relatively small number of texts are available from this period compared to the other subperiods of COHA. (see Appendix III, The composition of COHA) In the 20th century, the gradualness of the change is immediately noticeable since it is during this period that the share of BPs increases constantly from 35.04% to 46.38%.

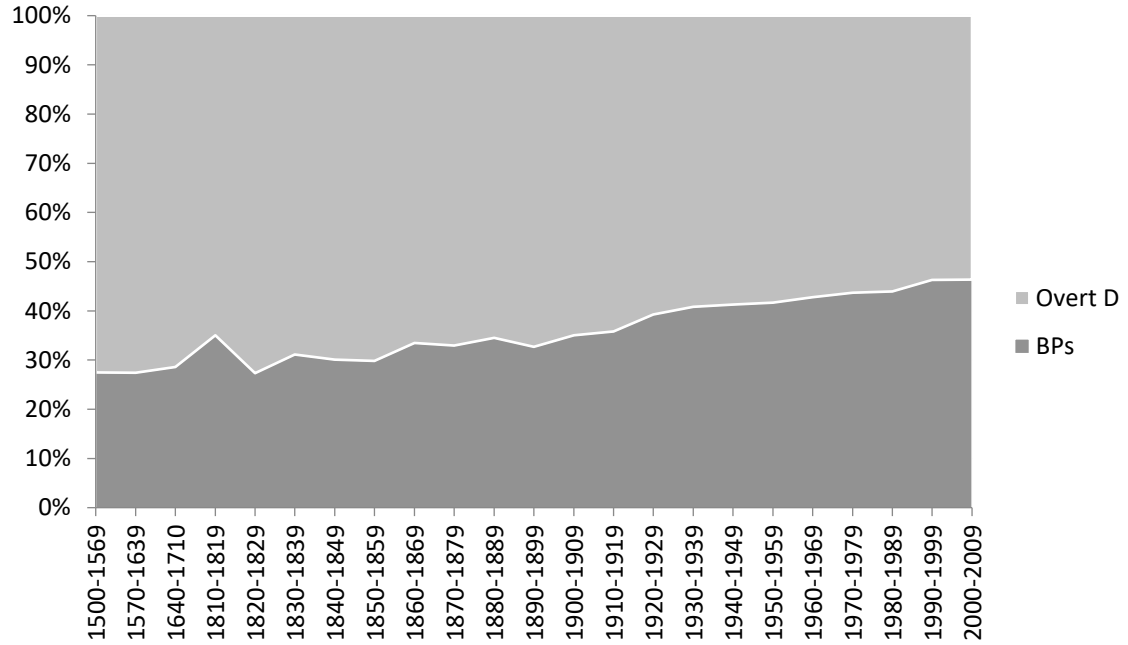


**Figure 10 The ratio of BPs and NPs with overt D between 1500 and 2009**

In sum, in this section it has been argued with the help of the PPCEME and COHA that - in relation to the overall number of NPs - the proportion of BP phrases has risen almost constantly, while the proportion of NPs with overt reference marking has dropped over the last 500 years.

This observation clearly does not comply with the original hypothesis of a trend towards obligatory reference marking in English. The result can nevertheless not be interpreted as direct evidence of an increasing use of BPs to express indefiniteness in ModE. For example, it could be the case that the increased frequency of bare nominals results from the fact that the share of non-referential BPs has risen, while the share of indefinite BPs has fallen or remained constant. Moreover, the role of fixed expressions involving BPs needs to be investigated as a potential increase in fixed expressions would have an impact on the overall number of BPs. Thus, the results presented in this section are only a first look into the diachronic development of BPs in ModE.

In the following sections, the development of BPs shall be examined in greater detail by viewing the BP construction on a more differentiated basis. In analogy to the theoretical framework developed in PART I of this thesis, some syntactic and semantic aspects of BPs in ModE will be investigated. More specifically, this means that differences in the development of argumental and non-argumental BPs and of indefinite and kind-referring BPs will be analysed. This qualitative analysis should shed light on the development of indefinite BPs and of plural NPs headed by *some* within the ModE period and thus give new insights into the significance of indefiniteness marking with plural nouns in PdE. Additionally, the diachronic development of fixed expressions containing BPs needs to be taken into account since idiomatic constructions cannot be equated with BPs as they are understood in this thesis.

## 6.2. Bare Plurals in PPCEME and COHA: a database

To assess differences in the distribution and interpretation of BPs between PPCEME and COHA by means of a qualitative analysis, it was necessary to establish a database first, as it would have been unrealistic to consider a total of 943,913 BPs in the analysis. To select a random sample of manageable size from both PPCEME and COHA and to compare the results of the syntactic and semantic analysis has proven to be a feasible strategy.

### 6.2.1. Criteria for collecting the sample

In the present study a selection of plural common nouns serves as sample of the entire collection of NPs listed in Table 5. To ensure the representativeness of the sample, as well as the feasibility of the analysis, the following criteria were followed in the selection of the nouns.

a) high frequency common noun in both PPCEME and COHA
b) manageability of the number of hits
c) usage of the noun not restricted to idiomatic contexts

In the end, the nouns *books* and *horses* were chosen. They are both high frequency words in PPCEME and COHA, but the number of hits is still manageable for analysis. Crucially, in the EModE data the different spelling variations of the two nouns had to be taken into account. *Books* is also spelt *bookes*, *boks*, *bokes* and *bockes* in PPCEME. Likewise, *horses* is sometimes spelt *horse* or *horsses*. Both nouns are used in a variety of contexts whereby idiomatic contexts are not predominant. It is important to note that, as in the previous section, the searches were restricted, so that they would only retrieve nouns with/without overt D in sentence-initial position and in post-verb position (see Appendix VII for a list of search queries).

**Table 7 The ratio of BPs and NPs with overt D in PPCEME and COHA**

| | PPCEME | | | COHA | | |
|---|---|---|---|---|---|---|
| | all | books | horses | all | books | horses |
| overt det. | 6,373<br>72.22% | 28<br>65.12% | 31<br>73.81% | 1,460,940<br>60.81% | 3,592<br>50.36% | 5,354<br>73.42% |
| BPs | 2,452<br>27.78% | 15<br>34.88% | 11<br>26.19% | 941,461<br>39.19% | 3,541<br>49.64% | 1,938<br>26.58% |

What is striking about the ratio NPs with overt D and BPs is that the share of BPs increases both with *books* (from 34.88% to 49.64%) and with *horses* (from 26.19% to 26.58%), as shown in Table 7. This is in line with the overall numbers, although the share of BPs with *horses* in COHA is unexpectedly low (26.58%).

### 6.2.2. Sorting out the sample

For the analysis, only the bare uses of *books* and *horses* were taken into account. However, the sample of 5,505 BPs (26 in PPCEME and 5,479 in COHA) was not yet the definitive one. The elimination of irrelevant material still needed to be carried out, which resulted in the exclusion of the following items from the sample:

a) sentence fragments and grammatically incorrect sentences
b) sentences with unclear meaning due to lack of context
c) repeated occurrences of the same text passage
d) citations from other time periods
e) erroneously tagged bare nouns
f) bare nouns in genitive case or in indirect object position

ad a) In order to analyse BPs following Carlson (1977), the presence of a subject and predicate is indispensable. Sentence fragments, found, for instance, in stage directions and movie scripts, were thus excluded. Also, grammatically incorrect sentences that complicate the analysis were eliminated.

ad b) The interpretation of isolated sentences can be problematic. Sentences are often ambiguous between different interpretations so that sufficient context is needed to analyse them correctly.

ad c) Occasionally a sentence from a text appears twice in the search results in COHA for no obvious reason. The duplicate sentences were thus removed.

ad d) Since the time periods that are investigated are clearly defined (1500-1710 and 1810-2009) it is obvious that textual evidence from other time periods was not taken into account in the analysis. This is especially the case with texts that contain citations from older texts.

ad e) The automatic POS-tagging of corpora naturally causes tagging errors. In the discussion about overall figures, erroneous tags could not be eliminated due to the large set of hits. A sample of 5,505 tokens, however, allows to identify any errors in the POS-tagging.

ad f) In order to facilitate the analysis, only bare nouns in subject, direct object or predicate position were included in the final sample.

This way the sample was reduced to 4,641 tokens, as the following table shows.

**Table 8 The sample of BPs from PPCEME and COHA**

| | PPCEME | | COHA | |
|---|---|---|---|---|
| | books | horses | books | horses |
| BPs | 13 | 11 | 2,980 | 1,637 |

### 6.2.3. Fixed expressions

One important observation made during the creation of the sample concerns the role of fixed expressions that contain bare nominals. In COHA, 56 of the 2,980 instances of *books* being used bare involve fixed expressions. For *horses*, the same is true in 107 of the 1,634 instances in which it is found bare. Figure 11 illustrates that the share of fixed expressions with *books* amounts to 1.88%, compared to 6.55% with *horses*.

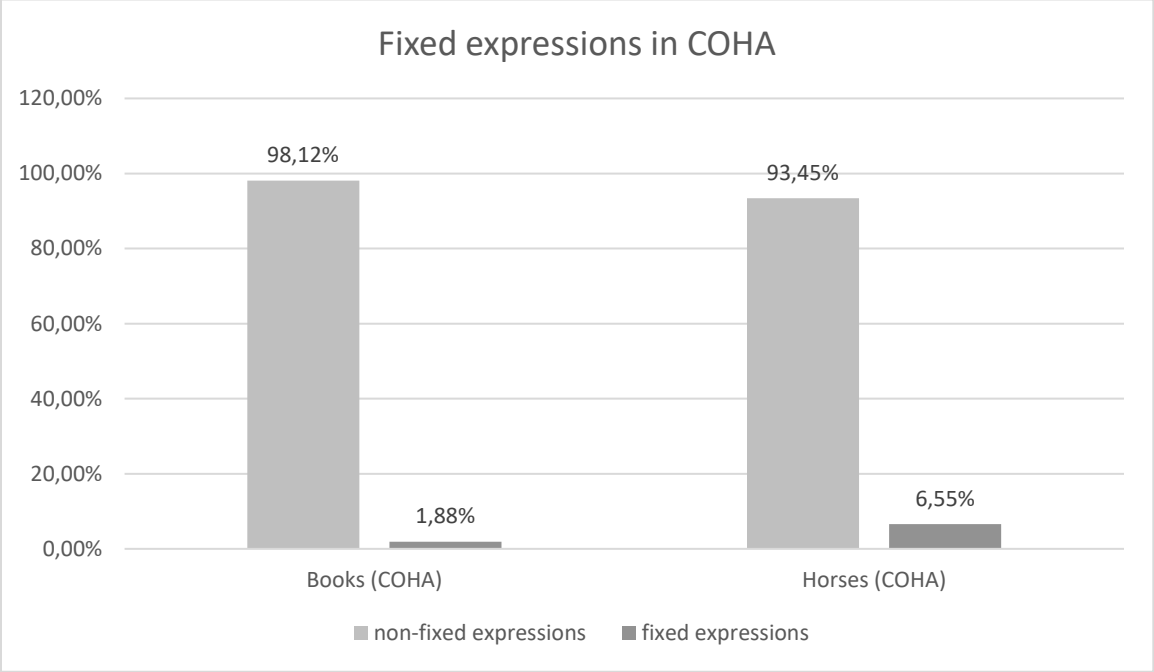


**Figure 11 Fixed vs. non-fixed expressions in COHA**

*Books* is used in the following fixed expression, following the OED (2018a: s.v. *book*):

– *to keep books* (ex. 49)

*Horses* appears in a variety of expressions (OED 2018b: s.v. *horse*):

– *to change horses* (ex. 50), *to change/ swap horses in midstream/ while crossing a stream*, *to hitch horses* and *horses of a different color*

(49) I work in an office. I *keep books* [my emphasis].

(COHA, fiction, 1990)

(50) He landed at Southampton and rode from there to Canterbury without a stop, except to *change horses* and for hurried meals [my emphasis].

(COHA, NF, 1950)

In the sample from PPCEME, no fixed expressions were identified. This observation does not necessarily give rise to the assumption that the use of BPs in fixed expressions has increased over the last 500 years as the lack of fixed expressions in PPCEME might also be attributed to the small size of the sample.

In order to investigate whether the share of fixed expressions has changed in the course of the LModE period, the ratio of fixed and non-fixed expressions was identified in the subperiods of COHA (Table 9).

**Table 9 Fixed expressions vs. non-fixed expressions in the subperiods of COHA**

| | COHA | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | books | | | | horses | | | |
| | non-fixed | | fixed | | non-fixed | | fixed | |
| 1810 | 1 | 100% | 0 | 0% | 0 | 0% | 0 | 9% |
| 1820 | 37 | 100% | 0 | 0% | 15 | 93.75% | 1 | 6.25% |
| 1830 | 67 | 98.53% | 1 | 1.47% | 47 | 88.68% | 6 | 11.32% |
| 1840 | 111 | 100% | 0 | 0% | 50 | 94.34% | 3 | 5.66% |
| 1850 | 126 | 98.44% | 2 | 1.56% | 59 | 85.51% | 10 | 14.49% |
| 1860 | 84 | 97.67% | 2 | 2.33% | 52 | 86.67% | 8 | 13.33% |
| 1870 | 106 | 96.36% | 4 | 3.64% | 66 | 89.19% | 8 | 10.81% |
| 1880 | 133 | 94.33% | 8 | 5.67% | 53 | 82.81% | 11 | 17.19% |
| 1890 | 162 | 97.59% | 4 | 2.41% | 78 | 97.5% | 2 | 2.50% |
| 1900 | 154 | 99.35% | 1 | 0.65% | 89 | 93.68% | 6 | 6.32% |
| 1910 | 270 | 98.90% | 3 | 1.10% | 80 | 93.02% | 6 | 6.98% |
| 1920 | 181 | 97.84% | 4 | 2.16% | 125 | 93.98% | 8 | 6.02% |
| 1930 | 157 | 96.91% | 5 | 3.09% | 102 | 92.73% | 8 | 7.27% |
| 1940 | 176 | 97.24% | 5 | 2.76% | 122 | 97.60% | 3 | 2.46% |
| 1950 | 158 | 96.93% | 5 | 3.07% | 95 | 95% | 5 | 5% |
| 1960 | 158 | 98.75% | 2 | 1.25% | 84 | 97.67% | 2 | 2.33% |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1970 | 171 | 97.16% | 5 | 2.84% | 80 | 95.24% | 4 | 4.76% |
| 1980 | 153 | 100% | 0 | 0% | 102 | 93.58% | 7 | 6.42% |
| 1990 | 250 | 98.43% | 4 | 1.57% | 111 | 93.28% | 8 | 6.72% |
| 2000 | 269 | 99.63% | 1 | 0.37% | 117 | 99.15% | 1 | 0.85% |

Figure 12 illustrates that the share of fixed expressions with *horses* has decreased between the 19th and the 20th century. With *books*, the share of fixed expressions has remained relatively stable.

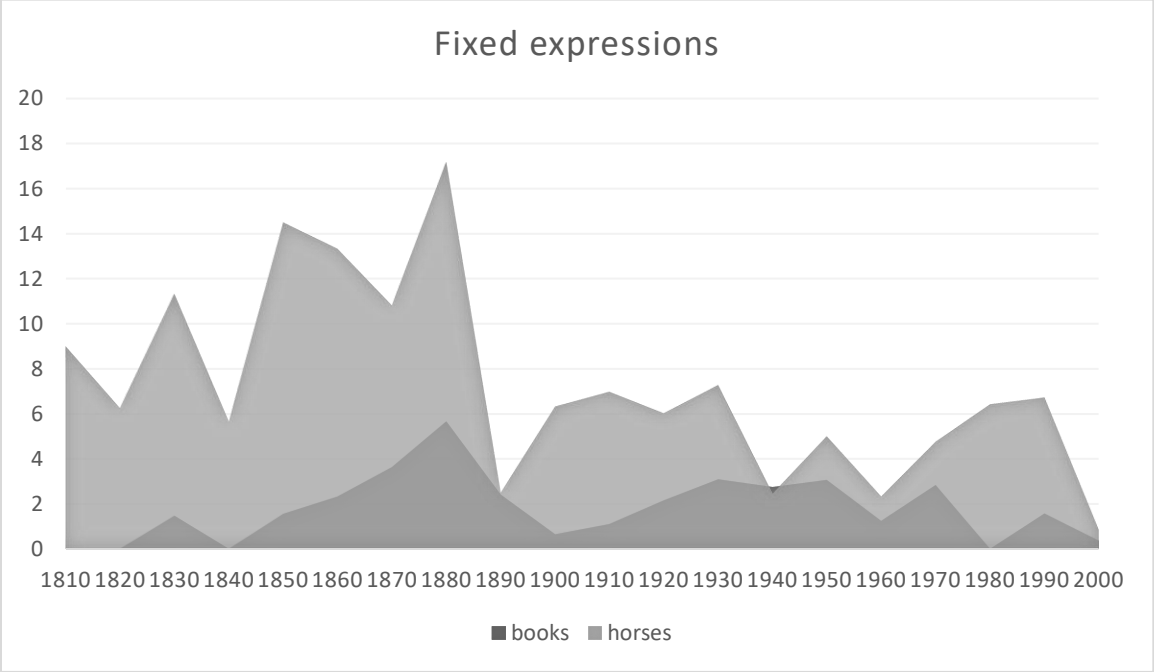


**Figure 12 The share of fixed expressions in percent of all BPs**

In short, the use of bare nouns has become established in certain fixed expressions. As an increased use of fixed expressions involving bare nouns obviously leads to an increase in the total number of BPs, it seems reasonable to conclude at the end of the analysis that fixed expressions do in fact play a central role in the increasing occurrence of BPs in the English language. However, the small size of the sample from PPCEME makes it difficult to draw meaningful conclusions from the results in this section.

### 6.2.4. Argumental vs. predicative BPs

An increased use of fixed expressions is not the only phenomenon that could explain the rise of BPs over the last 500 years. Another crucial factor that needs to be taken into account is the development of non-referential BPs (see sections 2.2.4. and 3.1.2.2.). Referentiality is said to be closely linked to argumenthood as only arguments (subjects and objects) are interpreted as DPs and considered to be referential. According to this theory, non-arguments (predicates and vocatives) are non-referential and analysed as simple NPs (Longobardi 1994, 2005). The difference between referential and non-referential NPs is in essence that the former, "denote particular entities in the universe of discourse" (Abraham, Stark & Leiss 2007: 5).

An increased occurrence of non-argumental BPs in COHA could therefore also have contributed to the rise of BPs in ModE. The diachronic development of non-argumental BPs thus deserves further investigation. Note that the fixed expressions listed in the previous section were discarded from the sample and will not be further considered in this study.

Table 10 shows the distribution of argumental versus non-argumental BPs in PPCEME and COHA. Interestingly, no examples of non-argumental BPs are found in PPCEME. Again, this could be attributed to the small size of the sample. In COHA *books* is used 104 times (ex. 51) and *horses* 62 times (ex. 52) in non-argumental position.

(51) Lastly, I am a writer, and my interests are *books* and sport [my emphasis].

(COHA, magazines, 1940)

(52) Those animals are *horses* [my emphasis].

(COHA, fiction, 1891)

**Table 10 Argumental vs. non-argumental BPs in PPCEME and COHA**

| | | PPCEME | | COHA | |
|---|---|---|---|---|---|
| | | books | horses | books | horses |
| arguments | subjects | 0 | 3 | 753 | 453 |
| | objects | 13 | 8 | 2,067 | 1,012 |
| | **TOTAL** | **13** | **11** | **2,820** | **1,465** |
| non-arguments | predicates | 0 | 0 | 104 | 62 |
| | vocatives | 0 | 0 | 0 | 0 |
| | **TOTAL** | **0** | **0** | **104** | **62** |

Next, the development of non-argumental BPs within the LModE period was traced. Table 11 shows the results for the subperiods of COHA.

**Table 11 Argumental vs. non-argumental BPs in the subperiods of COHA**

| | COHA | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | books | | | | horses | | | |
| | arguments | | non-arguments | | arguments | | non-arguments | |
| 1810 | 1 | 100% | 0 | 0% | 0 | 0% | 0 | 0% |
| 1820 | 36 | 97.30% | 1 | 2.70% | 15 | 100% | 0 | 0% |
| 1830 | 61 | 91.04% | 6 | 8.96% | 45 | 95.74% | 2 | 4.26% |
| 1840 | 100 | 90.09% | 11 | 9.91% | 49 | 94.23% | 3 | 5.77% |
| 1850 | 119 | 94.44% | 7 | 5.56% | 57 | 96.61% | 2 | 3.39% |
| 1860 | 81 | 96.43% | 3 | 3.57% | 47 | 90.38% | 5 | 9.62% |
| 1870 | 101 | 95.28% | 5 | 4.72% | 64 | 96.97% | 2 | 3.03% |
| 1880 | 129 | 96.99% | 4 | 3.01% | 51 | 96.23% | 2 | 3.77% |
| 1890 | 158 | 97.53% | 4 | 2.47% | 73 | 93.59% | 5 | 6.41% |
| 1900 | 148 | 96.10% | 6 | 3.90% | 84 | 94.38% | 5 | 5.62% |
| 1910 | 160 | 94.12% | 10 | 5.88% | 74 | 92.50% | 6 | 7.50% |
| 1920 | 172 | 95.03% | 9 | 4.97% | 123 | 98.40% | 2 | 1.60% |
| 1930 | 156 | 99.36% | 1 | 0.64% | 99 | 97.06% | 3 | 2.94% |
| 1940 | 171 | 97.16% | 5 | 2.84% | 115 | 94.26% | 7 | 5.74% |
| 1950 | 155 | 98.10% | 3 | 1.90% | 91 | 95.79% | 4 | 4.21% |
| 1960 | 152 | 96.20% | 6 | 3.80% | 80 | 95.24% | 4 | 4.76% |
| 1970 | 165 | 96.49% | 6 | 3.51% | 80 | 100% | 0 | 0% |
| 1980 | 148 | 96.73% | 5 | 3.27% | 101 | 99.02% | 1 | 0.98% |
| 1990 | 145 | 96.67% | 5 | 3.33% | 104 | 94.55% | 6 | 5.45% |
| 2000 | 262 | 97.40% | 7 | 2.60% | 112 | 95.73% | 5 | 4.27% |

There is clearly no trend towards an increased use of non-argumental BPs within the LModE period, as Figure 13 illustrates. One can even speak of a slight decrease in the share of predicative BPs between the 19$^{th}$ (average 4.54%) and the 20$^{th}$ century (average 3.88%).

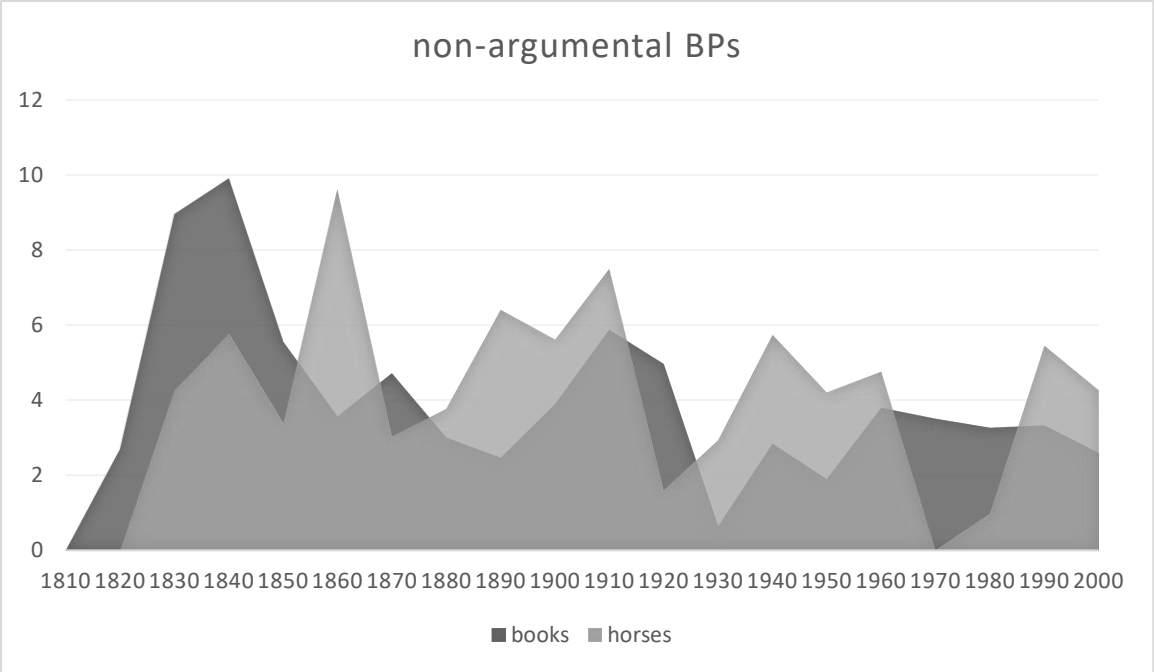


**Figure 13 Non-argumental BPs in COHA**

To sum up, this section suggests a rise in non-argumental BPs between EModE and LModE. However, such a conclusion can only be very tentative due to the small size of the sample from PPCEME. Within the LModE period, such a development cannot be confirmed as the average percentage of non-argumental BPs decreases from the 19th to the 20th century.

### 6.2.5. Generic vs. indefinite subject BPs

In the last two sections, it was noted that the use of fixed expressions involving BPs has become more frequent. Also, the share of non-argumental BPs has increased between EModE and LModE. However, due to small sample sizes, the results lack reliability. Further analyses are thus required.

Since the original hypothesis predicts a decrease in the share of indefinite BPs between EModE and LModE, it seems reasonable to focus exclusively on the frequencies of BPs with indefinite reference. For this purpose, they need to be separated from non-referential BPs. Non-argumental BPs are not alone in being considered non-referential. Generics also generally lack DP structure and are thus analysed as non-referential NPs (see section 3.1.2.2.).

Prior to the semantic analysis, the sample was further reduced as at this stage only subject BPs were taken into account (see section 6.4, for a discussion of object BPs). Since in PPCEME *books* and *horses* only occur three times in subject position, the sample of PPCEME was not

used. Rather, all BPs in sentence-initial position that were identified as subjects were analysed. As for COHA, the number of BPs in sentence-initial position (283,284 BPs) is definitely not manageable for analysis. Thus, the sample of COHA used in the last sections was used again. Finally, it is important to mention that examples with passive constructions and modal verbs were eliminated since the interpretation of these sentences often leads to ambiguity.

The remaining 48 BPs in PPCEME and 555 BPs in COHA were semantically analysed by checking every noun individually and deciding if it yields an indefinite (ex. 53) or a generic interpretation (ex. 54). The tests for NP genericity, summarized in section 3.2.4., served hereby as an aid to distinguish between the two readings.

(53) a. *Books* and papers lay scattered on a table [my emphasis].

(COHA, fiction, 1833)

b. *Horses* clattered by [my emphasis].

(COHA, fiction, 1930)

c. *Letters* from Rodolph made mention of great Treason, and of Provision of Men from beyond the Seas, to invade this Realm, according to his Order and Request, who was meant by Quarante [my emphasis].

(PPCEME, proceedings, trials, E2)

(54) a. *Books* are well written or badly written [my emphasis].

(COHA, fiction, 1997)

b. *Horses* are extremely intelligent [my emphasis].

(COHA, fiction, 1963)

c. *Men* often bring habitual hunger and thirst on themselves by custom [my emphasis].

(PPCEME, educational treatise, E3)

In Table 12, the ratio of existential and generic subject BPs in PPCEME and COHA is presented. It is striking that the share of existential BPs in COHA (*books* 29.87% & *horses* 51.90%) is higher than in PPCEME (all BPs, 27.08%). Considering that in the beginning of the empirical part (6.1.) an increase in the number of BPs between PPCEME and COHA was observed, this new result, which suggests an increase in indefinite BPs within the total number of subject BPs, further contradicts the original hypothesis.

**Table 12 Existential vs. generic subject BPs in PPCEME and COHA**

| | PPCEME | | COHA | | | |
|---|---|---|---|---|---|---|
| | all | % | books | % | horses | % |
| existential | 13 | 27.08% | 95 | 29.87% | 123 | 51.90% |
| generic | 35 | 72.92% | 223 | 70.13% | 114 | 48.10% |

Next, the development of indefinite subject BPs was more closely examined by comparing the results for the subperiods of PPCEME and COHA. Table 13 shows a sharp increase in indefinite BPs between the first (1500-1569) and the second (1570-1639) subperiod of PPCEME. This trend does not continue, though, as the share of indefinite BPs decreases slightly between the second and third (1640-1710) subperiod of PPCEME.

**Table 13 Existential vs. generic subject BPs in the subperiods of PPCEME**

| | E1 (1500-1569) | | E2 (1570-1639) | | E3 (1640-1710) | |
|---|---|---|---|---|---|---|
| | raw freq. | % | raw freq. | % | raw freq. | % |
| existential | 2 | 11.76% | 5 | 38.46% | 6 | 33.33% |
| generic | 15 | 88.24% | 8 | 61.54% | 12 | 66.67% |

Likewise, in COHA (Table 14 & Figure 14) the shares of indefinite BPs with *books* and *horses* fluctuates strongly, making it impossible to determine a trend towards an increasing or decreasing use of indefinite BPs within the LModE period.

**Table 14 Existential vs. generic subject BPs in COHA**

| | COHA | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | books | | | | horses | | | |
| | existential | | generic | | existential | | generic | |
| 1810 | 0 | 0% | 1 | 100% | 0 | 0% | 0 | 0% |
| 1820 | 0 | 0% | 5 | 100% | 1 | 50.00% | 1 | 50.00% |
| 1830 | 3 | 25.00% | 9 | 75.00% | 2 | 33.33% | 4 | 66.67% |
| 1840 | 5 | 23.81% | 16 | 76.19% | 5 | 71.43% | 2 | 28.57% |
| 1850 | 6 | 25.00% | 18 | 75.00% | 8 | 80.00% | 2 | 20.00% |
| 1860 | 3 | 27.27% | 8 | 72.73% | 1 | 16.67% | 5 | 83.33% |
| 1870 | 3 | 30.00% | 7 | 70.00% | 2 | 28.57% | 5 | 71.43% |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1880 | 2 | 15.38% | 11 | 84.62% | 2 | 28.57% | 5 | 71.43% |
| 1890 | 4 | 20.00% | 16 | 80.00% | 8 | 53.33% | 7 | 46.67% |
| 1900 | 8 | 47.06% | 9 | 52.94% | 3 | 33.33% | 6 | 66.67% |
| 1910 | 9 | 33.33% | 18 | 66.67% | 1 | 8.33% | 11 | 91.67% |
| 1920 | 4 | 18.18% | 18 | 81.82% | 11 | 52.38% | 10 | 47.62% |
| 1930 | 3 | 30.00% | 7 | 70.00% | 13 | 76.47% | 4 | 23.53% |
| 1940 | 6 | 31.58% | 13 | 68.42% | 11 | 64.71% | 6 | 35.29% |
| 1950 | 2 | 20.00% | 8 | 80.00% | 11 | 68.75% | 5 | 31.25% |
| 1960 | 5 | 33.33% | 10 | 66.67% | 7 | 58.33% | 5 | 41.67% |
| 1970 | 6 | 31.58% | 13 | 68.42% | 3 | 37.50% | 5 | 62.50% |
| 1980 | 7 | 35.00% | 13 | 65.00% | 3 | 37.50% | 5 | 62.50% |
| 1990 | 10 | 47.62% | 11 | 52.38% | 12 | 46.15% | 14 | 53.85% |
| 2000 | 9 | 42.86% | 12 | 57.14% | 12 | 38.71% | 19 | 61.29% |

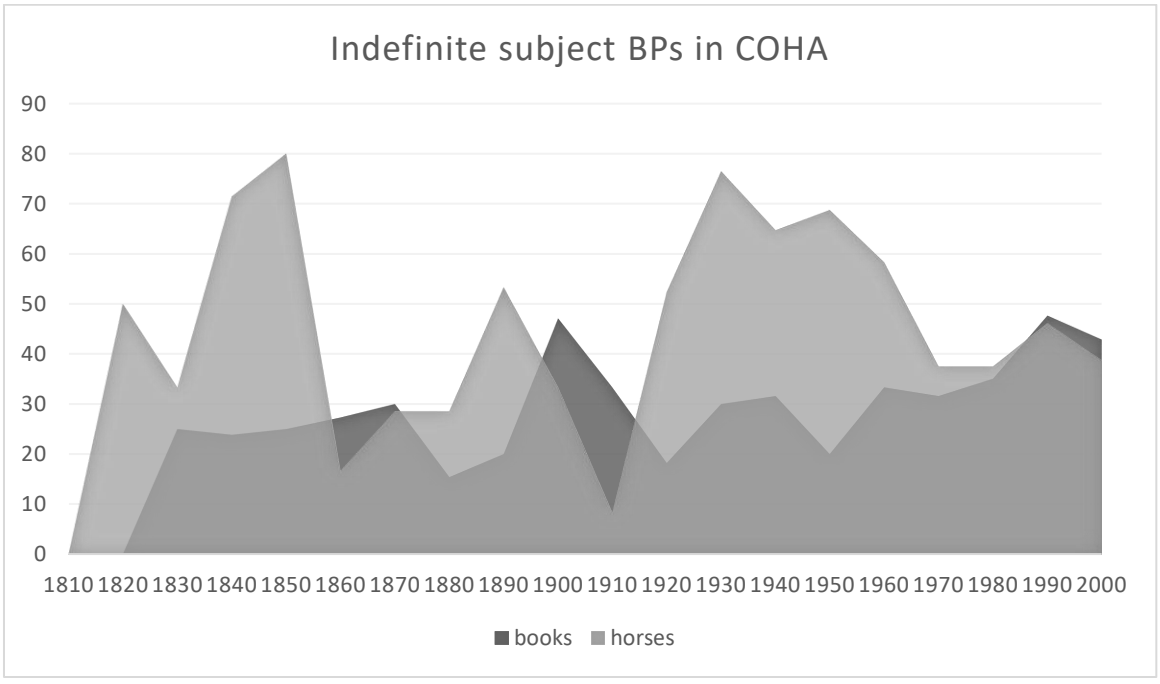


**Figure 14 The development of indefinite subject BPs within the LModE period**

Even though the results of the subperiods are not conclusive, the increasing use of indefinite subject BPs between EModE and LModE, noted in the beginning of this section, allows to formulate another hypothesis: A rise in indefinite BPs goes hand in hand with a decline in

'*some* + Pl. N' sequences. To check this hypothesis, the sample was extended by including indefinite subject NPs headed by *some* (ex. 55-57).

(55) *Some letters* say the eldest was already dead, tho wone can not say this is certain [my emphasis].

(PPCEME, letters, private, E3)

(56) *Some books* lay upon the table, arranged two by two [my emphasis];

(COHA, fiction, 1974)

(57) *Some horses* were grazing a little distance away and stood at gaze, to break and wheel and gallop away with flying manes and tails [my emphasis].

(COHA, fiction, 1922)

When the ratio of indefinite subject BPs and subject NPs headed by *some* was investigated from a diachronic point of view, the hypothesis could indeed be verified. Table 15 shows that the share of subject NPs headed by *some* has decreased in the course of the ModE period, while the share of indefinite subject BPs has increased sharply.

**Table 15 Indefinite subject BPs vs. subject NPs headed by *some* in PPCEME and COHA**

| | PPCEME | | COHA | | | |
|---|---|---|---|---|---|---|
| | all | % | books | % | horses | % |
| BPs | 48 | 69.57% | 318 | 93.26% | 237 | 93.68% |
| some + N | 21 | 30.43% | 23 | 6.74% | 16 | 6.32% |

The fact that the use of *some* as a marker of indefiniteness decreased between EModE and LModE raises the question as to whether this trend can also be observed within the LModE period. To clarify this issue, the development of '*some* + Pl. N' sequences in COHA was traced in more detail. Table 16 shows the results of three searches that were run on the subperiods of COHA.

**Table 16 The development of '*some*+Pl. N' sequences in LModE**

| | . some [nn2] | ! some [nn2] | ? some [nn2] | TOTAL | pmw |
|---|---|---|---|---|---|
| 1810 | 11 | 0 | 0 | 11 | 9.31 |
| 1820 | 123 | 2 | 3 | 128 | 18.48 |
| 1830 | 201 | 5 | 6 | 212 | 15.39 |

| | | | | | |
|---|---|---|---|---|---|
| 1840 | 219 | 2 | 9 | 230 | 14.33 |
| 1850 | 257 | 2 | 1 | 260 | 15.76 |
| 1860 | 251 | 5 | 10 | 266 | 15.53 |
| 1870 | 249 | 4 | 9 | 262 | 14.08 |
| 1880 | 319 | 10 | 12 | 341 | 16.34 |
| 1890 | 295 | 6 | 10 | 311 | 14.68 |
| 1900 | 410 | 6 | 8 | 424 | 18.81 |
| 1910 | 381 | 9 | 14 | 404 | 17.83 |
| 1920 | 467 | 5 | 20 | 492 | 19.19 |
| 1930 | 593 | 5 | 16 | 614 | 25.15 |
| 1940 | 792 | 3 | 24 | 819 | 33.92 |
| 1950 | 720 | 3 | 28 | 751 | 30.78 |
| 1960 | 893 | 6 | 31 | 930 | 38.87 |
| 1970 | 936 | 1 | 40 | 977 | 41.10 |
| 1980 | 1,057 | 3 | 25 | 1,085 | 43.09 |
| 1990 | 1,085 | 2 | 21 | 1,108 | 39.75 |
| 2000 | 956 | 6 | 22 | 984 | 33.38 |

Figure 15 clearly illustrates that the use of '*some* + Pl. N' sequences in sentence initial-position has become much more frequent in the course of the 20th century. This result suggests that the trend towards a decreasing use of *some* to mark indefiniteness with plural NPs has been reversing since the first half of the 20th century. In the second half of the 20th century, the share of '*some* + Pl. N' sequences has even more than doubled compared to the 19th century.

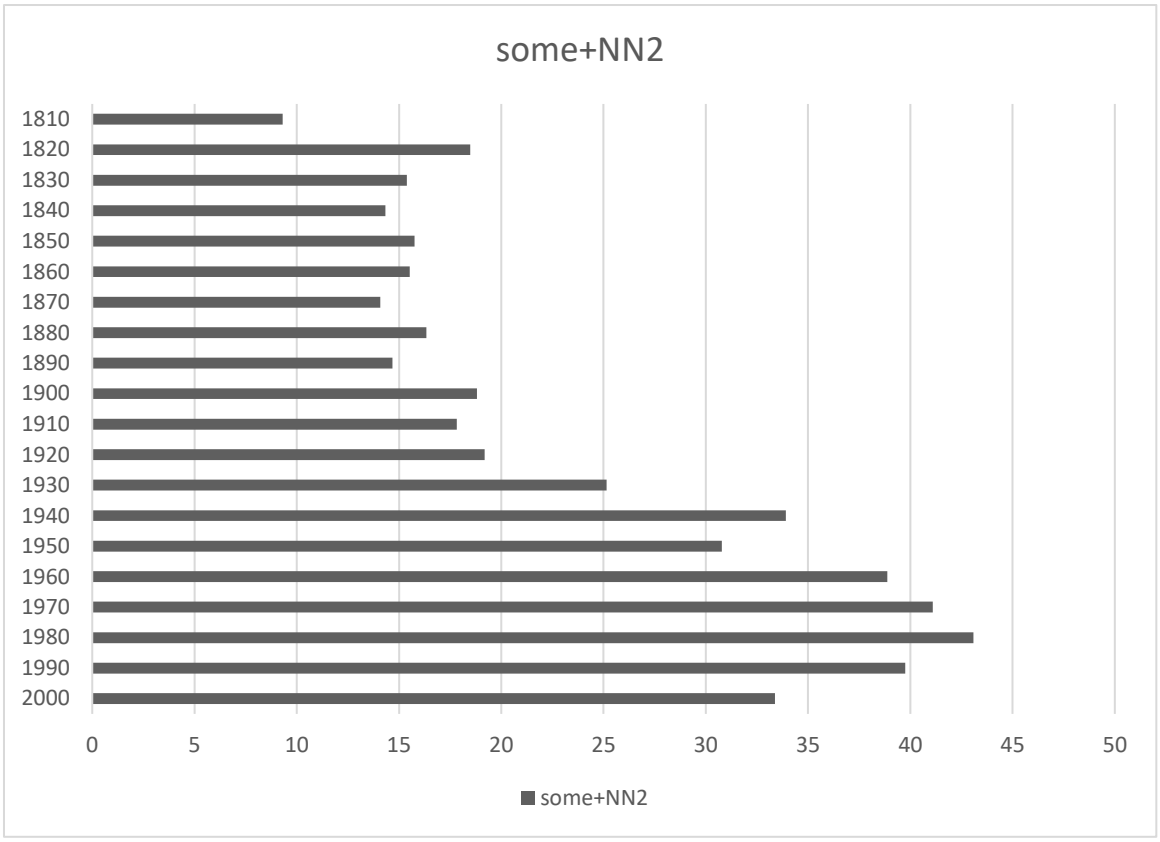


**Figure 15 The development of '*some*+Pl. N' sequences in sentence-initial position**

So far, the ratios of fixed vs. non-fixed BP expressions, argumental vs. non-argumental BPs and generic vs. indefinite subject BPs have been discussed. It was shown that the frequencies of BPs involving non-fixed expressions, of non-argumental BPs and of indefinite subject BPs have increased between EModE and LModE.

Nevertheless, it must be acknowledged that creating a sample of plural nouns for both corpora to investigate the development of BPs in greater detail did not necessarily yield compelling results. The huge difference in size between the two corpora makes it extremely difficult to select comparable samples. Whereas the sample of COHA allowed for discussion of the role of fixed expressions and the development of non-argumental BPs, the sample of PPCEME was too small to comment on these issues. As a consequence, extending the sample of BPs by analysing all subject BPs from PPCEME was therefore a useful decision. This way, the percentage of indefinite BPs within all subject BPs could be seen to have risen between EModE and LModE. Moreover, the declining use of '*some* + Pl. N' sequences in sentence-initial position is direct evidence of an increasing use of bare nouns to express indefiniteness in LModE.

Having placed the focus on indefinite subject BPs and NPs headed by *some* in sentence-initial position, it remains to be clarified how indefinite BPs in other positions have developed.

## 6.3. BPs in existential sentences

In the last section, a tendency towards increased use of indefinite BPs, rather than the structure '*some* + Pl. N', was observed in the LModE period. Since, compared to the sizes of the corpora, only a relatively small sample was used for analysis, it seems worthwhile to investigate some more instances of BPs. In particular, the investigation of indefinite BPs in existential sentences promises more conclusive results with regard to the development of indefinite BPs. Existential constructions play a significant role in Carlson's theory as they only accept S-level predicates (see section 3.2.1.). BPs following '*there* [be]' constructions have thus an existential reading.

For this study, it was not necessary to use a sample as it was possible to search for all existential sentences involving BPs in PPCEME and COHA and to compare the results to existential sentences involving nouns headed by *some* (see Appendix VIII for the results of each search query). In PPCEME the combinatorial pattern '*there* [be] + BP' occurs 76 times. This pattern occurs almost four times as often as '*there* [be] + *some* + Pl. N'. The structure in (58) is thus very common in EModE.

(58) Also *there are horses* that fight with Allegators or Crocodiles in Tancks or ponds of water where I also saw one Allegator kill 2. stone horses at one time [my emphasis].

(PPCEME, travelogue, E2)

The less frequent pattern '*there* [be] + *some* + Pl. N' has only 21 hits, one of which is as follows:

(59) for I am sure *there were some Strangers* came hither last Night [my emphasis];

(PPCEME, proceedings, trials, E3)

In COHA, the combination '*there* [be] + BP' is used proportionally even more frequently with 31,002 hits (ex. 60) compared to the pattern '*there* [be] + *some* + Pl. N' which occurs 3,049 times (ex. 61).

(60) *There were books* on a hanging shelf [my emphasis].

(COHA, fiction, 1948)

(61) *There were some trees* around it, and a little lawn in front [my emphasis].

(COHA, fiction, 1870)

An overview of the results is provided in Table 17. Crucially, the share of the pattern '*there* [be] + BP' amounts to 78.35% in PPCEME and 91.05% in COHA. This confirms the results of the analysis of BP subjects (see 6.2.5.), hence the increase of indefinite BPs between EModE and LModE.

**Table 17 BPs vs. NPs headed by *some* in existential sentences**

| | PPCEME | | COHA | |
|---|---|---|---|---|
| | raw freq. | % | raw freq. | % |
| BPs | 76 | 78.35% | 31,002 | 91.05% |
| some + N | 21 | 21.65% | 3,049 | 8.95% |

In the last section, it was also shown that the apparent decline of '*some* + Pl. N' sequences is not consistent in the LModE period. Particularly since mid-20th century, the usage of *some* in combination with plural nouns in the beginning of sentences has gained popularity. It is therefore only natural to ask whether a similar development is identifiable for existential sentences as well. Table 18 thus presents the results for the ratio of '*there* [be] + BP' and '*there* [be] + *some* + Pl. N' sequences in the subperiods of PPCEME and COHA. Surprisingly, a rise of '*there* [be] + *some* + Pl. N' sequences, comparable with the increase of subject NPs headed by *some* in the 20th century, is not clearly visible at any point during the ModE period. From the 19th (average 10.27%) to the 20th century (average 8.26%), the percentage of '*there* [be] + *some* + Pl. N' sequences has decreased. Within the 20th century, no particular trend is evident as the percentage of NPs headed by *some* in existential sentences fluctuates between 7.47% and 10.12%.

**Table 18 BPs vs. NPs headed by *some* in existential sentences (subperiods of PPCEME & COHA)**

| | PPCEME & COHA | | | |
|---|---|---|---|---|
| | BPs | | some + NN2 | |
| | raw freq. | % | raw freq. | % |
| 1500-1569 | 18 | 85.71% | 3 | 14.29% |
| 1570-1639 | 37 | 86.05% | 6 | 13.95% |
| 1640-1710 | 21 | 63.64% | 12 | 36.36% |
| 1810 | 63 | 100% | 0 | 0% |
| 1820 | 349 | 85.75% | 58 | 14.25% |

| | | | | |
|---|---|---|---|---|
| 1830 | 789 | 86.23% | 126 | 13.77% |
| 1840 | 937 | 86.92% | 141 | 13.08% |
| 1850 | 1,090 | 86.99% | 163 | 13.01% |
| 1860 | 1,206 | 89.27% | 145 | 10.73% |
| 1870 | 1,381 | 90.92% | 138 | 9.08% |
| 1880 | 1,573 | 90.14% | 172 | 9.86% |
| 1890 | 1,586 | 91.36% | 150 | 8.64% |
| 1900 | 1,795 | 89.88% | 202 | 10.12% |
| 1910 | 1,837 | 92.50% | 149 | 7.50% |
| 1920 | 1,932 | 92.53% | 156 | 7.47% |
| 1930 | 1,835 | 92.16% | 156 | 7.84% |
| 1940 | 1,930 | 91.82% | 172 | 8.18% |
| 1950 | 1,991 | 91.58% | 183 | 8.42% |
| 1960 | 1,952 | 90.75% | 199 | 9.25% |
| 1970 | 2,169 | 92.02% | 188 | 7.98% |
| 1980 | 2,385 | 93.20% | 174 | 6.80% |
| 1990 | 2,152 | 90.92% | 215 | 9.08% |
| 2000 | 2,050 | 92.68% | 162 | 7.32% |

Figure 16 illustrates once more that over the last 500 years the use of the combinatorial pattern '*there* [be] + BP' has increased almost steadily, while '*there* [be] + *some* + Pl. N' sequences have decreased.

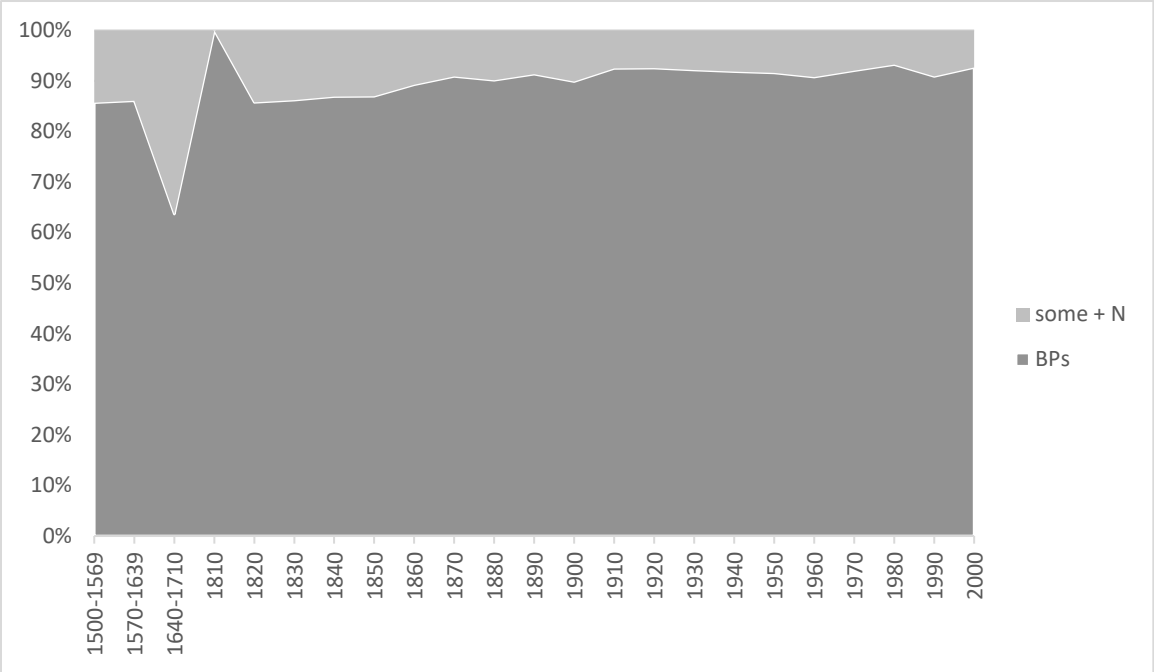


**Figure 16 The ratio of BPs and NPs headed by *some* in existential sentences**

In the previous as well as in the present section, similar patterns with regard to the development of indefinite BPs could be observed. In both subject position and existential sentences, the relative frequency of BPs with indefinite reference has increased and that of plural nouns headed by *some* decreased. The aim of the final section is to investigate whether this observed trend is also true for BPs in object position.

## 6.4. BPs in object position

BP objects generally tend to receive an existential interpretation. Only a small group of verbs, the so-called verbs of psychological state with experiencer subjects, select the universal reading of BPs (see section 3.2.1.). Since this study focuses on BPs with indefinite reference, BP objects with indefinite reference will be looked at in more detail in this section. The central question will be whether the share of sentences with existential BP objects increases or decreases in the ModE period. Will the same trend be evident as in the previous sections, i.e. an increase in BPs with indefinite reference over the last 500 years?

To analyse the frequency of 'S-level verb + BP' patterns and 'S-level verb + *some* + Pl. N' patterns in PPCEME and COHA, the transitive verbs *eat*, *hear* and *send* were selected. All three of them are used to express temporary states or to describe transitory activities and thus, following Carlson (1977), can be classified as S-level predicates. The results of the searches

have shown that these verbs are in general not used in idiomatic contexts, hence the findings will not be distorted by the occurrence of fixed expressions (see Appendix IX for the results of each search query).

In PPCEME 52 hits for '*eat/hear/send* + BP' (ex. 62-64), and 7 hits for '*eat/hear/send* + *some* + BP' (ex. 65-67) can be found. The share of objects headed by *some* amounts to 11.86%.

(62) We our selues are like those women which haue a longing to *eate coales*, and lime, and filth [my emphasis];

(PPCEME, sermon, E2)

(63) If he be more desirous, (as the most parte of children be) to *here thinges* marueilous and exquisite which hath in it a visage of some thinges incredible [my emphasis]

(PPCEME, educational treatise, E1)

(64) & soe I *sent papers* Into Westmorelande [my emphasis]

(PPCEME, biography, autobiography, E3)

(65) Here *eat some Plumbes* [my emphasis].

(PPCEME, drama, comedy, E2)

(66) for I have *hearde some Preachers* say, that there be some truthes which they would be loth to preach [my emphasis].

(PPCEME, sermon, E2)

(67) ffreinds were very much Concerned least they shoulde *sende some officers* to break uppe our meetinge [my emphasis]

(PPCEME, biography, autobiography, E3)

In COHA the pattern '*eat/hear/send* + BP' occurs 9,089 times (ex. 68-70), while '*eat/hear/send* + *some* + BP' occurs 283 times (ex. 71-73). The share of objects headed by *some* amounts to 3.02% and is thus clearly lower than in PPCEME.

(68) We drive some more, stop and *eat sandwiches*, chew and look at the country [my emphasis].

(COHA, fiction, 1948)

(69) "I thought I *heard horsemen* on the big road," said Runyon [my emphasis].

(COHA, fiction, 1931)

(70) I promised to *send letters* back with news from the eastern capitals, and, as far as I could, I gave him points of call to which he would send his own tidings and Ralf's about Arthur [my emphasis].

(COHA, fiction, 1973)

(71) 'Do *eat some biscuits*, Juliet.' [my emphasis]

(COHA, magazines, 2007)

(72) I *heard some horsemen* ride over the tracks, and also down the street, followed by the hurried footsteps of half a dozen men [my emphasis].

(COHA, fiction, 1896)

(73) So I phoned down to the hall porter, and said I wanted to *send some flowers* to somebody [my emphasis].

(COHA, fiction, 1960)

**Table 19 BPs vs. NPs headed by *some* in object position (PPCEME & COHA)**

| | PPCEME | | | | COHA | | | |
|---|---|---|---|---|---|---|---|---|
| | + BP | % | + some | % | + BP | % | + some | % |
| EAT | 6 | 85.71% | 1 | 14.29% | 1,365 | 97.36% | 37 | 2.64% |
| HEAR | 6 | 60.00% | 4 | 40.00% | 4,142 | 97.21% | 119 | 2.79% |
| SEND | 40 | 95.24% | 2 | 4.76% | 3,582 | 96.58% | 127 | 3.42% |
| **TOTAL** | **52** | **88.14%** | **7** | **11.86%** | **9,089** | **96.98%** | **283** | **3.02%** |

Table 19 shows that with every S-level verb tested in this section the share of the following object nouns headed by *some* is lower than in PPCEME. The findings thus clearly suggest an increase in indefinite BPs in object position between EModE and LModE. However, since the number of hits for all search queries which were run on PPCEME is very low, it is difficult to arrive at definite conclusions.

As the usage of '*some* + Pl. N' sequences in sentence-initial position has been rising since the beginning of the 20$^{th}$ century, it is natural to ask whether this is also the case for NPs in object position. Table 20 shows the results for six searches run on the subperiods of COHA. As can be seen, no trend towards an increased use of '*some* + Pl. N' sequences in object position can be observed in LModE. The share of '*some* + Pl. N' sequences as objects of *eat* remains steadily low (see Figure 17). The share of '*some* + Pl. N' sequences as objects of *hear* and *send* fluctuates strongly between 0% and almost 10%. But again, there does not seem to be a trend towards an increased use of '*some* + Pl. N' (see Figure 18 & 19).

**Table 20 BPs vs. NPs headed by some in object position (subperiods of COHA)**

| | COHA | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EAT | | | | HEAR | | | | SEND | | | |
| | +BP | | +some | | +BP | | +some | | +BP | | +some | |
| 1810 | 0 | 0% | 0 | 0% | 11 | 100% | 0 | 0% | 0 | 100% | 0 | 0% |
| 1820 | 9 | 100% | 0 | 0% | 51 | 96.23% | 2 | 3.77% | 37 | 92.50% | 3 | 7.50% |
| 1830 | 32 | 94.12% | 2 | 5.88% | 79 | 94.05% | 5 | 5.95% | 61 | 93.85% | 4 | 6.15% |
| 1840 | 21 | 100% | 0 | 0% | 94 | 97.92% | 2 | 2.08% | 81 | 100% | 0 | 0% |
| 1850 | 23 | 95.83% | 1 | 4.17% | 115 | 95.04% | 6 | 4.96% | 82 | 98.80% | 1 | 1.20% |
| 1860 | 43 | 100% | 0 | 0% | 121 | 96.03% | 5 | 3.97% | 106 | 92.98% | 8 | 7.02% |
| 1870 | 42 | 97.67% | 1 | 2.33% | 103 | 90.35% | 11 | 9.65% | 133 | 95% | 7 | 5% |
| 1880 | 36 | 97.30% | 1 | 2.70% | 130 | 93.53% | 9 | 6.47% | 126 | 94.03% | 8 | 5.97% |
| 1890 | 35 | 97.22% | 1 | 2.78% | 143 | 97.95% | 3 | 2.05% | 181 | 97.31% | 5 | 2.69% |
| 1900 | 60 | 93.75% | 4 | 6.25% | 184 | 95.83% | 8 | 4.17% | 203 | 97.13% | 6 | 2.87% |
| 1910 | 61 | 95.31% | 3 | 4.69% | 201 | 97.57% | 5 | 2.43% | 156 | 96.89% | 5 | 3.11% |
| 1920 | 74 | 96.10% | 3 | 3.90% | 249 | 98.03% | 5 | 1.97% | 236 | 97.93% | 5 | 2.07% |
| 1930 | 94 | 100% | 0 | 0% | 263 | 98.50% | 4 | 1.50% | 238 | 95.97% | 10 | 4.03% |
| 1940 | 84 | 97.67% | 2 | 2.33% | 310 | 98.41% | 5 | 1.59% | 268 | 95.71% | 12 | 4.29% |
| 1950 | 85 | 100% | 0 | 0% | 293 | 98.65% | 4 | 1.35% | 285 | 96.94% | 9 | 3.06% |
| 1960 | 83 | 100% | 0 | 0% | 295 | 97.36% | 8 | 2.64% | 234 | 96.30% | 9 | 3.70% |
| 1970 | 95 | 98.96% | 1 | 1.04% | 266 | 96.38% | 10 | 3.62% | 236 | 97.93% | 5 | 2.07% |
| 1980 | 118 | 94.4% | 7 | 5.60% | 349 | 97.21% | 10 | 2.79% | 270 | 95.74% | 12 | 4.26% |
| 1990 | 183 | 98.39% | 3 | 1.61% | 445 | 98.23% | 8 | 1.77% | 307 | 96.85% | 10 | 3.15% |
| 2000 | 187 | 95.90% | 8 | 4.10% | 440 | 98% | 9 | 2% | 342 | 97.71% | 8 | 2.29% |


**Figure 17 BPs vs. *some* + NPs as objects of *eat***

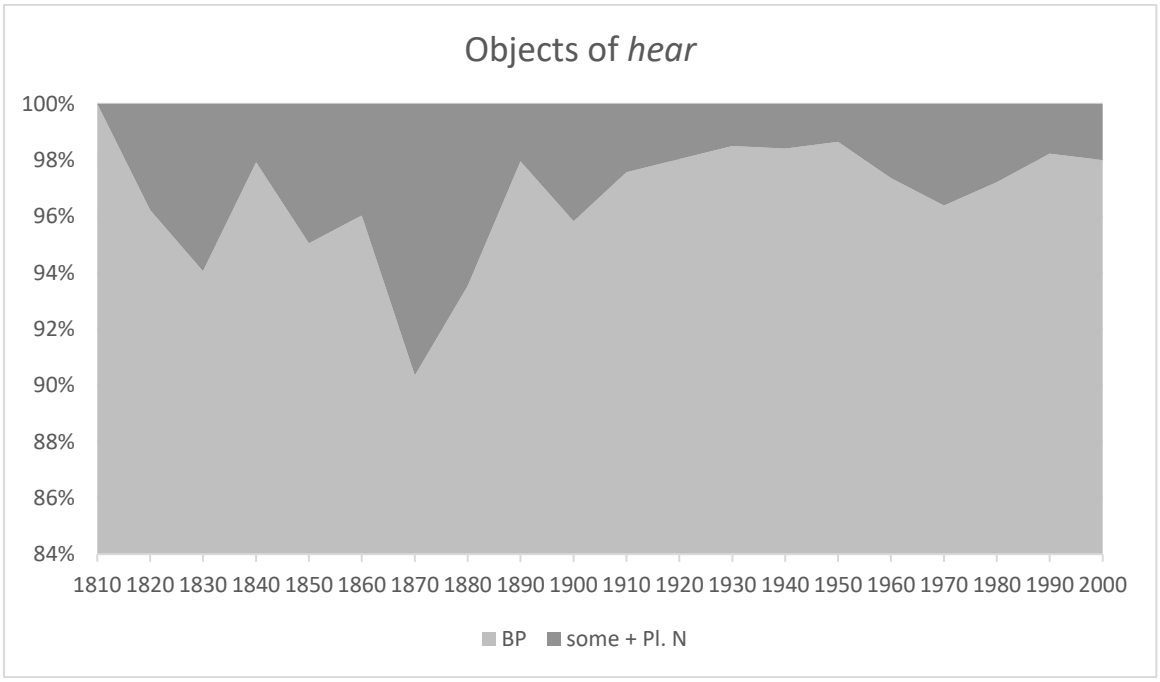


**Figure 18 BPs vs. *some* + NPs as objects of *hear***

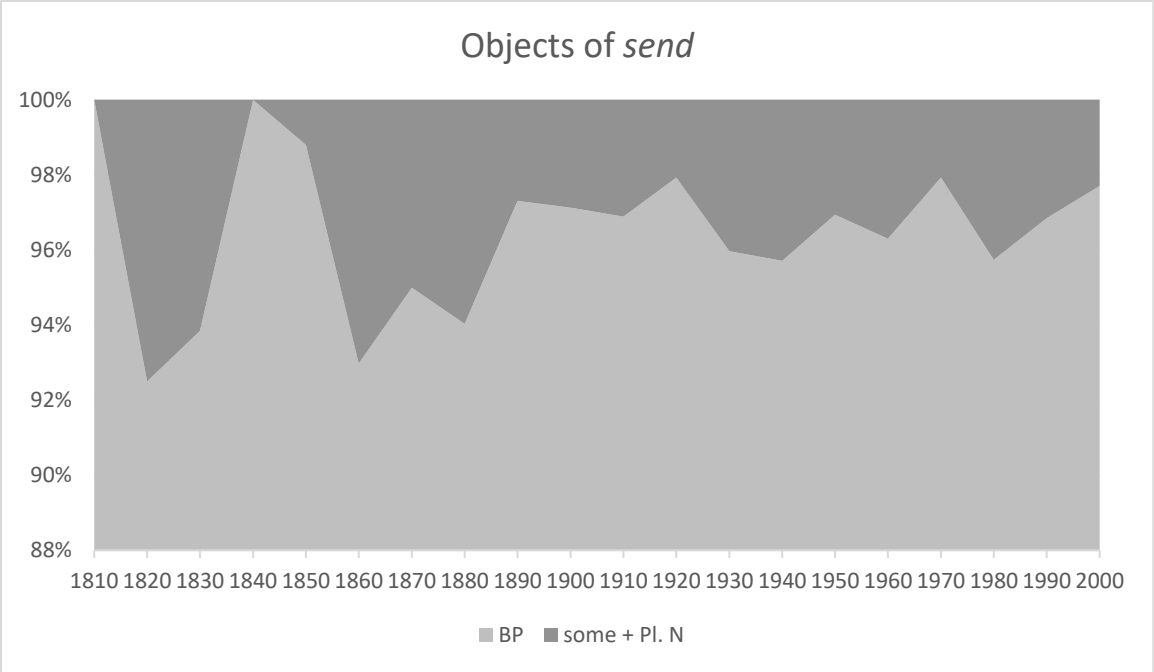


**Figure 19 BPs vs. NPs headed by *some* as objects of *send***

Summing up, this last section has shown that the observed trend towards an increased use of BPs with indefinite reference in sentence-initial position and in existential sentences is likewise consistent with the development of indefinite BPs in object position. Here too, the share of NPs headed by *some* has decreased, while the share of indefinite BPs has increased between EModE and LModE.

# 7. DISCUSSION OF THE RESULTS

The quantitative and qualitative analysis of plural NPs has shown that BP noun phrases with indefinite reference have become more frequent, whereas plural NPs headed by *some* have become less frequent between EModE and LModE. Irrespective of the syntactic position of BPs, the outcome of the analysis always refutes the central hypothesis of a trend towards obligatory indefiniteness marking with plural NPs. What does this unexpected result demonstrate?

The quantifying determiner *some* differs markedly from the definite article *the* and the indefinite article *a(n)* in terms of its diachronic development. As to the definite article, the OE demonstrative *sē* is reanalysed as the definite article *the* in the ME period and from this point assumes the function of definiteness marker (see 4.2.1.). Likewise, the OE numeral *one* is reanalysed as the indefinite article *a(n)*in ME and becomes obligatory as indefiniteness marker in EModE (see 4.2.2.). The development of *some* to a marker of indefiniteness, however, cannot be observed within the investigated time period.

Consequently, it seems that *some* continues to function solely as a marker of specificity in PdE. Its use is obligatory when plural nouns take wide scope or are interpreted as specific (see 2.2.3.2.). Crucially, Crisma (2015) points to parallels between the function of *a(n)* in late OE and that of *some* in PdE. In late OE, *a(n)* acts as a marker of specificity, hence, its presence is indispensable when a noun takes wide scope or receives a specific reading. The restricted occurrence of *a(n)* in late OE indicates that singular nouns with non-specific reference, as well as generics, occur bare at that time. The same applies to plural nouns in ModE.

Summing up, the results of the study and the above observations clearly provide no evidence for ongoing reanalysis of the quantifying determiner *some* as indefiniteness marker during the ModE period. Instead, the falling frequencies of NPs headed by *some,* as observed in sections 6.2.5., 6.3. and 6.4., point to an even more restricted occurrence of the determiner in PdE. Only in subject position does the rise in popularity of '*some* + Pl. N' sequences in the second half of the 20$^{th}$ century suggest a potential change in the use of *some* with plural nouns. To verify this hypothesis, however, further study would be required.

# 8. CONCLUSION

In the following sections, the most important findings of this thesis will be presented (8.1.). It is important to note, though, that a detailed summary of PART I is also provided in section 4.4. Finally, a short discussion about theoretical implications, limitations and directions for further research will close this thesis (8.2.).

## 8.1. Summary

To deal with the diachronic development of English BPs, some syntactic and semantic characteristics of the English NP and, in particular, of the BP construction were presented in the first part of this thesis. From a generative perspective, NPs, like all other phrases, are built according to the X-bar format, which implies that each position in the structure is either a head (X) or a maximal projection (XP). In the nominal domain, there has been considerable disagreement among scholars as to which word functions as a head. Since the introduction of the DP-hypothesis (Abney 1987), the functional category D rather than the lexical category N has increasingly been viewed as the head of NP among modern generativists (2.1.).

Chapter 2.2. showed that not all determiners are universal. Articles, for instance, are not an integral part of all languages. In English article usage is important, albeit not always obligatory. Mass singular nouns, proper names and generic nouns occur bare. Also, existential plural nouns with a non-specific interpretation may be used without a determiner. For a specific existential reading of plural nouns, the presence of weak *some* or *sm* is indispensable, though.

The role of D has remained a topic of debate. The delicacy of the issue is particularly evident when it comes to argumenthood. According to Szabolcsi (1987), Stowell (1989) and Longobardi (1994), arguments (subject, object) and non-arguments (predicate, vocative) have different structures. It is assumed that D is required to turn nominal expressions into arguments and, additionally, to specify the reference of the noun. This theory implies that non-arguments are non-referring, simple NPs. A question that naturally arises here concerns the analysis of BPs in argument position. Are they interpreted as DPs with null D or as simple NPs?

In chapter 3.1., two solutions to the problem were presented. Whereas Chierchia (1998) suggests that articleless NPs are as such either inherently argumental or predicative, which renders a DP layer superfluous, Longobardi (1994) claims that all arguments are DPs.

According to the latter, bare arguments are embedded in a fully developed DP structure with an empty D head yielding an existential interpretation of the BP. For generics, Longobardi has developed a theory of N-to-D-Movement where an existential interpretation is excluded due to the filling of the D position.

Great emphasis in this thesis was placed on the semantic interpretation of BPs (3.2.). According to Carlson (1977), who initiated the analysis of BPs, three primary readings of bare NPs – kind, generic and existential – can be distinguished. Crucially, Carlson claims that the decisive factor for the interpretation of BPs is the context of the sentence. Following Milsark (1974), Carlson assumes that 'stage-level predicates' select an existential reading of BPs, whereas 'individual-level predicates' only allow for the universal reading. Carlson's theory, according to which BPs uniformly refer to kinds, is in contrast to the later developed Ambiguity Approach (Wilkinson 1991; Diesing 1992b; Gerstner-Link & Krifka 1993; Kratzer 1995) which holds that BPs are either kind-referring or indefinites. The Neocarlsonian Approach (Carlson 1989; Chierchia 1998; Dayal 2004) tries to combine elements of both theories.

In chapter 4, the focus shifted to language change. First, some theoretical issues associated with generative historical syntax were discussed (4.1.). Most importantly, in the generative approach to language change (Lightfoot 1979) the focus of study is the I-language of speakers who start using new patterns and that of older speakers who do not. The key mechanism of change is reanalysis which refers to new associations of form and content. The generative approach to linguistic change is particularly criticized for its emphasis on abrupt change. Moreover, disagreement exists on the causation of change, whereby two competing approaches have emerged to address this question. Whereas Lightfoot (1991) considers external factors such as input as the source of change, Elly van Gelderen (2011) claims that change comes from the inside. The latter refers to language change as a cyclical change for which Feature Economy is responsible.

Chapter 4.2. provided an overview of the diachronic development of English determiners (*the, this, a(n), some)*. From a generative perspective, the development of the definite article *the* involves the reanalysis of the OE demonstrative *sē* as head and the concomitant loss of the deictic function of the demonstrative. The indefinite article *a(n)* developed from the numeral *one* whereby the number features were lost during this process. Both the definite and the indefinite article appear in their present form in the ME period. The development of *some*, however, deviates from familiar patterns of determiner development. There is no general

agreement as to whether *some* developed from OE adjectives or whether a separate category of quantifiers already existed in OE. It is clear, however, that *some* marks specific reference from the EModE period onwards.

Finally, the development of the DP was discussed (4.3.). There is no consensus as to whether DP already existed in OE. Some scholars (Yamato 1989; Osawa 2000) argue against the interpretation of OE nominals as DPs, claiming that only with the development of the article system and the related declining number of determinerless NPs was a new functional structure developed. Other scholars, such as Wood (2005) and Allen (2006), believe in the existence of a prenominal structure in OE and, thus, in the existence of DP at that stage of the English language.

Based on the theoretical considerations, it was hypothesized that the number of determinerless plural NPs has declined in the course of the ModE period, while in the same period the number of plural nouns headed by *some* has increased. This hypothesis was supported by the observed trend towards obligatory marking of definiteness and indefiniteness in the history of English. To verify the central hypothesis of this thesis, quantitative and qualitative analyses of plural NPs were conducted in the empirical part with the aim of analysing the syntactic and semantic behaviour of BPs thoroughly. The data originate from both the 1.7-million-word Penn-Helsinki Parsed Corpus of Early Modern English and the 400-million-word Corpus of Historical American English.

Chapter 6.1. focused on the general development of the BP construction in ModE. Frequencies of plural NPs with and without overt D in EModE were compared to those in LModE. It was shown that the proportion of BPs has increased significantly from EModE to LModE, while in the same time period the proportion of NPs with overt D has decreased.

To analyse the development of BPs not only quantitatively but also qualitatively, a database was created in 6.2. The analysis of BP samples from both corpora showed that fixed expressions play a vital role in LModE. However, this observation does not hold true for the EModE period as the sample from PPCEME does not contain any fixed expressions. Furthermore, the syntactic analysis of BPs demonstrated that in PPCEME the BPs from the sample only occur in argument position, while in COHA BPs appear both as arguments and non-arguments. Finally, the semantic analysis revealed that the share of indefinite BPs within

all subject BPs has risen between EModE and LModE, while the share of subject NPs headed by *some* has declined over the last 500 years. Interestingly, this trend reversed noticeably in the first half of the 20th century as the frequency of '*some* + Pl. N' sequences has sharply increased since the 1930s.

As the semantic analysis concentrated solely on subject BPs, further instances of BPs were investigated in 6.3. and 6.4. The analysis of BPs in existential sentences clearly confirmed the results from the analysis of subject BPs with indefinite reference. It was shown that the proportion of '*there* [be] + Pl. N' sentences has steadily increased, while the proportion of '*there* [be] + *some* + Pl. N' sequences has decreased in the course of the ModE period. The same pattern was observed for BP objects. The share of indefinite object BPs has increased, while the share of object NPs headed by *some* has decreased significantly between EModE and LModE.

## 8.2. Final thoughts

Having summarized the main findings of this study, this last section will address, on the one hand, some theoretical implications and, on the other hand, potential directions for future research resulting from the limitations of the analysis.

Essentially, this study achieved two key outcomes. First, it demonstrated that the relative frequency of bare plural NPs has increased between the EModE and LModE period, while the proportion of plural NPs with overt D has decreased within the same period of time. This finding is particularly surprising, considering the trend towards obligatory reference marking in the history of English. Secondly, it shed new light on the use of the quantifying determiner *some* with plural NPs within the investigated time period. Most importantly, it seems that *some* is not on the verge of becoming a marker of indefiniteness, but instead continues to function as a marker of specificity. It remains to be seen whether the signs of a rising popularity of *some* with plural NPs in the 20th century marks the potential beginning of an extended use of the quantifier.

It is important to mention that these conclusions are only tentative, considering that this study is subject to limitations. As already discussed in the introduction (1.4.), the corpora used for the quantitative and qualitative analysis of BPs differ significantly in size. This has proved very problematic during the analysis since the results from PPCEME are not as reliable as the

ones from COHA. At some stages of the qualitative analysis, the evaluation of the data from PPCEME was impossible due to the small size of the sample. This was the case when analysing the role of fixed expressions containing bare NPs as well as the ratio of argumental and non-argumental BPs. Thus, there remain many open questions, which means that there is a need for further research in this area. Using additional data from the EModE period would have an impact on the outcome of the study and produce even more conclusive results.

Another limitation is the decision taken during the syntactic analysis to restrict the searches, so that they would only retrieve nouns in sentence-initial and post-verb position. A broader investigation of plural NPs including other positions of BPs would yield new insights. In terms of the semantic analysis of BPs, sentences that contain passive constructions and modal verbs were completely ignored. The incorporation of those sentences in the analysis might change the outcome and would therefore be highly interesting for future research.

Lastly, this study exclusively focused on the comparison of bare plural NPs with plural NPs headed by *some*. However, other quantifying determiners that express indefinite quantities like *a few, lots of* and *several* would deserve attention as well. It may be the case that by looking at a wider range of determiners, the outcome of the study might be different. An extensive study of quantifiers heading plural NPs would therefore be an interesting future project.

The theoretical implications and limitations of the analysis described above demonstrate once again that this thesis offers only a first approach to the study of the diachronic development of BPs – an area of research which is underrepresented in the linguistic literature thus far.

# REFERENCES


Abney, Steven P. 1987. "The English Noun Phrase in its Sentential Aspect". PhD thesis, Department of Linguistics and Philosophy, MIT.

Abraham, Werner; Leiss, Elisabeth; Stark, Elisabeth (eds.). 2007. *Nominal Determination. Typology, context constraints, and historical emergence*. Amsterdam: Benjamins.

Ackles, Nancy. 1997. "Historical Syntax of the English Articles in Relation to the Count/Non-count Distinction". PhD thesis, University of Washington.

Alexiadou, Artemis; Haegeman, Liliane; Stavrou, Melita. 2007. *Noun Phrase in the Generative Perspective*. Berlin: Mouton de Gruyter.

Allen, Cynthia. 2006. "Possessives and Determiners in Old English". In Nevalainen, Terttu; Klemola, Juhani; Laitinen, Mikko (eds.). *Types of Variation: Diachronic, Dialectal and Typological Interfaces*. Amsterdam: Benjamins, 149-170.

Barber, Charles. 1993. *The English language: a historical introduction*. Cambridge: Cambridge University Press.

Barber, Charles. 1997. *Early Modern English*. Edinburgh: Edinburgh University Press.

Biber, Douglas; Gray, Bethany. 2011. "Grammatical change in the noun phrase: the influence of written language use". *English Language and Linguistics* 15(2), 223-250.

Breban, Tine. 2012. "Functional Shifts and the Development of English Determiners". In Meurman-Solin, Anneli; López-Couso María J.; Los, Bettelou (eds.). *Information structure and syntactic change in the history of English*. Oxford: Oxford University Press, 271-300.

Carlson, Anita. 1976. "A diachronic treatment of English quantifiers". MA thesis, McGill University.

Carlson, Gregory N. 1977. "A Unified Analysis of the English Bare Plural". *Linguistics and Philosophy* 1, 413-457.

Carlson, Gregory N. 1980. *Reference to kinds in English*. New York, NY: Garland.

Carlson, Gregory N. 1989. "On the semantic composition of English generic sentences". In Chierchia, Gennaro; Partee, Barbara; Turner, Raymond (eds.). *Properties, Types and Meaning*. Dordrecht: Kluwer, 167-192.

Carlson, Gregory N. 2003. "No lack of determination". In Cheng, Lisa; Sybesma, Rint (eds.). *The Second Glot International State-of-the-Article Book: The Latest in linguistics*. Berlin: Mouton de Gruyter, 149-178.

Cheng, Lisa Lai-Shen; Sybesma, Rint. 1999. "Bare and Not-So-Bare Nouns and the Structure of NP". *Linguistic Inquiry* 30(4), 509-542.

Chierchia, Gennaro. 1995. "Individual-level predicates as inherent generics". In Carlson, Gregory N.; Pelletier, Francis J. (eds.). *The Generic Book*. Chicago, IL: The University of Chicago Press, 176-223.

Chierchia, Gennaro. 1998. "Reference to Kinds across Languages". *Natural Language Semantics* 6, 339-405.

Clark, Robin; Roberts, Ian. 1993. "A Computational Model of Language Learnability and Language Change". *Linguistic Inquiry* 24(2), 299-345.

Crisma, Paola. 2011. "The emergence of the definite article in English: A contact-induced change?". In Sleeman, Petra; Harry, Perridon (eds.). *The Noun Phrase in Romance and Germanic: Structure, variation, and change.* Amsterdam: John Benjamins, 175-192.

Crisma, Paola. 2015. "The 'indefinite article' from cardinal to operator to expletive". In Gianollo, Chiara; Jäger, Agnes; Penka, Doris (eds.). *Language Change at the Syntax-Semantics Interface*. Berlin: De Gruyter Mouton, 125-152.

Davies, Mark. 2012. “Expanding horizons in historical linguistics with the 400-million word Corpus of Historical American English”. *Corpora* 7(2), 121-157.
Dayal, Veneeta. 2003. “Bare Nominals: Non-Specific and Contrastive Readings under Scrambling.” In Karimi, Simin (ed.). *Word order and scrambling*. Malden, MA: Blackwell, 67-90.
Dayal, Veneeta. 2004. “Number marking and (in)definiteness in kind terms”. *Linguistics and Philosophy* 27, 393-450.
Dayal, Veneeta, 2011. “Bare Noun Phrases”. In Maienborn, Claudia ; Heusinger, Klaus von; Portner, Paul (eds.). *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter, 1088-1208.
Delfitto, Denis; Schroten, Jan. 1991. “Bare Plurals and the Number Affix in DP”. *Probus* 3(2), 155-185.
Delfitto, Denis. 2005. “Bare Plurals”. In Everaert, Martin; Riemsdijk, Henk van (eds*.). The Blackwell Companion to Syntax.* Malden, MA: Blackwell, 214-258.
Diesing, Molly. 1992a. “Bare Plural Subjects and the Derivation of Logical Representation”. *Linguistic Inquiry* 23(3), 353-380.
Diesing, Molly. 1992b. *Indefinites*. Cambridge, MA: MIT Press.
Eynde, Frank van. 2006. “NP-Internal Agreement and the Structure of the Noun Phrase”. *Journal of Linguistics* 42(1), 139-186.
Fischer, Olga C.M.; Leek, Frederike C. van der. 1981. “Optional vs radical re-analysis: mechanisms of syntactic change. Review of Lightfoot, David W.1979. *Principles of diachronic syntax*. Cambridge. Cambridge University Press”. *Lingua* 55, 301-350.
Gelderen, Elly van. 2007. “The Definiteness Cycle in Germanic”. *Journal of Germanic Linguistics* 19(4), 275-308.
Gelderen, Elly van. 2011. *The Linguistic Cycle. Language Change and the Language Faculty*. Oxford: Oxford University Press.
Gerstner-Link, Claudia; Krifka, Manfred. 1993. „Genericity“. In Jacobs, Joachim; Stechow, Arnim von; Sternefeld, Wolfgang; Vennemann, Theo (eds.). *Syntax: An International Handbook of Contemporary Research. Vol. 1*. Berlin: de Gruyter, 966-978.
Giorgi, Alessandra; Longobardi, Giuseppe. 1991. *The syntax of Noun Phrases: Configuration, parameters and empty categories.* Cambridge: Cambridge University Press.
Giusti, Guiliana. 1991. “The categorial status of quantified nominals”. *Linguistische Berichte* 136, 438-452.
Giusti, Guiliana. 1997. “The categorial status of determiners”. In Haegeman, Lilliane (ed.). *The New Comparative Syntax*. New York, NY: Longman, 95-123.
Henry, Alison. 1997. “Viewing change in progress: the loss of V2 in Hiberno-English imperatives”. In Kemenade, Ans van; Vincent, Nigel (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 273-296.
Hope, Jonathan. 2003. *Shakespeare’s Grammar*. London: Bloomsbury.
Hopper, Paul J.; Traugott, Elizabeth C. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
Ioup, Georgette. 1975. “Some universal for quantifier scope”. In Kimball, John P. (ed.). *Syntax and Semantics 4*. San Diego, CA: Academic Press, 37-58.
Kemenade, Ans van; Vincent, Nigel. 1997. “Parameters and morphosyntactic change”. In Kemenade, Ans; Vincent, Nigel (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 1-25.
Kiparsky, Paul. 1997. “The rise of positional licensing”. In Kemenade, Ans van; Vincent, Nigel (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 460-494.

Kratzer, Angelika. 1995. “Stage-level and Individual-level predicates”. In Carlson, Gregory N.; Pelletier, Francis J. (eds.). *The Generic Book*. Chicago, IL: The University of Chicago Press, 125-175.
Krifka, Manfred. 2003. “Bare NPs: Kind-referring, Indefinites, Both, or Neither?”. In Young, Robert B.; Zhou, Yuping (eds.). *Proceedings of Semantics and Linguistic Theory (SALT) XIII*. Ithaca, NY: Cornell, 180-203.
Krifka, Manfred; Pelletier, Francis J.; Carlson, Gregory N.; Meulen, Alice ter; Link, Godehard; Chierchia, Gennaro. 1995. “Genericity: An introduction”. In Carlson, Gregory N.; Pelletier, Francis J. (eds.). *The Generic Book*. Chicago, IL: The University of Chicago Press, 1-124.
Kroch, Anthony. 2005. “Modeling language change and language acquisition”. ftp://babel.ling.upenn.edu/papers/faculty/tony_kroch/papers/lsa-forum.pdf (5 Feb. 2017).
Kroch, Anthony. 2011a. “The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)”. http://www.helsinki.fi/varieng/CoRD/corpora/PPCEME/index.html (30 Sept. 2017).
Kroch, Anthony. 2011b. “Background”. http://www.helsinki.fi/varieng/CoRD/corpora/PPCEME/background.html (30 Sept. 2017).
Kroch, Anthony. 2011c. “Basic structure”. http://www.helsinki.fi/varieng/CoRD/corpora/PPCEME/basic.html (30 Sept. 2017).
Kroch, Anthony; Santorini, Beatrice; Delfs, Lauren. 2016. “Corpus description, PPCEME”. https://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-3/index.html (2 Oct. 2017).
Langacker, Ronald W. 1977. “Syntactic reanalysis”. In Li, Charles N. (ed.). *Mechanisms of Syntactic Change*. Austin, TX: University of Texas Press, 57-139.
Leech, Geoffrey; Hundt, Marianne; Mair, Christian; Smith, Nicholas. 2009. *Change in Contemporary English.* Cambridge: Cambridge University Press.
Levin, Beth. 1993. *English Verb Classes and Alternations. A Preliminary Investigation.* Chicago, IL: The University of Chicago Press.
Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.
Lightfoot, David. 2006. “Cuing a New Grammar”. In Kemenade, Ans van; Los, Bettelou (eds.). *The Handbook of the History of English*. Malden, MA: Blackwell, 24-44.
Longobardi, Giuseppe. 1994. “Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form”. *Linguistic Inquiry* 25(4), 609-665.
Longobardi, Giuseppe. 2001. “The Structure of DPs: Some Principles, Parameters, and Problems”. In Baltin, Mark; Collins, Chris (eds.). *The handbook of contemporary syntactic theory*. Malden, MA: Blackwell, 562-604.
Longobardi, Giuseppe. 2005. “Toward a Unified Grammar of Reference*”. Zeitschrift für Sprachwissenschaft* 24, 5-44.
Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
Meillet, Antoine. 1921. “L’évolution des formes grammaticales [The evolution of grammatical forms]”. In Meillet, Antoine. *Linguistique historique et linguistique générale*. Paris : Champion, 130-148.
Milsark, Gary L. 1974. “Existential sentences in English”. PhD thesis, Department of Foreign Literatures and Linguistics, MIT.
Milsark, Gary L. 1977. “Toward an explanation of certain peculiarities of the existential construction in English”. *Linguistic Analysis* 3, 1-29.

OED Online. 2018a. “book, n.”. Oxford: Oxford University Press. http://www.oed.com/view/Entry/21412 (15 Dec. 2018).
OED Online. 2018b. “horse, n.”. Oxford: Oxford University Press. http://www.oed.com/view/Entry/88583 (15 Dec. 2018).
Osawa, Fuyo. 2000. “The Historical Emergence of DP in English”. *English Linguistics* 17(1), 51-79.
Philippi, Julia. 1997. “The Rise of the Article in Germanic Languages”. In Kemenade, Ans van; Vincent, Nigel (eds.). *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 62-93.
Radford, Andrew. 2003. *Syntax: A minimalist introduction*. Cambridge: Cambridge University Press.
Radford, Andrew. 2004. *English Syntax. An introduction*. Cambridge: Cambridge University Press.
Santorini, Beatrice. 2016. “Annotation manual for the Penn Historical Corpora and the York-Helsinki Corpus of Early English Correspondence”. http://www.ling.upenn.edu/ppche/ppche-release-2016/annotation/index.html (30 Sept. 2017).
Sommerer, Lotte. 2011. “Old English se: from demonstrative to article. A usage-based study of nominal determination and category emergence”. PhD thesis, Department of English and American Studies, University of Vienna.
Stowell, Timothy. 1989. “Subjects, specifiers, and X-bar theory”. In Baltin, Mark R.; Kroch, Anthony S. (eds.). *Alternative Conceptions of Phrase Structure*. Chicago, IL: The University of Chicago Press, 232-262.
Swart, Henriëtte de; Zwarts, Joost. 2009. „Less form – more meaning: Why bare singular nouns are special”. *Lingua* 119, 280-295.
Szabolcsi, Anna. 1987. “Functional categories in the noun phrase”. In Kenesei, István (ed.). *Approaches to Hungarian 2: Theories and analyses*. Szeged: JATE Press, 167-190.
UCREL CLAWS7 Tagset. http://ucrel.lancs.ac.uk/claws7tags.html (30 Sept. 2017).
Wilkinson, Karina J. 1991. “Studies in the Semantics of Generic NPs”. PhD thesis, University of Massachusetts.
Wood, Johanna L. 2005. “Is there a DP in Old English?”. In Salmons, Joseph C.; Dubenion-Smith, Shannon (eds.). *Historical Linguistics 2005: Selected Papers from the 17th International Conference on Historical Linguistics, Madison, Wisconsin, 31 July-5 August 2005. Current Issues in Linguistic Theory 284, no. 284*. Amsterdam: Benjamins, 167-187.
Wood, Johanna L. 2007. “Demonstratives and possessives: From Old English to present-day English”. In Abraham, Werner; Leiss, Elisabeth; Stark, Elisabeth (eds.). *Nominal Determination. Typology, context constraints, and historical emergence*. Amsterdam: John Benjamins, 339-361.
Yamamoto, Keiko. 1989. “The Historical Development of Determiners: A parametric approach”. *English Linguistics* 6, 1-17.
Zamparelli, Roberto. 2002. “Definite and bare kind-denoting noun phrases”. In Beyssade, Claire (ed.). *Romance languages and linguistic theory 2000: selected papers from ‘Going Romance’ 2000, Utrecht, 30 November-2-December*. Amsterdam: John Benjamins, 305-343.
Zamparelli, Roberto. 2005. “Introduction: some questions about (in)definiteness”. *Lingua* 115, 759-766.

**Corpora**

Davies, Mark. (2010-) *The Corpus of Historical American English (COHA): 400 million words, 1810-2009.* Available online at https://corpus.byu.edu/coha/.

Kroch, Anthony, Beatrice Santorini, and Lauren Delfs. 2004. The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME). Department of Linguistics, University of Pennsylvania. CD-ROM, first edition, release 3 (http://ling.upenn.edu/ppche-release-2016/PPCEME-RELEASE-3).

The Corpus of Late Modern English Texts (CLMETEV). Extended Version. 2006. Compiled by Hendrik De Smet. Department of Linguistics, University of Leuven.

# APPENDIX 1 The composition of PPCEME

**Table 21 Word count summary by text genre (Kroch 2011c)**

| ***Text genre*** | **Number of words** | **Percentage** |
|---|---|---|
| *Bible* | 134,275 | 7.7% |
| *Travelogue* | 125,337 | 7.2% |
| *Diary, private* | 123,106 | 7.0% |
| *Drama, comedy* | 120,428 | 6.9% |
| *Letters, private* | 116,915 | 6.7% |
| *Fiction* | 116,494 | 6.7% |
| *Law* | 115,863 | 6.6% |
| *Educational treatise* | 113,032 | 6.5% |
| *Handbook, other* | 112,419 | 6.4% |
| *History* | 108,706 | 6.2% |
| *Proceedings, trials* | 105,090 | 6.0% |
| *Sermon* | 97,400 | 5.6% |
| *Philosophy* | 85,107 | 4.9% |
| *Science, other* | 79,050 | 4.5% |
| *Letters, non-private* | 59,868 | 3.4% |
| *Biography, other* | 52,755 | 3.0% |
| *Science, medicine* | 41,786 | 2.4% |
| *Biography, autobiography* | 41,379 | 2.4% |
| ***Total*** | **1,749,010** | **100.0%** |

# APPENDIX 2 The tagset used for part-of-speech tagging of PPCEME

| | |
|---|---|
| ADJ | adjective |
| ADJR | adjective, comparative |
| ADJS | adjective, superlative |
| ADV | adverb |
| ADVR | adverb, comparative |
| ADVS | adverb, superlative |
| ALSO | the word ALSO (except when =AS) |
| BAG | BE, present participle |
| BE | BE, infinitive |
| BED | BE, past (including past subjunctive) |
| BEI | BE, imperative |
| BEN | BE, perfect participle |
| BEP | BE, present (including present subjunctive) |
| C | complementizer |
| CODE | non-text material (e.g., page numbers) |
| CONJ | coordinating conjunction |
| D | determiner |
| DAG | DO, present participle |
| DAN | DO, passive participle (verbal or adjectival) |
| DO | DO, infinitive |
| DOD | DO, past (including past subjunctive) |
| DOI | DO, imperative |
| DON | DO, perfect participle |
| DOP | DO, present (including present subjunctive) |
| ELSE | the word ELSE in the collocation OR ELSE |
| EX | existential THERE |
| FOR | infinitival FOR |
| FOR+TO | cliticized FOR+TO |
| FP | focus particle |
| FW | foreign word |
| HAG | HAVE, present participle |
| HAN | HAVE, passive participle (verbal or adjectival) |
| HV | HAVE, infinitive |
| HVD | HAVE, past (including past subjunctive) |
| HVI | HAVE, imperative |
| HVN | HAVE, perfect participle |
| HVP | HAVE, present (including present subjunctive) |
| ID | token identification number |
| INTJ | interjection |
| LB | line break |
| LS | list marker |
| MD | modal verb |
| MD0 | modal verb, untensed |
| META | special text material (e.g., stage directions); generally found only in drama |
| N | common noun, singular |
| N$ | common noun, singular, possessive |
| NEG | negation |
| NPR | proper noun, singular |

| | |
|---|---|
| NPR$ | proper noun, singular, possessive |
| NPRS | proper noun, plural |
| NPRS$ | proper noun, plural, possessive |
| NS | common noun, plural |
| NS$ | common noun, plural, possessive |
| NUM | cardinal number |
| NUM$ | cardinal number, possessive |
| ONE | the word ONE (except as focus participle) |
| ONE$ | ONE, possessive |
| OTHER | the word OTHER (except as conjunction) |
| OTHER$ | OTHER, nominal use, possessive |
| OTHERS | OTHER, nominal use, plural |
| OHTERS$ | OTHER, nominal use, plural possessive |
| P | preposition or subordinating conjunction |
| PRO | personal pronoun |
| PRO$ | possessive pronoun |
| Q | quantifier |
| Q$ | quantifier, possessive |
| QR | quantifier, comparative (MORE, LEAST) |
| QS | quantifier, superlative (MOST, LEAST) |
| RP | adverbial particle |
| SUCH | the word SUCH |
| TO | infinitival TO, TIL, and AT |
| VAG | present participle |
| VAN | passive participle (verbal or adjectival) |
| VB | infinitive, verbs other than BE, DO, HV |
| VBD | past (including past subjunctive) |
| VBI | imperative |
| VBN | perfect participle |
| VBP | present (including present subjunctive) |
| WADV | wh-adverb |
| WARD | the morpheme WARD |
| WD | wh-determiner |
| WPRO | wh-pronoun |
| WPRO$ | possessive wh-pronoun |
| WQ | WHETHER introducing indirect questions |
| X | tag for unknown part of speech |
| $ | possessive marker |
| + | joins constituent morphemes in compounds (Example: N+N mankind) |

**Punctuation tags**

| | |
|---|---|
| . | sentence-final punctuation |
| , | sentence-internal punctuation |
| ‘ | single quote |
| “ | double quote |

(Santorini 2016)

# APPENDIX 3 The composition of COHA

**Table 22 Word count summary by genre and decade (Davies 2012: 123)**

| *Decade* | Fiction | Magazines | Newspaper | NF Books | Total |
|---|---|---|---|---|---|
| *1810s* | 641,164 | 88,316 | 0 | 451,542 | 1,181,022 |
| *1820s* | 3,751,204 | 1,714,789 | 0 | 1,461,012 | 6,927,005 |
| *1830s* | 7,590,350 | 3,145,575 | 0 | 3,038,062 | 13,773,987 |
| *1840s* | 8,850,886 | 3,554,534 | 0 | 3,641,434 | 16,046,854 |
| *1850s* | 9,094,346 | 4,220,558 | 0 | 3,178,922 | 16,493,826 |
| *1860s* | 9,450,562 | 4,437,941 | 262,198 | 2,974,401 | 17,125,102 |
| *1870s* | 10,291,968 | 4,452,192 | 1,030,560 | 2,835,440 | 18,610,160 |
| *1880s* | 11,215,065 | 4,481,568 | 1,355,456 | 3,820,766 | 20,872,855 |
| *1890s* | 11,212,219 | 4,679,486 | 1,383,948 | 3,907,730 | 21,183,383 |
| *1900s* | 12,029,439 | 5,062,650 | 1,433,576 | 4,015,567 | 22,541,232 |
| *1910s* | 11,935,701 | 5,694,710 | 1,489,942 | 3,534,899 | 22,655,252 |
| *1920s* | 12,539,681 | 5,841,678 | 3,552,699 | 3,698,353 | 25,632,411 |
| *1930s* | 11,876,996 | 5,910,095 | 3,545,527 | 3,080,629 | 24,413,247 |
| *1940s* | 11,946,743 | 5,644,216 | 3,497,509 | 3,056,010 | 24,144,478 |
| *1950s* | 11,986,437 | 5,796,823 | 3,522,545 | 3,092,375 | 24,398,180 |
| *1960s* | 11,578,880 | 5,803,276 | 3,404,244 | 3,141,582 | 23,927,982 |
| *1970s* | 11,626,911 | 5,755,537 | 3,383,924 | 3,002,933 | 23,769,305 |
| *1980s* | 12,152,603 | 5,804,320 | 4,113,254 | 3,108,755 | 25,178,952 |
| *1990s* | 13,272,162 | 7,440,305 | 4,060,570 | 3,104,303 | 27,877,340 |
| *2000s* | 14,590,078 | 7,678,830 | 4,088,704 | 3,121,839 | 29,479,451 |
| ***Total*** | **207,633,395** | **97,207,399** | **40,124,656** | **61,266,574** | **406,232,024** |

# APPENDIX 4 The C7 tagset used for part-of-speech tagging of COHA

| | |
|---|---|
| APPGE | possessive pronoun, pre-nominal (e.g. my, your, our) |
| AT | article (e.g. the, no) |
| AT1 | singular article (e.g. a, an, every) |
| BCL | before-clause marker (e.g. in order (that), in order (to)) |
| CC | coordinating conjunction (e.g. and, or) |
| CCB | adversative coordinating conjunction (but) |
| CS | subordinating conjunction (e.g. if, because, unless, so, for) |
| CSA | as (as conjunction) |
| CSN | than (as conjunction) |
| CST | that (as conjunction) |
| CSW | whether (as conjunction) |
| DA | after-determiner or post-determiner capable of pronominal function (e.g. such, former, same) |
| DA1 | singular after-determiner (e.g. little, much) |
| DA2 | plural after-determiner (e.g. few, several, many) |
| DAR | comparative after-determiner (e.g. more, less, fewer) |
| DAT | superlative after-determiner (e.g. most, least, fewest) |
| DB | before determiner or pre-determiner capable of pronominal function (all, half) |
| DB2 | plural before-determiner (both) |
| DD | determiner (capable of pronominal function) (e.g. any, some) |
| DD1 | singular determiner (e.g. this, that, another) |
| DD2 | plural determiner (these, those) |
| DDQ | wh-determiner (which, what) |
| DDQGE | wh-determiner, genitive (whose) |
| DDQV | wh-ever determiner (whichever, whatever) |
| EX | existential there |
| FO | formula |
| FU | unclassified word |
| FW | foreign word |
| GE | germanic genitive marker – (‘ or’s) |
| IF | for (as preposition) |
| II | general preposition |
| IO | of (as preposition) |
| IW | with, without (as prepositions) |
| JJ | general adjective |
| JJR | general comparative adjective (e.g. older, better, stronger) |
| JJT | general superlative adjective (e.g. oldest, best, strongest) |
| JK | catenative adjective (able in be able to, willing in be willing to) |
| MC | cardinal number, neutral for number (two, three…) |
| MC1 | singular cardinal number (one) |
| MC2 | plural cardinal number (e.g. sixes, sevens) |
| MCGE | genitive cardinal number, neutral for number (two’s, 100’s) |
| MCMC | hyphenated number (40-50, 1770-1827) |
| MD | ordinal number (e.g. first, second, next, last) |
| MF | fraction, neutral for number (e.g. quarters, two-thirds) |
| ND1 | singular noun of direction (e.g. north, southeast) |
| NN | common noun, neutral for number (e.g. sheep, cod, headquarters) |

| | |
|---|---|
| NN1 | singular common noun (e.g. book, girl) |
| NN2 | plural common noun (e.g. books, girls) |
| NNA | following noun of title (e.g. M.A.) |
| NNB | preceding noun of title (e.g. Mr., Prof.) |
| NNL1 | singular locative noun (e.g. Island, Street) |
| NNL2 | plural locative noun (e.g. Islands, Streets) |
| NNO | numeral noun, neutral for number (e.g. dozen, hundred) |
| NNO2 | numeral noun, plural (e.g. hundreds, thousands) |
| NNT1 | temporal noun, singular (e.g. day, week, year) |
| NNT2 | temporal noun, plural (e.g. days, weeks, years) |
| NNU | unit of measurement, neutral for number (e.g. in, cc) |
| NNU1 | singular unit of measurement (e.g. inch, centimetre) |
| NNU2 | plural unit of measurement (e.g. ins., feet) |
| NP | proper noun, neutral for number (e.g. IBM, Andes) |
| NP1 | singular proper noun (e.g. London, Jane, Frederick) |
| NP2 | plural proper noun (e.g. Browns, Reagans, Koreas) |
| NPD1 | singular weekday noun (e.g. Sunday) |
| NPD2 | plural weekday noun (e.g. Sundays) |
| NPM1 | singular month noun (e.g. October) |
| NPM2 | plural month noun (e.g. Octobers) |
| PN | indefinite pronoun, neutral for number (none) |
| PN1 | indefinite pronoun, singular (e.g. anyone, everything, nobody, one) |
| PNQO | objective wh-pronoun (whom) |
| PNQS | subjective wh-pronoun (who) |
| PNQV | wh-ever pronoun (whoever) |
| PNX1 | reflexive indefinite pronoun (oneself) |
| PPGE | nominal possessive personal pronoun (e.g. mine, yours) |
| PPH1 | 3rd person sing. neuter personal pronoun (it) |
| PPHO1 | 3rd person sing. objective personal pronoun (him, her) |
| PPHO2 | 3rd person plural objective personal pronoun (them) |
| PPHS1 | 3rd person sing. subjective personal pronoun (he, she) |
| PPHS2 | 3rd person plural subjective personal pronoun (they) |
| PPIO1 | 1st person sing. objective personal pronoun (me) |
| PPIO2 | 1st person plural objective personal pronoun (us) |
| PPIS1 | 1st person sing. subjective personal pronoun (I) |
| PPIS2 | 1st person plural subjective personal pronoun (we) |
| PPX1 | singular reflexive personal pronoun (e.g. yourself, itself) |
| PPX2 | plural reflexive personal pronoun (e.g. yourselves, themselves) |
| PPY | 2nd person personal pronoun (you) |
| RA | adverb, after nominal head (e.g. else, galore) |
| REX | adverb introducing appositional constructions (namely, e.g.) |
| RG | degree adverb (very, so, too) |
| RGQ | wh-degree adverb (how) |
| RGQV | wh-ever degree adverb (however) |
| RGR | comparative degree adverb (more, less) |
| RGT | superlative degree adverb (most, least) |
| RL | locative adverb (e.g. alongside, forward) |
| RP | prep. adverb, particle (e.g. about, in) |
| RPK | prep. adv., catenative (about in be about to) |
| RR | general adverb |
| RRQ | wh-general adverb (where, when, why, how) |

| | |
|---|---|
| RRQV | wh-ever general adverb (wherever, whenever) |
| RRR | comparative general adverb (e.g. better, longer) |
| RRT | superlative general adverb (e.g. best, longest) |
| RT | quasi-nominal adverb of time (e.g. now, tomorrow) |
| TO | infinitive marker (to) |
| UH | interjection (e.g. oh, yes, um) |
| VB0 | be, base form (finite i.e. imperative, subjunctive) |
| VBDR | were |
| VBDZ | was |
| VBG | being |
| VBI | be, infinitive (To be or not… It will be…) |
| VBM | am |
| VBN | been |
| VBR | are |
| VBZ | is |
| VD0 | do, base form (finite) |
| VDD | did |
| VDG | doing |
| VDI | do, infinitive (I may do… To do…) |
| VDN | done |
| VDZ | does |
| VH0 | have, base form (finite) |
| VHD | had (past tense) |
| VHG | having |
| VHI | have, infinitive |
| VHN | had (past participle) |
| VHZ | has |
| VM | modal auxiliary (can, will, would, etc.) |
| VMK | modal catenative (ought, used) |
| VV0 | base form of lexical verb (e.g. give, work) |
| VVD | past tense of lexical verb (e.g. gave, worked) |
| VVG | -ing participle of lexical verb (e.g. giving, working) |
| VVGK | -ing participle catenative (going in be going to) |
| VVI | infinitive (e.g. to give… It will work…) |
| VVN | past participle of lexical verb (e.g. given, worked) |
| VVNK | past participle catenative (e.g. bound in be bound to) |
| VVZ | -s form of lexical verb (e.g. gives, works) |
| XX | not, n't |
| ZZ1 | singular letter of the alphabet (e.g. A, b) |
| ZZ2 | plural letter of the alphabet (e.g. A's, b's) |

(UCREL CLAWS7 Tagset)

# APPENDIX 5 Plural NPs with overt D in sentence-initial and post-verb position

**Pl. NPs in sentence-initial position (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| *_ID* *_D* *_NS* | 490 | 158 | 184 | 148 |
| *_ID* *_SUCH* *_NS* | 11 | 5 | 4 | 2 |
| *_ID* *_WD* *_NS* | 15 | 5 | 4 | 6 |
| *_ID* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_N$* *_NS* | 1 | 1 | 0 | 0 |
| *_ID* *_NPR$* *_NS* | 3 | 3 | 0 | 0 |
| *_ID* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_PRO$* *_NS* | 152 | 33 | 70 | 49 |
| *_ID* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_ID* *_NUM* *_NS* | 42 | 15 | 9 | 18 |
| *_ID* *_Q(R)/(S)/($)* *_NS* | 0 | 0 | 0 | 0 |
| **TOTAL** | **714** | **220** | **271** | **223** |

**Auxiliary/lexical verb + Pl. NP (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| *_BAG* *_D* *_NS* | 3 | 0 | 2 | 1 |
| *_BAG* *_SUCH* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_WD* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_N$* *_NS* | 1 | 0 | 1 | 0 |
| *_BAG* *_NPR$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_PRO$* *_NS* | 1 | 1 | 0 | 0 |
| *_BAG* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_BAG* *_NUM* *_NS* | 15 | 0 | 11 | 4 |
| *_BAG* *_Q(R)/(S)/($)* *_NS* | 4 | 0 | 3 | 1 |
| *_BE(D)/(I)/(N)/(P)* *_D* *_NS* | 202 | 73 | 69 | 60 |
| *_BE(D)/(I)/(N)/(P)* *_SUCH* *_NS* | 14 | 6 | 4 | 4 |
| *_BE(D)/(I)/(N)/(P)* *_WD* *_NS* | 1 | 0 | 1 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_N$* *_NS* | 0 | 0 | 0 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NPR$* *_NS* | 8 | 2 | 6 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NS$* *_NS* | 1 | 0 | 1 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |

| | | | | |
|---|---|---|---|---|
| *_BE(D)/(I)/(N)/(P)* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_PRO$* *_NS* | 92 | 31 | 36 | 25 |
| *_BE(D)/(I)/(N)/(P)* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NUM* *_NS* | 275 | 75 | 139 | 61 |
| *_BE(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *_NS* | 175 | 56 | 46 | 73 |
| *_DA(G)/(N)* *_D* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_SUCH* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_WD* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_N$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_NPR$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_PRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_DA(G)/(N)* *_NUM* *_NS* | 3 | 1 | 1 | 1 |
| *_DA(G)/(N)* *_Q(R)/(S)/($)* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_D* *_NS* | 41 | 12 | 18 | 11 |
| *_DO(D)/(I)/(N)/(P)* *_SUCH* *_NS* | 10 | 5 | 3 | 2 |
| *_DO(D)/(I)/(N)/(P)* *_WD* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_N$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NPR$* *_NS* | 1 | 0 | 1 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_PRO$* *_NS* | 19 | 3 | 12 | 4 |
| *_DO(D)/(I)/(N)/(P)* *_WPRO$* *_NS* | 1 | 0 | 0 | 1 |
| *_DO(D)/(I)/(N)/(P)* *_NUM* *_NS* | 1 | 0 | 1 | 0 |
| *_DO(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *_NS* | 39 | 13 | 16 | 10 |
| *_HA(G)/(N)* *_D* *_NS* | 4 | 2 | 2 | 0 |
| *_HA(G)/(N)* *_SUCH* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_WD* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_N$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_NPR$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_HA(G)/(N)* *_PRO$* *_NS* | 5 | 1 | 1 | 3 |
| *_HA(G)/(N)* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |

| | | | | |
|---|---|---|---|---|
| *_HA(G)/(N)* *_NUM* *_NS* | 6 | 2 | 3 | 1 |
| *_HA(G)/(N)* *_Q(R)/(S)/($)* *_NS* | 8 | 2 | 3 | 3 |
| *_HV(D)/(I)/(N)/(P)* *_D* *_NS* | 53 | 19 | 22 | 12 |
| *_HV(D)/(I)/(N)/(P)* *_SUCH* *_NS* | 13 | 6 | 2 | 5 |
| *_HV(D)/(I)/(N)/(P)* *_WD* *_NS* | 3 | 1 | 1 | 1 |
| *_HV(D)/(I)/(N)/(P)* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_N$* *_NS* | 1 | 1 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NPR$* *_NS* | 1 | 0 | 1 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_PRO$* *_NS* | 96 | 26 | 34 | 36 |
| *_HV(D)/(I)/(N)/(P)* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NUM* *_NS* | 99 | 39 | 38 | 22 |
| *_HV(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *_NS* | 162 | 39 | 60 | 63 |
| *_MD* *_D* *_NS* | 21 | 7 | 11 | 3 |
| *_MD* *_SUCH* *_NS* | 2 | 1 | 1 | 0 |
| *_MD* *_WD* *_NS* | 1 | 0 | 0 | 1 |
| *_MD* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_N$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_NPR$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_NPRS$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_NS$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_ONE$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_OTHERS$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_PRO$* *_NS* | 16 | 7 | 5 | 4 |
| *_MD* *_WPRO$* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_NUM* *_NS* | 0 | 0 | 0 | 0 |
| *_MD* *_Q(R)/(S)/($)* *_NS* | 14 | 6 | 7 | 1 |
| *_VB(D)/(I)/(N)/(P)* *_D* *_NS* | 1,494 | 441 | 600 | 453 |
| *_VB(D)/(I)/(N)/(P)* *_SUCH* *_NS* | 96 | 23 | 38 | 35 |
| *_VB(D)/(I)/(N)/(P)* *_WD* *_NS* | 45 | 17 | 7 | 21 |
| *_VB(D)/(I)/(N)/(P)* *_$* *_NS* | 0 | 0 | 0 | 0 |
| *_VB(D)/(I)/(N)/(P)* *_N$* *_NS* | 2 | 0 | 1 | 1 |
| *_VB(D)/(I)/(N)/(P)* *_NPR$* *_NS* | 5 | 2 | 3 | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NPRS$* *_NS* | 1 | 0 | 0 | 1 |
| *_VB(D)/(I)/(N)/(P)* *_NS$* *_NS* | 29 | 8 | 17 | 4 |
| *_VB(D)/(I)/(N)/(P)* *_ONE$* *_NS* | 1 | 0 | 0 | 1 |
| *_VB(D)/(I)/(N)/(P)* *_OTHER$* *_NS* | 0 | 0 | 0 | 0 |
| *_VB(D)/(I)/(N)/(P)* *_OTHERS$* *_NS* | 1 | 0 | 1 | 0 |
| *_VB(D)/(I)/(N)/(P)* *_PRO$* *_NS* | 1,623 | 452 | 692 | 479 |
| *_VB(D)/(I)/(N)/(P)* *_WPRO$* *_NS* | 5 | 2 | 0 | 3 |
| *_VB(D)/(I)/(N)/(P)* *_NUM* *_NS* | 348 | 87 | 188 | 73 |
| *_VB(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *_NS* | 597 | 187 | 214 | 196 |
| **TOTAL** | **5,659** | **1,656** | **2,323** | **1,680** |

**Total number of Pl. NPs with overt D (PPCEME)**

| | raw number | pmw | % of total number of NPs |
|---|---|---|---|
| NPs with overt D | 6,373 | 3,552.377077050853 | 9.245611489917307% |

**Pl. NPs in sentence-initial position (COHA)**

| **SEARCH TERM** | **HITS** | **1810** | **1820** | **1830** | **1840** | **1850** | **1860** | **1870** | **1880** | **1890** | **1900** | **1910** | **1920** | **1930** | **1940** | **1950** | **1960** | **1970** | **1980** | **1990** | **2000** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ! [at*] [nn2*] | 3,839 | 46 | 74 | 250 | 335 | 228 | 286 | 223 | 241 | 236 | 194 | 222 | 227 | 201 | 160 | 216 | 176 | 91 | 166 | 138 | 129 |
| ! [d*] [nn2*] | 1,772 | 29 | 75 | 99 | 163 | 142 | 164 | 126 | 130 | 118 | 98 | 93 | 85 | 74 | 60 | 62 | 59 | 52 | 49 | 36 | 58 |
| ! [appge*] [nn2*] | 2,127 | 44 | 95 | 189 | 243 | 157 | 170 | 113 | 147 | 127 | 108 | 85 | 129 | 48 | 66 | 70 | 81 | 69 | 62 | 52 | 72 |
| ! [ge*] [nn2*] | 158 | 0 | 0 | 0 | 2 | 14 | 12 | 7 | 15 | 19 | 13 | 42 | 19 | 3 | 2 | 3 | 3 | 1 | 0 | 1 | 2 |
| ! [m*] [nn2*] | 498 | 1 | 6 | 15 | 45 | 14 | 27 | 27 | 30 | 28 | 27 | 27 | 35 | 25 | 24 | 35 | 26 | 27 | 27 | 26 | 26 |
| . [at*] [nn2*] | 218,828 | 507 | 3,608 | 8,091 | 8,672 | 8,875 | 9,020 | 9,965 | 10,730 | 12,298 | 11,766 | 12,392 | 13,877 | 14,150 | 14,231 | 13,796 | 13,225 | 13,043 | 14,400 | 13,170 | 13,012 |
| . [d*] [nn2*] | 81,129 | 149 | 1,483 | 2,631 | 3,045 | 2,872 | 2,855 | 3,180 | 3,607 | 3,520 | 3,985 | 4,221 | 4,957 | 4,961 | 5,581 | 5,440 | 5,462 | 5,312 | 5,732 | 6,281 | 5,855 |
| . [appge*] [nn2*] | 68,797 | 131 | 1,795 | 2,521 | 3,019 | 2,900 | 2,652 | 2,599 | 2,721 | 3,012 | 3,058 | 3,442 | 4,197 | 4,266 | 4,225 | 3,707 | 4,013 | 4,074 | 4,591 | 5,700 | 6,174 |
| . [ge*] [nn2*] | 937 | 0 | 6 | 8 | 28 | 33 | 50 | 57 | 78 | 53 | 95 | 94 | 138 | 63 | 50 | 47 | 43 | 32 | 21 | 16 | 25 |
| . [m*] [nn2*] | 20,657 | 30 | 198 | 1,184 | 559 | 556 | 705 | 609 | 730 | 826 | 930 | 893 | 1,527 | 1,502 | 1,451 | 1,500 | 1,626 | 1,433 | 1,420 | 1,325 | 1,653 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ? [at*] [nn2*] | 5,895 | 12 | 81 | 120 | 224 | 149 | 176 | 156 | 183 | 176 | 439 | 396 | 473 |
| ? [d*] [nn2*] | 3,556 | 14 | 71 | 86 | 147 | 103 | 138 | 104 | 131 | 210 | 239 | 263 | 313 |
| ? [appge*] [nn2*] | 1,879 | 13 | 73 | 80 | 94 | 67 | 84 | 85 | 68 | 130 | 123 | 74 | 136 |
| ? [ge*] [nn2*] | 140 | 0 | 1 | 2 | 1 | 7 | 9 | 8 | 10 | 23 | 20 | 10 | 3 |
| ? [m*] [nn2*] | 1,171 | 4 | 5 | 14 | 23 | 10 | 11 | 20 | 25 | 96 | 69 | 70 | 134 |
| **TOTAL** | **411,383** | **980** | **7,571** | **15,290** | **16,600** | **16,127** | **16,359** | **17,279** | **18,846** | **21,087** | **22,401** | **26,081** | **25,193** |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ? [at*] [nn2*] | 298 | 322 | 486 | 473 | 459 | 469 | 455 | 390 | 431 | 439 | 396 | 176 | 183 |
| ? [d*] [nn2*] | 201 | 253 | 263 | 313 | 240 | 226 | 241 | 211 | 210 | 239 | 263 | 102 | 131 |
| ? [appge*] [nn2*] | 126 | 128 | 142 | 136 | 102 | 95 | 110 | 91 | 130 | 123 | 74 | 58 | 68 |
| ? [ge*] [nn2*] | 2 | 2 | 8 | 3 | 2 | 4 | 9 | 7 | 23 | 20 | 10 | 12 | 10 |
| ? [m*] [nn2*] | 60 | 72 | 103 | 134 | 118 | 99 | 113 | 108 | 96 | 69 | 70 | 17 | 25 |
| **TOTAL** | **27,693** | **27,522** | **27,470** | **25,193** | **25,635** | **25,769** | **26,778** | **26,100** | **26,081** | **22,401** | **21,087** | **20,602** | **18,846** |

**Auxiliary/lexical verb + Pl. NP (COHA)**

| **SEARCH TERM** | **HITS** | **1810** | **1820** | **1830** | **1840** | **1850** | **1860** | **1870** | **1880** | **1890** | **1900** | **1910** | **1920** | **1930** | **1940** | **1950** | **1960** | **1970** | **1980** | **1990** | **2000** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [vb*] [at*] [nn2*] | 52,732 | 189 | 1,139 | 2,157 | 2,666 | 2,525 | 2,472 | 2,422 | 2,786 | 2,953 | 3,020 | 3,051 | 3,280 | 2,903 | 2,951 | 2,867 | 2,902 | 2,989 | 3,225 | 3,101 | 3,134 |
| [vb*] [d*] [nn2*] | 27,186 | 49 | 459 | 993 | 1,155 | 1,102 | 1,050 | 1,191 | 1,403 | 1,412 | 1,670 | 1,531 | 1,646 | 1,580 | 1,677 | 1,597 | 1,655 | 1,670 | 1,716 | 1,823 | 1,807 |
| [vb*] [appge*] [nn2*] | 12,684 | 100 | 447 | 639 | 809 | 756 | 651 | 633 | 681 | 708 | 691 | 652 | 679 | 630 | 608 | 594 | 584 | 599 | 632 | 706 | 885 |
| [vb*] [ge*] [nn2*] | 230 | 0 | 1 | 1 | 7 | 9 | 6 | 5 | 14 | 16 | 31 | 32 | 12 | 23 | 8 | 13 | 13 | 13 | 9 | 9 | 8 |
| [vb*] [m*] [nn2*] | 19,537 | 25 | 248 | 567 | 613 | 627 | 689 | 746 | 903 | 1,069 | 1,176 | 1,154 | 1,521 | 1,384 | 1,293 | 1,362 | 1,204 | 1,099 | 1,209 | 1,266 | 1,382 |
| [vd*] [at*] [nn2*] | 5,469 | 15 | 78 | 167 | 231 | 227 | 169 | 207 | 240 | 253 | 367 | 305 | 355 | 272 | 331 | 360 | 297 | 317 | 400 | 447 | 431 |

| 339 | 223 | 2 | 109 | 1,473 | 1,408 | 1,025 | 1 | 1,394 | 131 | 46 | 53 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 322 | 196 | 2 | 98 | 1,372 | 1,246 | 909 | 0 | 1,253 | 133 | 60 | 57 | 0 |
| 262 | 158 | 2 | 50 | 1,386 | 1,224 | 859 | 1 | 1,029 | 152 | 54 | 50 | 0 |
| 217 | 130 | 1 | 42 | 1,272 | 1,113 | 773 | 1 | 866 | 161 | 40 | 43 | 1 |
| 221 | 136 | 0 | 45 | 1,264 | 1,094 | 776 | 3 | 1,026 | 158 | 43 | 55 | 3 |
| 203 | 114 | 1 | 48 | 1,284 | 1,016 | 791 | 2 | 947 | 153 | 50 | 46 | 1 |
| 164 | 116 | 1 | 35 | 1,240 | 983 | 834 | 3 | 816 | 159 | 46 | 41 | 1 |
| 193 | 108 | 0 | 31 | 1,244 | 907 | 784 | 2 | 763 | 147 | 61 | 34 | 5 |
| 200 | 115 | 0 | 35 | 1,323 | 906 | 821 | 4 | 706 | 182 | 55 | 45 | 0 |
| 197 | 116 | 2 | 19 | 1,127 | 819 | 780 | 1 | 535 | 168 | 49 | 70 | 0 |
| 204 | 92 | 1 | 17 | 1,169 | 800 | 743 | 5 | 476 | 163 | 59 | 51 | 2 |
| 189 | 98 | 0 | 15 | 1,082 | 756 | 741 | 4 | 471 | 147 | 53 | 81 | 0 |
| 139 | 95 | 0 | 8 | 1,094 | 804 | 675 | 5 | 443 | 144 | 74 | 62 | 2 |
| 138 | 92 | 1 | 7 | 1,024 | 645 | 635 | 1 | 364 | 132 | 55 | 71 | 1 |
| 121 | 103 | 0 | 7 | 912 | 613 | 646 | 1 | 300 | 149 | 67 | 94 | 3 |
| 105 | 105 | 0 | 9 | 922 | 622 | 609 | 2 | 266 | 159 | 86 | 106 | 0 |
| 119 | 111 | 0 | 2 | 918 | 525 | 553 | 0 | 285 | 180 | 115 | 140 | 0 |
| 88 | 89 | 0 | 2 | 737 | 425 | 489 | 0 | 239 | 166 | 58 | 108 | 0 |
| 75 | 56 | 0 | 1 | 367 | 269 | 282 | 0 | 94 | 85 | 58 | 88 | 0 |
| 15 | 12 | 0 | 0 | 67 | 32 | 74 | 0 | 12 | 24 | 20 | 40 | 0 |
| 3,511 | 2,265 | 13 | 580 | 21,277 | 16,207 | 13,799 | 36 | 12,285 | 2,893 | 1,149 | 1,335 | 20 |
| [vd*] [d*] [nn2*] | [vd*] [appge*] [nn2*] | [vd*] [ge*] [nn2*] | [vd*] [m*] [nn2*] | [vh*] [at*] [nn2*] | [vh*] [d*] [nn2*] | [vh*] [appge*] [nn2*] | [vh*] [ge*] [nn2*] | [vh*] [m*] [nn2*] | [vm*] [at*] [nn2*] | [vm*] [d*] [nn2*] | [vm*] [appge*] [nn2*] | [vm*] [ge*] [nn2*] |

| 4 | 5,587 | 1,968 | 5,212 | 4 | 624 | 8,780 | 2,040 | 12,568 | 2 | 2,444 | 9,805 | 3,973 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 5,225 | 1,970 | 4,590 | 3 | 528 | 8,106 | 1,935 | 11,186 | 9 | 2,079 | 10,150 | 3,920 |
| 9 | 4,994 | 1,909 | 3,277 | 2 | 365 | 7,951 | 1,955 | 9,000 | 5 | 1,876 | 9,310 | 3,509 |
| 6 | 3,774 | 1,309 | 3,089 | 6 | 304 | 6,688 | 1,736 | 7,117 | 4 | 1,666 | 8,620 | 3,369 |
| 8 | 3,499 | 1,232 | 2,934 | 8 | 362 | 6,821 | 1,716 | 7,495 | 8 | 1,791 | 8,374 | 3,334 |
| 8 | 3,604 | 1,175 | 2,897 | 10 | 355 | 7,718 | 1,703 | 8,321 | 7 | 1,886 | 8,768 | 3,322 |
| 3 | 3,362 | 1,240 | 2,843 | 4 | 295 | 7,677 | 1,678 | 8,962 | 5 | 1,770 | 8,883 | 3,259 |
| 9 | 3,367 | 1,142 | 2,711 | 5 | 264 | 7,237 | 1,542 | 8,167 | 4 | 1,525 | 8,879 | 3,206 |
| 7 | 3,501 | 1,135 | 2,845 | 17 | 272 | 7,328 | 1,509 | 7,828 | 14 | 1,304 | 9,241 | 3,169 |
| 8 | 3,485 | 1,103 | 2,656 | 6 | 221 | 6,170 | 1,443 | 7,295 | 17 | 838 | 8,664 | 2,890 |
| 1 | 3,203 | 1,008 | 2,463 | 2 | 217 | 6,330 | 1,449 | 7,442 | 4 | 788 | 8,164 | 2,722 |
| 4 | 3,176 | 876 | 2,666 | 10 | 165 | 6,248 | 1,388 | 6,661 | 1 | 822 | 7,917 | 2,444 |
| 4 | 3,367 | 1,042 | 2,666 | 5 | 207 | 5,180 | 1,356 | 6,051 | 1 | 519 | 7,580 | 2,618 |
| 3 | 2,827 | 942 | 2,443 | 5 | 281 | 4,980 | 1,261 | 6,049 | 5 | 577 | 6,784 | 2,278 |
| 4 | 3,050 | 959 | 2,676 | 2 | 169 | 4,657 | 1,093 | 5,483 | 0 | 501 | 6,347 | 2,219 |
| 2 | 2,925 | 870 | 2,568 | 1 | 125 | 4,520 | 1,151 | 5,216 | 1 | 353 | 6,864 | 2,259 |
| 1 | 3,365 | 1,084 | 2,891 | 1 | 117 | 4,278 | 1,105 | 4,931 | 1 | 340 | 6,434 | 2,192 |
| 0 | 3,125 | 901 | 2,651 | 1 | 182 | 3,442 | 790 | 3,963 | 1 | 308 | 6,078 | 1,909 |
| 0 | 1,429 | 479 | 1,392 | 0 | 61 | 1,656 | 440 | 2,225 | 1 | 109 | 3,189 | 1,144 |
| 0 | 364 | 137 | 471 | 0 | 9 | 173 | 39 | 169 | 0 | 6 | 638 | 206 |
| 92 | 67,229 | 22,481 | 55,941 | 92 | 5,123 | 115,940 | 27,329 | 136,129 | 90 | 21,502 | 150,689 | 53,942 |
| [vm*] [m*] [nn2*] | [vv0*] [at*] [nn2*] | [vv0*] [d*] [nn2*] | [vv0*] [appge*] [nn2*] | [vv0*] [ge*] [nn2*] | [vv0*] [m*] [nn2*] | [vvd*] [at*] [nn2*] | [vvd*] [d*] [nn2*] | [vvd*] [appge*] [nn2*] | [vvd*] [ge*] [nn2*] | [vvd*] [m*] [nn2*] | [vvi*] [at*] [nn2*] | [vvi*] [d*] [nn2*] |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [vvi*] [appge*] [nn2*] | 122,067 | 638 | 3,087 | 5,357 | 5,917 | 5,918 | 5,571 | 5,749 | 6,201 | 6,187 | 6,198 | 6,476 | 7,056 |
| [vvi*] [ge*] [nn2*] | 161 | 0 | 0 | 1 | 4 | 2 | 6 | 7 | 11 | 7 | 12 | 13 | 25 |
| [vvi*] [m*] [nn2*] | 9,566 | 16 | 97 | 216 | 218 | 254 | 269 | 291 | 328 | 406 | 436 | 571 | 630 |
| [vvz*] [at*] [nn2*] | 29,544 | 145 | 468 | 1,084 | 1,251 | 1,190 | 1,194 | 1,194 | 1,328 | 1,375 | 1,493 | 1,568 | 1,820 |
| [vvz*] [d*] [nn2*] | 10,703 | 41 | 173 | 389 | 480 | 449 | 486 | 460 | 492 | 442 | 555 | 566 | 580 |
| [vvz*] [appge*] [nn2*] | 21,762 | 149 | 362 | 690 | 869 | 762 | 790 | 755 | 939 | 913 | 857 | 1,052 | 1,418 |
| [vvz*] [ge*] [nn2*] | 59 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 4 | 9 | 7 | 7 |
| [vvz*] [m*] [nn2*] | 5,908 | 3 | 44 | 93 | 103 | 117 | 109 | 132 | 190 | 184 | 277 | 275 | 349 |
| **TOTAL** | **1,049,557** | **3,878** | **20,103** | **38,106** | **44,012** | **43,785** | **43,639** | **45,394** | **49,658** | **51,944** | **54,367** | **55,932** | **61,940** |

| 9,567 | 8,684 | 7,034 | 6,399 | 6,260 | 6,514 | 6,641 | 6,613 | 7,056 | 6,476 | 6,198 | 6,187 | 6,201 | 5,749 | 5,571 | 5,918 | 5,917 | 5,357 | 3,087 | 638 | 122,067 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 6 | 12 | 15 | 10 | 9 | 4 | 13 | 25 | 13 | 12 | 7 | 11 | 7 | 6 | 2 | 4 | 1 | 0 | 0 | 161 |
| 893 | 826 | 718 | 639 | 738 | 709 | 689 | 622 | 630 | 571 | 436 | 406 | 328 | 291 | 269 | 254 | 218 | 216 | 97 | 16 | 9,566 |
| 2,463 | 2,526 | 1,969 | 1,838 | 1,862 | 1,632 | 1,570 | 1,574 | 1,820 | 1,568 | 1,493 | 1,375 | 1,328 | 1,194 | 1,194 | 1,190 | 1,251 | 1,084 | 468 | 145 | 29,544 |
| 887 | 958 | 752 | 692 | 679 | 564 | 513 | 545 | 580 | 566 | 555 | 442 | 492 | 460 | 486 | 449 | 480 | 389 | 173 | 41 | 10,703 |
| 2,230 | 2,192 | 1,389 | 1,573 | 1,394 | 1,172 | 1,142 | 1,114 | 1,418 | 1,052 | 857 | 913 | 939 | 755 | 790 | 762 | 869 | 690 | 362 | 149 | 21,762 |
| 2 | 5 | 3 | 6 | 1 | 5 | 1 | 4 | 7 | 7 | 9 | 4 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 59 |
| 857 | 773 | 514 | 377 | 443 | 385 | 352 | 331 | 349 | 275 | 277 | 184 | 190 | 132 | 109 | 117 | 103 | 93 | 44 | 3 | 5,908 |
| **83,766** | **78,682** | **68,971** | **60,574** | **60,448** | **62,213** | **62,200** | **59,945** | **61,940** | **55,932** | **54,367** | **51,944** | **49,658** | **45,394** | **43,639** | **43,785** | **44,012** | **38,106** | **20,103** | **3,878** | **1,049,557** |

**Total number of NPs with overt D (COHA)**

| | raw number | pmw | % of total number of NPs |
|---|---|---|---|
| NPs with overt D | 1,460,940 | 3,596.319132142078 | 8.179282454315488% |

# APPENDIX 6 BPs in sentence-initial and post-verb position

**BPs in sentence-initial position (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| *_ID* *_NS* | 180 | 54 | 55 | 71 |

**Auxiliary/lexical verb + BP (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| *_BAG* *_NS* | 30 | 4 | 15 | 11 |
| *_BE(D)/(I)/(N)/(P)* *_NS* | 530 | 167 | 228 | 135 |
| *_DA(G)/(N)* *_NS* | 3 | 0 | 0 | 3 |
| *_DO(D)/(I)/(N)/(P)* *_NS* | 35 | 15 | 6 | 14 |
| *_HA(G)/(N)* *_NS* | 19 | 4 | 7 | 8 |
| *_HV(D)/(I)/(N)/(P)* *_NS* | 175 | 60 | 63 | 52 |
| *_MD* *_NS* | 15 | 3 | 6 | 6 |
| *_VB(D)/(I)/(N)/(P)* *_NS* | 1,465 | 403 | 599 | 463 |
| **TOTAL** | **2,272** | **656** | **924** | **692** |

**Total number of BPs (PPCEME)**

| | raw number | pmw | % of total number of NPs |
|---|---|---|---|
| BPs | 2,452 | 1,366.770530822013 | 3.55723197446685% |

**BPs in sentence-initial position (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ! [nn2*] | 8,800 | 107 | 123 | 547 | 658 | 371 | 723 | 638 | 708 | 452 | 434 | 531 | 539 | 451 | 304 | 496 | 425 | 299 | 368 | 296 | 330 |
| . [nn2*] | 258,552 | 759 | 2,415 | 7,328 | 6,925 | 6,193 | 8,060 | 8,287 | 9,407 | 9,049 | 10,108 | 11,325 | 16,409 | 18,021 | 18,772 | 18,023 | 18,616 | 18,534 | 20,430 | 23,969 | 25,922 |
| ? [nn2*] | 15,932 | 62 | 153 | 345 | 326 | 171 | 445 | 423 | 566 | 391 | 872 | 1,229 | 1,841 | 1,041 | 1,125 | 1,348 | 1,208 | 1,436 | 1,236 | 908 | 806 |
| **TOTAL** | **283,284** | **928** | **2,691** | **8,220** | **7,909** | **6,735** | **9,228** | **9,348** | **10,681** | **9,892** | **11,414** | **13,085** | **18,789** | **19,513** | **20,201** | **19,867** | **20,249** | **20,269** | **22,034** | **25,173** | **27,058** |

**Auxiliary/lexical verb + BP (COHA)**

| **SEARCH TERM** | **HITS** | **1810** | **1820** | **1830** | **1840** | **1850** | **1860** | **1870** | **1880** | **1890** | **1900** | **1910** | **1920** | **1930** | **1940** | **1950** | **1960** | **1970** | **1980** | **1990** | **2000** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [vb*] [nn2*] | 122,777 | 390 | 1,815 | 3,783 | 4,557 | 4,535 | 4,916 | 5,137 | 6,126 | 6,089 | 6,692 | 6,607 | 7,647 | 7,402 | 7,661 | 7,571 | 7,568 | 8,018 | 8,712 | 8,791 | 8,760 |
| [vd*] [nn2*] | 9,465 | 11 | 51 | 112 | 166 | 184 | 203 | 260 | 292 | 301 | 445 | 490 | 562 | 576 | 544 | 628 | 621 | 708 | 830 | 1,199 | 1,282 |
| [vh*] [nn2*] | 47,572 | 191 | 687 | 1,287 | 1,521 | 1,399 | 1,578 | 1,576 | 1,944 | 1,950 | 2,158 | 2,279 | 2,704 | 2,778 | 2,818 | 2,910 | 3,191 | 3,236 | 3,916 | 4,599 | 4,850 |
| [vm*] [nn2*] | 4,308 | 44 | 148 | 229 | 324 | 266 | 278 | 200 | 361 | 241 | 260 | 211 | 295 | 140 | 151 | 143 | 185 | 178 | 223 | 222 | 209 |
| [vv0*] [nn2*] | 123,115 | 363 | 1,371 | 2,762 | 2,991 | 2,914 | 3,521 | 3,305 | 4,012 | 3,832 | 4,637 | 4,743 | 6,614 | 6,984 | 7,801 | 7,962 | 8,684 | 9,060 | 10,577 | 14,883 | 16,099 |
| [vvd*] [nn2*] | 119,686 | 125 | 1,214 | 2,510 | 2,979 | 3,171 | 3,955 | 3,893 | 4,012 | 4,571 | 5,140 | 5,239 | 7,218 | 7,899 | 8,537 | 8,750 | 8,318 | 8,449 | 9,891 | 11,497 | 12,318 |
| [vvi*] [nn2*] | 188,773 | 482 | 2,091 | 4,319 | 4,684 | 5,230 | 5,420 | 5,819 | 7,245 | 7,062 | 8,137 | 9,150 | 10,927 | 11,766 | 12,467 | 12,622 | 12,783 | 13,577 | 15,706 | 19,395 | 19,891 |
| [vvz*] [nn2*] | 42,481 | 82 | 339 | 885 | 949 | 1,041 | 1,062 | 1,244 | 1,417 | 1,305 | 1,820 | 1,877 | 2,195 | 2,211 | 2,336 | 2,472 | 2,803 | 3,082 | 3,725 | 5,689 | 5,947 |
| **TOTAL** | **658,177** | **1,688** | **7,716** | **15,887** | **18,171** | **18,740** | **20,933** | **21,434** | **25,409** | **25,351** | **29,289** | **30,596** | **38,162** | **39,756** | **42,315** | **43,058** | **44,153** | **46,308** | **53,580** | **66,275** | **69,356** |

**Total number of BPs (COHA)**

| | raw number | pmw | % of total number of NPs |
|---|---|---|---|
| BPs | 941,461 | 2,317.545009696232 | 5.27090464955598% |

# APPENDIX 7 Sample of nouns

**'Books' with overt D in sentence-initial position or in post-verb position (PPCEME)**

| SEARCH TERM | HITS |
|---|---|
| *_ID* *_D* *books/bookes/boks/bokes/bockes*_NS* | 1 |
| *_ID* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 1 |
| *_ID* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_ID* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BAG* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |

| | |
|---|---|
| *_BE(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 1 |
| *_DA(G)/(N)* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DA(G)/(N)* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |

| | |
|---|---|
| *_HV(D)/(I)/(N)/(P)* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 2 |
| *_HV(D)/(I)/(N)/(P)* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 2 |
| *_MD* *_D* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 1 |
| *_MD* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_MD* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_D* *books/bookes/boks/bokes/bockes*_NS* | 7 |
| *_VB(D)/(I)/(N)/(P)* *_SUCH* *books/bookes/boks/bokes/bockes*_NS* | 1 |
| *_VB(D)/(I)/(N)/(P)* *_WD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_N$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NPR$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NPRS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_ONE$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_OTHER$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_OTHERS$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_PRO$* *books/bookes/boks/bokes/bockes*_NS* | 10 |
| *_VB(D)/(I)/(N)/(P)* *_WPRO$* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NUM* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *books/bookes/boks/bokes/bockes*_NS* | 2 |
| **TOTAL** | **28** |

**'Books' used bare in sentence-initial position or in post-verb position (PPCEME)**

| **SEARCH TERM** | **HITS** |
|---|---|
| *_ID* *books/bookes/boks/bokes/bockes*_NS* | 1 |
| *_BAG* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *books/bookes/boks/bokes/bockes*_NS* | 0 |

| | |
|---|---|
| *_DA(G)/(N)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HA(G)/(N)* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *books/bookes/boks/bokes/bockes*_NS* | 2 |
| *_MD* *books/bookes/boks/bokes/bockes*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *books/bookes/boks/bokes/bockes*_NS* | 12 |
| **TOTAL** | **15** |

**'Horses' with overt D in sentence-initial position or in post-verb position (PPCEME)**

| **SEARCH TERM** | **HITS** |
|---|---|
| *_ID* *_D* *horses/horse/horsses*_NS* | 2 |
| *_ID* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_PRO$* *horses/horse/horsses*_NS* | 1 |
| *_ID* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_ID* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 1 |
| *_BAG* *_D* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_PRO$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_BAG* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_D* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_ONE$* *horses/horse/horsses*_NS* | 0 |

| | |
|---|---|
| *_BE(D)/(I)/(N)/(P)* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_PRO$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_D* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_PRO$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_DA(G)/(N)* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_D* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_PRO$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_D* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_PRO$* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_WPRO$* *horses/horse/horsses*_NS* | 0 |

| | |
|---|---|
| *_HA(G)/(N)* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 1 |
| *_HV(D)/(I)/(N)/(P)* *_D* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_PRO$* *horses/horse/horsses*_NS* | 2 |
| *_HV(D)/(I)/(N)/(P)* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *_NUM* *horses/horse/horsses*_NS* | 1 |
| *_HV(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_D* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_PRO$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_MD* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_D* *horses/horse/horsses*_NS* | 2 |
| *_VB(D)/(I)/(N)/(P)* *_SUCH* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_WD* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_N$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NPR$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NPRS$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NS$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_ONE$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_OTHER$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_OTHERS$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_PRO$* *horses/horse/horsses*_NS* | 20 |
| *_VB(D)/(I)/(N)/(P)* *_WPRO$* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_NUM* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *_Q(R)/(S)/($)* *horses/horse/horsses*_NS* | 1 |
| **TOTAL** | **31** |

**'Horses' used bare in sentence-initial position or in post-verb position (PPCEME)**

| SEARCH TERM | HITS |
|---|---|
| *_ID* *horses/horse/horsses*_NS* | 1 |
| *_BAG* *horses/horse/horsses*_NS* | 0 |
| *_BE(D)/(I)/(N)/(P)* *horses/horse/horsses*_NS* | 1 |
| *_DA(G)/(N)* *horses/horse/horsses*_NS* | 0 |
| *_DO(D)/(I)/(N)/(P)* *horses/horse/horsses*_NS* | 0 |
| *_HA(G)/(N)* *horses/horse/horsses*_NS* | 0 |
| *_HV(D)/(I)/(N)/(P)* *horses/horse/horsses*_NS* | 1 |
| *_MD* *horses/horse/horsses*_NS* | 0 |
| *_VB(D)/(I)/(N)/(P)* *horses/horse/horsses*_NS* | 8 |
| **TOTAL** | **11** |

**'Books' with overt D in sentence-initial position or in post-verb position (COHA)**

| SEARCH TERM | HITS |
|---|---|
| ! [at*] *books*.[nn2*] | 4 |
| ! [d*] *books*.[nn2*] | 5 |
| ! [appge*] *books*.[nn2*] | 7 |
| ! [ge*] *books*.[nn2*] | 0 |
| ! [m*] *books*.[nn2*] | 0 |
| . [at*] *books*.[nn2*] | 449 |
| . [d*] *books*.[nn2*] | 268 |
| . [appge*] *books*.[nn2*] | 249 |
| . [ge*] *books*.[nn2*] | 2 |
| . [m*] *books*.[nn2*] | 35 |
| ? [at*] *books*.[nn2*] | 12 |
| ? [d*] *books*.[nn2*] | 17 |
| ? [appge*] *books*.[nn2*] | 4 |
| ? [ge*] *books*.[nn2*] | 1 |
| ? [m*] *books*.[nn2*] | 2 |
| [vb*] [at*] *books*.[nn2*] | 130 |
| [vb*] [d*] *books*.[nn2*] | 117 |
| [vb*] [appge*] *books*.[nn2*] | 39 |
| [vb*] [ge*] *books*.[nn2*] | 0 |
| [vb*] [m*] *books*.[nn2*] | 29 |
| [vd*] [at*] *books*.[nn2*] | 22 |
| [vd*] [d*] *books*.[nn2*] | 7 |
| [vd*] [appge*] *books*.[nn2*] | 10 |
| [vd*] [ge*] *books*.[nn2*] | 0 |
| [vd*] [m*] *books*.[nn2*] | 3 |
| [vh*] [at*] *books*.[nn2*] | 79 |
| [vh*] [d*] *books*.[nn2*] | 84 |
| [vh*] [appge*] *books*.[nn2*] | 71 |
| [vh*] [ge*] *books*.[nn2*] | 0 |
| [vh*] [m*] *books*.[nn2*] | 26 |

| | |
|---|---|
| [vm*] [at*] *books*.[nn2*] | 2 |
| [vm*] [d*] *books*.[nn2*] | 3 |
| [vm*] [appge*] *books*.[nn2*] | 2 |
| [vm*] [ge*] *books*.[nn2*] | 0 |
| [vm*] [m*] *books*.[nn2*] | 1 |
| [vv0*] [at*] *books*.[nn2*] | 220 |
| [vv0*] [d*] *books*.[nn2*] | 120 |
| [vv0*] [appge*] *books*.[nn2*] | 188 |
| [vv0*] [ge*] *books*.[nn2*] | 0 |
| [vv0*] [m*] *books*.[nn2*] | 10 |
| [vvd*] [at*] *books*.[nn2*] | 343 |
| [vvd*] [d*] *books*.[nn2*] | 173 |
| [vvd*] [appge*] *books*.[nn2*] | 408 |
| [vvd*] [ge*] *books*.[nn2*] | 0 |
| [vvd*] [m*] *books*.[nn2*] | 109 |
| [vvi*] [at*] *books*.[nn2*] | 485 |
| [vvi*] [d*] *books*.[nn2*] | 251 |
| [vvi*] [appge*] *books*.[nn2*] | 399 |
| [vvi*] [ge*] *books*.[nn2*] | 1 |
| [vvi*] [m*] *books*.[nn2*] | 38 |
| [vvz*] [at*] *books*.[nn2*] | 90 |
| [vvz*] [d*] *books*.[nn2*] | 41 |
| [vvz*] [appge*] *books*.[nn2*] | 72 |
| [vvz*] [ge*] *books*.[nn2*] | 0 |
| [vvz*] [m*] *books*.[nn2*] | 19 |
| **TOTAL** | **3,592** |

**‘Books’ used bare in sentence-initial position or in post-verb position (COHA)**

| **SEARCH TERM** | **HITS** |
|---|---|
| ! *books*.[nn2*] | 16 |
| . *books*.[nn2*] | 754 |
| ? *books*.[nn2*] | 25 |
| [vb*] *books*.[nn2*] | 329 |
| [vd*] *books*.[nn2*] | 12 |
| [vh*] *books*.[nn2*] | 137 |
| [vm*] *books*.[nn2*] | 4 |
| [vv0*] *books*.[nn2*] | 706 |
| [vvd*] *books*.[nn2*] | 459 |
| [vvi*] *books*.[nn2*] | 848 |
| [vvz*] *books*.[nn2*] | 251 |
| **TOTAL** | **3,541** |

**‘Horses’ with overt D in sentence-initial position or in post-verb position (COHA)**

| **SEARCH TERM** | **HITS** |
|---|---|
| ! [at*] *horses*.[nn2*] | 11 |

| | |
|---|---|
| ! [d*] *horses*.[nn2*] | 1 |
| ! [appge*] *horses*.[nn2*] | 9 |
| ! [ge*] *horses*.[nn2*] | 0 |
| ! [m*] *horses*.[nn2*] | 4 |
| . [at*] *horses*.[nn2*] | 869 |
| . [d*] *horses*.[nn2*] | 84 |
| . [appge*] *horses*.[nn2*] | 206 |
| . [ge*] *horses*.[nn2*] | 0 |
| . [m*] *horses*.[nn2*] | 71 |
| ? [at*] *horses*.[nn2*] | 10 |
| ? [d*] *horses*.[nn2*] | 0 |
| ? [appge*] *horses*.[nn2*] | 2 |
| ? [ge*] *horses*.[nn2*] | 0 |
| ? [m*] *horses*.[nn2*] | 5 |
| [vb*] [at*] *horses*.[nn2*] | 92 |
| [vb*] [d*] *horses*.[nn2*] | 28 |
| [vb*] [appge*] *horses*.[nn2*] | 24 |
| [vb*] [ge*] *horses*.[nn2*] | 1 |
| [vb*] [m*] *horses*.[nn2*] | 41 |
| [vd*] [at*] *horses*.[nn2*] | 4 |
| [vd*] [d*] *horses*.[nn2*] | 0 |
| [vd*] [appge*] *horses*.[nn2*] | 4 |
| [vd*] [ge*] *horses*.[nn2*] | 0 |
| [vd*] [m*] *horses*.[nn2*] | 1 |
| [vh*] [at*] *horses*.[nn2*] | 87 |
| [vh*] [d*] *horses*.[nn2*] | 29 |
| [vh*] [appge*] *horses*.[nn2*] | 48 |
| [vh*] [ge*] *horses*.[nn2*] | 0 |
| [vh*] [m*] *horses*.[nn2*] | 56 |
| [vm*] [at*] *horses*.[nn2*] | 6 |
| [vm*] [d*] *horses*.[nn2*] | 1 |
| [vm*] [appge*] *horses*.[nn2*] | 3 |
| [vm*] [ge*] *horses*.[nn2*] | 0 |
| [vm*] [m*] *horses*.[nn2*] | 0 |
| [vv0*] [at*] *horses*.[nn2*] | 264 |
| [vv0*] [d*] *horses*.[nn2*] | 33 |
| [vv0*] [appge*] *horses*.[nn2*] | 322 |
| [vv0*] [ge*] *horses*.[nn2*] | 0 |
| [vv0*] [m*] *horses*.[nn2*] | 18 |
| [vvd*] [at*] *horses*.[nn2*] | 608 |
| [vvd*] [d*] *horses*.[nn2*] | 42 |
| [vvd*] [appge*] *horses*.[nn2*] | 917 |
| [vvd*] [ge*] *horses*.[nn2*] | 0 |
| [vvd*] [m*] *horses*.[nn2*] | 84 |

| | |
|---|---|
| [vvi*] [at*] *horses*.[nn2*] | 674 |
| [vvi*] [d*] *horses*.[nn2*] | 70 |
| [vvi*] [appge*] *horses*.[nn2*] | 482 |
| [vvi*] [ge*] *horses*.[nn2*] | 1 |
| [vvi*] [m*] *horses*.[nn2*] | 41 |
| [vvz*] [at*] *horses*.[nn2*] | 56 |
| [vvz*] [d*] *horses*.[nn2*] | 9 |
| [vvz*] [appge*] *horses*.[nn2*] | 28 |
| [vvz*] [ge*] *horses*.[nn2*] | 0 |
| [vvz*] [m*] *horses*.[nn2*] | 8 |
| **TOTAL** | **5,354** |

**'Horses' used bare in sentence-initial position or in post-verb position (COHA)**

| **SEARCH TERM** | **HITS** |
|---|---|
| ! *horses*.[nn2*] | 8 |
| . *horses*.[nn2*] | 462 |
| ? *horses*.[nn2*] | 8 |
| [vb*] *horses*.[nn2*] | 149 |
| [vd*] *horses*.[nn2*] | 8 |
| [vh*] *horses*.[nn2*] | 131 |
| [vm*] *horses*.[nn2*] | 1 |
| [vv0*] *horses*.[nn2*] | 270 |
| [vvd*] *horses*.[nn2*] | 379 |
| [vvi*] *horses*.[nn2*] | 443 |
| [vvz*] *horses*.[nn2*] | 79 |
| **TOTAL** | **1,938** |

# APPENDIX 8 NPs with overt D and BPs in existential sentences

**BPs (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| *_EX* *_BE* *_NS* | 76 | 18 | 37 | 21 |

**NPs headed by *some* (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| *_EX* *_BE* some_Q* *_NS* | 19 | 1 | 6 | 12 |
| *_EX* *_BE* som_Q* *_NS* | 1 | 1 | 0 | 0 |
| *_EX* *_BE* somme_Q* *_NS* | 1 | 1 | 0 | 0 |
| **TOTAL** | **21** | **3** | **6** | **12** |

**BPs (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [ex*] [vb*] [nn2] | 31,002 | 63 | 349 | 789 | 937 | 1,090 | 1,206 | 1,381 | 1,537 | 1,586 | 1,795 | 1,837 | 1,932 | 1,835 | 1,930 | 1,991 | 1,952 | 2,169 | 2,385 | 2,152 | 2,050 |

**NPs headed by *some* (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| [ex*] [vb*] some [nn2] | 3,049 | 0 | 58 | 126 | 141 | 163 | 145 | 138 | 172 | 150 | 202 | 149 | 156 | 156 | 172 | 183 | 199 | 188 | 174 | 215 | 162 |

# APPENDIX 9 NPs with overt D and BPs in object position

**'Eat' + *some* + NN2 (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| eat_VB* some_Q* *_NS* | 1 | 0 | 1 | 0 |

**'Eat' + BP (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| eat_VB* *_NS* | 1 | 0 | 1 | 0 |
| eate_VB* *_NS* | 3 | 1 | 2 | 0 |
| ete_VB* *_NS* | 2 | 2 | 0 | 0 |
| **TOTAL** | **6** | **3** | **3** | **0** |

**'Hear' + *some* + NN2 (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| heard_VB* some_Q* *_NS* | 1 | 0 | 0 | 1 |
| hearde_VB* some_Q* *_NS* | 1 | 0 | 1 | 0 |
| heare_VB* some_Q* *_NS* | 1 | 0 | 1 | 0 |
| here_VB* some_Q* *_NS* | 1 | 0 | 1 | 0 |
| **TOTAL** | **4** | **0** | **3** | **1** |

**'Hear' + BP (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| heard_VB* *_NS* | 2 | 0 | 0 | 2 |
| heare_VB* *_NS* | 2 | 1 | 1 | 0 |
| hears_VB* *_NS* | 1 | 0 | 0 | 1 |
| here_VB* *_NS* | 1 | 1 | 0 | 0 |
| **TOTAL** | **6** | **2** | **1** | **3** |

**'Send' + *some* + NN2 (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| send_VB* some_Q* *_NS* | 1 | 0 | 0 | 1 |
| sende_VB* some_Q* *_NS* | 1 | 0 | 0 | 1 |
| **TOTAL** | **2** | **0** | **0** | **2** |

**'Send' + BP (PPCEME)**

| SEARCH TERM | HITS | E1 | E2 | E3 |
|---|---|---|---|---|
| send_VB* *_NS* | 11 | 4 | 6 | 1 |
| sende_VB* *_NS* | 3 | 3 | 0 | 0 |
| sendeth_VB* *_NS* | 1 | 0 | 1 | 0 |
| sends | 3 | 0 | 0 | 3 |
| **TOTAL** | **18** | **7** | **7** | **4** |

**'Eat' + *some* + NN2 (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EAT. [vv0*] some.[d*] [nn2] | 9 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 2 | 2 |
| EAT. [vvd*] some.[d*] [nn2] | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 2 | 0 | 1 | 0 | 0 | 0 | 4 | 1 | 1 |
| EAT. [vvi*] some.[d*] [nn2] | 16 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 5 |
| EAT. [vvz*] some.[d*] [nn2] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **TOTAL** | **37** | **0** | **0** | **2** | **0** | **1** | **0** | **1** | **1** | **1** | **4** | **3** | **3** | **0** | **2** | **0** | **0** | **1** | **7** | **3** | **8** |

**'Eat' + BP (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| EAT. [vv0*] [nn2] | 393 | 0 | 6 | 13 | 11 | 3 | 11 | 8 | 12 | 10 | 13 | 20 | 20 | 21 | 18 | 24 | 28 | 25 | 42 | 56 | 52 |
| EAT. [vvd*] [nn2] | 374 | 0 | 1 | 6 | 3 | 4 | 15 | 7 | 6 | 10 | 17 | 16 | 18 | 27 | 23 | 33 | 23 | 34 | 35 | 49 | 47 |
| EAT. [vvi*] [nn2] | 511 | 0 | 2 | 13 | 7 | 13 | 14 | 26 | 16 | 14 | 27 | 21 | 30 | 42 | 36 | 23 | 25 | 32 | 34 | 62 | 74 |
| EAT. [vvz*] [nn2] | 87 | 0 | 0 | 0 | 0 | 3 | 3 | 1 | 2 | 1 | 3 | 4 | 6 | 4 | 7 | 5 | 7 | 4 | 7 | 16 | 14 |
| **TOTAL** | **1,365** | **0** | **9** | **32** | **21** | **23** | **43** | **42** | **36** | **35** | **60** | **61** | **74** | **94** | **84** | **85** | **83** | **95** | **118** | **183** | **187** |

**'Hear' + *some* + NN2 (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HEAR. [vv0*] some.[d*] [nn2] | 17 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 3 | 0 | 0 | 1 | 1 | 1 | 3 | 4 | 0 |
| HEAR. [vvd*] some.[d*] [nn2] | 67 | 0 | 1 | 0 | 1 | 1 | 3 | 7 | 5 | 2 | 6 | 3 | 1 | 4 | 4 | 2 | 5 | 9 | 5 | 2 | 6 |
| HEAR. [vvi*] some.[d*] [nn2] | 33 | 0 | 0 | 4 | 1 | 5 | 2 | 2 | 3 | 1 | 2 | 2 | 1 | 0 | 1 | 1 | 2 | 0 | 2 | 1 | 3 |
| HEAR. [vvz*] some.[d*] [nn2] | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| **TOTAL** | **119** | **0** | **2** | **5** | **2** | **6** | **5** | **11** | **9** | **3** | **8** | **5** | **5** | **4** | **5** | **4** | **8** | **10** | **10** | **8** | **9** |

**'Hear' + BP (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HEAR. [vv0*] [nn2] | 651 | 6 | 12 | 18 | 25 | 15 | 22 | 15 | 30 | 20 | 36 | 23 | 22 | 27 | 49 | 37 | 41 | 40 | 59 | 71 | 83 |
| HEAR. [vvd*] [nn2] | 2,001 | 1 | 18 | 36 | 34 | 69 | 57 | 52 | 53 | 76 | 90 | 113 | 146 | 146 | 148 | 148 | 136 | 131 | 158 | 210 | 179 |
| HEAR. [vvi*] [nn2] | 1,270 | 4 | 19 | 24 | 31 | 27 | 36 | 31 | 40 | 41 | 48 | 55 | 66 | 76 | 101 | 87 | 103 | 79 | 118 | 134 | 150 |
| HEAR. [vvz*] [nn2] | 220 | 0 | 2 | 1 | 4 | 4 | 6 | 5 | 7 | 6 | 10 | 10 | 15 | 14 | 12 | 21 | 15 | 16 | 14 | 30 | 28 |
| **TOTAL** | **4,142** | **11** | **51** | **79** | **94** | **115** | **121** | **103** | **130** | **143** | **184** | **201** | **249** | **263** | **310** | **293** | **295** | **266** | **349** | **445** | **440** |

**‘Send’ + *some* + NN2 (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SEND. [vv0*] some.[d*] [nn2] | 17 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 1 | 0 | 1 | 0 | 1 | 2 | 2 | 2 | 2 | 0 | 1 | 1 | 1 |
| SEND. [vvd*] some.[d*] [nn2] | 38 | 0 | 1 | 0 | 0 | 0 | 2 | 1 | 5 | 3 | 4 | 1 | 3 | 3 | 1 | 2 | 2 | 2 | 2 | 3 | 3 |
| SEND. [vvi*] some.[d*] [nn2] | 71 | 0 | 2 | 4 | 0 | 1 | 4 | 5 | 2 | 2 | 1 | 4 | 1 | 5 | 9 | 5 | 5 | 3 | 9 | 6 | 3 |
| SEND. [vvz*] some.[d*] [nn2] | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **TOTAL** | **127** | **0** | **3** | **4** | **0** | **1** | **8** | **7** | **8** | **5** | **6** | **5** | **5** | **10** | **12** | **9** | **9** | **5** | **12** | **10** | **8** |

**‘Send’ + BP (COHA)**

| SEARCH TERM | HITS | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SEND. [vv0*] [nn2] | 448 | 0 | 2 | 5 | 14 | 9 | 13 | 9 | 12 | 20 | 14 | 16 | 20 | 26 | 40 | 43 | 23 | 24 | 31 | 49 | 78 |
| SEND. [vvd*] [nn2] | 1,495 | 0 | 16 | 22 | 34 | 32 | 36 | 62 | 45 | 73 | 99 | 61 | 103 | 111 | 106 | 113 | 119 | 101 | 129 | 112 | 121 |
| SEND. [vvi*] [nn2] | 1,382 | 0 | 17 | 27 | 24 | 35 | 49 | 52 | 60 | 80 | 81 | 70 | 99 | 78 | 113 | 104 | 84 | 91 | 93 | 108 | 117 |
| SEND. [vvz*] [nn2] | 257 | 0 | 2 | 7 | 9 | 6 | 8 | 10 | 9 | 8 | 9 | 9 | 14 | 23 | 9 | 25 | 8 | 20 | 17 | 38 | 26 |
| **TOTAL** | **3,582** | **0** | **37** | **61** | **81** | **82** | **106** | **133** | **126** | **181** | **203** | **156** | **236** | **238** | **268** | **285** | **234** | **236** | **270** | **307** | **342** |

# APPENDIX 10 Abstract in English and German

**Abstract**

This thesis aims to investigate the diachronic development of English bare plurals and of plural indefinites with overt reference marking within the Modern English period. To shed light on this issue, quantitative and qualitative analyses of plural NPs in the *Penn-Helsinki Parsed Corpus of Early Modern English* and the *Corpus of Historical American English* were conducted. In particular, the frequencies of bare plurals and of plural nouns headed by *some* (in sentence-initial and post-verb position) were compared. The results reveal that within the investigated time period the occurrences of bare plurals with indefinite reference has increased, while the opposite development is observed for plural NPs with overt reference marking. It can thus be concluded that the trend towards obligatory reference marking in the history of English does not seem to apply to plural indefinites.

**Zusammenfassung**

Ziel dieser Diplomarbeit ist es, die diachrone Entwicklung von artikellosen Pluralnomen und von indefiniten Pluralnomen mit Artikel innerhalb der letzten 500 Jahren zu untersuchen. Um diese Thematik zu beleuchten, wurden quantitative und qualitative Analysen von pluralen Nominalphrasen im *Penn-Helsinki Parsed Corpus of Early Modern English* und dem *Corpus of Historical American English* durchgeführt. Im Besonderen wurden die Häufigkeiten von artikellosen Pluralnomen (am Satzanfang und nach Verben) mit Pluralnomen, welchen der Quantifikator *some* vorangestellt ist, verglichen. Die Ergebnisse zeigen, dass innerhalb der untersuchten Zeitspanne die Zahl der artikellosen indefiniten Pluralnomen angestiegen ist. Demgegenüber ist bei den pluralen Nominalphrasen mit Artikel eine gegenteilige Entwicklung erkennbar. Hieraus lässt sich schließen, dass ein Trend nach obligatorischer Artikelverwendung in der Geschichte des Englischen bei pluralen Nominalphrasen nicht festgestellt werden kann.