

Emottaget:

Drink tea, coffee or chocolate, or some other cooling infusion!

A Corpus Study of the Words *Tea*, *Coffee*, and *Chocolate*
1640-1920

Tove Andersson

Pro gradu-avhandling i Engelska språket och
litteraturen

Handledare: Brita Wårvik

Fakulteten för humaniora, psykologi och teologi

Åbo Akademi

(2019)

ÅBO AKADEMI – FAKULTETEN FÖR HUMANIORA, PSYKOLOGI OCH TEOLOGI

Abstrakt för avhandling pro gradu

Ämne: Engelska	
Författare: Tove Andersson	
Arbetets titel: "Drink tea, coffee or chocolate, or some other cooling infusion!" A Corpus Study of the Words <i>Tea</i> , <i>Coffee</i> , and <i>Chocolate</i> :1640-1920	
Handledare: Brita Wårvik	Handledare:
<p>Abstrakt:</p> <p>This is a historical corpus study of the words <i>tea</i>, <i>coffee</i>, and <i>chocolate</i>, examining the occurrence and semantic changes between 1640 and 1920. The uses of these words are also investigated in a sociohistorical and cultural context: in what ways do the uses of these words by men and women differ, and what does the use of these words in texts say about the English society between 1640 and 1920?</p> <p>The concordance software WordSmith Tools has been used to find occurrences of the words including spelling variations and compounds in a historical corpus of 12.2 million words. The findings were sorted into categories of meaning varying from 'produce' to 'social events', verbs and colours, among others.</p> <p>The analysis has two parts: first, meaning and prevalence, and secondly, sociolinguistic aspects, genre, and the impact of the empire on life in Britain. Over all, it was found that <i>tea</i> was in markedly more frequent use than <i>coffee</i> and <i>chocolate</i> even though it took longer for tea to become a staple commodity in Britain. The growth in occurrences of the words fits well with the cultural and social history of the goods they refer to. In comparison with <i>coffee</i> and <i>chocolate</i>, <i>tea</i> is more often used to denote something other than just the actual drink, such as a mealtime or a social event or even as a verb. However, even though there is a growing number of different compounds based on the words <i>tea</i>, <i>coffee</i>, and <i>chocolate</i> over the period of the study, it is only the most basic ones that remain popular. Likewise, even though the meanings of the words <i>tea</i>, <i>coffee</i>, and <i>chocolate</i> widen, it is the most typical semantic meanings that are most frequent even in the later periods.</p> <p>While the corpora were not created for sociolinguistic studies, some conclusions can still be drawn. Most interestingly, it was found that men tended to use the word <i>coffee</i> more than women, and more often in connection to places outside the home. Correspondingly, <i>tea</i> was more frequently used by and in relation to women in a domestic setting than <i>coffee</i> and <i>chocolate</i>.</p>	
Nyckelord: lexicology, semantic change, corpus linguistics, historical linguistics, sociolinguistics, gender, genre	
Datum: 1.5.2019	Sidantal: 110
Abstraktet godkänt som mognadsprov:	

Table of Contents

List of tables.....	0
List of tables in Swedish for the Swedish summary.....	0
1. Introduction.....	1
2. Background	2
2.1 Lexicology and semantics.....	2
2.1.1 Lexical borrowing.....	2
2.1.2. Semantic changes.....	3
2.1.3. Compounds	4
2.2 Historical sociolinguistics	5
2.2.1 Regarding materials	6
2.2.2 Genre and gender	7
2.2.3 Attitudes.....	9
2.3 Cultural and social history of tea, coffee, and chocolate	11
2.3.1 Coffee and coffee-houses.....	11
2.3.2 Tea.....	13
2.3.3 Chocolate	15
2.4 The words <i>tea</i> , <i>coffee</i> , and <i>chocolate</i>	15
3. Material.....	16
3.1. Corpora	16
3.2 Adjustments of text periodisation and corpus size.....	19
4. Methods.....	20
4.1 The <i>OED</i>	22
4.2 Finding Compounds.....	24
4.3 Spelling	25
4.3.1 Spelling of compounds.....	25
4.3.2 Spelling of <i>tea</i>	25
4.3.3 Spelling of <i>coffee</i>	26
4.3.4 Spelling of <i>chocolate</i>	26
4.4 Calculation of frequencies	27

5. Analysis part 1: Meanings and prevalence	28
5.1. Semantic classification.....	28
5.1.1 Semantic classification of <i>tea</i>	29
5.1.2 Semantic classification of <i>coffee</i>	32
5.1.3 Semantic classification of <i>chocolate</i>	33
5.2. Frequencies of <i>tea</i> , <i>coffee</i> , and <i>chocolate</i> in the corpus.....	35
5.3. Meanings of <i>tea</i> , <i>coffee</i> and <i>chocolate</i>	38
5.3.1 <i>Tea</i>	38
5.3.2 <i>Coffee</i>	42
5.3.3 <i>Chocolate</i>	43
5.4. Discussion of frequencies and meanings	44
6. Analysis part 2: Sociolinguistic and cultural aspects.....	45
6.1 Tea	49
6.1.1 Periods 1 (1650-1709) and 2 (1710-1779).....	50
6.1.2 Periods 3 (1780-1849) and 4 (1850-b1920).....	53
6.2 Coffee.....	57
6.3 Chocolate	61
7. Discussion and conclusion	63
Svensk sammanfattning.....	66
1. Inledning och bakgrund	66
2. Material och metod	67
3. Analys del 1: betydelser och förekomst.....	68
4. Analys del 2: sociolingvistiska och kulturella aspekter	69
5. Sammanfattning	71
Appendices.....	72
Appendix 1.1: Distribution of <i>tea</i>	72
Appendix 1.2: Distribution of <i>coffee</i>	74
Appendix 1.3: Distribution of <i>chocolate</i>	76
Appendix 2.1: Numbers of occurrences of <i>tea</i>	78
Appendix 2.2: Numbers of occurrences of <i>coffee</i>	80
Appendix 2.3: Numbers of occurrences of <i>chocolate</i>	81

Appendix 3.1.....	82
Period 1: 1640-1709.....	82
Appendix 3.3.....	85
Period 3: 1780-1849.....	85
Appendix 3.4.....	87
Period 4: 1850-1920.....	87
Appendix 4.1 Distribution across genres in period 1	90
Appendix 4.2 Distribution across genres in period 2	91
Appendix 4.3 Distribution across genres in period 3	93
Appendix 4.4 Distribution across genres in period 4.....	95
Bilaga 5: Tabeller för den svenska sammanfattningen	97
Bibliography	101
Primary sources.....	101
Secondary sources.....	101

List of tables

Table 1: Corpora, diachronic	17
Table 2: Time Periods.....	18
Table 3: Frequency of the words tea, coffee, and chocolate, over time, balanced with size of corpus.....	35
Table 4: Frequency within time periods	36
Table 5: Frequency over time, comparing periods	37
Table 6: Genders in <i>CLMET</i>	46
Table 7: Genders in the whole corpus; authors who use either of the words.....	47
Table 8: Genders' shares of the occurrences of the words	47
Table 9: The words' share of use by gender	48
Table 10: Genders' use of genres	48
Table 11: Relative frequency before and after 1784.....	50
Table 12: Frequency of each of the words before and after 1784, and over time	50
Table 13: Gender representation in selected works, chronologically, periods 3 and 4, <i>tea</i>	54
Table 14: Gender representation, sorted	56
Table 15: Comparisons in periods 3 and 4, <i>tea</i>	56

List of tables in Swedish for the Swedish summary

Tabell 1: Korpusar, diakronisk översikt	97
Tabell 2: Tidsperioder med korpusar och antal ord	97
Tabell 3: Förekomst av orden <i>tea</i> , <i>coffee</i> och <i>chocolate</i> över tid	98
Tabell 4: Sammanställning av betydelser hos orden <i>tea</i> , <i>coffee</i> och <i>chocolate</i> samt kategorier av sammansatta ord.....	98
Tabell 5: Mäns och kvinnors andel av ordens förekomst i period 3	99
Tabell 6: Ordens andel av användningen hos män och kvinnor i period 3	99
Tabell 7: Representation av män och kvinnor i valda verk, i kronologisk ordning; perioderna 3 och 4, <i>tea</i>	100

1. Introduction

This thesis is a historical corpus study of the words *coffee*, *tea*, and *chocolate*. These words all relate to the same area of life, namely food and recreation, and were borrowed into the English language at roughly the same time - late 16th century to early 17th century. It is a continuation of my bachelor's thesis "Tea, Tea, Coffee, Tea, or Chocularitee – A Corpus Study of the Words *Tea*, *Coffee*, and *Chocolate*" (2014), in which I presented the overall change in the uses of the words *tea*, *coffee*, and *chocolate* during the period 1640-1920. The primary material is the same as here. I looked for the most frequent meanings of the words as well as which other words they combined with to make new compound words.

The aim in my bachelor's thesis was to find out how the meanings of the words *coffee*, *tea*, and *chocolate* had changed since they were borrowed into English, and how these changes related to socio-historical development. The purpose of this study is to look closer at how the uses of the words have developed and changed since they were borrowed, and more specifically, to see if there is anything to be said about differences of use of these words that are related to gender of the author and genre of the text. To that end, I will also take a closer look at the culture surrounding tea, coffee, and chocolate, as well as particularly in the case of tea, the role it played in the development of Britain's role in the world and Britishness and thus, locate the study more firmly in the historical context.

The substances coffee, tea, and chocolate can all appear in various forms: the word *coffee* can refer to the beans or the final drink; *tea*, likewise, can refer to the dried leaves or the drink prepared from them; and *chocolate* can refer to the powder made from the cocoa bean, the drink made from the powder, or finally, pieces of solid chocolate. However, meanings of words can also be widened and become more abstract. Thus, in the case of *tea*, the word is not used for the product only, but also for the meal including, or not including tea, or the time of day when this meal is often taken. Furthermore, verbs can be derived from nouns (and vice versa), and this has happened with the word *tea*, achieving the meaning 'to drink tea; to take the meal called tea, to have one's tea', or 'to supply or regale with tea; to entertain at tea; to give a tea to' (*OED*, s.v. *tea*, v.). The same applies to the word *coffee* (*OED*, s.v. *coffee*, v.).

The results of my earlier study suggest that *tea* is more often used by women, while *coffee* is more often used by men. How is this reflected in genres? If there are differences reflected in genres, what is the case with fictional characters? Do male or female writers more

often put the words *tea* or *coffee* in the mouths of men or women? Do women more often than men use the different meanings and compounds - i.e., do they use the words more creatively? For instance, the results of my BA-thesis suggested that the location is more important than the actual drink when it comes to *coffee* than is the case with *tea*. This is certainly related to genre and quite possibly also gender, even though to a certain extent the choice of genre also depends on the gender of the author.

While my BA-thesis could also be followed up by a study of a more recent corpus - Simon Horobin lists several more modern uses of the word *tea* in his blog post in the OED blog, such as *tea-head* and *tea-pad* referring to drug trade - I have chosen to delve deeper into the patterns of my original material.

2. Background

This chapter presents background and methodology from different fields that are all relevant to the present study, none of which is primary to the others. I will begin by looking at the linguistic features that are involved in semantic change: borrowing and changes, as well as different properties of compounds. After that, I will take up issues that are related to a larger context: sociolinguistics and specifically historical sociolinguistics and some of the challenges in the field, and questions concerning genre and gender as well as language attitudes. Finally, I will look at the historical and social culture surrounding the substances tea, coffee, and chocolate, as well as the backgrounds of the words *tea*, *coffee*, and *chocolate* in the English language.

2.1 Lexicology and semantics

The strictly linguistic matters that are relevant to this study all belong to lexicology and semantics: word borrowing, semantic change, and compound words.

2.1.1 Lexical borrowing

One of the main reasons behind lexical borrowings is the need, in the recipient language, for new terms for new phenomena in the world. Explorers and colonizers constantly met with new phenomena and often borrowed the words from the local languages. When new items and concepts from abroad were brought back to the homeland, foreign words often followed (Campbell 1998: 59). This is the case with *tea*, *coffee*, and *chocolate*.

There are other reasons for borrowing such as prestige and degradation, and all the intricacies of how and why foreign words replace native words, or start being used for one aspect of a thing whereas the use of the native word is restrained to another aspect are quite interesting. Examples of prestige could be *pork* and *veal* being borrowed from French, which was a more prestigious language in the centuries following the Norman conquest, while *pig* and *calf* remained in the language to denote the living animals. However, since this thesis deals with three words that came together with the three substances they refer to, the relevant explanation for lexical borrowing in this case is the need for new words for new things in life.

2.1.2. Semantic changes

A semantic change means a change in “the concepts associated with a word” (Campbell 1998: 255), that is, the same word starts to be used to refer to new things which in one way or another may be related to the original meaning of the word, but this kind of relationship between meanings does not necessarily have to exist. There are many descriptions and definitions of different kinds of semantic changes; here the focus will be on the ones most relevant to the material in this study. In general, the shifts in meaning seem to go from concrete to abstract, even though there are cases where the result is a more specific meaning.

Changes of scope, widening and narrowing of meanings, could be argued to be the major kind of semantic change, while all other definitions of semantic change are subtypes of them (Campbell 1998: 266). Widening means that the scope of a word’s meaning has a greater range of referents that it can refer to: the word “can be used in more contexts than were appropriate for it before the change” (ibid: 256). Campbell uses the word *salary* to illustrate this; the meaning of this particular word has gone from “a soldier’s allotment of salt”, via “a soldier’s wages in general”, to “wages in general” (ibid: 255-256). The opposite process, narrowing (e.g. *meat* originally meaning ‘food in general’ and *deer* meaning ‘animal’), is not relevant for the material in this study.

Other relevant types of semantic change are metaphors and metonymy. There are many ways of describing what these two concepts are, and I will supplement Campbell’s descriptions with those of Friedrich Ungerer and Hans-Jörg Schmid (1996) and David Lee (2001). The lines between these two categories can sometimes be blurred, Ungerer and Schmid (1996: 128) point out, but generally, a metaphor is mapping across different cognitive models, while metonymies are mapping within one model.

Metaphor means using a word for another phenomenon that in some way reminds of

the original meaning. Campbell lists, among others, the following metaphors for ‘drunk’: *blasted, shredded, wasted*, on one hand, and *pissed* and *soaked* on the other, where in the first category the metaphors refer to being somehow damaged, and the other to the involvement of liquids (1998: 258). Ungerer and Schmid (1996: 155) describe metaphors as based on notions of similarity or comparison between literal and figurative meanings of expressions. Lee (2001: 6) assists in making Ungerer and Schmid’s reasoning clear: a metaphor is “a device that involves conceptualising one domain of experience in terms of another.”

Metonymy involves the inclusion of meanings which are related to the original word in the non-linguistic world. Very relevant for this study, Campbell mentions the word *tea*, which has taken on the meaning of ‘evening meal’ (1998: 259), as well as the meaning ‘to take tea’. Ungerer and Schmid (1996: 115) describe metonymy as the nearness between the literal meaning and the figurative counterpart. They also list several contiguity-relations in metonymies (ibid: 116), of which at least “container for content” could easily be expected to be found in the material of this study.

Polysemy is one result of and a prerequisite for semantic change. Words do not always go completely from only having one specific meaning to only having another specific meaning. Instead, a single word can acquire several meanings, which Stockwell and Minkova describe as a “consequence of the change from concrete to abstract meaning - i.e., figurative use of language” (2001: 147). At the same time, polysemy is a prerequisite for semantic change: without words acquiring any additional meanings, the meanings could not change completely. Additionally, the different meanings of the same word can be close to each other or distant: Carter’s example (1987:11) is *line*, which can refer to drawing, fishing, or railways.

Finally, Campbell points out that words which are related semantically tend to “undergo parallel semantic shifts” (1998: 270). With this in mind, it would seem likely that the words *tea*, *coffee*, and *chocolate* have undergone similar shifts in similar ways.

2.1.3. Compounds

Compounds are the combinations of two or more free morphemes put together to form a third (Bauer 1983:11). Compounds can belong to most parts of speech, consist of a varying number of elements, and of components of varying parts of speech, and, finally, be spelt in different ways. Adams (1973: 59), while trying to differentiate between compounds and noun phrases, points out that the spelling of compounds is rather arbitrary, and they can be written

as a single word, as two words, or with a hyphen, “sometimes regardless of the degree of unity that may be felt between the elements”. Likewise, Bauer (1998: 19) makes no difference in the treatment of compounds spelt in different ways. “Since *girl friend*, *girl-friend*, *girlfriend* all behave identically in everything except spelling, it seems more sensible to include them all under the same heading”.

“Knowledge of the world”, as Adams (1973: 63) puts it, is often useful to interpret compounds even though most classifications of compounds are based on grammatical relations between the combined parts. That is, it does not necessarily suffice to base the interpretation of every compound on its grammatical characteristics and thus, it is hard to classify compounds in a definitive and tidy manner. ‘Knowledge of the world’ is particularly important when categorising compounds into semantic categories.

Grammatically speaking, most compounds consist of two nouns (Bauer 1983: 202). They can, however, be almost any combination of nouns, verbs, and adjectives. Also, the largest groups of compounds are nouns in themselves, even though there are also a great deal of adjectives and some verbs. This can be seen from simply looking at for instance Adams’s table of contents (1973).

Where semantics is concerned, most compounds over all are what Bauer (1983: 30, 203ff) calls endocentric, that is, the first element modifies the second element, which is the grammatical head, and the whole compound is a hyponym of the grammatical head - for instance *beehive*, where *bee* specifies what kind of *hive* we are dealing with. Exocentric compounds on the other hand are metaphorical or synecdochic; the compound is not a hyponym of the grammatical head. Bauer’s example is *highbrow*. Appositional compounds are hyponyms of both the grammatical head and the modifier; *maidservant* (Bauer’s example) is both a kind of maid and a kind of servant. The final group in this semantic division of compounds is the copulative compounds: “the elements name separate entities which combine to form the entity denoted by the compound” (ibid: 31). The grammatical head is unclear in these cases - an example would be *Alsace-Lorraine*.

2.2 Historical sociolinguistics

Sociolinguistic studies about present day topics and present day language have the advantage of having constant and/or continuous access to the informants and speech communities they investigate. This is not the case with historical sociolinguistics where an extensive length of time can separate the object of study from those who study it. “Historical linguists cannot

safely rely on their modern intuitions about the range of linguistic variation available and acceptable to the people they investigate” (Nevalainen & Raumolin-Brunberg 2003: 3). The social evaluations that we hold true today do not necessarily make sense with historical data, and the longer the time that has passed, the less can we be sure that we can use our intuition about historical data. Present-day sociolinguistic studies may be carried out in the native dialect or sociolect of the researcher, with an advantage of deep inside insights, but the historical sociolinguist cannot claim to have such a position and such insights.

As descriptive linguistics became the most prominent branch of linguistics in the 20th century, as well as sociolinguistics growing to become its own field within linguistics there was an increasingly strong belief that speech is by far the most important form of language and that written language is only a surrogate or reflection of spoken language (Romaine 1982: 14). However, Romaine argues that the nature of writing is independent and that writing works as a medium of language use, and because of this, the function of writing is not “merely to record the spoken language; writing has an independent existence” (ibid: 15). At the time that Romaine wrote her book, sociolinguistics was still a very new field, and she was one of the first to consider it in a historical perspective: because of this, she needed to justify the combination of sociolinguistic methods with historical (written) material.

Romaine also discusses the relation between literary and non-literary written language, and based on Halliday, she concludes that “there exist linguistic methods of analysis which are already valid; if written language is an instance of language, then the same techniques apply to all instances of language. In other words, linguistic theory should not need to be *extended* to cover all relevant cases; it should just apply to these cases as it has done elsewhere” (Romaine 1982: 17). I think this is important to keep in mind seeing as my material in this study will consist of texts from several different genres, including literary works.

2.2.1 Regarding materials

One general problem with historical sociolinguistics is the material. The material that we have access to is limited even without considering the absence of spoken material, and a lot of the material has been published at some point. Even in the cases where we have access to the most original copy of a text, it may well be that it is a published version and therefore, it has been edited, often by someone other than the author. The later the version that we use, the more has it been edited. For this reason, searching for e.g. spelling variation in my material is

as good as futile, because only a very small part of my whole corpus has not had its spelling edited.

In general, it is also a problem that the written material was produced by such a small share of the people due to the fact that a lot of people were illiterate, even though in the period that is covered in my study, the situation is notably better than a few hundred years earlier. This is less of a problem in my study since my major questions are concerned with semantics, as that is one of the many areas of language use that are more difficult to standardize than spelling.

Despite these problems, judged as serious by sociolinguists like Labov, according to Romaine (1982: 122), Romaine herself does not agree that historical material is inherently bad for a sociolinguistic analysis. Despite being fragmentary and incomplete, historical material can be valid in its own right and can be useful for both synchronic and diachronic studies. We just need to remember the limitations that our material sets for us as researches, and the fact that we cannot expect to come to the same kind of conclusions or build similar theories from a historical sociolinguistic study as one can reach in a present-day, speech-based sociolinguistic study.

Romaine (1982: 107) discusses the problem of sampling, ‘random’ sampling and sample size. She points out that the idea of using random samples comes from a stricter science of sociology but even though sociolinguistics pioneers such as Labov and Trudgill claim to use random samples they do not, and Romaine claims that such a thing as a purely random sample in a statistical sense is very hard to achieve in sociolinguistics. At the same time, she argues that while we should pay attention to methods used in the selection of material (informants, or, as in her case as well as mine, texts), there is an inherent quality of randomness to any set of language data. In my case, the randomness lies in the selection of historical corpora available to me at my university, which will be presented in more detail in chapter 3.

2.2.2 Genre and gender

Genres are, of course, an important parameter in this study: speech-like genres like drama and letters are more likely to foster innovations than typical written genres, such as legal documents (Romaine 1982, as cited in Nevalainen and Raumolin-Brunberg 2003:2). My data consist of a very varied material, including letters and two plays, as well as newsletters, essays and polemical texts. A large part of my material consists of novels, which would be somewhere in

the middle between speech-like and typical written texts. Novels in the nineteenth century are likely to be stylistically much more complex and formal than some more modern novels as well as less experimental, but at the same time they may include a lot of characters' direct speech, which may be rendered closer to the spoken language in the area and at the time of the novels' settings. Yet, as Romaine points out (1982: 125), the norms for reporting speech may have varied over time or between genres.

Nevalainen and Raumolin-Brunberg (2003:7) point out that one possible source for historical sociolinguistic material is contemporary comments on usage, and one such source for comments alongside handbooks and debate entries, although not originally written as an explicit comment on linguistic features, is dramatic dialogue. Nevalainen and Raumolin-Brunberg deal with material that mostly predates mine (1410-1681 in their case versus 1640-1920 in my case), and so does Romaine, but the possibilities are the same. However, Nevalainen and Raumolin-Brunberg also point out that the trustworthiness of such comments is not complete; it may be the case that an author only had a superficial idea of a certain dialect included in a play or a novel, and it may also be that the use of different varieties of English was exaggerated for humorous purposes.

Romaine (1982: 114) discusses types of text and there she does not use the word genre. Still, the rough division of the texts in her study is very similar to the division I can present in my material, even though my material is much larger than hers, covering a much longer period of time as well as wider geographical spread. If I understand Romaine correctly, she sees type of text, or genre, as a way of representing sociolinguistic ideas of style in a setting where style of speech cannot be studied. Interestingly enough, the writing style in a historical document may be more natural in its context than recorded interviews where an informant is interviewed by a linguist he or she does not know, a situation which is known to influence the formality of speech (ibid: 118). She suggests that since there are no phonological cues available to define styles in a historical context, the method which remains is frequency of specific items in the texts, and thus, the different types or genres of text could be arranged on a stylistic continuum according to the frequency of the item one is studying (ibid: 120).

Historical sociolinguistics is not just historical linguistics. Historical sociolinguists need to be able to take into account not only linguistic matters, but also social matters, and to draw on social historians' consensus of the time period they study. They "should be able to rely on social historians when exploring, for example, the connection between social status and language change" (Nevalainen and Raumolin-Brunberg 2003: 10) in any period of history.

Since this study is about word meanings and the widening thereof, it is not so much a

question of archaic forms that have been replaced by newer ones, which is an important part of the studies of Nevalainen and Raumolin-Brunberg as well as Romaine. Therefore, I will not look at the four questions that Nevalainen and Raumolin-Brunberg emphasise in their introduction (2003: 5-6):

- How and when did certain new forms get replaced by those used today
- How did the new form gain popularity
- Who promoted the new forms, and what was the social status of these people
- How did the evaluation of the new forms change over time?

However, it can still be useful to look at the range of external factors listed (2003:2): age, social status, gender, domicile, and, in the case of letters, and relationship between correspondents.

2.2.3 Attitudes

Orientalism became an increasingly important feature of English culture in the 18th century (Bailey 2010: 187), and the consumption of tea and coffee fit well into this scheme. At the same time, however, attitudes in Britain toward the foreign over all were not very friendly in the 18th century. As we will see in chapter 2.3, there was a suspicion towards China as well as the Arab world, which led to a certain degree of hostility in some circles towards both tea and coffee. Coffee drinkers were viewed with apprehension and the popularity of tea was seen as a danger to the British economy. It was not only countries far away and products imported from these places that were seen this way – concerns were raised also about imports from France.

The case of France is of course related to the general animosity between Britain and France in the 18th century: “For much of the 18th century France was ‘the enemy’” (Beal 2012: 142). It was not just foreign commodities that were viewed with such suspicion. Attitudes towards the French language were ambivalent bordering on contradictory. Beal points out that “on the one hand, knowledge of French is a mark of education and gentility but, on the other hand, overuse of French words and pronunciations [...] is seen at best as a ‘fantastical badge of foreign foppery’ and at worst as unpatriotic” (2012: 143). Bailey (2010: 186) illustrates this same idea by the help of Horace Walpole who coined the term *Frenchism* for loanwords on etiquette, which filled a gap but were ridiculed.

English linguists competed against their French colleagues to collect a dictionary of their respective languages. There were no general dictionaries of the English language as late

as in the 1720s while the French Academy as well as its Italian counterpart had created their own relatively extensive ones (Mugglestone 2012: 175), even though the Royal Society had let form a committee as early as in the Restoration era to consider the English language with the aim of making English “sufficiently copious to enable serious intellectual work” in English instead of Latin (Bailey 2010: 182).

The need for dictionaries and other books on language was connected to nationalism, since a good language was seen as both honourable and useful to a nation, and dictionaries were needed to keep up the quality of the language and “tackling that ever-vexed issue of linguistic change” (Mugglestone 2012: 176). Hickey (2010: 1) says that the 18th century was indeed a period when prescriptivism was established, and goes on to point out that during this time there was a growing number of professions that had to deal with questions of language (ibid: 3). The books on lexicology, grammar and pronunciation were mostly aimed at the middle classes (ibid: 8-9), which also shows that there was a growing middle class interested in writing and in social movement. This, especially a general growth in numbers of writers and texts, is something that one might expect to see traces of in my material.

Samuel Johnson is perhaps the most famous lexicologist thanks to his *Dictionary* from 1755, which is a work where his own attitudes to different things and people shine through quite clearly. For a large part of the 18th century, many of the English dictionaries that were produced concentrated on polysyllabic words with a heavy Latinate influence (which would, in later dictionaries, be described as obsolete), and decidedly less on “simple”, more clearly English words (Mugglestone 2010: 311). Johnson’s importance seems to lie in the fact that he also included simple words and listing, for instance, 64 senses of the word ‘go’ (ibid: 316), despite the fact that he was criticised for it during his lifetime.

Still, it was only in the late 1800s that a project concerning a much more definitive dictionary was started, that is, what eventually became the *Oxford English Dictionary*, which also serves as a source of both facts and choice of methodology for this study. The starting point for the compilation of the *OED* was a very inclusive and ambitious one, with an aim to include everything even remotely English, also the vocabulary of English colonists living in India and South Africa, as well as rural parts of the British isles (Mugglestone 2012: 180). At the same time, however, there are examples of descriptions from the first edition of the *OED* that show a tendency even stronger than that of Johnson’s personal attitudes of nationalistic supremacy. Mugglestone gives the example of the word *petticoat*, which, alongside the description of an undergarment for women is “applied also to the rudimentary garment worn by women among primitive or uncivilized peoples” (ibid: 182). The extent to which this

tendency is prevalent in the first edition of the *OED*, could, according to Mugglestone, actually be argued to be much worse than Johnson's attitudes even towards Scotland, as in his description of *oats*: "A grain, which in England is generally given to horses, but in Scotland supports the people" (Johnson 1785 (2): image 206).

2.3 Cultural and social history of tea, coffee, and chocolate

Tea, coffee, and chocolate are products that originate in different parts of the world and were brought to Europe at roughly the same time. While tea originates in China and Japan and was first brought to Western Europe by the Dutch in 1610 and around 1645 to England (Sabeti 2010: 86-91), coffee, which originates in Africa and was a popular drink among the Arabs, seems to have arrived to England a little earlier.

The first coffee house in London was opened in 1652, two years after the very first one in England had been opened in Oxford (Griffiths 2007: 352), and although that is a few years later than the first arrival of tea, some time was probably needed for coffee to be popular enough for a coffee house to open and work. Tea, in fact, was introduced to a wider public in England at the coffee houses, after having been a luxury product for royalties and nobility (Sabeti 2010: 91). Chocolate, on the other hand, was brought to Europe from its native Central America by the Spanish, above all by returning Spanish Franciscan missionaries. The first chocolate houses were also established in London in the late 1650s (Moss & Badenoch 2009: 31, 35).

Eventually, a gender split became clear: gentlemen took coffee and tea at coffee-houses, as well as chocolate at the chocolate-houses as long as those lasted (Moss & Badenoch 2009: 37), while discussing politics and business as well as literature and philosophy; in fact, the clubs of the 18th century were born out of the coffee houses (Roberts 1996: 49). Ladies, on the other hand, generally took their tea at home, after Thomas Twining in 1717 had started selling dry tea in his shop (Griffiths 2007: 354). After a period in the early 19th century when coffee-houses had declined and most people took their tea at home, the mid-19th century came with more elaborate meals at tea time, including social visits (Sabeti 2010: 102-106).

2.3.1 Coffee and coffee-houses

The common starting point is that all the three drinks tea, coffee, and chocolate were seen as

beneficial to the health and they were mainly taken for this reason rather than for their tastes in the 17th century (Moss & Badenoch 2009: 37). For instance, a doctor called Thomas Willis stated “I do frequently prescribe this drink sooner than any thing else for their cure, and therefore am wont to send the sick to the coffee houses sooner than to the apothecaries shops” (as cited in Cowan, 2005: 25), and he also described the effects of coffee as “anti-hypnotick ... highly efficacious or the driving away the *Narcosis* or stupyfyingness” (as cited in Markman 2006 (4): ix).

There was also hope that a wider consumption of coffee would replace the widespread use of alcohol and the “too prevalent” drunkenness (Cowan 2005: 26). The latter was the case not just in general, but quite specifically under the puritan regime of the English Republic (Markman 2006 (1): xxvi). As a matter of fact, coffee and the Levant rituals around coffee were initially likened to the consumption of alcoholic drinks in Europe and to the culture and rituals surrounding European taverns and alehouses (Cowan 2005: 18), and this may have been something that increased the hopes of coffee serving as a replacement for alcohol.

However, all three drinks quickly came to constitute an important part of social life. According to Varey, “[d]uring his time in London, [Jonathan] Swift frequently ate on his own but drank his tea or coffee or chocolate in company” (1996: 47). As opposed to the culture around more traditional alehouses and taverns, the culture around coffeehouses, originally based on a group of elite gentleman intellectuals, as these had been the people who introduced coffee to England, grew to be associated with learnedness to the extent that coffeehouses were referred to as *penny universities*. Some of the coffeehouses catered to a very select group of people, often intellectuals, while other were more open to the public, and as with other establishments, they could be open at almost any time of the day (Cowan 2005: 79-87). Markman (2006 (4): x) compares the attending of coffeehouses to academic research centres and symposia, where scholars could debate and exchange ideas.

Coffeehouse culture was to a certain degree a symbol for the urban experience of English cities, London especially (Markman 2006 (1): xii). Drinking coffee was not the most important part of the coffeehouse culture even though the phenomenon has its name from the drink - it played a large role in both the social and commercial life of the city and primarily the coffeehouses were places for conversation and entertainment. According to Markman (2006 (2): vii), coffeehouses were more common in London and after 1688, they “seemed to represent the indigenous spirit of liberty in the English constitution”, serving as places where people of different background could meet and discuss in a spirit of equality. But the picture described of coffeehouses in Markman’s work is not simple: according to the estimations there were over

two thousand of them in London in the 18th century and it is not surprising to find great variety among them.

The re-establishment of theatres after the Restoration coincided with the growth of the coffeehouse-culture, according to Markman (2006 (3): vii-viii). This led to a certain overlap in the social space of theatres and coffeehouses, especially in the vicinity of the two theatres that were built in the late 1600s, and this added a literary tone to these coffeehouses. Markman even goes as far to say that these particular coffeehouses were evidence for a growing professionalisation of literary criticism (ibid: ix), but also a common setting for dramatic comedies and satires.

Coffeehouses did not only serve coffee, but also tea and chocolate in a variety of combinations, as well as the more traditionally European alcoholic beverages. Coffee, however, was not generally served at other drinking establishments beside coffeehouses (Cowan, 80-82). Coffeehouses became a place where new products were introduced and new innovative consumption habits were developed (Cowan 2005: 83). One of the most important products introduced to the public in coffeehouses was tea. All through the 17th century the English preferred coffee over tea and chocolate, but from the 1720s onward, tea gained more and more popularity especially among women, who drank tea in the homes while gentlemen frequented the gentlemen's clubs which had grown out of the coffeehouses of the previous century.

2.3.2 Tea

While the acceptance of tea was slow and the British preferred coffee in the 17th century (Saber 2010: 91), those who could afford it drank tea in the 18th century (Varey 1996: 37). In 1784, the Commutation Act cut taxes on tea from 119 % to 12.5 % and the result was, among other things, that tea became affordable to all social classes. However, there still were class differences: in the 19th century, the higher classes would enjoy huge tea receptions with various kinds of entertainments, while the lower and middle classes took high-tea, which was "a good hearty meal" (Saber 2010: 106) with all kinds of hot and cold, savoury and sweet dishes. A meal in the afternoon, mostly referred to as *afternoon tea*, became a tradition when gas lighting had resulted in longer working days and postponed the evening meal (Griffiths 2007: 356).

One possible reason for tea being moved into the homes to a much larger extent than coffee is that tea is much easier to prepare, and more economical since the tea leaves can be steeped more than once, if need be (Fromer 2008: 13). The latter, in combination with the lowered taxes, was probably vital in the spread of the consumption of tea also to the lower social

classes.

In her book *A Necessary Luxury* (2008), Julie Fromer argues that tea in the Victorian age was one of the corner stones of English society as well as the British Empire. Tea played great roles in matters of nation, class, and gender, in terms of how these were perceived and how they were created. Fromer bases her study on different kinds of texts from the Victorian time: apart from novels where tea plays a central part, she also analyses what was said in contemporary histories of tea.

In Victorian England, the English society was considered to be built on tea because tea was the one thing that functioned as a bridge between many opposing forces such as “men/women, middle class/lower class, labor/leisure, necessity/luxury, England/Orient, home/empire, ideal/real” (Fromer 2008: 16). Tea represented the comfort of home, hospitality, spiritual nourishment, and connection to others – within the family, with guests, and to other English citizens throughout the country as well as representing home and Englishness out in the Empire (ibid: 31). Tea also came to have a highly gendered symbolism: while men connect over other substances such as coffee, tobacco and alcohol (or in the case of coffeehouses some hundred years earlier, all of these as well as intellectual debate), women chose tea when they were among themselves as well as when they were in the company of men. This, of course, put tea in a more domestic and private setting than coffee, and thus, tea was seen as something that helped men in their transition from work to home on a daily basis (ibid: 22).

The huge demand for tea in Britain was an important factor in the growth of the British Empire. In the early days of tea in England, it had been looked down upon to rely on a foreign product to such a degree. Since tea was found to grow naturally in India, however, the fact that tea had already become a domesticated product in England, Victorian imperialists could justify to themselves the British expansion into India (Fromer 2008: 27, 41). And by drinking tea, the English people were told to envision themselves as contributing to the growth of the empire.

Fromer argues that in the Victorian age, even the English moral was considered to be based on a constant rehearsal of the habit of drinking tea. Tea drinking was claimed to replace vices such as alcoholism (as had previously been argued about coffee, as mentioned in 2.3.1) and violence with values of domestic economy and good taste (Fromer 2008: 21, 63). It was even put forth in newspapers at the time that the chemical components of tea made the food a person had consumed last longer, and thus, it was highly recommendable that poor people should drink high quality tea to save money on food (ibid: 60). Drinking tea, it was argued, was a way of becoming a better English person, simultaneously building one’s character, one’s relations to other people, and the Empire, all in a very desirable middle-class manner.

2.3.3 Chocolate

If the cultures surrounding coffee and tea are slightly complicated and gendered, then so it can be argued to be the case with chocolate as well. Consumption of chocolate was kept low in Europe due to wars and revolution during the late 18th century in Central- and South America, where it was grown (Moss & Badenoch 2009: 54-55). Chocolate also, more than coffee and tea it seems, retained the position of medicine in much of the 19th century (ibid: 60). There were, however, two important inventions in the 19th century: van Houten's cacao press from the 1820s which allowed the separation of cocoa butter from the cocoa bean (ibid: 58), and the 1879 Swiss invention of milk chocolate (ibid: 61). These two inventions allowed chocolate to take the solid form that is most common today (i.e., fairly late in relation to most of the material to be analysed in the thesis).

Solid chocolate came to be marketed as "healthy nourishment for children" (ibid: 67) during the 19th century, something a good mother would provide, and by doing so also contribute to the good fortunes of the nation. On the other hand, the drink hot chocolate or cocoa still existed and came to be marketed as the perfect "nutritious foodstuff and meal replacement" (ibid: 63) for the working classes which would "boost industry and production" (ibid: 72) and thus, it was targeted more toward men than women.

2.4 The words *tea*, *coffee*, and *chocolate*

According to the OED, the word *tea*, probably from the Amoy dialect of Chinese, was first used in English around 1655. English most likely borrowed the word from Dutch (Campbell 1998: 59). The leaves of the tea-plant were mentioned as early as 1598 by the name *chaa*, from Mandarin Chinese. The many spelling variants listed in OED are *tay*, *tey*, *té*, *thé*, *the*, *tee*, *thea*, and finally, *tea*. The most important meanings listed are 'the leaves of the tea-plant', 'a drink made by infusing these leaves in hot water', 'the plant from which tea is obtained', 'a meal or social entertainment at which tea is served', 'a general name for infusions made in the same way as tea'. Finally, the OED lists several attributive and combining usages, and as a separate entry, the verb *to tea*: 'to supply or regale with tea; to entertain at tea; to give a tea to', and 'to drink tea; to take the meal called tea, to have one's tea'.

The word *coffee* most likely originates in the Arabic word *qahwah*, and has probably found its way to the European languages through Turkish in about the year 1600. The first entry in the OED is from 1598 in the form *Chaoua*; the spelling variants listed are *caoua*,

chaoua, cahve, coava, coave, cahu, coho, kauhi, kauhe, cauwa, coffa, cafffa, capha, caphe, cauphe, cophie, coffi(e), coffey, coffea, coffy, coffe, cophee, caufee, and coffee. Meanings are ‘a drink made by infusion or decoction from the seeds of a shrub’, ‘a light repast at which coffee is taken, or a final course at dinner consisting of coffee’, ‘a shade or tint of the colour of coffee’, ‘the seeds or 'berries' [...] from which the drink is made’, and ‘the trees or shrub from which coffee is obtained’. There are also several general and special combinations, as well as a separate entry for the verb: ‘to drink coffee’, and ‘to entertain at coffee’.

The background of the word *chocolate* seems to be rather confused and something of a mix up of two different Mexican words referring to different kinds of products made from cacao only and from cacao and another tree; thence the confusion in English of *cacao*, *cocoa*, and *chocolate*. The word *chocolate* was borrowed to English from Spanish (Campbell 1998: 58). Spelling variants listed in the OED are *chocolata*, *-latte*, *-letta*, *-lat*, *chocaletto*, *-latte*, *chocolate*, *jacolatte*, *jacolat*, and *jocolat*. The meanings listed for *chocolate* are ‘a beverage made from the seeds of cacao-tree’ and the beverage ‘made by dissolving chocolate cake in boiling water or milk’. The OED also lists ‘a paste or cake composed of the seeds of the cacao-fruit’, ‘chocolate colour’ and ‘chocolate-coloured’, as well as several attributive and combining usages. There is a separate entry in the OED for *cocoa*; which, according to the OED, is ‘the powder produced by crushing and grinding the [cacao] seeds’ and ‘a common beverage made from this powder or from the prepared seeds’.

3. Material

The primary material for my study consists of six historical corpora: *Parsed Corpus of Early English Correspondence* (PCEEC), *The Helsinki Corpus of Early Modern English* (HCME), *Newdigate Newsletters* (NEWD), *Century of Prose Corpus* (COPC), *The Corpus of Late Modern English Texts* (CLMET), and *The Corpus of Late Modern English Prose* (CLMEP). The texts are from almost any thinkable genre, ranging from private letters, treatises and essays to novels – however, with a substantial weight on fiction and letters.

3.1. Corpora

Table 1 gives an overview of the different corpora used in this study. As can be seen, some of them overlap each other in time. The numbers of words vary a great deal, which is perhaps not surprising, considering the huge growth of writing and publishing over the centuries. For

this study, the six corpora have been sorted into four time periods: 1640-1709, 1710-1779, 1780-1849, and 1850-1920.

Table 1

Number of words in the corpora (million) and the time periods which they cover

<i>Name of corpus</i>	PCEEC	HCET	NEWDI	COPC	CLMET	CLMEP
<i>Time period</i>						
1600						
1650						
1700	0.555	0.17	1.06			
1750				0.497	2.096	
1800					3.74	
1850						
1900					3.982	0.1
1950						

This division of the time periods is mine; originally, the time periods used for *The Helsinki Corpus of Early Modern English* and *The Corpus of Late Modern English Texts* were 1640-1710, 1710-1780, 1780-1850 and 1850-1920 where the years 1710, 1780 and 1850 are included in two periods each. De Smet (2004) does not mention this at all in his presentation of *The Corpus of Late Modern English Texts*, and thus, every text in the corpus should be examined to find out whether 1780-1850 refers to 1780-1849 or something else. What I did here was to check the texts where any of the three words *coffee*, *tea*, or *chocolate* appear. In period 1, there are texts including the words written in the time span 1642-1696; in period 2 1711-1776; in period 3 1780-1848, and period 4 1857-1920. Most of the other corpora fitted within these period borders used in the two corpora mentioned above; I will discuss the *COPC* in more detail below.

Altogether, the entire corpus used in this study consists of 12,200,000 words. The division of the corpora over the periods is illustrated in Table 2 below.

Table 2: Time periods with corpora and number of words

	Name of corpora	Total number of words in corpora. millions of words	Total number of words in time period
Period 1: 1640-1709	PCEEC	0.56	
	HCET	0.17	
	NEWDI	1.06	
	COPC until 1709	0.14	1.927
Period 2: 1710-1779	CLMET A	2.10	
	COPC 1710-1779	0.35	2.441
Period 3: 1780-1849	CLMET B	3.74	
	COPC after 1780*	0.01	3.75
Period 4: 1850-1920	CLMET C	3.98	
	CLMEP	0.10	4.082
		Total:	12.2

Parsed Corpus of Early English Correspondence (PCEEC) consists of letters from the years 1410-1681. For this study, out of the whole 2.2 million word corpus, only the latest period is used, i.e., from 1640 onwards with about 555,000 words (PCEEC: 2006).

The Helsinki Corpus of English Texts (HCET), 1.5 million words) consists of various kinds of texts: e.g. letters, treatises, scientific texts, and fiction. It is divided into three periods: Old English (starting around the year 730), Middle English, and Early Modern English, each of which is also divided into sub-periods. For this study, only the period 1640-1710, i.e. the youngest sub-period of the Early Modern English collection, is used, contributing with 170,000 words.

Newdigate Newsletters (NEWDI) consists of 2,100 newsletters with a total of 1.06 million words, written in the period 1673-1692.

Century of Prose Corpus (COPC) has a similar variety of texts as the *HCET*, written by 120 different authors between 1680 and 1780 with a total word count of 497,000. In this corpus, the spelling has been normalised, and therefore it is not useful in the tracing of spelling variation through time. That, however, does not render it useless for investigating changes in word meaning. *COPC* is the only corpus that crosses the borders between the time periods that the other corpora fit into, as it ranges from 1680 to 1780. Therefore it had to be split into two parts: one before and excluding 1710 and one including 1710 and after, where the numbers of words had to be calculated separately. *COPC* in itself consists of two

parts: part A consists of text samples of roughly 5000 words each, representing specific authors, and part B of genre-based samples of roughly 2000 words each (Milic 1990: 27-30).

The Corpus of Late Modern English Texts (CLMET) consists mostly of literary texts from the period 1710-1920, and has a total of 9.8 million words. The corpus is divided into three parts of 70 years each: 1710-1780 (2.096 million words), 1780-1850 (3.74 million words), and 1850-1920 (3.982 million words). Note that this is not the newest version of the corpus, which has 34 million words (Diller: 2011), but the new version was not available at the time the coding of this project was already finished.

The Corpus of Late Modern English Prose (CLMEP) consists of letters from the years 1861-1919 and a word count of 100,000 words.

3.2 Adjustments of text periodisation and corpus size

In checking the dates of all the texts, I noticed a discrepancy between the information in different sources regarding one of the writers, Samuel Baron. His text had, in the process of splitting up the *COPC*, previously been filed in Period 2, due to the dates given in the text files (1732). However, according to the description of the corpus (Milic 1990), the date should be 1700, thus placing it before the period boundary of 1710. However, according to Winterbottom (2013), Baron enclosed the manuscript for the text in a letter sent in 1686, which means that even though the text was not published until 1732 (which corresponds with the date in the corpus file), it must have been written at the very latest in 1686. This places the text firmly in period 1.

Similarly, I also noticed that despite the fact that the *COPC* is supposed to include texts written in the century between 1680 and 1780, there is one text that was written after 1780 and published in 1790: Edward Gibbon's *Memoirs of My Life*. This occasioned a similar transferral of two of Gibbon's texts from period 2 to period 3. The other text (a part of *History of the Decline and Fall of the Roman Empire* which was written in 1781) does not include the words *tea*, *coffee*, or *chocolate*, but I moved it along with the memoirs in order to aim for a more correct word count in each of the periods. I have not, however, checked every text in the *COPC* which do not include any of the three words to see if they were written after 1779, so there may still be texts that are included in the wrong period.

The transferral of Baron's text consisting of roughly 2000 words from period 2 to period 1 does not notably change the balance between the two periods, and the approximate sizes of the periods remain the same, and the same is the case for the transferral of Gibbon's

two texts consisting of roughly 5000 words each from period 2 to period 3. Yet since the *PCEEC* has already previously been counted as having 555,000 words, it could not be justified not to correct the number of words of the older part of the *COPC* to 142,000 instead of 140,000, and similarly, the newer part of *COPC* from 360,000 to 355,000 thus also correcting the fault resulting from previous rounding of the number of words from 357,000 to 360,000. As a result, the size of the whole corpus also shrank from 12.185 million words to 12.182 million words.

This more strict treatment of the *COPC* led to a checking of the more detailed numbers of words for the *CLMET* as well, which in the end actually enlarged the corpus to a final total of 12.2 words. More details can be found in table 2.

4. Methods

Kennedy (1998: 1) describes a corpus as a "body of written or transcribed speech which can serve as a basis for linguistic analysis and description", and adds that since the 1970s, the technical innovations have allowed corpora to take the form of computerized databases which can be coded in several ways: tags reflecting grammatical and semantic groups of the words (Baker 2014: 8-9). Kennedy also points out that corpus linguistics is "not an end in itself", but sees it as a "source of evidence for improving descriptions of the structure and use of language".

By using computer software it is quicker and more accurate to identify, count and sort the words, collocations and grammatical structures found in a corpus (Kennedy 1998: 204), and there are certain routine procedures that are involved in a corpus study. For instance, one can make word lists of all the words in a corpus including numbers of frequency, and these lists can easily be sorted alphabetically or according to other parameters (ibid: 244-250).

Perhaps the most useful method is the creation of concordances, which means using a software to search for specific words in the corpus and make a file containing each occurrence of the word, or node, presented in the centre with a predetermined or easily altered amount of context on either side (Deignan 1999: 20, Kennedy 1998: 251). The most usual form of a concordance is the *KWIC Concordance*, which stands for *Key Word in Context Concordance*. Like word lists, the concordance can be sorted in various ways, e.g. right-sorted, i.e., alphabetically based on the first word to the right after the search word, *node* (ibid: 256), or based on the names of the text files in the corpus and in the order of occurrence within each text. Every way of sorting has its uses; e.g., if you are interested in

seeing who uses the words the most, and every author has his or her own text file, the latter method of sorting is useful; if you are studying collocations, you would choose the former method. For this end, I have used a software called WordSmith 4 (Scott 2004).

There are a variety of corpora created for different purposes, consisting of a wide variety in the selection of texts, either in full or in smaller excerpts from texts. Some corpora, reference corpora, are more general and aim to include a bit of everything, while others are specialized in e.g. spoken language, regional varieties, or specific genres. However, the distinction, Baker says (2014: 10) is linear rather than binary.

The point of corpora is that they are collections of "naturally occurring data" as Deignan (1999: 19) puts it, from the "real world" (Baker 2014: 7). The technical procedures allow studies in more detail and more reliable statistics. Also, Deignan (1999: 23) points out, corpus linguistics can provide "less intuitively obvious information" about linguistic features. A good thing about corpus studies is that one can avoid accusations of cherry-picking one's examples in order to prove a point, Baker (2014: 10) says. It is not inherently impossible to cherry-pick examples even in a corpus study, especially if corpora are used in a qualitative analysis, but the corpus is still there to prove either the researcher or the critique right or wrong.

As an approach within linguistics, corpus linguistics is empirical much like social sciences and bases generalizations on samples Baker says (2014: 7), but at the same time he points out that it is a "misperception that it is a purely quantitative approach". Another thing that he mentions is that the word *corpus* does not always refer to an actual corpus; many authors use the word for their data set regardless of the size or methods (ibid: 6). Also, while a corpus can offer data to analyse, it "does not always yield explanations for language patterns" (ibid: 197) and reminds that one often needs to take the context into consideration in order to understand our findings.

Teubert (2004: 104) discusses the importance of corpus linguistics in lexicology and semantics, and how suitable corpus linguistics are for these specific fields of linguistic studies. He points out that while there are branches of linguistics that do not particularly care for corpus linguistics at all, they are often not very much concerned with semantic change. One important distinction can be found in Teubert (2004: 112) - the one between corpus based findings and corpus driven findings. The first is the case when everything that is being said about the results is validated by evidence in a corpus. In the latter case, the findings are intellectually processed and analysed after having been extracted from the corpus with the methods of corpus linguistics - in other words, what Baker (above) said is often the case with

corpus studies especially if one aims to make claims along more sociolinguistic lines. Chapter 5 in this study is to a large extent corpus based, while chapter 6 is corpus driven.

4.1 The *OED*

When I looked for the spelling variants and meanings of the words *tea*, *coffee*, and *chocolate*, as well as the compounds associated with each of the words listed in the *OED*, I used the spelling variants as search words in WordSmith 4 (Scott: 2004) to find as close to all occurrences of the words as possible. The WordSmith concordances were used for finding the instances in the corpora, and the occurrences were categorised according to period and the frequencies and distributions of the different meanings of the words.

The categories of meaning that are used in this study are loosely based on the meanings listed for the word *tea* in the *OED*. The most important meanings listed in the *OED* are as follows:

1. a. The leaves of the tea-plant
2. a. A drink made by infusing these leaves in hot water
3. The plant from which tea is obtained
4. a. A meal or social entertainment at which tea is served
5. Used as a general name for infusions made in the same way as tea (sense 2)

OED: s.v. *tea*, n.

1. trans. To supply or regale with tea; to entertain at tea; to give tea to
2. intr. To drink tea; esp. to take the meal called tea, to have one's tea

OED: s.v. *tea*, v.

The basic categories in my categorisation are 'produce' (including the leaves as in *OED*'s sense 1 and the plant as in *OED*'s sense 3), 'drink', 'meal', 'mealtime', and the word being used as a verb. In addition to these, a quick glance at the concordances proved the need for almost any combination of 'produce', 'drink', 'meal', 'mealtime', and verb: in some cases a few of these meanings could be excluded while the meaning still remained unclear.

In the case of *coffee*, I chose to keep the same categories to the extent that I found cases that fitted (several combinations of categories were omitted in my tables for *coffee*). The *OED* meanings are the following:

1. a. A drink made by infusion or decoction from the seeds of a shrub (see 3)
1. b. A light repast at which coffee is taken (cf. *Tea*); or a final course at

dinner consisting of coffee

1. c. A shade or tint of the colour of coffee
2. The seeds or 'berries', either raw or roasted; or the powder made by grinding the roasted seeds, from which the drink is made.
3. The tree or shrub from which coffee is obtained.
4. a. The name has been commercially applied to various substances or preparations used as imitations of coffee, or substitutes for it

OED: s.v. coffee, n.

1. intr. To drink coffee.
2. trans. To entertain at coffee.

OED: s.v. coffee, v.

For *chocolate* I also added the meanings 'solid sweet' and 'colour'. The OED meanings for *chocolate* are the following:

1. A beverage made from the seeds of the cacao-tree
2. A paste or cake composed of the seeds of the cacao-fruit
4. a. Chocolate colour
4. b. as adj. Chocolate-coloured

OED: s.v. chocolate, n.

The compounds were grouped only loosely based on the OED. All three words have attributive combinations and special combinations; I have here cited the main subheadings from the OED for the word *tea*:

9. Attrib. And Comb. a. Of, pertaining or relating to, dealing or connected with tea as a commodity; containing or intended to contain tea; of or pertaining to the tea-plant or its cultivation
- b. Objective and obj. gen. ns. and adj.; instrumental and parasynthetic adjs.; also simulative adj.
- c. Special combs.

OED: s.v. tea, n.

Here, I found the categories in the OED somewhat wide and abstract, so I settled for a more concrete categorisation: utensils for preparation or serving; storage; related to plants; locations; meals or events; people, merchandise and transport; proper nouns; and finally, some others that do not fit into any of these categories. I started with the grouping of the compounds concentrating on *tea*, and therefore, these groups may not suit the needs perfectly when it comes to *chocolate*; only two of the groups are used in that case and one of them is 'other' containing an array of different words. That is, the category 'other' was created for those of the compounds based on *tea* that did not fit into any of the other categories, but hardly any of the compounds based on *chocolate* fit into any of the more informative

categories, nor do they form groups of their own. Therefore, most of them fall under the rather uninformative title ‘other’.

Finally, a quick note regarding the choice of words to include in the study: the confusion concerning the background of the word *chocolate*, as described in chapter 2.4, would render it worthwhile to take the words *cocoa* and *cacao* into consideration in this study, along with *chocolate*. However, the attempt turned out less fruitful than could have been expected based on the information in the OED for the simple reason that the word *cacao* does not appear even once in any of the corpora, and even though the word *cocoa* appears 174 times in all corpora, only about a dozen of the instances refer to chocolate; the vast majority of the instances refers to coconuts (in these cases spelt *cocoa-nut*) and the tree that the coconut grows on.

4.2 Finding Compounds

A general problem with any corpus program is that a search for *tea* will find every instance of the independent word *tea*, and compounds that are spelt either with a hyphen or as two words, even though in the latter case, the whole compound is not necessarily marked as a search result. On the other hand, sometimes the word following either an independent case of *tea* or a compound spelt with or without a hyphen may be marked as a search result even though it has nothing to do with it. The search for *tea* will not find compounds which have been written as one word, e.g. *teapot*. On the other hand using an asterisk in the search (e.g., *tea**) will include words such as *teach* and *tear*. In CLMET alone, this search finds 3923 cases, and in all of the corpora the number of these cases is 4628, which is too much data to go through manually. It is also worth noting that a search for **tea* in the material does not add anything new, i.e., there are no cases of *hightea* or *afternoontea* spelt as one word.

To find the compounds spelt as single words, I did a normal WordSmith search in Concord for *tea**. The concordance list with the marked search results and their environments was copied and pasted into a Word document. That document was then used in the WordSmith feature Wordlist to create a word list which could be sorted alphabetically and where every word starting with *tea* could easily be found. This is not necessarily the most efficient way to go about the problem, but it is easier than to read through all the 4628 cases. However, after finding out which compound words exist in the material, I had to search separately for them again in Concord, because the information about what text they appear in originally was lost in the process, but that was also easier than to look for every compound

suggested by the OED separately or every compound otherwise found in the material.

4.3 Spelling

In a corpus based study on lexicography it is important to consider the spelling of the words that are being examined. When dealing with historical material, one must remember that often there was a lack of strict spelling-rules, perhaps especially so with new loan-words. Another aspect is that of artistic expression in literary works, which will be seen in this study. This section presents my decisions regarding spellings, both regarding the search procedure and the presentation in the thesis.

4.3.1 Spelling of compounds

Keeping in mind what Adams (1973) and Bauer (1983) say about the spelling of compounds (cf. 2.1.3), I have treated all *tea spoons*, *tea-spoons*, and *teaspoons* the same way in this thesis. This non-differential treatment of compounds spelt in different ways is a reason for the endeavour to search for as many of the compounds as possible described above. Knowing that there could be *teaspoons* that could not be found by the easiest search procedures led to an increase in the numbers of instances to analyse.

The equal treatment of differently spelt compounds has, however, resulted in slight inconsistencies of spelling in the tables of occurrences: I have simply chosen the spelling of each word that I have come across first in the list of occurrences in Wordsmith, which is a result of the default sorting, based on the file names of each of the texts. The spelling used for reference to the compound need thus not be the most frequent spelling. Spelling the compounds with a hyphen is most frequent, but while there is e.g. *coffee-seller* there is also *coffee woman*, as well as *tea-spoon* and *teaspoonful*.

4.3.2 Spelling of *tea*

The only corpora that contain words similar to the various spelling variants of the word *tea* as they are listed in the OED are the CLMET and the PCEEC. While searching for these variants, however, I found that most of them do not refer to tea but to something else. Firstly, most of the cases of *Tay* refer to the Scottish river by that name. Still, there are two cases that definitely refer to the substance tea: the two compounds *tay-time* and *tay-boy* (both from

Thackery's *Vanity Fair* from the late 1840s, i.e., period 3). The two cases of *tee* in the CLMET do not seem to have anything to do with *tea*, and finally, the same is the case with *thea*, which appears to be a personal name also in the case where it is spelt without a capital initial letter. Finally, *tee* occurs once in the PCEEC: *I have scarce bought any thinge for my selfe but an Indian Brewhouse for tee* (Arundel, 1642) - which at the same time is one of only two cases of *tea* from first period in my material.

Concerning the spellings of the word *tea*, there is one listed in the OED that is more or less impossible to search for using a concord program: *the*. Even though it seems likely that at least some of the older texts have the word *tea* spelt that way, it would be impossible to find among the masses of definite articles. This could be a topic for a more computer science-oriented study, but not for the present one.

4.3.3 Spelling of *coffee*

The CLMET, the Newdigate newsletters, and the PCEEC are the only corpora that offer any spelling variants, and still, only a few of the variants listed in the OED can be found at all. The CLMET and the PCEEC offer *coffy*, found e.g. in Thackeray's *Vanity Fair*, and it refers to coffee, but it seems to be a case where the author uses misspelling to characterise a novel character; the sentence following this spelling variant of *coffee* contains the word *accadent*. Both words are spelt with capitals only. We may also note that this is the same text that contains an alternative spelling version of the word *tea*. A case from the PCEEC clearly refers to the drink: *what good efficacy Coffy has for the taking away of headache* (PCEEC: Conway1664).

Caffa, suggested by the OED as a spelling variant to *coffee*, occurs a couple of times in the CLMET, but it exclusively refers to a geographic name. In the Newdigate newsletters and the PCEEC there are eleven of cases of *coffe* referring to coffee, for example: *to grant 6 months liberty to the Coffe house to vend Coffe under the condition mentioned in the proclamation*. Many of the words spelt that way also refer to cough: *Poor Tommy has bin very hott and his Coffe very troublesom* (PCEEC: Browne 1681).

4.3.4 Spelling of *chocolate*

Since *chocolate* is the word, out of the three examined in this study, that occurs by far the least often, it is perhaps not surprising that there are only two instances of alternative

spellings. However, considering the confused background of the word *chocolate*, it is also a little surprising that there are only two alternatives to the standard form. One of them is *chocolat*, in the French phrase *Chocolat a la Vanille* in *The Woman in White* by Collins, dated 1860 or in period 4. The other is *chocolato*, which is found in a letter from 1664.

4.4 Calculation of frequencies

One of the purposes of this study is a quantitative overview of the uses of the words *tea*, *coffee*, and *chocolate*. Therefore, one of the main foci is on comparing the frequencies of these words from various points of view.

Firstly, the total numbers of findings in each period must be balanced with the size of the corpus, i.e. the total number of words in each respective period, and expressed in numbers of findings per one million words. This is because of the fairly low numbers of occurrences, a total 1594 cases in a corpus of more than 12 million words, and because of the fairly varying distribution across the time periods; this also makes it possible to avoid uncomfortably many decimals. The low number of occurrences over all as well as the varying distribution are, naturally, characteristic of a study that deals with specified content words that are new to the language in the time period in question. These results will be presented in Table 3 in Chapter 5.

Secondly, the frequencies of each of the words within the time periods are compared to each other, as are the frequencies of the independent words and the compound words, the focus being on the separate words. Moreover, the frequency of each word over time, as well as the independent and compound words are compared, and here the focus is on how the occurrences are spread out over the time periods. These results will be presented in Tables 4 and 5 in Chapter 5.

Thirdly, the instances of the categories of different meanings are compared to each other, both independent words and compounds separately as well as independent words and compounds together. Furthermore, the instances of the different meanings have been calculated including the meanings of the morphemes *tea*, *coffee*, and *chocolate* in the compounds. These details will be presented in the tables in Appendix 1, while Appendix 2 contains the numbers of occurrences of each of the meanings of the words as well as each of the compound words.

As to the shares of texts in different genres and texts that contain the words *tea*, *coffee*, and *chocolate*, it should be pointed out that the texts that I use as the foundation for the

proportions are the texts where any of the words *tea*, *coffee*, or *chocolate* occur, or where more than one of them occur. All three words do not appear in all texts, and all texts do not contain occurrences of all the words, but when calculating shares of each word, all the texts in the corpus are counted. Of course, the texts in the corpora are samples to begin with, so it might well be that some of the texts that do not appear in my WordSmith searches do contain the words in other parts than the ones represented in the corpora.

5. Analysis part 1: Meanings and prevalence

This chapter presents the frequencies and distributions of the different meanings of the words *tea*, *coffee*, and *chocolate*, as independent words as well as the compounds based on these three words.

5.1. Semantic classification

While this thesis analyses every instance of the words *coffee*, *tea*, and *chocolate* in my material and tries to fit each of them into a category of meaning, it is important to remember that this is not easily done. The number of unclear categories with two or more alternative meanings shows this very clearly. Still, while it would be nice if each instance of these words could easily be put into clear-cut categories of their own, it is the unclear ones that are interesting. They show that with these words and many of the possible meanings, there are no clear indicators to show what they mean. It could even be argued that there is a kind of continuum with 'the drink' at one end, 'a meal' in the middle, and 'a mealtime' at the other end. One could even add 'produce' before 'the drink', based on the fact that the drink is made of the produce.

The numbers of cases that have a clear, definite meaning are relatively few, with the exception of the most common one, the drink itself. This is the case for all three words, even though it is most obvious in the case of *tea*, partly because of the overall much larger number of occurrences. However, in every case where the words could refer to something other than the drink itself, I have decided to mark this by sorting the cases into a mixed category even though the more likely meaning is the drink itself. The following examples from a letter by Jane Austen show that the word can refer to a meal or the drink (example 1), and to the drink or the produce (example 2):

(1) We paid an additional shilling for our Tea. (Austen, Jane: letters)

(2) Our Tea and sugar will last a great while. (ibid)

Deciding between the meanings 'the drink itself' and 'a meal' or 'a mealtime' based on grammatical environment is almost impossible to do. Here the lexical information given close to the words *tea*, *coffee*, or *chocolate* has to be taken into consideration. There are basically three kinds of cases where it is clear that *tea* refers to a drink: when it is the object of verbs such as *drink* or *pour* (example 3), when it occurs together with *a cup of* or *a dish of* (example 4), and when it is listed as one of several food-stuffs at a meal (example 5). The last kind should not be confused with cases where *tea* refers to a meal and the rest of the food-stuffs are listed, which can indicate that *tea* refers to a meal consisting of several things, sometimes without the drink tea even being present, as in example 6.

(3) Jones desired the favour of Mrs Whitefield's company to drink tea with him... (Fielding, Henry, 1749: *The History of Tom Jones, a Foundling*)

(4) ...with a cup of tea before her. (Brontë, Emily, 1847: *Wuthering Heights*)

(5) They chatted until Amy came up the stairs with tea and bacon.
(Bennett, Arnold, 1908: *The Old Wives' Tale*)

(6) ...gripped by a desire to eat eggs at breakfast or tea... (ibid.)

It could be expected that the grammatical environments of the words *coffee*, *tea*, and *chocolate* would help to decide on their exact meanings. However, as can be seen in example 7, this is not always the case even when trying to identify the verbs, which could intuitively be expected to be easier than finding and recognising different nominal meanings.

(7) The two sat down to tea in Bathsheba's parlour... (Hardy, Thomas, 1874: *Far from the Madding Crowd*)

5.1.1 Semantic classification of *tea*

As stated before, it is not the cases where the meaning is hard to define that are the most interesting, but the unclear cases. First, however, I will mention two cases that are fairly straightforward, which involve the verb *to tea*. As could be seen in example 7, the grammatical environments are not always very helpful even for deciding whether a word is a verb or a noun. Example 8 is one of the only two cases where *tea* is very likely to have the meaning 'to tea' in the whole corpus:

(8) Frank stayed to tea (Rutherford, Mark, 1896: *Clara Hopgood*)

The combination of the verb *stay* and the word *to*, in this case used as an infinitive marker and not a preposition per se, renders it most likely that *tea* is a verb; if it were a noun referring to a meal or a mealtime, the preposition *for* would be used instead. This is supported by a couple of WordSmith-searches. A search for *stay* to* in the *CLMET* shows that a vast majority of the instances (49 out of 58) is followed by a verb, while *stay* for* is always followed by a noun phrase of different kinds, close to half of which express time. In both of these searches, it is the cases that are all related to food that could be followed by either verbs or nouns simply because the words that collocate with the phrases here look the same in both forms: *stay to tea* as in example 8 above, *stay to breakfast* (example 9), and *stay for tea* (example 10) where it is unclear whether *tea* refers to a meal or a mealtime. At the same time, however, Fielding also used a clear noun following *stay to* (example 11), so this method of distinguishing verbs and nouns from each other is in no way complete. Still, the share of verbs following the construction *stay to* in the *CLMET* is 83 % and thus fairly convincing.

- (9) strongly pressed her to stay to breakfast (Fielding, Henry, 1759:
Amelia)
(10) You know Miss Insull is staying for tea? (Bennett, Arnold, 1908: *The Old Wives' Tale*)
(11) desired the favour of her to stay to supper (Fielding, Henry, 1759:
Amelia)

A WordSmith search for *stop* to* shows that this phrase is most often followed by a verb as well (130 out of 176, i.e., 74 %). Most of the noun phrases that appear in that position occur if the phrase is *(put) a (adverb) stop* to*, and that is something quite different. There are no cases in the *CLMET* where *stop* to* is followed by nouns outside the construction mentioned above, but there are two cases where the word class of word following *stop* to* is unclear from its form: again, these words are related to food: *tea* - the second case in my material counted as a verb (example 12) and *lunch* (example 13):

- (12) Don't you see that you've frightened him away? I meant him to stop to *tea*. (Forster, E. M., 1910: *Howard's End*)
(13) If you'll stop to lunch I'll prove you this time travelling up to the hilt (Wells, Herbert George, 1898: *The Time Machine*)

Like with example 8 it is the suggestion that *to* is used as an infinitive marker that denotes *tea* as a verb. With examples 8 and 12 it can be questioned what is proof for what: is the word *tea* a verb because it follows an infinitive marker, or is *to* an infinitive marker and

not a preposition because it is followed by a potential verb form? With the background searches explained here, however, it seems rather safe to say that these two cases are indeed verbs. A search for *stop* for* in the *CLMET* finds no cases where this combination is followed by verbs.

OED's examples of word forms created based on the verb *to tea* - *teaing* and *teaer* - do not appear in my material.

Some of the instances that have gone into the categories 'meal OR verb' and 'meal OR mealtime OR verb' also have expressions with *to* but are not as clearly verbs as the two instances mentioned above. The fact that many of the instances of *to tea* follow verbs expressing motion strengthens the impression that *to* is, in these cases, a preposition.

- (14) We only met another Mr. Evelyn, his cousin, whose wife came to Tea... (Austen, Jane: letters)
 (15) James called on us just as we were going to tea... (ibid)

In example 14, it could be speculated that the capital letter in *Tea* could indicate that it is a noun. In example 15, *we were going to* could be an expression of intention, in which case *to tea* would be a verb, but it could also refer to a physical movement, in which case *to* would be a preposition and *tea* a meal. *Come to tea* is a phrase that is often sorted into this category, as it is quite impossible to decide whether *tea* functions as a verb or a noun, or in case it is a noun, whether it refers to a meal or mealtime. If *tea* were to be replaced by *dinner* or *dine* in these cases, there is no way to decide which to choose. This is also the case with example 16:

- (16) They sat down to tea in the bar... (Dickens, Charles, 1841: *Barnaby Rudge*).

Both *They sat down to dinner in the bar* and *They sat down to dine in the bar* work equally well.

In the same way as the above analysed examples, all of the following show the same kind of ambiguity as to whether the instances of the word *tea* are nouns or verbs, and in case they are nouns, whether they refer to the drink or the meal or mealtime:

- (17) ...when a servant entered and summoned the company to tea with the ladies (Fielding, Henry, 1751: *Amelia*)
 (18) A friend is to be with us to tea on my account (Burns, Robert, 1780-96: Letters)
 (19) And then to come home and find nobody in to tea! (Brontë, Anne, 1847: *Agnes Grey*)
 (20) When she comes to tea here she does not speak a word (Thackery, W.

M., 1847-48: *Vanity Fair*)

(21) Would real ladies have asked him to tea? (Forster, E.M., 1910: *Howards End*)

It is understandable that the meanings ‘drink’, ‘meal’ and ‘mealtime’ are difficult to distinguish. The meaning ‘produce’ might seem easier, however, and in many cases there is no problem in doing that: there are 21 cases of *tea* referring to the dry leaves or other forms of the produce before it has been made into the drink. Nevertheless, there are also cases that are unclear: first and foremost, there is a fine line between ‘produce’ and ‘drink’ in the cases where the brewing procedure is described. One case that could possibly go into this category, but has been sorted into the very last option, ‘unclear’, is example 22. Also, it can be hard to figure out in what state tea is sold (example 23): is it a restaurant-type establishment or more of a grocery store or does the woman sell drinks to be consumed there or to be brought home? In example 24, the same ambiguity as in example 23 prevails.

(22) ‘I insist on extracting tea before evening service’ (Forster, E. M., 1908: *A Room with a View*)

(23) a good old woman who sells wine, coffee, and tea there (Hazlitt, William, 1821-1822: *Table Talk*).

(24) the celebrated tea men, Forsyte and Treffry, whose tea, like no other man’s tea, had a romantic aroma. (Galsworthy, John, 1906: *The Man of Property*)

5.1.2 Semantic classification of *coffee*

Over all, the unclear cases of the word *coffee* are fewer than those of the word *tea*. In the early periods there are rather few cases of *coffee* as an independent word: in period 1 there are two cases referring to the produce and two referring to the drink. There is one slightly unclear case that refers to either the drink or the produce, even though ‘the produce’ is the more likely meaning:

(25) at ye E India house the Cheife Comodityes sett up were Coffee & Druggs. (Newdigate Newsletter: October 3, 1682)

Already in period 2, the majority of the independent cases of *coffee* refers to the drink. However, there is also a growing number of unclear cases. In period 2, there are cases where *coffee* either refers to a meal or a mealtime (example 26), the drink or a meal (example 27), and to the produce or the drink (example 28):

(26) After coffee I took my leave (T. G. Smollet, T. G., 1771: *The*

Expedition of Humphrey Clinker)

(27) Here is a fellow now (said I) falls fast asleep over his coffee
(*Westminster Magazine* (1776))

(28) lace, furs, jewels, furniture of all sorts, sugar, rum, tea, chocolate
and coffee (Smollet, T. G., 1771: *The Expedition of Humphrey Clinker*)

In example 28 the presence of things such as furs and jewels, on the one hand, seems to indicate that it is a list of imported products or such and thus, it would seem logical that the tea, chocolate and coffee would be in their raw state; on the other hand, however, rum is also a drink, and so could the other three be.

In period 3, there are a couple of interesting cases:

(29) the single meal which I could venture to order was coffee or tea. (De Quincey, Thomas, 1822: *Confessions of an English Opium-Eater*)

(30) like a magician disguised in a Welsh wig and a suit of coffee colour
(Dickens, Charles, 1848: *Dombey and Son*)

In example 29, the author has clearly marked that *coffee* can be used to denote a whole meal, even though it remains unclear exactly what that meal could be expected to include. In example 30, it could be argued that *coffee* could be interpreted as an adjective. In any case *coffee* specifies the colour of the suit, but it is unclear whether it refers to the colour of the produce or the drink. Perhaps that is not even important, the colour of both is about the same. *Coffee colour* could perhaps even be treated as compound, like the adjectives *chocolate-coloured* to be mentioned later.

The other unclear cases both in period 3 and period 4 follow along the same lines as those in the previous periods and a closer look at more of them would be repetitious. However, it is noteworthy to point out that according to the OED, *coffee* has had the option to mean 'to drink coffee' or 'to entertain at coffee', but there are no instances of this in my material.

5.1.3 Semantic classification of *chocolate*

In periods 1 and 2, the meanings of *chocolate* are restricted to 'produce' and 'drink', for the simple reason that it is not until the later years of period 3 that the technical inventions required for the production of solid chocolate sweets had taken place. That said, it is not always evident from the lexical surroundings that *chocolate* cannot refer to a solid sweet. For instance, from a modern day perspective, example 31 could very well refer to solid sweets. However, in quite a few of the cases, *chocolate* co-occurs with the verb *to drink* and those

cases are quite clear.

(31) I should be very glad you would bring some chocolate along with you (Hatton, 1677)

In period 3 the cases that refer to ‘the drink’ are in a vast majority, since there are not even any compounds in that period. Two of the cases may refer to solid sweets:

(32) they regaled me with chocolate and sweet-meats (Borrow, George, 1842: *The Bible in Spain*)

(33) they prefer their chocolate and biscuits (ibid)

However, the surrounding cases of *chocolate* in the same text refer to the drink. Furthermore, the text is about Spain and the invention of solid chocolate sweets may not have reached that part of Europe so early, even though chocolate originally came to Europe through Spain.

In period 4 there are four cases that have to be solid sweets for different reasons. In the case of example 34, it would seem unlikely (and if not that, at least rather rude) for someone to throw either cocoa powder or liquid chocolate to children. In the case of example 35, the word *chocolate* in the plural, which also gives it away as referring to a solid sweet.

(34) We threw pennies & chocolate down to the (bambina) (Dowson, Ernest, 1899: letters).

(35) and they were found drugged with cigarettes or chocolates and their whole property missing (Chesterton, G. K., 1914: *The Wisdom of Father Brown*)

In period 4, cases that refer to a colour start to appear, either as adjectives in their own right (examples 36 and 37), or as descriptions of colour using the *of*-genitive, cases which have been sorted into the category “produce”. Finally, in period 4, there is an increase in unclear cases, like example 40, which is a case that may seem strange to modern-day readers in how it presupposes that chocolate is indeed something nourishing. Bearing in mind what was said about the development and marketing of chocolate in the early 1900, however, it still remains unclear whether this case of *chocolate* refers to a solid sweet or a drink.

(36) Its gloomy chocolate, picked out with light green, took his fancy (Galsworthy, John, 1906: *The Man of Property*)

(37) I will see if we can match the colour (dark chocolate) (Grossmith, 1894: *The Diary of Nobody*).

(38) It was the colour of chocolate (Bennett, Arnold, 1908: *The Old Wives' Tale*)

(39) very much the appearance and colour of chocolate (Huxley, Aldous, 1894: *Discourses*).

(40) Let me send you up some more chocolate. I'm sure you're hungry
(Bennett, Arnold, 1902: *The Grand Babylon Hotel*)

5.2. Frequencies of *tea*, *coffee*, and *chocolate* in the corpus

The overall frequencies of the words *tea*, *coffee*, and *chocolate* balanced with the size of the corpus are displayed in table 3.

Table 3 shows that the relative frequency of *coffee* is the most stable throughout the material; it is only in period 2 that *coffee* has a lower share compared to the other periods: only 27 occurrences per million words compared to between roughly 35 and 40 occurrences per million words. *Chocolate* has a much lower frequency over all, but grows steadily over time. *Tea*, finally, is almost non-existent in period 1, overtakes *coffee* in period 2 only because *coffee* is less frequent in that period, to eventually be by far the most common of the three in periods 3 and 4.

Table 3: Frequency of *tea*, *coffee*, and *chocolate* (N), over time, occurrences per million word; balanced with size of corpus

Period	millions of words	Tea		Coffee		Chocolate		All words	
		N	frequency	N	frequency	N	frequency	N	frequency
Period 1	1.927	2	1.04	65	33.73	5	2.59	72	37.36
Period 2	2.441	74	30.32	65	26.63	11	4.51	150	61.45
Period 3	3.75	405	108.00	147	39.20	22	5.87	574	153.07
Period 4	4.082	606	148.46	163	39.93	28	6.86	797	195.25
All	12.2	1087	89.10	440	36.07	66	5.41	1593	130.57

Table 4 shows the proportional growth of *tea* and decline of *coffee* and *chocolate*, compared to each other, over time in the material. While *tea* only occurs twice in the first period, in the second it suddenly makes up almost half of the cases of these three words of that period. It grows to 70 per cent in the third period, to finally reach more than 75 per cent of all the cases in period 4. *Coffee*, on the other hand, declines from over 90 per cent in the first period to only 20 in the fourth. *Chocolate*, finally, has a fairly low share in throughout the time covered in the study, but nonetheless it declines from about 8 per cent to 3 per cent.

Most of these figures correspond to the knowledge about when and how tea, coffee, and chocolate were brought to England and how their popularity spread (cf. 2.3). However, the numbers may seem somewhat extreme for period 1 especially, and that is probably due to the selection of texts in the various corpora of that period. We know that tea was introduced

in England during that time, and that Samuel Pepys wrote about it in his diary, but the word has not happened to make it into many of the texts that were picked out for any of the corpora used in this study.

Table 4: Frequency of tea, coffee, and chocolate as independent words and compounds within periods

	TEA		COFFEE		CHOCOLATE		all words	
	N	%	N	%	N	%	N	%
Period 1								
independent	2	18.18%	5	45.45%	4	36.36%	11	100.00%
compounds	0	0.00%	60	98.36%	1	1.64%	61	100.00%
all	2	2.78%	65	90.28%	5	6.94%	72	100.00%
Period 2								
independent	58	69.88%	17	20.48%	8	9.64%	83	100.00%
compounds	16	23.88%	48	71.64%	3	4.48%	67	100.00%
all	74	49.33%	65	43.33%	11	7.33%	150	100.00%
Period 3								
independent	286	73.15%	83	21.23%	22	5.63%	391	100.00%
compounds	119	65.03%	64	34.97%	0	0.00%	183	100.00%
all	405	70.56%	147	25.61%	22	3.83%	574	100.00%
Period 4								
independent	434	74.06%	131	22.35%	21	3.58%	586	100.00%
compounds	172	81.52%	32	15.17%	7	3.32%	211	100.00%
all	606	76.04%	163	20.45%	28	3.51%	797	100.00%
Total								
independent	780	72.83%	236	22.04%	55	5.14%	1071	100.00%
compounds	307	58.81%	204	39.08%	11	2.11%	522	100.00%
all	1087	68.24%	440	27.62%	66	4.14%	1593	100.00%

a	d					
b	e					
c	f					
a	total number of independent words in each period					
b	total number of compounds in each period					
c	all cases of each word in each period					
d	share of independent words within each period, in per cent					
e	share of compounds within each period, in per cent					
f	share of all cases of each word within each period, in per cent					

It is also worth pointing out that the actual numbers of occurrences of each word rise in each new period even though the proportional shares go down. This concerns *coffee* and *chocolate* especially: the number of cases of *chocolate* doubles from period 1 to period 2 and from period 2 to period 3. Almost the same concerns *coffee*: the number of cases rises slightly

from period 1 to period 2 and more than double from period 2 to period 3. From period 3 to period 4 both seem to stabilize somewhat. *Tea*, on the other hand, has sprung from almost non-existence to a number that almost equals *coffee* and *chocolate* together in period 2, multiplied in period 3, and grown further, but not as dramatically as earlier, in period 4.

Table 5, finally, shows how the occurrences of the words independently in each period compare to each other. While we already know that the numbers of occurrences for all the words rise over time, in table 5 we can see the exact proportions of the findings of each word within each period related to the total numbers of occurrences of each word.

Table 5: Frequency of tea, coffee, and chocolate as independent words and compounds over time

	Period 1		Period 2		Period 3		Period 4		Total	
	N	%	N	%	N	%			N	%
TEA										
independent	2	0.26%	58	7.44%	286	36.67%	434	55.64%	780	100.00%
compounds	0	0.00%	16	5.21%	119	38.76%	172	56.03%	307	100.00%
all	2	0.18%	74	6.81%	405	37.26%	606	55.75%	1087	100.00%
COFFEE										
independent	5	2.12%	17	7.20%	83	35.17%	131	55.51%	236	100.00%
compounds	60	29.41%	48	23.53%	64	31.37%	32	15.69%	204	100.00%
all	65	14.77%	65	14.77%	147	33.41%	163	37.05%	440	100.00%
CHOCOLATE										
independent	4	7.27%	8	14.55%	22	40.00%	21	38.18%	55	100.00%
compounds	1	9.09%	3	27.27%	0	0.00%	7	63.64%	11	100.00%
all	5	7.58%	11	16.67%	22	33.33%	28	42.42%	66	100.00%
all words										
independent	11	1.03%	83	7.75%	391	36.51%	586	54.72%	1071	100.00%
compounds	61	11.69%	67	12.84%	183	35.06%	211	40.42%	522	100.00%
all	72	4.52%	150	9.42%	574	36.03%	797	50.03%	1593	100.00%

a	d								
b	e								
c	f								
a	number of independent words								
b	number of compounds								
c	total number of the word, including compounds								
d	share of independent words of all occurrences in each period, in per cent								
e	share of compounds of all occurrences in each period, in per cent								
f	share of all words of all occurrences in each period, in per cent								

More than half of all cases of *tea* are found in period 4, while barely a tenth of a percent is found in period 1. *Chocolate* is a little more evenly spread, with 40 percent in period 4 and 8 in period 1. *Coffee*, again, is found much more evenly across the whole period, even though like with *tea* and *chocolate*, the highest share is in period 4 (37 percent).

However, as much as around 15 percent of all the instances of *coffee* are found already in period 1 and period 2 each, and this is the clearest difference from *tea* and *chocolate*. In period 2 *coffee* and *chocolate* have almost the same share of all the occurrences of these words, while in period 3, all the three words have a fairly similar share.

The transferral of one text from period 2 to period 1 resulted in the doubling of the number of occurrences of the word *tea* in period 1. It is perhaps not surprising that such an occurrence is found in a text written by a man who seems to have been more or less a native to what is modern Hanoi. Seeing as it is such an early occurrence, it is only natural that it was written outside of England. The man himself seems to have been elusive and I have not been able to find years of either birth or death, but Winterbottom's (2013) primary sources seem trustworthy on the fact that the manuscript existed in 1686.

5.3. Meanings of *tea*, *coffee* and *chocolate*

The following sections present the occurrences and frequencies of the meanings of *tea*, *coffee* and *chocolate*. The tables describing the more detailed distribution of the different meanings of each of the words, as well as the distribution of the groups of compounds are found in the appendices 1.1, 1.2, and 1.3, because of the size of these tables. Appendices 2.1, 2.2, and 2.3 present the occurrences of each of the meanings as well as the occurrences of each of the compounds, and can be useful as references when reading the sub-sections especially on the compounds based on each of the words *tea*, *coffee*, and *chocolate*.

5.3.1 *Tea*

As can be seen in appendix 2.1, the most common meaning of the word *tea* is, unsurprisingly, the drink itself: 42 per cent of the independent cases of *tea* in the whole corpus, and 30 per cent of the independent and compound words combined. Taking into account all the cases where the word could refer to 'the drink' but could also refer to 'meal', 'mealtime', or be verb, in different constellations, the number is of course even higher than what is shown in the table - up to 75 per cent of the independent words and 54 per cent of all the cases. The words with the only likely meaning 'the drink', if the meanings of the morpheme *tea* in the compounds are also taken into account, make out 53 per cent, and all the words that could mean 'the drink', including the morpheme *tea* of the compounds, make out 77 per cent.

Noteworthy is that the share of the instances with the only likely meaning 'the drink'

decline over time, from 50 per cent in period 1 (which is a number that one must not pay too much attention to since there are only two cases of *tea* in period 1) to 35 per cent in period 4. What increases the most at the cost of clear cases of ‘the drink’ is the category ‘meal OR drink’. On the other hand, most of the morphemes *tea* in the compounds have a fairly clear meaning and most of the mixed categories are unchanged when the meanings of the morphemes are added to the meanings of the independent words. As for the compounds, by far the most of them are some kind of utensils for preparing or serving tea. About a fourth of the total number of compounds based on *tea* refer to meals or events.

Over all, the tables show an increase over time in variety both where independent word and compound word frequencies are concerned.

5.3.1.1 Compounds based on *tea*

Not only is *tea* the most frequently used out of the three words dealt with in this study, it is also the one with the widest range of compounds. Even though there are no compounds in period 1, which corresponds well with the fact that there are only two cases of *tea* at all in period 1, the wide range is present already in period 2.

The compounds referring to utensils used for preparing and serving tea are one of the more numerous groups: in period 2, there are four different words (*tea-things*, *tea-kettle*, *tea-equipage*, and *tea-cup*) used in total eight cases. All of these are still present in period 3, used altogether 20 times, but in period 3 there are also six new compounds referring to utensils (*tea-ladle*, *tea-tray*, *tea-pot*, *tea-spoon*, *tea-board*, and *tea-urn*); all compounds referring to utensils are used 76 times in period 3. The most frequent are *tea-pot* (19 occurrences) and *tea-spoon* (14 occurrences). In period 4, three of the compounds from period 2 are still present (*tea-things*, *tea-kettle*, and *tea-cup*), as are four of the ones that were new to period 3 (*tea-tray*, *tea-pot*, *tea-spoon*, and *tea-urn*), and finally, there are four compounds new to period 4 (*tea-cozy*, *tea-set*, *tea-bell*, and *tea-service*). In period 4, the most frequently used one is one that was present already in period 2: *tea-cup* (25 occurrences).

The category ‘storage’ could perhaps be included in the category ‘utensils’, but if they were, we would not automatically notice that compounds referring to the storage of tea do not appear until period 3: *tea-caddy* is the most frequent, with five occurrences, of the four compounds (the other three being *tea-chest*, *tea-canister*, and *tea-basket*) in this category. Overall it is a much smaller category than ‘utensils’, with a total of only ten occurrences.

There is a fairly even spread of compounds related to plants. Three of them refer directly to specific parts of the actual tea-plant: *tea-seed*, and *tea-tree*, occurring once each in

period 2, and *tea-leaf* (most often in the plural: *tea-leaves*), occurring once in period 3 and three times in period 4. Furthermore there are seven compounds referring to various kinds of either actual teas or other herbs used the same way as tea, such as, on the one hand *senna-tea* and *gunpowder-tea*, and on the other, *dandelion tea* and *cowslip tea*. These are most frequent in period 3.

Compounds with *tea* do not as frequently refer to locations as those with *coffee* do (see chapter 5.3.2.1). Nevertheless, there are six compounds referring to locations, most of them appearing in periods 3 and 4: *tea-room*, *tea-table*, *tea-parlour*, *tea-garden*, *tea-house*, and *tea-shop*. *Tea-house* only occurs once, compared to the 125 times *coffee-house* appears. *Tea-room* is used the most, but still it only occurs three times (to compare with a total of 20 occurrences of *coffee-room*). *Tea-table* is a word that I had originally categorised as a utensil, as it refers to a piece of furniture you serve tea on, but in this study I chose to concentrate on the fact that it is also something you sit at when you have your tea.

Apart from ‘utensils’, the largest group of compounds with *tea* is ‘meals or events’, although none of them appear in periods 1 or 2. Even in period 3 there are only five different compounds in this category: *tea-time*, *tea-party*, *tea-hour*, *tea-drinking*, and *rosemary-tea breakfast*. The last one is a very specific one and is used only once in the corpus. In period 4, however, the number of different compounds in this category has risen to 15 (in addition to *tea-time* and *tea-party* from period 3: *tea-meeting*, *tea-business*, *high-tea*, *afternoon-tea*, *lounge-tea*, *five o’clock tea*, *meat-tea*, *beef-tea*, *book-tea*, *nursery-tea*, *musical tea*, *Friday-tea*, and *tea-drinking afternoon*). The ones that are most frequently used are, perhaps not surprisingly, *tea-time* and *tea-party* (26 and 18 occurrences respectively). Other ones that are widely used in period 4 are *high-tea*, *afternoon-tea*, and *five o’clock tea* (5, 5, and 6 occurrences respectively).

In this category there are also two compounds that at first glance could belong in other categories: in the case of *tea-business*, it becomes clear from the context that the only occurrence in the whole material (example 41) does not necessarily refer to strict business but rather some kind of vague event related to tea. *Tea-drinking*, too, could easily mean a lot of different things and fill a lot of different grammatical functions, but three of them clearly refer to various kinds of social events, as proved by for instance the extract in example 42.

(41) ...he would promptly climb out of the boat, and sit on the bank, till the whole tea business was over. (Jerome, Jerome K., 1889: *Three Men in a Boat*)

(42) There was a great tea-drinking held in the Kirkgate of Irvine. (Galt,

John, 1821: *The Ayrshire Legatees*)

Finally, the ever increasing importance of tea in British society is shown by a number of compounds referring to social events that include tea: *book-tea* (example 43) and *musical tea*, for instance (example 44).

(43) Sawston, with its semi-detached houses and snobby schools, its book teas and bazaars, was certainly petty and dull... (Forster, E. M., 1905: *Where Angels Fear to Tread*)

(44) Mrs. Liversedge, with a sense of the fitness of things, had given a musical tea in his honour. (Galsworthy, John, 1904: *The Island Pharisees*)

The reason for the rather varied category ‘people, merchandise and transport’ is that it started out as two categories: ‘merchandise and transport’ and ‘people’, but many of the compounds referring to people refer to people that in fact deal with tea as merchandise, such as *tea-merchant*, and *tea-man* (*tea-men*) that refers to the founders of a company that imported tea. To this category belong also words such as *tea-drinker*, *tea-boy*, *tea-tax*, and *tea-clipper*.

There are a few proper nouns: one of them referring to a historical event, the Boston Tea Party, and two referring to specific locations: *Little Original Gold Tea Pot* is a shop or tea-room in Thackeray’s *Vanity Fair*, and *Montpelier Tea Gardens* is mentioned in Hazlit’s 1821-1822 *Table Talk*.

The last category, ‘other’, with 14 items, is perhaps unsurprisingly a group of compounds that do not have anything else in common but the fact that they do not fit into any other category. For instance, *tea-grounds*, which is not mentioned in the *OED*, but occurs once in a 1862 letter by J. R. Green where he is discussing a cake in example 45. There is also a verb phrase with an *ing*-participle to be found twice that is sorted into this category (example 46), cf. Example 42.

(45) opinions divided, Orthodox say "tea grounds," the Neologians "slate-pencil" nibbled while waiting for grub. (Green, J. R., 1862: letter)

(46) The ‘squire said, he would lay a tea-drinking on his head (T. G. Smollet 1771: *The Expedition of Humphrey Clinker*).

In the category ‘other’ there is also a group of five compounds based on *tea-table*: *tea-table conversation*, *tea-table chat*, *tea-table story* and *tea-table hospitality*. Similarly, we find *tea-sipping gossip*, *teacup time*, and *teacup entertainment*. The measurements *teacupful* and *teaspoonful* also come in into this category, and so do *tea-gown* and *tea-craving*.

5.3.2 *Coffee*

As can be seen in appendix 2.2, *coffee*, like *tea*, refers to a great part to the drink: 68 per cent of the independent cases of *coffee* refer clearly to ‘the drink’, and 37 per cent of all cases of *coffee* in the corpus. If the mixed categories are included, 92 per cent of the independent words refer to ‘the drink’, and the share is 42 per cent of the whole corpus. The unclear cases of *coffee* are not as numerous as those of *tea*. If the meanings of the morpheme *coffee* are counted together with the meanings of the independent words, as with *tea* in 4.1 above, it turns out that 76 per cent of all the cases of *coffee* refer to ‘the drink’. Furthermore, if the mixed categories are included, the share of cases referring to ‘the drink’ rises to 92 per cent. Even though many of the compounds as a whole refer to locations, the morpheme *coffee* refers to ‘the drink’, because the drink is per se the purpose of coffee houses and other such places.

There are a few further important differences from *tea*: firstly, there are many more instances of the word *coffee* in period 1 than there are of *tea*, and secondly, the shares of compounds are completely different. In the whole corpus, 74 per cent of the compounds (35 per cent of all the instances of *coffee*) refer to locations. In period 1 the proportions are as high as 82 per cent and 76 per cent respectively, and they decrease over time to 60 per cent and 12 per cent respectively in period 4.

5.3.2.1 Compounds based on *coffee*

The clearest trend among the compounds based on *coffee* is that in the earliest parts of the material, *coffee-house* is in an overwhelming majority, and over time, it gives way to a slightly greater variety of compounds. In period 1, as many as 82 per cent (49 cases) of all the compounds refer to locations, and one of those cases is *coffee-room*; the rest are *coffee-house*. There are a few mentions of people dealing with coffee, ranging from simply *coffee man* to the more specified *coffee housekeeper*. While the most frequent group of compounds for *tea* is utensils, there are less of those concerning coffee: in period 1, there is only one instance of a *coffee grate*. In period 2 there are four cases of *coffee-pot*, but *coffee-house* is still in majority.

In period 3 there is a wider variety of utensils used for preparing and serving coffee. In period 3 there are also three cases of *coffee-coloured* as well as a few more examples of more specific drinks involving coffee: already in period 2 there is *milk coffee*, and in period 3 we find the phrase *coffee and milk* as well as the compound *breakfast-coffee*. The

colour-aspect is the only one of these that appears also in period 4. As for locations, *coffee-shop* and *coffee-stall* make appearances only in period 4, as does *coffee stall-keeper*.

A compound in period 3 that must be treated as something of a special case is *coffee-house politician*, which appears in an essay by William Hazlitt: “Essay IV on Coffee-House Politicians” from the 1821-1822 collection “Table Talk: Essays on Men and Manners”. Even though there are four cases of *coffee-house politician* in the material and they make out a majority of the compounds based on *coffee* referring to people in period 3, it seems somewhat unlikely that it is a word that was very commonly used even at the time.

5.3.3 *Chocolate*

As we already know, *chocolate* is the word out the three in this study that is by far the least common. However, it is still possible to discern some trends, and it turns out that some of these trends are similar to those of the words *tea* and *coffee*. As can be seen in appendix 2.3, ‘the drink’ is by far the most common meaning of the independent words: 34 out of 55 cases of independent words have the clear meaning ‘the drink’.

That makes it 62 per cent of the cases of independent words, 52 per cent of all the cases, and 78 per cent and 65 per cent respectively if the mixed categories are added, that have the meaning ‘the drink’. If the meanings of the morpheme *chocolate* in the compounds are taken into account, the share of ‘the drink’ is 58 per cent, and 73 per cent if the mixed categories are added. All other categories of meaning have fairly equal, low shares, but ‘colour’ is one that grows notably when the meanings of the morphemes in the compounds are added: from around 5 per cent to around 12.

5.3.3.1 Compounds based on *chocolate*

When *chocolate* and compounds are concerned, there is not very much to say. In the first two periods, there are a total of four cases, all of which are the same: *chocolate-house*. In period 3 there are, as mentioned above, no compounds at all.

In period 4 there are seven compounds that either refer to colour (of skin, clothing, and interior design) or specific sweets or drinks. *Chocolate-coloured* occurs twice, and the other compounds referring to colour are *chocolate-clad*, *chocolate-toned*, and *chocolate-and-yellow*. Note that one of the cases sorted into the category ‘colour’ also contains *yellow*, but that is phrased *the one in the dining-room was yellow and chocolate* (Rutherford 1893: *Catherine Furze*). In that case, since there are no hyphens, the two colours

listed separately and independently from each other, while the case treated as a compound is treated as such because of the hyphens, even though hyphens are not a definitive and clear-cut way of distinguishing compounds from noun phrases. Finally, *the box of chocolate drops* (Collins 1861: *The Moonstone*) most likely refers to a specific kind of solid chocolate sweet, and so does the French phrase *Chocolat a la Vanille* (ibid 1859-1860: *The Woman in White*).

5.4. Discussion of frequencies and meanings

The over-all trend in the material of this study, as far as the independent cases go, is a widening of the meanings of the words *tea*, *coffee*, and *chocolate*. As speculated in chapter 2.2.2, in relation to Campbell's point about parallel semantic shifts, there are certain similarities in the developments of the three words, but perhaps not as much as expected. The differences can, to a certain degree, be explained by the fact that the numbers of occurrences of these words vary so much - there simply is not as much room for variety in the meanings of a word that occurs 66 times, as in the meanings of a word that occurs 1088 times.

Tea and *coffee* have the most in common with each other, since *chocolate* never acquired the meanings 'meal' or 'mealtime' like *tea* and *coffee* did. On the other hand, chocolate was developed into a solid sweet, which is not the case with tea or coffee. Metonymy is the most prominent semantic relationship between the different meanings in particular of *tea* and *coffee*, but to a certain degree also of the meanings of the word *chocolate*. As mentioned before, many of the meanings of the word *tea* could be placed along a continuum with 'the produce' at one end and 'mealtime' at the other, and the same applies to the word *coffee*, even though there are not as many cases of *coffee* as of *tea* that refer to meals or mealtimes. Similarly, it could be argued that a similar continuum could be applied to the word *chocolate*, but instead of 'meal' and 'mealtime' at the one end, these meanings would remain in the field of actual food-stuffs and the meaning here would be 'solid sweet'.

As for metaphoric uses of the words in this study, it is only the word *chocolate* that is used on its own in a metaphoric sense, when it is used to describe a colour. The word *coffee* is also used to describe a colour, but not on its own - the phrase used in such cases is e.g. *coffee-coloured*. It is interesting to note that it is only the word *chocolate* that is used in a strictly metaphoric sense despite the fact that *chocolate* is by far the least widely used of the three words in the material. *Tea* does not refer to colour in the material of this study, but that does not, of course, eliminate the possibility of the word being used in that meaning. Still, the

absence of such a use may simply be explained by the fact that the colour of tea varies more than the colour of coffee or chocolate, and perhaps the colour or colours of tea are not as prominent in other areas of life.

The distribution of the compounds that occur in the material across parts of speech corresponds to the overall distribution of compounds across parts of speech, perhaps even a little more clearly so than could be expected based on what Bauer says about compounds (see chapter 2.2.3). Out of all the 115 different compound words in the material, only one is a verb (*tea-drinking* in the participle), and five are adjectives which are all related to the colour meaning of *coffee* and *chocolate*. The rest, 109 compounds or a total of 95 per cent, are nouns.

Most of the compounds in the material are endocentric, i.e., the first element modifies the second, the grammatical head. *Tea* and *coffee* act both as the first element (e.g. *tea-cup*, *tea-garden*, and *coffee-house*) and as the grammatical head in the compounds (*herb tea*, *afternoon-tea*, and *milk coffee*), even though they are far less frequent than the cases where they are first elements of the compounds - only 36 out of 115 different compounds have the element *tea* or *coffee* as grammatical heads. In the case of *chocolate*, there are no compounds in my material that have *chocolate* as the grammatical head.

The compounds based on *chocolate* differ from the ones based on *tea* and *coffee* in several ways. A majority of them are adjectives and even though most of these are also endocentric, the only compound that can be classified as copulative in the whole study is *chocolate-and-yellow* - that is, the colour or pattern described is equally combined of a chocolate-brown and a yellow colour. Finally, there is one that can be defined as an exocentric compound: *Little Original Gold Tea Pot* - it seems to be the proper name of a tea-house or a tea-room, and as such, it does not actually refer to any kind of pot at all.

6. Analysis part 2: Sociolinguistic and cultural aspects

Since this study is based on corpora that were not collected with the object of creating material that could be used for sociolinguistic studies, the representation of female authors was not something that the creators took into particular consideration. Therefore, while one part of the CLMET has a fair share of women writers, the other periods within the CLMET and the other corpora I have used do not. Furthermore, the proportions of female writers within the periods seem to have little connection to the actual numbers of female writers in Britain at the times that the corpora represent. One might perhaps have expected the number

of female writers to grow over time, both when looking at the actual number of female writers as well as their share of all writers, but this is not the case in my material.

Table 6 shows the gender representation in the largest single corpus. I chose to limit this to the CLMET for practical reasons; it is the corpus that has the best and easiest available information about each text, and thus I did not have to go through a lot of material to find out which texts were written by men and women in every corpus. The CLMET also has the advantage over the other corpora that none of the texts consist of parts written by either men or women, as do for instance the letter compilations in both CLMEP and PCEEC. This, of course, results in the lack of period 1 in this sample. This table shows the complete number of authors of both genders, as well as the number of authors who use either of the words.

Table 6: Genders in the CLMET			male	female	total
Period 2	total number of authors		14	1	15
			93.33%	6.67%	100.00%
	number of authors who use either of the words		9	0	9
			100.00%	0.00%	100.00%
Period 3	total number of authors		21	8	29
			72.41%	27.59%	100.00%
	number of authors who use either of the words		15	6	21
			71.43%	28.57%	100.00%
Period 4	total number of authors		27	1	28
			96.43%	3.57%	100.00%
	number of authors who use either of the words		26	1	27
			96.30%	3.70%	100.00%

As already mentioned, and seen clearly in Table 6, it is only period 3 which has a more notable presence of women authors. It is worth pointing out that of these eight women, six use either of the words, equalling 75 %, while the corresponding number for the male writers is 71 %. In period 2, however, the one woman in the corpus has used neither of the words, while in period 4, the one woman in the corpus has used one or several of the words.

Table 7 shows the gender representation in all of the corpora among the authors who have used either of the words. Here, I have added a separate column for the few texts where I have not managed to find out if the writer is a man or a woman, but this concerns only periods 1 and 2. Here we can see that the distribution in period 3 is almost the same in both CLMET only and the whole corpus, while periods 2 and 4 differ a little bit and,

of course, period 1 is included. Here, period 2 has one woman represented, and period 4 has a total of four women. This raises the share of women writers in period 4 with quite a lot even though also the number of male authors using either of the words has increased. This, in fact, justifies the inclusion of more corpora than the CLMET, even though the CLMET is by far the largest and also easiest to use.

Table 7: Genders in the whole corpus;
number of authors using either of the words

	male	female	unknown	total
Period 1	8 72.73%	2 18.18%	1 9.09%	11 100.00%
Period 2	15 78.95%	1 5.26%	3 15.79%	19 100.00%
Period 3	16 72.73%	6 27.27%		22 100.00%
Period 4	30 88.57%	4 11.43%		34 100.00%

Tables 8 to 10 concentrate on period 3 because, as seen above, the share of female writers in the other periods does not allow for any worthwhile comparisons to be done.

Table 8: Genders' use of genres in Period 3

Genres	Number		Total
	Men	Women	
Fiction	6	6	12
Fiction/Children	1	1	2
Fiction/Autobiography	1		1
Fiction/Parody	1		1
Fiction/Polemics		1	1
Prose	1		1
Essays	1		1
Letters	2	1	3
Letters/Travel		1	1
Travel	3		3
History	2		2
Biography	1		1
	19	10	29

Table 8 shows that fiction is by far the most popular genre for both men and women.

Out of the four subgenres within fiction, there are texts by both men and women in fiction for children. Aside from these, there are female authors behind letters and letters from travels, but not behind autobiography, general travel writing, history, essays and prose. That is, out of this small sample of individual texts, it seems that women wrote in a narrower set of genres than men, but were comparably productive in the fiction genre as men.

In tables 9 and 10 we move on to more detail about the genders' use of the words *tea*, *coffee*, and *chocolate* in period 3.

Table 9: Genders' share of occurrences of the words in period 3

	TEA		COFFEE		CHOCOLATE	
Men	202	49.88%	113	76.87%	12	54.55%
Women	203	50.12%	34	23.13%	10	45.45%
Total	405	100.00%	147	100.00%	22	100.00%

Table 9 confirms the hypothesis that women do not use the word *coffee* as much as men. As for *tea* and *chocolate*, the shares are fairly even, with women using the word *tea* a little more often and men using the word *chocolate* a little more often. Here, of course, that fact that the numbers of actual uses are 12 versus 10 limits the usefulness of these percentages.

Table 10: The words' share of use by each gender in period 3

	Men		Women		Total	
TEA	202	61.77%	203	82.19%	405	70.56%
COFFEE	113	34.56%	34	13.77%	147	25.61%
CHOCOLATE	12	3.67%	10	4.05%	22	3.83%
All words	327	100.00%	247	100.00%	574	100.00%

Table 10 shows the same thing as Table 9 but from the opposite perspective: how often each gender has used the words. Taking into account the large differences in total numbers of occurrences for each word, it is not surprising that both men and women have used the word *tea* most often, and yet, there is a clear difference in just how often: almost 62 per cent of the times the men have used either of these words, their choice was *tea*; for women, it is 82 per cent. Following the previous table, the difference in the use of the word *chocolate* here is not very big, but *coffee* is clearly what the men have used instead of *tea*. 34

per cent of the times, the men have used *coffee* instead of *tea* or *chocolate*, while the corresponding share for women is only close to 14 per cent.

The rest of this chapter is divided according to the same principle as chapter 5; that is, there is a sub-chapter for each of the words *tea*, *coffee*, and *chocolate*. This order was chosen because *tea* is the most prominent and *chocolate* the least prominent of the words. However, this results in a slight discrepancy in chronology, since, as we have already seen, there are very few cases of the word *tea* in period 1 while *coffee* is much more evenly spread out across the whole timespan of the thesis.

As seen in Table 8, there is a larger focus on fictional works in my corpus in period 3 – or at the least, fictional works are more often the texts where the word *tea* is found. There are, of course, also texts in the corpora that do not mention the word *tea*, or the samples of the texts that have been chosen for the corpora do not contain the word even though it might be mentioned somewhere else in the complete work. In period 2, only 34 per cent of the texts are fictional, while 61 per cent of the texts are fictional in period 3. In period 4, the share has risen to 69 per cent. The shares of the cases of *tea* in fictional texts are 74 per cent in period 2, 79 per cent in period 3, and 85 per cent in period 4.

6.1 Tea

The year 1784 when tax on tea was dramatically lowered, is almost exactly in the middle of the time span of this study. In fact, it is four years into the third period, which means that out of the texts in period 3, it is only William Beckford's *Dreams, Walking Thoughts, and Incidents* (1783) that was written and/or published before the great change in taxes on tea. Some of the letters of Robert Burns were written before that, but none of the letters in the material contain the word *tea* before 1784. This means that out of the 1087 instances of *tea* in the whole material, only 79 occur before 1784 and the rest, a staggering 1008 cases (93 per cent), occur after the tax change. By comparison, the corresponding distribution of *coffee* is 32 per cent before and 68 per cent after 1784. Even though the over-all numbers of occurrences are higher for all words after 1784, it would seem that the lowering of the taxes on tea that year had a profound impact on the spread of the drinking of tea, mirrored in the use of the word *tea* in texts.

As for the relative frequencies, some recalculations had to be made where some of the letters of Robert Burns as well as Beckford's text were moved to the first half of the material and the number of occurrences were balanced against the total number of words before and

after 1784. As seen in Tables 11 and 12, the greatest difference is actually in the numbers of *coffee* when the material is divided at 1784 compared to when the material is divided at 1780. This supports the hypothesis that 1784 really is the most important year when it comes to the use of tea and, by extension, the use of the word *tea*.

Table 11: Relative frequency of the words *tea*, *coffee*, and *chocolate* (N), before and after 1784, occurrences per million words; balanced with size of corpus and with comparison of periods 1 and 2

Period	millions of words	Tea		Coffee		Chocolate		All words	
		N	frequency	N	frequency	N	frequency	N	frequency
Before 1784	4.463	79	17.70	141	31.59	17	3.81	237	53.10
Periods 1 and 2	4.368	76	17.40	130	29.76	16	3.66	222	50.82
After 1784	7.737	1008	130.28	299	38.65	49	6.33	1356	175.26
Periods 3 and 4	7.832	1011	129.09	310	39.58	50	6.38	1371	175.05

Table 12: Frequency of *tea*, *coffee*, and *chocolate* (N), before and after 1784, and frequency over time

	TEA		COFFEE		CHOCOLATE		All words	
	N	%	N	%	N	%	N	%
Before 1784	79	33.33%	141	59.49%	17	7.17%	237	100.00%
		7.27%		32.05%		25.76%		14.88%
After 1784	1008	74.34%	299	22.05%	49	3.61%	1356	100.00%
		92.73%		67.95%		74.24%		85.12%
All time	1087	68.24%	440	27.62%	66	4.14%	1593	100.00%
		100.00%		100.00%		100.00%		100.00%

N	
a	
b	

N	number of occurrences
a	share of all cases of each word within the periods before and after 1784, in per cent
b	share of each word of all occurrences before and after 1784, in per cent

Chapter 6.1 is further divided into two parts where the first part discusses periods 1 and 2, and the second part periods 3 and 4. Because of the tax reduction and the consequences thereof in the distribution of the word *tea*, it is natural to divide the discussion there, because it enables a different approach to the material, that is, more author- and genre oriented one.

6.1.1 Periods 1 (1650-1709) and 2 (1710-1779)

In period 1 there are only two cases of *tea*. One is in a letter, probably written in Antwerp

(example 47), and the other in a travel account from the English East Indies (example 48). Both cases probably refer to the drink, even though especially example 47 is a bit unclear.

(47) have scarce bought any thinge for my selfe but an Indian Brewhouse for tee, which hath beene very good Black Lack worke (William Howard 1642: letter to his mother)

(48) But if they are invited by their superiors or equals, then they entertain them as they find occasion, either with tea or meat, not omitting betel, which is always the first and last part of the regale. (Baron, Samuel 1686: *A Description of the Kingdom of Tonqueen*)

By the middle of period 2, Horace Walpole seems to consider tea quite universal in England, even though this was up to fifty years before the lowering of the taxes (here, the discussion is actually about a possible further duty on tea). Example 49 is also interesting in that the first mention of the word *tea* is most likely to refer to the produce, because it seems more likely that an extra duty on tea would be carried out on the produce that is bought rather than the number of cups served in the homes, while the second mention of the word more likely refers to the drink or even the whole meal culture based on the drink.

(49) They have talked of a new duty on tea, to be paid by every housekeeper for all the persons in their families; but it will scarce be proposed. Tea is so universal, that it would make a greater clamour than a duty on wine. (Walpole, Horace. 1743: letter to Horace Mann)

Already in period 2, there are clear cases of the word *tea* referring to a wider range of concepts than just the drink or the dry tea leaves. In example 50, Goldsmith clearly refers to a meal by the word *tea*. Smollet (example 51), on the other hand, while talking about the drink called *tea*, points out other noteworthy aspects: he shows that tea could be drunk rather late in the evenings; that tea was already at that time supposed to help keep you awake; and there was also already a notion of good and bad quality tea. These are all aspects discussed by Fromer (2008). In regard to the first point, she points out that tea could be drunk at almost any time during day and night, even though most ritualised “tea tables” at breakfast or in the afternoon (between lunch and dinner, dinner sometimes being served at midnight), sometimes late afternoon preceding supper, sometimes between 5 and 8 p.m., and sometimes after dinner as an opportunity for men and women to socialise again after dinner and the men’s sojourn with e.g. tobacco and perhaps liquors (Fromer 2008: 14). While Fromer discusses the Victorian era, it is interesting to note that the tendency was there more than fifty years earlier.

(50) If the cakes at tea eat short and crisp, they were made by Olivia.

(Goldsmith, Oliver. 1766: *The Vicar of Wakefield*)

(51) (...) while the other half is drinking hot water, under the denomination of tea, till nine or ten o'clock at night, to keep them awake for the rest of the evening. (Smollet, Tobias. 1771: *The Adventures of Humphrey Clinker*)

There are authors in my material that draw upon the internationality of tea. Kinglake (example 52) mentions the sense of fellowship it creates with the people in its area of origin, and Southey (example 53) discusses a variety of foreign products and their arrival in Europe and England.

(52) The love of tea is a glad source of fellow-feeling between the Englishman and the Asiatic. (Kinglake, Alexander. 1844: *Eothen, or Traces of Travel Brought Home from the East*)

(53) And yet in that age you profited slowly by the commodities which the eastern and western parts of the world afforded. Gold, pearls, and spices were your first imports. For the honour of science and of humanity, medicinal plants were soon sought for. But two centuries elapsed before tea and potatoes—the most valuable products of the East and West--which have contributed far more to the general good than all their spices and gems and precious metals--came into common use; nor have they yet been generally adopted on the Continent, while tobacco found its way to Europe a hundred years earlier; and its filthy abuse, though here happily less than in former times, prevails everywhere. (Southey, Robert. 1829: *Sir Thomas Moore, or Colloquies on Society*.)

Compounds based on *tea* referring to locations are not very popular at any point during the period of this study, which is interesting seeing as coffee-houses are so prominent, especially in earlier texts. The variety of compounds based on *tea* referring to locations is wider than for *coffee*, but while 40 out of the 63 compounds based on *coffee* refer to a location, only 14 out of 119 compounds based on *tea* do. At the same time, there are 33 instances of *coffee-house* (and 7 of *coffee-room*), but only 1 of *tea-house*, 2 of *tea-room*, 1 of *tea-parlour*, 1 of *tea-garden*, and 9 of *tea-table*. It is interesting to note that the only *tea-house* referred to is found in a parody (example 54), and the only *tea-parlour* is described as “unheroic” (example 55), and is probably used metaphorically in one way or another, since Cleopatra and tea usually would not be combined.

(54) That John and Mrs. Bull, from ale and tea-houses, May shout Huzza for Punch's Apotheosis! (Smith, James and Horace Smith. 1812: *Rejected Addresses*)

(55) Hast not thou thyself perchance seen diademed Cleopatra, daughter of the Ptolemies, pleading, almost with bended knee, in unheroic tea-parlour, or dimlit retail-shop, to inflexible gross Burghal Dignitary, for

leave to reign and die; being dressed for it, and moneyless, with small children;--while suddenly Constables have shut the Thespian barn, and her Antony pleaded in vain? (Carlyle, Thomas. 1837: *The French Revolution*)

The fact that compounds based on *tea* that refer to locations are not as common as is the case with *coffee* may simply be explained by the fact that tea was to a much larger extent taken at home and not so much out of the home, except when you were invited to tea in other peoples' homes, and only the very rich had separate rooms for the consumption of tea in any of the word's senses.

6.1.2 Periods 3 (1780-1849) and 4 (1850-1920)

Period 3 is the part of the material that has the largest representation of female writers. Fromer analyses two of them in her book: Elizabeth Gaskell (*Mary Barton*, 1848), and Emily Brontë (*Wuthering Heights*, 1847). Among other texts, she uses these to illustrate her thesis about tea and its role in the English society of the Victorian era. To compare, I looked through some of the other texts in my material to see if there are any clear differences. Here I chose some of the texts that had the largest number of occurrences, and tried to pick an equal number of texts written by male and female authors. In all of these cases I limited the investigation to the concordance lines of about 28 words surrounding the main word, and counted the cases that do refer to women by use of personal names, nouns or feminine pronouns and those that do not. The categories here are not references to men and references to women, but references to women on the one hand and "other", that is, without clear references to women on the other hand.

In some cases, the whole line consists of what a character is saying, and it may be the words of a woman even though the mention of the character's gender is only mentioned outside the space of the concordance line and the character is referred to in the concordance line only by the first person pronoun *I*. In these cases I tried to look for clues outside the concordance lines, but if I could not find any clear signs I marked the cases as having no reference to women.

Table 13 gives an overview of these results of these investigations. It shows periods 3 and 4 above each other; female novelists in red and male novelists in black; female non-fiction in green and male non-fiction in blue. The texts are sorted according to year of publication.

Table 13: Gender representation in selected works, in chronological order: Periods 3 and 4

References to women in close proximity to the word *tea*

Author	Title	Year of publication	Number of references to women	Share of total	Number of cases without references to women	Share of total	Total number of cases
Austen, Jane*	Letters to her sister Cassandra	1796-1817	29	94%	2	6%	31
Austen, Jane	Letters to her sister Cassandra	1796-1817	19	61%	12	39%	31
Austen, Jane, average	Letters to her sister Cassandra	1796-1817	24	77%	7	23%	31
De Quincey, Thomas	<i>Confessions of an English Opium Eater</i>	1821	2	10%	19	90%	21
Brontë, Emily	<i>Wuthering Heights</i>	1847	22	73%	8	27%	30
Gaskell, Elizabeth	<i>Mary Barton</i>	1848	55	89%	7	11%	62
Dickens, Charles	<i>Dombey and Son</i>	1848	13	42%	18	58%	31
Thackeray, William	<i>Vanity Fair</i>	1848	29	78%	8	22%	37
Hughes, Thomas	<i>Tom Brown's Schooldays</i>	1857	7	19%	30	81%	37
Yonge, Charlotte	<i>The Clever Woman of the Family</i>	1865	10	100%	0	0%	10
Booth, William	<i>In Darkest England</i>	1890	2	13%	13	87%	15
Gosse, W. E.	<i>Father and Son</i>	1907	3	25%	9	75%	12
Forster, E. M.	<i>Howards End</i>	1910	28	74%	10	26%	38
Bell, Gertrude*	Letters	1917	12	92%	1	8%	13
Bell, Gertrude	Letters	1917	3	23%	10	77%	13
Bell, Gertrude, average	Letters	1917	7.5	58%	5.5	42%	13
							337

Key:

Fiction, female authors

3

Fiction, male authors

4

Non-fiction, female authors

2

Non-fiction, male authors

3

Period 3

6

Period 4

6

*Female authors' letters: personal pronouns referring to women, i.e., the author herself

Female author's letters: personal pronouns not included in references to women

In the case of the letters that were included in this sub-study, I have calculated the share in two ways: including the personal pronouns as referring to the woman writing the letters, or omitting them, only counting the cases where there are other references to women. There are also averages for both of the letter-writers.

To begin with a literary text also included in Fromer's study, my sample from *Wuthering Heights* includes 30 cases of *tea*, and out of these, 8 (27 per cent) do not directly refer to situations where women are present. My sample from *Mary Barton* is the one in period 3 in my material with the most cases of *tea*, a total of 62, out of which only 7 (11 per cent) appear without any direct reference to women.

There are 37 instances of *tea* that are found in the sample from William Thackeray's *Vanity Fair* (1848) and 38 instances in the sample from E. M. Forster's *Howards End* (1910). In *Vanity Fair* there are 8 cases (22 per cent) where there were no references to women. In *Howards End*, the corresponding numbers are 10 and 26 per cent. However, in one of the cases where no women are mentioned by name, noun or feminine pronoun, I know that the person speaking and using the first person pronoun *I* is in fact a woman. Since it does not show from the line included in the concordance, I have counted it as neutral, however.

In period 4, there are only two novels by female authors, only one of which contains a number of occurrences of the word *tea* that is worth mentioning: Charlotte Yonge's *The Clever Woman of the Family* (1865). Here, all 10 cases of *tea* are included in situations where women are present. In order to try to keep the number of texts equal, I also added Charles Dickens's *Dombey and Son* (1848) where 18 (58 per cent) of the cases of *tea* do not have any references to women in their close vicinity. An interesting point with the sample from *Dombey and Son* is that most of the female references occur where the word *tea* is on its own, and only two of them occur together with compounds.

The texts mentioned above are all novels; I did the same for non-fictional texts, including one on religion, which proposes basic social care for poor people: William Booth's *In Darkest England and the Way Out* (1890), and a memoir, W. E. Gosse's *Father and son, a Study of Two Temperaments* (1907). There are 15 cases of *tea* in the former, and in fact, only 2 of them refer directly to women. On the other hand, only 5 of the 15 cases refer directly to men, so the majority of the cases are really unclear on this point. In the latter, there are 12 cases of *tea*, and only 3 refer directly to women being present, so 9 (75 per cent) do not.

The letters were written by Jane Austen (1796-1817) and Gertrude Bell (1917) and my procedure in regard to these letters was explained above. In the case where personal pronouns are counted as referring to women, Austen has only 2 cases out of 31 (6 per cent) where women are not at all included, and the corresponding number for Bell is 1 out of 13 (8 per cent). If, however, the personal pronouns referring to Austen and Bell themselves are excluded, there are 12 cases (39 per cent) in Austen's letters and 10 (77 per cent) in Bell's that do not refer to women other than themselves.

Table 14 contains the same information as Table 13, but sorted according to share of references to women. Its purpose is to illustrate how clearly female authors also refer more often to women in their writings. Austen and Bell only fall lower down on the scale in the cases where I have not counted personal pronouns as references to women, which is arguably an unnecessary precaution since it is likely that they are in fact referring to themselves when they are using first person pronouns in their letters.

Table 14: Gender representation in selected works, sorted according to share of references to women

References to women in close proximity to the word *tea*

Author	Title	Year of publication	Number of references to women	Share of total	Number of cases without references to women	Share of total	Total number of cases
Yonge, Charlotte	<i>The Clever Woman of the Family</i>	1865	10	100%	0	0%	10
Austen, Jane*	Letters to her sister Cassandra	1796-1817	29	94%	2	6%	31
Bell, Gertrude*	Letters	1917	12	92%	1	8%	13
Gaskell, Elizabeth	<i>Mary Barton</i>	1848	55	89%	7	11%	62
Thackeray, William	<i>Vanity Fair</i>	1848	29	78%	8	22%	37
Forster, E. M.	<i>Howards End</i>	1910	28	74%	10	26%	38
Brontë, Emily	<i>Wuthering Heights</i>	1847	22	73%	8	27%	30
Austen, Jane	Letters to her sister Cassandra	1796-1817	19	61%	12	39%	31
Dickens, Charles	<i>Dombey and Son</i>	1848	13	42%	18	58%	31
Gosse, W. E.	<i>Father and Son</i>	1907	3	25%	9	75%	12
Bell, Gertrude	Letters	1917	3	23%	10	77%	13
Hughes, Thomas	<i>Tom Brown's Schooldays</i>	1857	7	19%	30	81%	37
Booth, William	<i>In Darkest England</i>	1890	2	13%	13	87%	15
De Quincey, Thomas	<i>Confessions of an English Opium Eater</i>	1821	2	10%	19	90%	21

381

Table 15 shows general comparison during these periods: in general, it seems as if female writers include women more often in relation to tea than male authors do. But women are also more often mentioned in relation to tea in fictional texts than in non-fictional texts. Finally, it seems as if period 3 was a golden era for women to be mentioned close to the word *tea* in texts, even though this certainly is related to the fact that there were more female authors in my material of period 3 than period 4.

Table 15: Comparisons in Periods 3 and 4

	Number of references to women	Share of total	Number of cases without references to women	Share of total	Total number of cases	Share of total
	Male vs Female authors					
Male	84	44%	107	56%	191	57%
Female	118.5	81%	27.5	19%	146	43%
	Fiction vs non-fiction				337	
Fiction	164	67%	81	33%	245	73%
Non-fiction	38.5	42%	53.5	58%	92	27%
	Period 3 vs Period 4				337	
Period 3	145	68%	67	32%	212	63%
Period 4	57.5	46%	67.5	54%	125	37%
					337	

6.2 Coffee

The gender- and genre-based comparisons I made regarding the word *tea* were deemed doable because of the overall fairly large number of instances the word *tea*, by both female and male authors as well, especially in periods 3 and 4. With *coffee* reaching only less than half the number of instances compared to *tea*, it is not possible to see any similar trends. However, there is still something to be said about *coffee*.

In the early periods, we see examples of the perceived health benefits of coffee, as well as a reference to the English coffeehouse-culture. In relation to the health benefits of coffee, Henry More wrote in a 1664 letter to a female friend suffering poor health a recommendation to try coffee (example 56). P. D. Chesterfield (1756) wrote a warning in a letter to his son about the fact that the coffeehouses, not even the English one, in Paris are not as reputable as the ones in, e.g., London (example 57).

(56) [a] Docteur here, though not of Physic but of Law, was telling me what good efficacy Coffy has of taking away of the headache. I know not whether your Ladship has as yett ever made the experiment. (More, Henry. 1644: Letter)

(57) I must insist upon your never going to what is called the English coffee-house at Paris, which is the resort of all the scrub English, and also of the fugitive and attainted Scotch and Irish; party quarrels and drunken squabbles are very frequent there; and I do not know a more degrading place in all Paris. Coffee-houses and taverns are by no means creditable at Paris. (Chesterfield, P. D. 1756: Letter)

In period 1, the vast majority of the instances of the word *coffee* are found in the Newdigate Newsletters. This shows quite clearly that this early, coffee was not yet a very common thing among the people - if we omit the number of occurrences in the Newdigate Newsletters, the frequency of the words *coffee* and *chocolate* is almost the same in period 1. Seeing as the topics of these newsletters were political, military and diplomatic matters, both domestic and international, it is a little impressive that they mention coffee and coffeehouses as many times as 58 in the years 1674 to 1692, which are the years that are included in the corpus. Most of the time the coffeehouses are not necessarily the main topic of discussion; rather, they are mentioned in relation to where something of importance has happened, as in examples 58 and 59:

(58) Mr Holborn was wounded soe yt being carryed into a Coffee house he dyed (London, December 29, 1681)

(59) Yesterday Capt Walcup of the Kings Regemt & one mr Dives being at A Coffee house A quarrel Arose & they fought the latter tis said is dead (London: March 27, 1686)

By period 2, a slight majority (23 cases of 66) of *coffee* are found in literary texts, but letters and periodicals are not far behind (18 and 15 cases respectively). The one author to use the word most often is Horace Walpole in his letters (11 cases), but Henry Fielding is not far behind (with 9 cases). Walpole discusses having coffee, a coffee-pot (example 60), and experiences different things in coffeehouses (example 61), while Fielding seems to centre quite a few of his scenes in coffeehouses (example 62).

(60) You must either let me pay for them, or I will demand my unfortunate coffee-pot again, which has put you upon ruining yourself (Walpole, Horace, in a letter to Sir Horace Mann: March 25, 1743)

(61) all the morn one makes parties in masque to the shop and coffee-houses, and all the evening to the operas and balls. (Walpole, Horace, in a letter to Richard West: February 27, 1740)

(62) The next day, early in the morning, the colonel came into to the coffee-house and sent for his friend, who lodged but at a little distance. (Fielding, Henry, 1751: *Amelia*)

In period 3, the distribution of the word *coffee* is fairly even across genres and gender.

Since the genre fiction is so much larger in number of texts, it is not surprising that 66 of 146 instances of *coffee* are found there; however, the word is present in almost all texts within the genre. Only four out of 22 individual authors in period 3 do not use the word *coffee* in the texts that are sampled in my corpus, and out of those four, two are women and two are men.

More than a third of the eight instances of *coffee* in Dickens's works refer to the colour brown (example 63). The same share of instances are *coffee-houses*, and only two actually refer directly to the drink itself.

(63) ... like a magician disguised in a Welsh wig and a suit of coffee colour ... (Dickens, Charles. 1848: *Dombey and Son*.)

William Hazlitt discusses what he refers to as *coffee-house politicians* in one of his essays. While this text is a bit later than the golden era of coffeehouses described in chapter 2.3.1, it still fits in the more satirical descriptions of coffeehouse-cultures that e.g. Markman (2006 (1 and 2)) refers to. Hazlitt compares men of real education with a coffee-house politician, whom he describes as uncertain and deliberately misleading, in example 61:

(64) Men of education and men of the world order this matter better. They know what they have to say on a subject, and come to the point at once. Your coffee-house politician balances between what he heard last and what he shall say next; and not seeing his way clearly, puts you off with circumstantial phrases, and tries to gain time for fear of making a false step. (Hazlitt, William. 1822: *Table Talk*.)

By period 4, the fiction genre vastly outnumbers the other genres in numbers of texts, and it is, thus, not surprising that 125 cases of 163 are found in the fictional texts. Here, however, there are also texts within the fiction genre which have many more instances of *coffee* than the others, for instance *Howards End* by Forster which has 19 in all, or George Gissing's *New Grub Street* with 14. Most of the other texts have less than 5 cases.

In period 4, there are only two cases of *coffeehouse*, and both are in texts about history in one way or another: a novel with a fair emphasis on history (example 65), and a biography (example 66). This shows that by this time, the coffeehouse-culture had indeed died out, or at least been replaced by something else. On the other hand, there are more cases of *coffee-room* than before, which might indicate that the middle class was richer by this time and had a habit of drinking coffee in their homes (example 67). In Gissing's text there are also examples of other locations for selling coffee than coffeehouses (examples 68 and 69), showing that coffee might have become something that people had begun to drink in a little more of a hurry or more in passing than in the 18th century.

- (65) It was during Anne's reign [...] the clever gentlemen, who were called the wits, used to meet and talk at coffeehouses, and read newspapers, and discuss plays and poems (Yonge, Charlotte Mary, 1865: *The Clever Woman of the Family*)
- (66) when a boy at school at Chester, used to break bounds that he might watch Handel smoking his pipe in the Exchange coffee house (Butler, Samuel, 1903: *The Way of All Flesh*)
- (67) ...who had put his wife to bed, and was sitting with her in the empty coffee-room (Forster, E. M., 1910: *Howards End*)
- (68) and presently were seated in one of the boxes of a small coffee-shop (Gissing, George, 1891: *New Grub Street*)
- (69) When he had reached Camden Town railway-station he was attracted by a coffee-stall (ibid)

In William Booth's *In Darkest England and the Way Out* (1890), *coffee* is used in more positive circumstances than *tea*. *Coffee* is used together with the words *much* and *many*, listed together with other items that people have in abundance, and even showing that someone has found work. *Tea*, on the other hand, is described to be the only thing people have eaten, or something that they have but lack the means to prepare a meal out of. Could it be that tea is considered to be the very basis of life, the last thing that people give up, but coffee is more of a sign of affluence? A comparison between examples 70-73 and examples 74-77 supports this idea.

- (70) So much coffee, so much bread, so much shelter
- (71) so many coffee beans passing through a coffee mill
- (72) We have coffee tins, meat tins, salmon tins, and tins ad nauseam.
- (73) He found work and now owns a coffee-stall in Billingsgate Market
- (74) A cup of tea and a bit of bread yesterday, and the same to-day, is all he has had.
- (75) Sometimes the neighbours take them in to tea. It is really a mystery how they live.
- (76) she had not eaten any food up to twelve o'clock, not even a cup of tea to drink.
- (77) bought a penn'orth each of tea and ditto of sugar (produces from the same pocket) but can't get any place to make the tea; was hoping to get into a lodging house where he could borrow a teapot, but he had no money.

The use of the word *coffee* in my material seems to reflect the background history in chapter 2.3.1 fairly well. It is used in situations where it is described as being good for the health, even though some writers have their doubts about the institutions called coffeehouses. Towards the end of the period of this study, the coffeehouses have grown out of fashion in favour for drinking coffee at home or more on the go, yet coffee remains more of a luxury

product than tea during this period.

6.3 Chocolate

Something that stands out when going through the concordances of the word *chocolate* is how much more often this word is listed with a lot of other closely related victuals and commodities. These may be related in that they are used in similar ways, or have similar colonial background. This is most obvious in the first two periods, especially in period 2, since period 1 only has five examples which makes it hard to draw any general conclusions.

Example 78 is from a private letter dating from period 1. Examples 79 to 81 are all found in non-fiction: debate, essays, and histories. The lists are different, and the purposes of the texts differ, but the common thread is that they all seem to refer to something that is not directly to do with the author but to the wealth or habits of someone else. This indicates that chocolate was still a very luxurious product for the most part of the 18th century, which is in agreement with what was said in chapter 2.3.3.

(78) and I told you I had sent [?] brother [?] of good chocolate and 6 rollittos of excellent tobacco. (Hatton)

(79) new members of this house, who do not, more than once a-day, drink tea, cofffee, or chocolate, or some other cooling and diluting infusion (Johnson, Samuel, 1740-1741: *Parliamentary Debates*)

(80) consumed in Hungary and Poland, and there is some demand there for sugar, chocolate. and tobacco, of that new quarter of the world. (Smith, Adam, 1766: *An Inquiry into the Nature and Causes of the Wealth of the Nations*)

(81) stuffs, silks, lace, furs, jewels, furniture of all sorts, sugar, sum, tea, chocolate and coffee; in a word, not only every mode of the extravagant luxury... (Smollet, Tobias George, 1752: *A Complete History of England*)

As for the healthiness and moral aspects of chocolate, opinions vary throughout the material. Aphra Behn is optimistic about the health benefits of chocolate (example 82), while Chesterfield warns that it can make you fat, although the connotations of fatness were a bit different than today (example 83). Gay seems to have some misgivings about the moral soundness of chocolate-houses as institutions (example 84).

(82) [I] counsel him every morning to go to the coffee-house and drink some chocolate; and above all things advise him to desist from tobacco and drying things, or any other things that are too cooling for the kidneys (Behn, Aphra. 1686?: *The Ten Pleasures of Marriage*).

(83) ... taking as little as you can of those things that would make you fat.

Drink no chocolate; take your coffee without cream... (Chesterfield, P. D. S. 1746: Letters)
 (84) Marybone and the chocolate-houses are his Undoing (Gay, John. 1728: *The Beggar's Opera*).

Moss and Badenoch (2009: 67, 72-73) point out that there is a split in how different kinds of chocolate were used and marketed towards men and women. Due to the fact that the solid chocolate was invented as late as in the 1820s, there are a lot of cases in my material where the word *chocolate* cannot refer to a solid sweet simply because the texts they occur in are too old. There are two cases in period 3 that otherwise even could be ambiguous (examples 85 and 86):

(85) Next morning, the Padre gave us chocolate in his apartment
 (Beckford, William, 1783: *Dreams, Waking Thoughts, and Incidents*)
 (86) Could you give me four penn'orth of chocolate? (Smith, James and Horace Smith, 1812: *Rejected Addresses*)

However, as both of these examples are from before the late 1820s, they have to refer to the drink chocolate and not solid chocolate sweets. Not even in period 4, are there any clear indication of the gender split mentioned: women as well as men are involved with both the drink and the solid sweets.

Most authors who use the word chocolate use it only once or twice per text, or more precisely, it only appears once or twice in the samples of the texts in the corpora. One of the few authors who makes more use of the word is Maria Edgeworth in her *The Parent's Assistant* from 1796. 10 out of 22 cases in period 3 are found in this text. Interestingly enough, in this case it is the man of the house, presumably the father of Lucy and not the girl herself, who drinks the chocolate in all ten cases:

(87) Lucy. Whilst you are drinking your chocolate, papa?
 Just. No, no, no—I never see anybody till I have done my chocolate,
 darling. (He tastes his chocolate.)

Edgeworth, Maria 1796: *The Parent's Assistant*

The only other author who uses the word *chocolate* to the same extent as Edgeworth is Arnold Bennett, but in his case, the cases are divided between two different texts: *The Grand Babylon Hotel* (five cases) and *The Old Wive's Tale* (four cases). In *The Grand Babylon Hotel* there are three cases where it is clear that the chocolate is a drink consumed at breakfast by a man and a woman (example 88), but some of the others are more ambiguous (examples 89 and 90), although example 89 is most likely to refer to the drink as well.

(88) Nella and the Prince were breakfasting [...] upon chocolate that had been [...] hastily brewed for them. ‘I never tasted such excellent chocolate,’ exclaimed the Prince. [...] the Hôtel Wellington is not celebrated for its chocolate. (Bennet, Arnold. 1902: *The Grand Babylon Hotel*)

(89) a Negro boy appeared with a tray of chocolate. Nella took it and threwed it [...] overboard. (ibid.)

(90) Let me send you up some more chocolate. I’m sure you’re hungry. (ibid)

The word *chocolate* appears comparatively seldom in the material, which in itself serves as support for the view that chocolate remained a luxury product to an even greater extent than coffee. Many of the instances of the word in my material are found in texts that are set outside of England, or in essays on society or economy, compared to the vast numbers of the word *tea* that are found in fiction and letters. The gender difference in the marketing strategies for the drink chocolate and the solid chocolate bar is not visible in my material, but it may be explained by timing: maybe the most insistent campaigns were relatively late, so the tendency has not made it to my material which ends in 1920?

7. Discussion and conclusion

In the words of Teubert (2004: 104): “What corpus linguists make out as the meaning of words, can, thus, never be more than an approximation”. This certainly is the case with this study. For instance, the *CLMET* has grown substantially since the version used for this study and could easily render other results than those presented here. This concerns both the main parts of the study - the word meanings and the prevalence of the words and their different meanings, and the sociolinguistic and cultural aspects.

Alongside the growing number of meanings of the words *tea*, *coffee*, and *chocolate*, is a levelling out of these meanings: some of them disappear in my material. That, of course, is not to say they would not be found in other texts from the same times, or even elsewhere in the full texts rather than the selected parts of the texts in this corpus, since absence of evidence does not equal evidence of absence. Over all, the most common meaning of *tea* and *coffee* especially - the words referring to the drinks - grows faster than the cases which are unclear or clearly mean something else.

The sociolinguistic comparisons I have made in chapter 6 are not without problems. On the one hand, the number of samples is often too small to provide trustworthy numbers to compare, but on the other hand, including more of the texts from period 4 would skew the

relation between male and female writers, as there are no more female authors in period 4 than the two I have included, apart from two further collections of letters where some of the cases of *tea* may be found in letters written by women, but the numbers of occurrences in these collections are fairly low to begin with. Still, some patterns can be deductible from these choices of texts.

One of Julie Fromer's main foci in her study is the gender perspective on the tea-table in the Victorian era. In my material, the Victorian era stretches over the latter parts of period 3 and the earlier parts of period 4. Fromer pointed out that women are very prevalent in tea-table scenes in Victorian novels and that men on their own get together over other substances. Combining this with her readings of other texts from the same time she found a fairly sound basis for claiming this on a more general level in society than just in novels. My very basic study of the samples from novels and other texts from the same period of time, as well as a few older ones and a few later ones, agree to at least some extent.

The shares of cases of *tea* that include women in their close vicinity in the texts and those that do not vary depending on how the texts are grouped together, but in slightly more cases than not, women are more often present than not. When comparing female and male authors, it is the female authors who by far more often include women and exclude men at the tea-scenes. When comparing fictional texts and non-fictional texts, women are more present in tea-related situations in fictional texts than in non-fictional. It also seems as if women are more present in the texts of period 3 than in period 4, which also goes together with the fact that there are so many more female authors in period 3, as well as the fact that women mostly wrote fiction and are only represented in one other genre beside that – letters.

Fromer's second topic of research was that of class, and her third topic that of patriotism and building the nation. While I have not had the chance to put very much effort into these topics with my material, I found one striking aspect. William Booth's *In the Darkest of England* shows how widespread the consumption of tea was in the late 19th century by listing everything that poor people had consumed in one day: it varies from something to eat with a cup of tea or "not even a cup of tea to drink". Booth's work is a report on the lives of the poor in England, but at the same time he shows how deeply ingrained the notion of tea as the foundation of a sheltered home life was. Gertrude Bell, on the other hand, shows the other end of the spectrum as she writes her letters from Iraq only some thirty years later. She takes her tea in her house or with generals, princes, French nuns. Fresh bread was "a rare luxury", but her teas could also include dates and oranges, and what is depicted here is Englishness in a very imperialist setting.

Maybe the fact that the word *tea* in English seems to have been associated with French *thé* (see Johnson 1785 (2)) – even though tea did more likely come to England via the Dutch and not the French – in combination with the French conception of tea salons is a reason for tea to be something to be wary about? A complementary search for the word *salon* in my material could be enlightening, even though I cannot recall finding it in relation to the word *tea* anywhere in the concordances. Nationalism in general, and a suspicion of all things foreign – specifically French on the one hand and Levantine/Oriental on the other – is reason enough for criticism and wariness towards all three substances, especially the consumption of large amounts thereof. It is indeed a bit ironic, however, that even though the OED lists *the* as a possible spelling variant of the word *tea*, it is one that cannot easily be found in WordSmith due to the fact any cases of *tea* spelled like that will be mixed up with a huge number of definite articles. Therein could lie the topic for an interdisciplinary study within corpus linguistics and computer science: how to find specific cases of a word that might exist but risks being mixed up with something infinitely more common.

Svensk sammanfattning

Drick te, kaffe eller choklad, eller någon annan svalkande dekokt!

En korpusstudie av orden tea, coffee och chocolate 1640-1920

1. Inledning och bakgrund

Denna avhandling är en korpusbaserad undersökning inom historisk lingvistik som analyserar betydelseförändringar hos och sociolingvistiska skillnader i användningen av de engelska orden *tea*, *coffee* och *chocolate* från det att de lånades in i engelska språket fram till 1920. Avhandlingen utgör en fördjupning av min kandidatavhandling (2014) och förstärker undersökningens plats i en historisk kontext genom att analysera skillnader i användningen hos kvinnor och män och förekomsten i olika textgenrer samt vilken roll te, kaffe och choklad spelade i det brittiska samhället under den berörda tiden.

Det handlar alltså om orden för tre olika substanser som härrör sig till samma område i livet, det vill säga mat och rekreation. Både orden och själva substanserna kom till Europa och England runt i stort sett samma tid från olika delar av världen, men de tre orden kan syfta på flera olika former av substanserna och därutöver också användas för måltider, tidpunkter, evenemang, färger, och som verb (framför allt *tea*). Min tidigare studie antydde en skillnad i användning hos män och kvinnor: kvinnor verkade oftare förknippas med te, medan män oftare verkade använda ordet *coffee*.

Denna studie beaktar fenomen från flera olika ämnesområden, och inget av dessa är viktigare än något annat. De grundläggande språkvetenskapliga fenomenen som beaktas är lånord (orsaker till att språk lånar ord från andra språk), betydelseförändring (utvidgning och inskränkning av ordens betydelse, samt olika processer som är relevanta för detta såsom metonymi, metafor och polysemi) och sammansatta ord.

Inom sociolingvistiken som i sig är ett relativt nytt område är den historiska sociolingvistiken ännu nyare, och ifrågasätts ofta eftersom den grundläggande utmaningen är att materialet är begränsat och skriftligt, samt att människorna vars språk studeras inte är tillgängliga för forskaren på samma sätt som i forskning som görs på nutida språkvarianter. En nutida forskare kan inte heller lita på sin intuition om vad som anses språkligt korrekt när materialet är historiskt (Nevalainen & Raumolin-Brunberg 2003: 3). Romaine (1982: 14-15) poängterar dock att skriftligt språk kan vara ett självständigt medium och inte enbart en representation av talspråk, och därmed är också det skriftliga språket från historisk tid värt att

studera ur ett sociologiskt perspektiv. Det viktigaste är att hålla i minnet vilka begränsningar som finns i historiskt språkmaterial: det skriftliga material som har överlevt genom århundraden är begränsat till sin mängd, och ofta handlar det om utgivna texter som ofta redigerats av någon annan än författaren. Dessutom skrevs texter av en liten andel av befolkningen, och även om situationen är bättre redan under den tidsepok jag undersöker i denna studie finns det tydliga brister även här.

Den historiska och kulturella bakgrunden som behandlas gäller teets, kaffets och chokladens geografiska ursprung och framför allt deras ankomst till Europa och England i samband med upptäcktsresor och kolonier. Alla tre drycker ansågs i början ha hälsofrämjande effekter men blev snabbt också en viktig del av det sociala livet. Stor vikt läggs vid den så kallade kaffehuskulturen som uppstod i England under slutet av 1600-talet och sägs ha spelat en stor roll i att sprida innovationer och nya idéer (Markman 2006 (1-4)). Samtidigt är det också här den största könsskillnaden finns: det var män som besökte kaffehus för att dricka kaffe och debattera, medan kvinnor höll sig hemma för att dricka te till en sådan grad att teet så småningom sågs som en av grundpelarna i det engelska hemmet och samhället (Fromer 2008).

2. Material och metod

Korpusen som använts består av följande sex historiska korpusar: *Parsed Corpus of Early English Correspondence* (PCEEC), *The Helsinki Corpus of Early Modern English* (HCME), *Newdigate Newsletters* (NEWD), *Century of Prose Corpus* (COPC), *The Corpus of Late Modern English Texts* (CLMET), och *The Corpus of Late Modern English Prose* (CLMEP). Hela korpusen omfattar 12,2 miljoner ord och är indelad i fyra tidsperioder om 70 år: period 1 1640-1709, period 2 1710-1779, period 3 1780-1849 och period 4 1850-1920. Tabellerna 1 och 2 i bilaga 5 ger en överblick över korpusarna och tidsperioderna. I hela korpusen finns texter från flera olika genrer: skönlitterära texter, brev, vetenskapliga texter och prosa.

När jag påbörjade den fördjupade studien gjorde jag några korrigeringar gällande tidsbestämmelsen av enskilda texter i korpusen, vilket innebär att alla fynd inte stämmer överens med dem i kandidatavhandlingen. På detaljnivå handlar det om små skillnader, men detta ledde till en noggrannare beräkning av antalet ord i hela korpusen, och därför skiljer sig även det totala antalet ord från den tidigare studien (12,185 miljoner ord).

Korpusstudier innebär, kortfattat, att med hjälp av datormjukvara analysera stora mängder text, i mitt fall genom att leta upp alla fall av tre specifika ord. Det kan sägas finnas

två typer av korpusundersökningar (Teubert 2004): korpusbaserad, där alla resultat valideras direkt av fynden i korpusen, och korpusdriven, där fynden tolkas i en större kontext. Denna studie är en kombination av båda.

3. Analys del 1: betydelser och förekomst

Tabell 3 i bilaga 5 ger en överblick över de tre ordens förekomst under tidsperioden 1640-1920. Det mest i ögonen fallande är den kraftiga tillväxten av ordet *tea*. Tabell 4 i bilaga 5 sammanfattar de betydelseskategorier som använts.

I allmänhet var inte var lätt att entydigt bestämma en specifik betydelse för varje fall. Många fall finns därför listade i grupper som kombinerar de olika möjliga betydelserna på olika sätt. Detta gäller alla tre ord, men speciellt ordet *tea*. I exempel 1 kan ordet *tea* syfta på drycken, en måltid, eller rent av vara ett verb ifall man tolkar *to* som en infinitivmarkör i stället för en preposition. Också ordet *coffee* har använts i betydelsen 'måltid', och även om det inte alltid är uppenbart att så är fallet innebär tydliga fall (som exempel 2) att tvetydiga fall med rätta kan tolkas som tvetydiga.

(1) The two sat down to tea in Bathsheba's parlour... (Hardy, Thomas, 1874: *Far from the Madding Crowd*)

'De två satte sig ner för att dricka te/ för måltiden te i Bathshebas salong...'

(2) the single meal which I could venture to order was coffee or tea. (De Quincey, Thomas, 1822: *Confessions of an English Opium-Eater*)

'den enda måltid jag kunde drista mig till att beställa var kaffe eller te.'

Den allmänna tendensen är föga överraskande att betydelsen hos orden *tea*, *coffee* och *chocolate* vidgats under perioden 1640-1920. Utvecklingen för de tre orden är i viss mån parallell, men den stora skillnaden i förekomsten av ord gör att en fullständig parallellitet inte ens kan förväntas. *Tea* och *coffee* har mest gemensamt: *chocolate* tar aldrig betydelsen måltid, medan chokladen å andra sidan utvecklades också till en fast sötsak. Av de processer som påverkar ords betydelseförändring är metonymi den vanligaste i denna studie, men ordet *chocolate* används också metaforiskt då det på egen hand hänvisar till en färg.

4. Analys del 2: sociolingvistiska och kulturella aspekter

Fördelning av manliga och kvinnliga författare är mycket ojämn i mitt material, eftersom korpusen inte sammanställdes med avsikten att möjliggöra sociolingvistiska studier. Det är endast i period 3 (1780-1849) som det finns en relativt hög andel kvinnliga författare, och då är andelen ändå under 28 procent. Därför är det denna period jag delvis har koncentrerat mig på i en del av analysen, exempelvis då det gäller vilken genre orden *tea*, *coffee* och *chocolate* förekommer i. I det lilla samplet texter i period 3 i mitt material har kvinnor i huvudsak skrivit skönlitteratur och brev, medan män också skrivit i andra genrer som historia, biografier, prosa och essäer. När det gäller skönlitteratur är dock fördelningen mellan män och kvinnor jämn: både män och kvinnor har sex stycken texter i detta sample.

Då det kommer till användningen av orden *tea*, *coffee* och *chocolate* bekräftas min hypotes om könsfördelning: den stora skillnaden är att *coffee* används betydligt oftare av män än av kvinnor som å sin sida använder *tea* och *chocolate* lite oftare än män. De exakta förhållandena framgår ur tabeller 5 och 6 i bilaga 5.

I period 1 och 2 framgår det ur texterna som nämner ordet *tea*, trots betydligt lägre antal fynd, att te redan då ansågs vara en central del av det engelska samhället dels genom kommentarer om att te kan konsumeras nästan vilken tid på dygnet som helst och dels genom att använda teet som något som skapar gemenskap inom landet men också mellan England/Storbritannien och folken i teets ursprungsländer. Den låga förekomsten av sammansatta ord med *tea* som syftar på platser är dock iögonenfallande med tanke på hur ofta *coffee-house* förekommer. Det finns bara ett fall av *tea-house* och ett fall av *tea-parlour* och de verkar snarast användas i parodiskt syfte, vilket också stärker hypotesen om att kaffe faktiskt var något som konsumerades utanför hemmet medan te konsumerades hemma.

Period 3 och 4 lämpar sig bättre för sociolingvistiska undersökningar och därför har jag med inspiration av Julie Fromers *A Necessary Luxury* gjort en sammanställning över några utvalda texter ur dessa perioder för att undersöka förekomsten av ordet *tea* i närheten av referenser till kvinnor. Jag valde bland de texter som hade flest fall av ordet *tea*, men eftersträvade samtidigt en någorlunda jämn fördelning av texter av män och kvinnor samt skönlitterära och övriga texter. Jag analyserade sedan varje fall av ordet *tea* för att hitta referenser till kvinnor eller avsaknad av tydlig referens till kvinnor i samband med dessa. I vissa fall är det en kvinnlig karaktär i en roman som använder ordet *tea* eller på något sysslar med något te-relaterat, eller i de brev som är skrivna av kvinnor räcker det också att skribenten använt det personliga pronomenet *I* i närheten av *tea*.

Det finns en tydlig sammanställning av detta i tabell 7 i bilaga 5, men den övergripande trenden är att kvinnliga författare och skribenter i högre utsträckning tydligt kopplar ihop kvinnor och te. Också två manliga författares textsampel innehåller dock en hög andel kopplingar: i Thackerarys *Vanity Fair* finns 78 % av fallen av *tea* i närheten av referenser till kvinnor, och i Forsters *Howards End* är den motsvarande siffran 73 %. Iögonenfallande är i de fall andelen referenser till kvinnor är under 25 % är författaren manlig (undantaget brev av Gertrude Bell beroende på om man räknar det personliga pronomenet *I* som en referens till den kvinnliga skribenten eller inte). Det framkommer dessutom att det är vanligare med kopplingen mellan *tea* och referenser till kvinnor i skönlitterära texter än i andra typer av texter, samt att denna koppling är vanligare i period 3 än i period 4, vilket kan bero på att det i mitt material i allmänhet finns fler texter av kvinnor i period 3 än i period 4.

Eftersom antalet fall av orden *coffee* och *chocolate* är så mycket lägre än *tea* är det inte lönsamt att göra likadana analyser av dessa två ord som av *tea*, men några saker kan ändå nämnas. I början finns de flesta fallen av *coffee* i en samling personliga nyhetsbrev, vilket visar att kaffe inte var en allmän produkt för folket, men också hur centrala de s.k. kaffehusen var för viktiga händelser. Ännu i period 2 är förekomsten av *coffee-house* vanlig, men ordet *coffee* förekommer då också oftare i skönlitterära texter. I period 4 finns endast två fall av *coffee-house*, i texter som behandlar historiska teman, och i kombination med nya ord som *coffee-room* och *coffee-stall* tyder detta på att kaffehuskulturen ersatts dels av att man också drack kaffe i hemmet eller mer i förbifarten på väg någonstans.

Den intressantaste texten är William Booths samhällskritiska *In Darkest England and the Way Out* (1890) som genomgående använder kaffe för att belysa en positiv situation (exempel 3) och te för att exemplifiera hur dåligt folk har det (exempel 4). Denna användning antyder att te uppfattas som något grundläggande som man sent skulle ge upp medan kaffe är ett tecken på (relativ) rikedom.

(3) So much coffee, so much bread, so much shelter

‘Så mycket kaffe, så mycket bröd, så mycket skydd’

(4) A cup of tea and a bit of bread yesterday, and the same to-day, is all he has had

‘En kopp te och en bit bröd igår, och det samma idag, är allt han har haft att äta’

Samtidigt som Booth visar hur teet är en grundläggande stapelvara i det brittiska samhället finns i mitt material också exempel på hur teet placerar Storbritannien i en imperialistiskt sammanhang: Gertrude Bells brev från Irak från 1910-talet visar hur tekonsumtionen är en gemensam nämnare i alla olika sammanhang hon rör sig i, vare sig det är i sällskap av generaler, prinsar eller franska nunnor.

Chocolate förekommer minst av de tre orden, så det är ännu svårare att hitta tydliga tendenser i användningen, men det finns några genomgående fenomen. *Chocolate* används ofta i listor över olika saker som används på samma sätt eller har en liknande kolonial bakgrund, också tillsammans med *coffee* och *tea*. Ett tema som ofta hänvisas till är den påstådda hälsoaspekten hos choklad. I mina bakgrundsstudier stötte jag på påståendet att olika typer av choklad under 1800-talet marknadsfördes på olika sätt till män och kvinnor, men i mitt material finns inga belägg för en sådan könsfördelning, vilket kan bero på att så liten del av mitt material är från efter utvecklingen av choklad i fast form. Ordet används ofta i sammanhang utanför England eller i essäer om samhälle eller ekonomi och inte i lika hög grad i skönlitteratur som *tea*. Detta, samt den jämförelsevis låga förekomsten av *chocolate* stöder tesen att choklad förblev en lyxprodukt länge än både te och kaffe.

5. Sammanfattning

Man måste komma ihåg att texterna i min korpus inte är fullständiga verk utan i de flesta fall sampel ur längre verk. Orden förekommer inte i alla texter i korpusen, men ingenting utesluter att orden har förekommit i andra delar av de fullständiga verken som inte finns med i korpusen. Teubert (2004) poängterar också att den betydelse som en korpuslingvist lägger i ett ord bara kan vara ungefärlig, och eftersom åtminstone *CLMET* har förstorats avsevärt från den version som använts i denna studie skulle en ny undersökning kunna ge helt annorlunda resultat än de som presenterats här. Detta gäller såväl ordens förekomst och betydelse som de sociolingvistiska och kulturella aspekterna.

Man måste också komma ihåg att brist på bevis inte är det samma som bevis på att något inte finns. Även om mitt material antyder att ordens betydelser utjämnas mot slutet av den undersökta perioden samt att vissa av de sammansatta ord som förekom tidigt i materialet inte dyker upp senare betyder det inte att de helt försvunnit ur bruk. Vad man kan konstatera angående betydelsen av framför allt orden *tea* och *coffee* är att förekomsten den mest grundläggande betydelsen, dryckerna, växer snabbare än andra betydelser, i takt med att användningen överlag också ökar.

Appendices

Appendix 1.1: Distribution of *tea*

TEA meanings	Period 1			Period 2			Period 3			Period 4			All		
	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
produce				7	12.07%	9.46%	21	7.34%	5.19%	6	1.38%	0.99%	34	4.36%	3.13%
drink	1	50.00%	50.00%	31	53.45%	41.89%	140	48.95%	34.57%	153	35.25%	25.25%	325	41.67%	29.90%
meal				1	1.72%	1.35%	20	6.99%	4.94%	44	10.14%	7.26%	65	8.33%	5.98%
mealtime							4	1.40%	0.99%	3	0.69%	0.50%	7	0.90%	0.64%
meal OR mealtime				4	6.90%	5.41%	16	5.59%	3.95%	33	7.60%	5.45%	53	6.79%	4.88%
meal OR drink	1	50.00%	50.00%	10	17.24%	13.51%	57	19.93%	14.07%	132	30.41%	21.78%	200	25.64%	18.40%
verb										2	0.46%	0.33%	2	0.26%	0.18%
meal OR verb				1	1.72%	1.35%				28	6.45%	4.62%	29	3.72%	2.67%
meal OR drink OR verb				1	1.72%	1.35%	14	4.90%	3.46%	11	2.53%	1.82%	26	3.33%	2.39%
produce OR drink				3	5.17%	4.05%	4	1.40%	0.99%	8	1.84%	1.32%	15	1.92%	1.38%
meal OR mealtime OR drink							10	3.50%	2.47%	10	2.30%	1.65%	20	2.56%	1.84%
unclear										4	0.92%	0.66%	4	0.51%	0.37%
	2	100.00%	100.00%	58	100.00%	78.38%	286	100.00%	70.62%	434	100.00%	71.62%	780	100.00%	71.76%
TEA compounds	Period 1			Period 2			Period 3			Period 4			All		
	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
Utensils for preparation or serving				8	50.00%	10.81%	66	55.46%	16.30%	66	38.37%	10.89%	140	45.60%	12.88%
Storage							9	7.56%	2.22%	2	1.16%	0.33%	11	3.58%	1.01%
Related to plants				3	18.75%	4.05%	7	5.88%	1.73%	5	2.91%	0.83%	15	4.89%	1.38%
Locations				1	6.25%	1.35%	14	11.76%	3.46%	18	10.47%	2.97%	33	10.75%	3.04%
Meals or events							15	12.61%	3.70%	64	37.21%	10.56%	79	25.73%	7.27%
People, merchandise and transp							4	3.36%	0.99%	6	3.49%	0.99%	10	3.26%	0.92%
Other				3	18.75%	4.05%	2	1.68%	0.49%	11	6.40%	1.82%	16	5.21%	1.47%
Proper nouns				1	6.25%	1.35%	2	1.68%	0.49%				3	0.98%	0.28%
				16	100.00%	21.62%	119	100.00%	29.38%	172	100.00%	28.38%	307	100.00%	28.24%
	2			74		100.00%	405		100.00%	606		100.00%	1087		100.00%
meanings of independent words and compounds separately as well as together															

TEA	Period 1		Period 2		Period 3		Period 4		ALL	
meanings	N	%	N	%	N	%	N	%	N	%
produce			9	12.16%	33	8.15%	15	2.48%	57	5.24%
drink	1	50.00%	45	60.81%	244	60.25%	287	47.36%	577	53.08%
meal			2	2.70%	21	5.19%	69	11.39%	92	8.46%
mealtime					4	0.99%	5	0.83%	9	0.83%
meal OR mealtime			4	5.41%	16	3.95%	33	5.45%	53	4.88%
meal OR drink	1	50.00%	10	13.51%	57	14.07%	132	21.78%	200	18.40%
verb							2	0.33%	2	0.18%
meal OR verb							28	4.62%	28	2.58%
meal OR drink OR verb			1	1.35%	14	3.46%	11	1.82%	26	2.39%
produce OR drink			3	4.05%	5	1.23%	10	1.65%	18	1.66%
meal OR mealtime OR drink					10	2.47%	10	1.65%	20	1.84%
drink OR verb					1	0.25%			1	0.09%
unclear							4	0.66%	4	0.37%
	2	100.00%	74	100.00%	405	100.00%	606	100.00%	1087	100.00%
meanings including the meaning of the morpheme <i>tea</i> in the compounds										

Appendix 1.2: Distribution of *coffee*

COFFEE meanings	Period 1			Period 2			Period 3			Period 4			All		
	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
produce	2	40.00%	3.08%	2	11.76%	3.08%	7	8.43%	4.76%	6	4.58%	3.68%	17	7.20%	3.86%
drink	2	40.00%	3.08%	10	58.82%	15.38%	48	57.83%	32.65%	101	77.10%	61.96%	161	68.22%	36.59%
meal					0.00%	0.00%	1	1.20%	0.68%		0.00%	0.00%	1	0.42%	0.23%
meal OR mealtime				1	5.88%	1.54%	1	1.20%	0.68%		0.00%	0.00%	2	0.85%	0.45%
meal OR drink				2	11.76%	3.08%	13	15.66%	8.84%	16	12.21%	9.82%	31	13.14%	7.05%
produce OR drink	1	20.00%	1.54%	2	11.76%	3.08%	8	9.64%	5.44%	6	4.58%	3.68%	17	7.20%	3.86%
meal OR mealtime OR drink							5	6.02%	3.40%	2	1.53%	1.23%	7	2.97%	1.59%
	5	100.00%	7.69%	17	100.00%	26.15%	83	100.00%	56.46%	131	100.00%	80.37%	236	100.00%	53.64%
COFFEE compounds	Period 1			Period 2			Period 3			Period 4			All		
	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
Utensils for preparation or serving	1	1.67%	1.54%	4	8.33%	6.15%	8	12.50%	5.44%	9	28.13%	5.52%	22	10.78%	5.00%
Storage							1	1.56%	0.68%	1	3.13%	0.61%	2	0.98%	0.45%
Related to plants	1	1.67%	1.54%		0.00%	0.00%	2	3.13%	1.36%	1	3.13%	0.61%	4	1.96%	0.91%
Locations	49	81.67%	75.38%	42	87.50%	64.62%	41	64.06%	27.89%	19	59.38%	11.66%	151	74.02%	34.32%
Meals or events								0.00%	0.00%	1	3.13%	0.61%	1	0.49%	0.23%
People, merchandise and transp	8	13.33%	12.31%				6	9.38%	4.08%		0.00%	0.00%	14	6.86%	3.18%
Other	1	1.67%	1.54%	2	4.17%	3.08%	6	9.38%	4.08%	1	3.13%	0.61%	10	4.90%	2.27%
	60			48	100.00%	73.85%	64	100.00%	43.54%	32	100.00%	19.63%	204	100.00%	46.36%
	65			65		100.00%	147		100.00%	163		100.00%	440		100.00%
meanings of independent words and compounds separately as well as together															

COFFEE	Period 1		Period 2		Period 3		Period 4		ALL	
meanings	N	%	N	%	N	%	N	%	N	%
produce	4	6.15%	2	3.08%	16	10.88%	11	6.75%	33	7.50%
drink	53	81.54%	58	89.23%	100	68.03%	120	73.62%	331	75.23%
meal				0.00%	1	0.68%			1	0.23%
meal OR mealtime			1	1.54%	1	0.68%			2	0.45%
meal OR drink			2	3.08%	13	8.84%	16	9.82%	31	7.05%
produce OR drink	8	12.31%	2	3.08%	11	7.48%	14	8.59%	35	7.95%
meal OR mealtime OR drink					5	3.40%	2	1.23%	7	1.59%
	65	100.00%	65	100.00%	147	100.00%	163	100.00%	440	100.00%
meanings including the meaning of the morpheme <i>coffee</i> in the compounds										

Appendix 1.3: Distribution of *chocolate*

CHOCOLATE	Period 1			Period 2			Period 3			Period 4			All		
meanings	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
produce	3	75.00%	60.00%	1	12.50%	9.09%							4	7.27%	6.06%
drink	1	25.00%	20.00%	5	62.50%	45.45%	20	90.91%	90.91%	8	38.10%	28.57%	34	61.82%	51.52%
solid sweet										4	19.05%	14.29%	4	7.27%	6.06%
produce OR drink				2	25.00%	18.18%							2	3.64%	3.03%
drink OR solid sweet							2	9.09%	9.09%	3	14.29%	10.71%	5	9.09%	7.58%
colour										3	14.29%	10.71%	3	5.45%	4.55%
produce OR drink OR sweet										2	9.52%	7.14%	2	3.64%	3.03%
produce OR sweet										1	4.76%	3.57%	1	1.82%	1.52%
	4	100.00%	80.00%	8	100.00%	72.73%	22	100.00%	100.00%	21	100.00%	75.00%	55	100.00%	83.33%
CHOCOLATE	Period 1			Period 2			Period 3			Period 4			All		
compounds	N	%	%	N	%	%	N	%	%	N	%	%	N	%	%
Locations	1	100.00%	20.00%										1	9.09%	1.52%
Other				3	100.00%	27.27%				7	100.00%	25.00%	10	90.91%	15.15%
	1	100.00%	20.00%	3	100.00%	27.27%				7	100.00%	25.00%	11	100.00%	16.67%
	5	100.00%		11	100.00%		22	100.00%		28	100.00%		66	100.00%	
meanings of independent words and compounds separately as well as together															

CHOCOLATE	Period 1		Period 2		Period 3		Period 4		ALL	
meanings	N	%	N	%	N	%	N	%	N	%
produce	3	60.00%	1	9.09%					4	6.06%
drink	2	40.00%	8	72.73%	20	90.91%	8	28.57%	38	57.58%
solid sweet				0.00%			5	17.86%	5	7.58%
produce OR drink			2	18.18%					2	3.03%
drink OR solid sweet					2	9.09%	4	14.29%	6	9.09%
colour							8	28.57%	8	12.12%
produce OR drink OR sweet							2	7.14%	2	3.03%
produce OR sweet							1	3.57%	1	1.52%
	5	100.00%	11	100.00%	22	100.00%	28	100.00%	66	100.00%
meanings including the meaning of the morpheme <i>chocolate</i> in the compounds										

Appendix 2.1: Numbers of occurrences of *tea***TEA**

meanings	Period 1	Period 2	Period 3	Period 4	Total
produce		7	21	6	34
drink	1	31	140	153	325
meal		1	20	44	65
mealtime			4	3	7
meal OR mealtime		4	16	33	53
meal OR drink	1	10	57	132	200
verb				2	2
meal OR verb		1		28	29
meal OR drink OR verb		1	14	11	26
produce OR drink		3	4	8	15
meal OR mealtime OR drink			10	10	20
unclear				4	4
	2	58	286	434	780

TEA

compounds	Period 1	Period 2	Period 3	Period 4	Total
<i>Utensils for preparation or serving</i>					
tea-things		2	7	3	12
tea-kettle		3	5	5	13
tea-equipage		1	1		2
tea-cup		2	7	25	34
tea-ladle			1		1
tea-tray			11	4	15
tea-pot			19	14	33
tea-spoon			14	8	22
tea-board			1		1
tea-urn			1	2	3
tea-cosy (cozy)				2	2
tea-set				1	1
tea-bell				1	1
tea-service				1	1
		8	67	66	141

Storage

tea-caddy			5		5
tea-chest			1		1
tea-canister			2		2
tea-basket				2	2
			8	2	10

Related to plants

tea-seed		1			1
tea-tree		1			1
tea-leaf (leaves)			1	3	4
senna-tea		1			1
gunpowder-tea			2		2
Dandelion tea			1		1
herb tea			2		2
green tea			1		1
cowslip tea				1	1
hollyhock tea				1	1
		3	7	5	15

Locations

tea-room			2	1	3
tea-table		1	9	15	25

tea-parlour			1		1
tea-garden			1	1	2
tea-house			1		1
tea-shop				1	1
		1	14	18	33

Meals or events

tea-time			6	20	26
tea-party			2	16	18
tea-hour			2		2
tea-drinking			4		4
tea-meeting				1	1
tea-business				1	1
high-tea				5	5
afternoon-tea				5	5
lounge tea				1	1
five o'clock tea				6	6
meat-tea				3	3
beef-tea				1	1
book tea				1	1
rosemary tea breakfast			1		1
nursery tea				1	1
musical tea				1	1
Friday tea				1	1
tea-drinking afternoon				1	1
			15	64	79

People, merchandize and transport

tea-drinker			1		1
tea-maker			1		1
tea-boy			1		1
tea-merchant				2	2
tea-man (men)				1	1
tea-selling (noun)			1		1
tea-tax				1	1
tea-ship				1	1
tea-clipper				1	1
			4	6	10

Other

tea-table conversation		1			1
tea-table chat		1			1
tea-drinking (progressive verb)		1			1
tea-sipping gossip			1		1
tea-gown				1	1
tea-craving				1	1
teacupful				1	1
teacup entertainment				1	1
teacup time				1	1
morning teacup				1	1
teaspoonful			1	2	3
tea-grounds				1	1
tea-table story				1	1
tea-table hospitality				1	1
		3	2	11	16

Proper nouns

Boston Tea-Party		1			1
Little Original Gold Tea Pot			1		1
Montpelier Tea Gardens			1		1
		1	2		3

	16	119	172	307
2	74	405	606	1087

Appendix 2.2: Numbers of occurrences of *coffee*

COFFEE

meanings	Period 1	Period 2	Period 3	Period 4	Total
produce	2	2	7	6	17
the drink	2	10	48	101	161
meal			1		1
meal OR mealtime		1	1		2
drink OR meal		2	13	16	31
produce OR drink	1	2	8	6	17
drink OR meal OR mealtime			5	2	7
	5	17	83	131	236

COFFEE

compounds	Period 1	Period 2	Period 3	Period 4	Total
<i>Utensils for preparation or serving</i>					
coffee-pot		4	1	2	7
coffee-mill			2	2	4
coffee-bowl			1		1
coffee-things			1		1
coffee-cup			2	4	6
coffee-urn			1	1	2
coffee grate	1				1
	1	4	8	9	22

Storage

coffee-chest			1		1
coffee-tin				1	1
			1	1	2

Related to plants

coffee-tree			1		1
coffee-beans			1	1	2
coffee berry (berries)	1				1
	1		2	1	4

Locations

coffee-house	48	41	34	2	125
coffee-room	1	1	7	11	20
coffee shop				2	2
coffee-stall				4	4
	49	43	40	19	151

People

coffee-seller			1		1
coffee-house politician			4		4
coffee-bearer			1		1
coffee stall-keeper				1	1
coffee housekeeper	1				1
coffee woman	1				1
coffee man	6				6

8		6	1	15
---	--	---	---	-----------

Other

coffee-coloured			3	1	4
coffee house door	1				1
coffeehouse box		1			1
coffee-room-table			1		1
milk coffee		1			1
coffee and milk			1		1
breakfast-coffee			1		1
	1	2	6	1	10
	60	49	63	32	204

65	66	146	163	440
----	----	-----	-----	------------

Appendix 2.3: Numbers of occurrences of *chocolate***CHOCOLATE**

meanings	Period 1	Period 2	Period 3	Period 4	Total
produce	3	1			4
the drink	1	5	20	8	34
solid sweet				4	4
produce OR drink		2			2
drink OR solid sweet			2	3	5
colour				3	3
produce OR drink OR sweet				2	2
produce OR sweet				1	1
	4	8	22	21	55

CHOCOLATE

compounds	Period 1	Period 2	Period 3	Period 4	Total
<i>Locations</i>					
chocolate house	1	3			4
	1	3			4

Other

chocolate drops				1	1
chochlate-coloured				2	2
chocolate-toned				1	1
chocolate-and-yellow				1	1
chocolate-clad				1	1
chocolat a la vanille				1	1
				7	7
	1	3		7	11

5	11	22	28	66
---	----	----	----	-----------

Appendix 3.1 Period 1: 1640-1709

Author	Work	Year published	Genre	<i>tea</i>		<i>coffee</i>		<i>chocolate</i>	
				w/t	w/g	w/t	w/g	w/t	w/g
Pepys, Samuel	Penny Merriments	1687	Fiction	1		2			
Behn, Aphra	The Ten Pleasures of Marriage	1682	Fiction: Play			2		1	1
Vanbrugh, John	The Relapse		Fiction: Play			1	5	1	1
Secretary of State's office	The Newdigate Newsletters		Newletters			58	58		
Howard, William	Letters	1642	Letters						
More, Henry	Letters	1664	Letters			1			
(Hatton, Charles)	Letters	1676	Letters						3
Lyttelton, Charles		1664	Letters					1	
Lyttelton, Charles		1672	Letters					1	
Hatton, Frances		1677	Letters					1	
Petty, William	Letters	1679	Letters		1	1	2		
Baron, Samuel	<i>A Description of the Kingdom of Tonqueen</i>	1686?	Travels	1	1				
				2	2	65	65	5	5

w/t: number of occurrences per textfile

w/g: number of occurrences per genre

Appendix 3.2 Period 2: 1710-1779

Author	Work	Year published	Age of author	Genre	tea		coffee		chocolate	
					w/t	w/g	w/t	w/g	w/t	w/g
Fielding, Henry	<i>The History of Tom Jones, a Foundling</i>	1749	42	Fiction	6					
Fielding, Henry	<i>Amelia</i>	1751	44	Fiction	25		9			
Goldsmith, Oliver	<i>The Vicar of Wakefield</i>	1766	36	Fiction	6		1			
Hearne, Mary	<i>The Lover's Week</i>	1718		Fiction	3					
Smollett, Tobias George	<i>The Adventures of Peregrine Pickle</i>	1751	30	Fiction	1		3		1	
Smollett, Tobias George	<i>The Expedition of Humphrey Clinker</i>	1771	50	Fiction	11		5		2	
Sterne, Laurence	<i>The Life and Opinions of Tristram Shandy</i>	1759-1767	46	Fiction	3		4			
Sterne, Laurence	<i>A Sentimental Journey through France and Italy</i>	1768	55	Fiction		55	1	23		3
Gay, John	<i>The Beggar's Opera</i>	1728	43	Fiction/Play (Opera)			1	1	1	1
Steele, Richard	<i>The Spectator?</i>	1711	39	Periodical			9			
unknown	<i>The Free-Thinker</i>	1718		Periodical			1			
unknown	<i>Gray's Inn journal</i>	1753		Periodical			1		1	
unknown	<i>Westminster Magazine</i>	1776		Periodical			4	15		1
Atterbury, Francis	<i>English Advice to the Freeholders of England</i>	1714	51	Debate			1			
Burke, Edmund	<i>On Conciliation with America</i>	1775	46	Debate	2		1			
Johnson, Samuel	<i>Parliamentary Debates</i>	1740-1741	31	Debate	1	3	1	3	1	1
Berkeley, George	<i>Essay Towards Preventing the Ruin of Great Britain</i>	1721	36	Essay			1			
Burke, Edmund	<i>Essay on Taste</i>	1759	30	Essay	1		1			

Goldsmith, Oliver	<i>Enquiry into the Present State of Polite Learning</i>	1759	29	Essay			1			
Smith, Adam	<i>An Inquiry into the Nature and Causes of the Wealth of the Nations</i>	1766	43	Essay	4	5	1	4	1	1
Chesterfield, Philip Dormer Stanhope	<i>Letters to his Son</i>	1746	52	Letters	3		7		1	
Walpole, Horace	<i>Letters</i>	1735-1748	18-31	Letters	5	8	11	18	1	2
Smollett, Tobias George	<i>Travels through France and Italy</i>	1766	45	Travel	1				1	
Wood, Robert	<i>The Ruins of Balbec</i>	1757		Travel		1	1	1		1
Smollett, Tobias George	<i>A Complete History of England</i>	1752	31	History	1	1				
Goldsmith, Oliver	<i>Life of Richard Nash</i>	1762	32	Biography	1	1			1	1
					74	74	65	65	11	11

w/t: number of occurrences per textfile

w/a: number of occurrences per author

Appendix 3.3 Period 3: 1780-1849

Author	Work	Year published	Age of author	Genre	tea		coffee		chocolate	
					w/t	w/g	w/t	w/g	w/t	w/g
Austen, Jane	<i>Sense and Sensibility</i>	1811	36	Fiction	5		1			
Austen, Jane	<i>Pride and Prejudice</i>	1813	38	Fiction	6		4			
Brontë, Anne	<i>Agnes Grey</i>	1847	27	Fiction	13		1			
Brontë, Anne	<i>The Tenant of Wildfell Hall</i>	1848	28	Fiction	39		9			
Brontë, Emily	<i>Wuthering Heights</i>	1847	29	Fiction	30		2			
Dickens, Charles	<i>Barnaby Rudge</i>	1841	29	Fiction	6		3			
Dickens, Charles	<i>A Christmas Carol in Prose</i>	1843	31	Fiction	4		1			
Dickens, Charles	<i>Dombey and Son</i>	1848	36	Fiction	31		4			
Galt, John	<i>The Ayrshire Legatees</i>	1821	42	Fiction	25		9			
Galt, John	<i>Annals of the Parish</i>	1821	42	Fiction	28		1			
Gaskell, Elizabeth	<i>Mary Barton</i>	1848	38	Fiction	62		2			
Thackeray, William Makepeace	<i>Vanity Fair</i>	1848	37	Fiction	37	286	29	66	4	4
Edgeworth, Maria	<i>The Parent's Assistant</i>	1796	29	Fiction/children	12				10	
Marryat, Frederick	<i>Masterman Ready</i>	1841	49	Fiction/children	6	18	6	6		10
Hazlitt, William	<i>Liber Amoris</i>	1821	43	Fiction/Autobiography	9	9				
Smith, James and Horace Smith*	<i>Rejected Addresses</i>	1812	33 / 37	Fiction/parody	4	4			1	1
Wollstonecraft, Mary	<i>Maria; or, the Wrongs of Women</i>	1798	39	Fiction/Polemics	4	4				
Beckford, William	<i>Dreams, Waking Thoughts, and Incidents</i>	1783	23	Prose	3	3	11	11	1	1

Hazlitt, William	<i>Table Talk</i>	1822	44	Essays	7	7	9	9	1	1
Austen, Jane	<i>Letters to her Sister Cassandra and Others</i>	1796-1817	21 - 42	Letters	31		4			
Burns, Robert	<i>Letters</i>	1780-1896	23 - 47	Letters	4					
Byron, George Gordon	<i>Letters 1810-1813</i>	1810-1813	22 - 25	Letters		35	3	7		
Wollstonecraft, Mary	<i>Letters on Norway, Sweden, and Denmark</i>	1796	37	Letters/Travel	1	1	11	11		
Borrow, George Henry	<i>The Bible in Spain</i>	1843	40	Travel	2		12		5	
Darwin, Charles	<i>The Voyage of the Beagle</i>	1839	30	Travel	3		4			
Kinglake, Alexander William	<i>Eothen, or Traces of Travel Brought Home from the East</i>	1844	35	Travel	9	14	6	22		5
Carlyle, Thomas	<i>The French Revolution</i>	1837	42	History	2		11			
Southey, Robert	<i>Sir Thomas More</i>	1829	55	History	1	3	1	12		
Gibbon, Edward	<i>Memoirs of My Life</i>	1790*	53	Memoir			1	1		
De Quincey, Thomas	<i>Confessions of an English Opium-Eater</i>	1821	36	Autobiography	21	21	2	2		
					405	405	147	147	22	22

w/t: number of occurrences per textfile

w/g: number of occurrences per genre

*Included in COPC (1680-1780) despite the fact that all parts of the texts were written after 1780 and published as late as 1790.

Appendix 3.4 Period 4: 1850-1920

Author	Work	Year published	Age of author	Genre	tea		coffee		chocolate	
					w/t	w/g	w/t	w/g	w/t	w/g
Bennett, Arnold	<i>The Grand Babylon Hotel</i>	1902	35	Fiction			2		5	
Bennett, Arnold	<i>The Old Wives' Tale</i>	1908	41	Fiction	67		13		4	
Blackmore, Richard	<i>Lorna Doone, A Romance of Exmoor</i>	1869	44	Fiction			1			
Carroll, Lewis	<i>Alice's Adventures in Wonderland</i>	1865	33	Fiction	25					
Carroll, Lewis	<i>Through the Looking Glass</i>	1871	39	Fiction	6		1			
Carroll, Lewis	<i>Sylvie and Bruno</i>	1889	57	Fiction	19		3			
Chesterton, Gilbert Keith	<i>The Wisdom of Father Brown</i>	1914	40	Fiction	7		4		1	
Collins, William Wilkie	<i>The Woman in White</i>	1859-1860	35	Fiction	17				1	
Collins, William Wilkie	<i>The Moonstone</i>	1868	44	Fiction	11		4		1	
Forster, Edward Morgan	<i>Where Angels Fear to Tread</i>	1905	26	Fiction	18		2			
Forster, Edward Morgan	<i>A Room with a View</i>	1908	29	Fiction	26		2			
Forster, Edward Morgan	<i>Howards End</i>	1910	31	Fiction	38		19			
Galsworthy, John	<i>The Island Pharisees</i>	1904	37	Fiction	4		8			
Galsworthy, John	<i>The Man of Property</i>	1906	39	Fiction	24		3		2	
Gissing, George	<i>New Grub Street</i>	1891	34	Fiction	21		14			
Gissing, George	<i>The Odd Women</i>	1893	36	Fiction	19		3			
Grossmith, George, and Weedon Grossmith*	<i>The Diary of a Nobody</i>	1894		Fiction	15				1	
Haggard, Henry Rider	<i>She</i>	1887	31	Fiction	3				1	
Hardy, Thomas	<i>A Pair of Blue Eyes</i>	1873	33	Fiction	14		2		1	
Hardy, Thomas	<i>Far from the Madding Crowd</i>	1874	34	Fiction	12		1		1	
Hope, Anthony	<i>The Prisoner of Zenda</i>	1894	31	Fiction	8		3			
Hope, Anthony	<i>Rupert of Hentzau</i>	1898	35	Fiction			1		1	

Hughes, Thomas	<i>Tom Brown's Schooldays</i>	1857	35	Fiction	37		11		
Jerome, Jerome K.	<i>Three Men in a Boat</i>	1889	30	Fiction	36		4		1
Jerome, Jerome K.	<i>They and I</i>	1909	50	Fiction	21		5		
Kipling, Rudyard	<i>The Jungle Book</i>	1894	29	Fiction	1				
Kipling, Rudyard	<i>Captains Courageous</i>	1897	32	Fiction	3		10		2
Meredith, George	<i>The Adventures of Harry Richmond</i>	1870	42	Fiction	1		1		1
Meredith, George	<i>The Amazing Marriage</i>	1895	67	Fiction	8		3		
Rutherford, Mark	<i>Catherine Furze</i>	1893	62	Fiction	22		4		1
Rutherford, Mark	<i>Clara Hopgood</i>	1896	65	Fiction	12		1		
Wells, Herbert George	<i>The War of the Worlds</i>	1897	31	Fiction	2				
Wells, Herbert George	<i>Mankind in the Making</i>	1902-1903	36-37	Fiction	1				
Yonge, Charlotte Mary	<i>The Clever Woman of the Family</i>	1865	42	Fiction	10	508		125	24
Yonge, Charlotte Mary	<i>Young Folk's History of England</i>	1873	50	Fiction/children	4	4	3	3	
Chesterton, Gilbert Keith	<i>What's Wrong with the World</i>	1912	38	Debate	7		2		
Bagehot, Walter	<i>The English Constitution</i>	1867	41	Debate	1		2		
Bagehot, Walter	<i>Physics and Politics</i>	1869	43	Debate		8	7	11	
Huxley, Thomas Henry	<i>Discourses</i>	1894	69	Essays	2				1
Pater, Walter Horatio	<i>Essays from 'The Guardian'</i>	1886-1890	47-51	Essays	1	3			1
Several				Letters	12		1		
Bell, Gertrude				Letters	13		4		
Dowson, Ernest				Letters	7				1
Green, John Richard				Letters	5				
Webb, Sidney and Webb, Beatrice*				Letters	4	41		5	1

Booth, William	<i>In Darkest England and the Way Out</i>	1890	61	Religion	15	15	15	15	
Churchill, Winston	<i>The River War, An Account of the Reconquest of the Sudan</i>	1899	25	History	1	1			2 2
Gosse, William Edmund	<i>Father and Son, a Study of Two Temperaments</i>	1907	58	Memoir	12		1		
Lord and Lady Amberly				Diary	3				
Butler, Samuel	<i>Note-Books</i>	1912	77	Notebooks	3	18	1	2	
Butler, Samuel	<i>The Way of All Flesh</i>	1903	68	Biography	8	8	2	2	
					606	606	163	163	28 28

w/t: number of occurrences per textfile

w/a: number of occurrences per author

Appendix 4.1 Distribution across genres in period 1

Appendix 4.1.1: Distribution of <i>tea</i> across genres in period 1					
	TEA				
	M		W		Total
Fiction					0 0.00%
Fiction/Children					0 0.00%
Fiction/Autobiography					0 0.00%
Fiction/Parody					0 0.00%
Fiction/Polemics					0 0.00%
Prose					0 0.00%
Essays					0 0.00%
Letters	1	50.00%			1 50.00%
Letters/Travel	1	50.00%			1 50.00%
Travel					0 0.00%
History					0 0.00%
Biography					0 0.00%
	2	100.00%	0	0.00%	2 100.00%

Appendix 4.1.2: Distribution of <i>coffee</i> across genres in period 1					
	COFFEE				
	M		W		Total
Fiction	2	3.17%			2 3.08%
Fiction/Children		0.00%			0 0.00%
Fiction/Autobiography		0.00%			0 0.00%
Fiction/Parody		0.00%			0 0.00%
Fiction/Polemics					0 0.00%
Fiction/Play	1	1.59%	2 100.00%		3 4.62%
Newsletters	58	92.06%			58 89.23%
Prose		0.00%			0 0.00%
Essays		0.00%			0 0.00%
Letters	2	3.17%			2 3.08%
Letters/Travel					0 0.00%
Travel		0.00%			0 0.00%
History		0.00%			0 0.00%
Biography		0.00%			0 0.00%
	63	100.00%	2 100.00%		65 100.00%

Appendix 4.1.3: Distribution of <i>chocolate</i> across genres in period 1					
	CHOCOLATE				
	M		W		Total
Fiction					0 0.00%
Fiction/Children					0 0.00%
Fiction/Autobiography					0 0.00%
Fiction/Parody					0 0.00%
Fiction/Polemics					0 0.00%
Fiction/Play	1	33.33%	1 50.00%		
Prose					0 0.00%
Essays					0 0.00%
Letters	2	66.67%	1 50.00%		3 60.00%
Letters/Travel					0 0.00%
Travel					0 0.00%
History					0 0.00%
Biography					0 0.00%
	3	100.00%	2 100.00%		5 60.00%

Appendix 4.2 Distribution across genres in period 2

Appendix 4.2.1: Distribution of <i>tea</i> across genres in period 2										
	TEA				Total		Number of texts			
	M		W				S/TG	N	S/N	
Fiction	52	73.24%	3	100.00%	55	74.32%	8	31%	55	74%
Fiction/Children										
Fiction/Autobiography										
Fiction/Parody										
Fiction/Polemics										
Fiction/Play							1	4%		
Journalistic							4	15%		
Debate	3	4.23%			3	4.05%	3	12%	3	4%
Prose										
Essays	5	7.04%			5	6.76%	4	15%	5	7%
Letters	8	11.27%			8	10.81%	2	8%	8	11%
Letters/Travel										
Travel	1	1.41%			1	1.35%	2	8%	1	1%
History	1	1.41%			1	1.35%	1	4%	1	1%
Biography	1	1.41%			1	1.35%	1	4%	1	1%
	71	100.00%	3	100.00%	74	100.00%	26	100%	74	100%
Appendix 4.2.2: Distribution of <i>coffee</i> across genres in period 2										
	COFFEE				Total		Number of texts			
	M		W				S/TG	N	S/N	
Fiction	23	35.38%			23	35.38%	8	31%	23	35%
Fiction/Children										
Fiction/Autobiography										
Fiction/Parody										
Fiction/Polemics										
Fiction/Play	1	1.54%			1	1.54%	1	4%	1	2%
Journalistic	15	23.08%			15	23.08%	4	15%	15	23%
Debate	3	4.62%			3	4.62%	3	12%	3	5%
Prose										
Essays	4	6.15%			4	6.15%	4	15%	4	6%
Letters	18	27.69%			18	27.69%	2	8%	18	28%
Letters/Travel										
Travel	1	1.54%			1	1.54%	2	8%	1	2%
History							1	4%		
Biography							1	4%		
	65	100.00%	0	0.00%	65	100.00%	26	100%	65	100%

Appendix 4.2.3: Distribution of *chocolate* across genres in period 2

	CHOCOLATE				Number of texts					
	M		W		Total		S/TG	N	S/N	
Fiction	3	27.27%			3	27.27%	8	31%	3	27%
Fiction/Children										
Fiction/Autobiography										
Fiction/Parody										
Fiction/Polemics										
Fiction/Play	1	9.09%			1	9.09%	1	4%	1	9%
Journalistic	1	9.09%			1	9.09%	4	15%	1	9%
Debate	1	9.09%			1	9.09%	3	12%	1	9%
Prose										
Essays	1	9.09%			1	9.09%	4	15%	1	9%
Letters	2	18.18%			2	18.18%	2	8%	2	18%
Letters/Travel										
Travel	1	9.09%			1	9.09%	2	8%	1	9%
History							1	4%		
Biography	1	9.09%			1	9.09%	1	4%	1	9%
	11	100.00%	0	0.00%	11	100.00%	26	100%	11	100%
	S/TG	Share of texts in genre								
	N	Number of occurrences								
	S/N	Share of occurrences								

Appendix 4.3 Distribution across genres in period 3

Appendix 4.3.1: Distribution of *tea* across genres in period 3

	TEA						Number of texts			
	M		W		Total		S/TG		N	S/N
Fiction	131	64.85%	155	76.35%	286	70.62%	12	41%	286	71%
Fiction/Children	6	2.97%	12	5.91%	18	4.44%	2	7%	18	4%
Fiction/Autobiography	9	4.46%			9	2.22%	1	3%	9	2%
Fiction/Parody	4	1.98%			4	0.99%	1	3%	4	1%
Fiction/Polemics			4	1.97%	4	0.99%	1	3%	4	1%
Prose	3	1.49%			3	0.74%	1	3%	3	1%
Essays	7	3.47%			7	1.73%	1	3%	7	2%
Letters	4	1.98%	31	15.27%	35	8.64%	3	10%	35	9%
Letters/Travel			1	0.49%	1	0.25%	1	3%	1	0%
Travel	14	6.93%			14	3.46%	3	10%	14	3%
History	3	1.49%			3	0.74%	1	3%	3	1%
Memoir							1	4%		
Autobiography	21	10.40%			21	5.19%	1	3%	21	5%
	202	100.00%	203	100.00%	405	100.00%	29	100%	405	100%

Appendix 4.3.2: Distribution of *coffee* across genres in Period 3

	COFFEE						Number of texts			
	M		W		Total		S/TG		N	S/N
Fiction	47	41.59%	19	55.88%	66	44.90%	12	41%	66	45%
Fiction/Children	6	5.31%			6	4.08%	2	7%	6	4%
Fiction/Autobiography							1	3%		
Fiction/Parody							1	3%		
Fiction/Polemics							1	3%		
Prose	11	9.73%			11	7.48%	1	3%	11	7%
Essays	9	7.96%			9	6.12%	1	3%	9	6%
Letters	3	2.65%	4	11.76%	7	4.76%	3	10%	7	5%
Letters/Travel			11	32.35%	11	7.48%	1	3%	11	7%
Travel	22	19.47%			22	14.97%	3	10%	22	15%
History	12	10.62%			12	8.16%	1	3%	12	8%
Memoir	1	0.88%			1	2.94%	1	4%	1	1%
Autobiography	2	1.77%			2	1.36%	1	3%	2	1%
	113	100.00%	34	100.00%	147	102.26%	29	100%	147	100%

Appendix 4.3.3: Distribution of *chocolate* across genres in Period 3

	CHOCOLATE						Number of texts			
	M		W		Total		S/TG		N	S/N
Fiction	4	33.33%			4	18.18%	12	41%	4	18%
Fiction/Children			10	100.00%	10	45.45%	2	7%	10	45%
Fiction/Autobiography							1	3%		
Fiction/Parody	1	8.33%			1	4.55%	1	3%	1	5%
Fiction/Polemics							1	3%		
Prose	1	8.33%			1	4.55%	1	3%	1	5%
Essays	1	8.33%			1	4.55%	1	3%	1	5%
Letters							3	10%		
Letters/Travel							1	3%		
Travel	5	41.67%			5	22.73%	3	10%	5	23%
History							1	3%		
Memoir							1	4%		
Autobiography							1	3%		
	12	100.00%	10	100.00%	22	100.00%	29	100%	22	100%
	S/TG	Share of texts in genre								
	N	Number of occurrences								
	S/N	Share of occurrences								

Appendix 4.4 Distribution across genres in period 4

Appendix 4.4.1: Distribution of *tea* across genres in period 4

	TEA						Number of texts			
	M		W		Total		S/TG		N	S/N
Fiction	498	86.01%	10	37.04%	508	83.83%	34	67%	508	84%
Fiction/Children			4	14.81%	4	0.66%	1	2%	4	1%
Fiction/Autobiography					0					
Fiction/Parody					0					
Fiction/Polemics					0					
Fiction/Play					0					
Journalistic					0					
Debate	8	1.38%			8	1.32%	3	6%	8	1%
Prose					0					
Essays	3	0.52%			3	0.50%	2	4%	3	0%
Letters	28	4.84%	13	48.15%	41	6.77%	5	10%	41	7%
Letters/Travel					0					
Travel					0					
History	1	0.17%			1	0.17%	1	2%	1	0%
Religion	15	2.59%			15	2.48%	1	2%	15	2%
Diary/Notebook/Memoir	18	3.11%			18	2.97%	3	6%	18	3%
Biography	8	1.38%			8	1.32%	1	2%	8	1%
	579	100.00%	27	100.00%	606	100.00%	51	100%	606	100%

Appendix 4.4.2: Distribution of *coffee* across genres in period 4

	COFFEE						Number of texts			
	M		W		Total		S/TG		N	S/N
Fiction	125	80.13%			125	76.69%	34	67%	125	77%
Fiction/Children			3	42.86%	3	1.84%	1	2%	3	2%
Fiction/Autobiography					0					
Fiction/Parody					0					
Fiction/Polemics					0					
Fiction/Play					0					
Journalistic					0					
Debate	11	7.05%			11	6.75%	3	6%	11	7%
Prose					0					
Essays					0		2	4%		
Letters	1	0.64%	4	57.14%	5	3.07%	5	10%	5	3%
Letters/Travel					0					
Travel					0					
History					0		1	2%		
Religion	15	9.62%			15	9.20%	1	2%	15	9%
Diary/Notebook/Memoir	2	1.28%			2	1.23%	3	6%	2	1%
Biography	2	1.28%			2	1.23%	1	2%	2	1%
	156	100.00%	7	100.00%	163	100.00%	51	100%	163	100%

Appendix 4.4.3: Distribution of *chocolate* across genres in period 4

	CHOCOLATE				Number of texts			
	M		W		S/TG		N	S/N
Fiction	24	85.71%		24	34	67%	24	86%
Fiction/Children				0	1	2%		
Fiction/Autobiography				0				
Fiction/Parody				0				
Fiction/Polemics				0				
Fiction/Play				0				
Journalistic				0				
Debate				0	3	6%		
Prose				0				
Essays	1	3.57%		1	2	4%	1	4%
Letters	1	3.57%		1	5	10%	1	4%
Letters/Travel				0				
Travel				0				
History	2	7.14%		2	1	2%	2	7%
Religion				0	1	2%		
Diary/Notebook/Memoir				0	3	6%		
Biography				0	1	2%		
	28	100.00%	0	0.00%	28	100.00%	51	100%
	S/TG	Share of texts in genre						
	N	Number of occurrences						
	S/N	Share of occurrences						

Bilaga 5: Tabeller för den svenska sammanfattningen

Tabell 1

Antal ord i korpusarna (miljoner) och tidsperioderna som korpusarna täcker

*Korpusens**namn* PCEEC HCET NEWDI COPC CLMET CLMEP*Tidsperiod*

1600

1650

0,555

0,17

1,06

1700

0,497

2,096

1750

1800

3,74

1850

1900

3,982

0,1

1950

Tabell 2: Tidsperioder med korpusar och antal ord

	Korpusarnas namn	Totalt antal ord i korpusen, i miljoner	Totalt antal ord i tidsperioden
Period 1: 1640-1709	PCEEC	0,56	
	HCET	0,17	
	NEWDI	1,06	
	COPC fram til 1709	0,14	1,927
Period 2: 1710-1779	CLMET A	2,10	
	COPC 1710-1779	0,35	2,441
Period 3: 1780-1849	CLMET B	3,74	
	COPC efter 1780*	0,01	3,75
Period 4: 1850-1920	CLMET C	3,98	
	CLMEP	0,10	4,082
Totalt:			12,2

Tabell 3: Förekomst av orden *tea*, *coffee* och *chocolate* (N) över tid, och förekomst per miljoner ord; balanserat mot korpusens storlek

Period	miljoner ord	Tea		Coffee		Chocolate		All words	
		N	förekomst	N	förekomst	N	förekomst	N	förekomst
Period 1	1,927	2	1,04	65	33,73	5	2,59	72	37,36
Period 2	2,441	74	30,32	65	26,63	11	4,51	150	61,45
Period 3	3,75	405	108,00	147	39,20	22	5,87	574	153,07
Period 4	4,082	606	148,46	163	39,93	28	6,86	797	195,25
Alla	12,2	1087	89,10	440	36,07	66	5,41	1593	130,57

Tabell 4: Sammanställning av betydelser hos orden *tea*, *coffee* och *chocolate* i korpusen, samt olika semantiska grupper av sammansatta ord.

TEA	COFFEE	CHOCOLATE
Tebusken, teblad	Kaffebusken, kaffeböner	Kakaoböner
Dryck	Dryck	Drycken
Måltid	Måltid	
Tid på dygnet (under vilken måltiden te ofta intas)		
Måltid eller tid på dygnet	Måltid eller tid på dygnet	
Måltid eller dryck	Måltid eller dryck	
Verb		
Måltid eller verb		
Måltid eller dryck eller verb		
Teblad eller dryck	Kaffeböner eller dryck	Kakaoböner eller dryck
Måltid, tid på dygnet eller dryck	Måltid, tid på dygnet eller dryck	
Oklar		Choklad i fast form Dryck eller fast form Färg Kakaoböner, dryck, eller fast form Kakaoböner eller fast form
SAMMANSATTA ORD		
Redskap för tillredning eller servering	Redskap för tillredning eller servering	
Förvaring	Förvaring	
Relaterat till växter	Relaterat till växter	
Platser	Platser	Platser
Måltider eller evenemang	Måltider eller evenemang	
Människor, handel och transport	Människor, handel och transport	
Övrigt	Övrigt	Övrigt
Egennamn		

Tabell 5: Mäns och kvinnors andel av ordens förekomst i period 3

	TEA		COFFEE		CHOCOLATE	
Män	202	49,88%	113	76,87%	12	54,55%
Kvinnor	203	50,12%	34	23,13%	10	45,45%
Totalt	405	100,00%	147	100,00%	22	100,00%

Tabell 6: Ordens andel av användningen hos män och kvinnor i period 3

	Män		Kvinnor		Totalt	
TEA	202	61,77%	203	82,19%	405	70,56%
COFFEE	113	34,56%	34	13,77%	147	25,61%
CHOCOLATE	12	3,67%	10	4,05%	22	3,83%
Alla ord	327	100,00%	247	100,00%	574	100,00%

Tabell 7: Representation av män och kvinnor i valda verk, i kronologisk ordning; perioderna 3 och 4

Referenser till kvinnor i närheten av ordet *tea*

Författare	Titel	År för publicering	Antal referenser till kvinnor	Andel av alla förekomster	Antal fall utan referenser till kvinnor	Andel av alla förekomster	Totalt antal förekomster av ordet <i>tea</i>
Austen, Jane*	Letters to her sister Cassandra	1796-1817	29	94 %	2	6 %	31
Austen, Jane	Letters to her sister Cassandra	1796-1817	19	61 %	12	39 %	31
Austen, Jane, average	Letters to her sister Cassandra	1796-1817	24	77 %	7	23 %	31
	<i>Confessions of an English Opium</i>						
De Quincey, Thomas	<i>Eater</i>	1821	2	10 %	19	90 %	21
Brontë, Emily	<i>Wuthering Heights</i>	1847	22	73 %	8	27 %	30
Gaskell, Elizabeth	<i>Mary Barton</i>	1848	55	89 %	7	11 %	62
Dickens, Charles	<i>Dombey and Son</i>	1848	13	42 %	18	58 %	31
Thackeray, William	<i>Vanity Fair</i>	1848	29	78 %	8	22 %	37
	<i>Tom Brown's Schooldays</i>						
Hughes, Thomas	<i>The Clever Woman of the Family</i>	1857	7	19 %	30	81 %	37
Yonge, Charlotte	<i>In Darkest England</i>	1865	10	100 %	0	0 %	10
Booth, William	<i>Father and Son</i>	1890	2	13 %	13	87 %	15
Gosse, W. E.	<i>Howards End</i>	1907	3	25 %	9	75 %	12
Forster, E. M.	Letters	1917	12	92 %	1	8 %	13
Bell, Gertrude*	Letters	1917	3	23 %	10	77 %	13
Bell, Gertrude, average	Letters	1917	7,5	58 %	5,5	42 %	13

337

Tolkning:	Antal
Skönlitteratur, kvinnliga författare	3
Skönlitteratur, manliga författare	4
Övriga texter, kvinnliga författare	2
Övriga texter, manliga författare	3
Period 3	6
Period 4	6

*Kvinnliga författares brev: personliga pronomen som syftar på författaren själv inkluderas
 Kvinnliga författares brev: personliga pronomen inkluderas inte i referenser till kvinnor

Bibliography

Primary sources

CLMET: *The Corpus of Late Modern English Texts*, 2004. De Smet, Hendrik. 2004. "A Corpus of Late Modern English Texts". *ICAME Journal* 29: 69-82. Available: <https://perswww.kuleuven.be/~u0044428/clmet.htm> [27 February, 2014]

CLMEP: *The Corpus of Late Modern English Prose*, 1994. Denison, David. 1994. "A corpus of late Modern English prose". In Merja Kytö, Matti Rissanen & Susan Wright (eds.). *Corpora across the centuries*. Amsterdam and Atlanta GA: Rodopi.

COPC: *Century of Prose Corpus*, 1989. Milic, Louis T. 1990. "A New Historical Corpus". *ICAME Journal* 14: 26-39.

HCET: *The Helsinki Corpus of English Texts*, text version. 1996. Kytö, Merja. 1996: *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts*. Available: <http://khnt.hit.uib.no/icame/manuals/HC/INDEX.HTM> [26 February 2014]

NEWDI: *Newdigate Newsletters*. 1994. Hines Jr., Philip. 1994. *Newdigate Newsletters*. Available: <http://khnt.hit.uib.no/icame/manuals/NEWDIGAT/INDEX.HTM> [21 February, 2014]

PCEEC: *Parsed Corpus of Early English Correspondence*, text version. 2006. Taylor, Ann and Arja Nurmi. 2006. *The Parsed Corpus of Early English Correspondence*. Available: http://www-users.york.ac.uk/~lang22/PCEEC-manual/corpus_description/index.htm [21 November, 2012]

Secondary sources

Adams, Valerie. 1973. *An Introduction to Modern English Word-Formation*. London: Longman Group Limited.

Bailey, Richard W. 2010. "Variation and change in eighteenth-century English". In Hickey, Raymond (ed). 2010: 182-199

Baker, Paul. 2014. *Using Corpora to Analyze Gender*. London: Bloomsbury.

Bauer, Laurie. 1983. *English Word-formation*. Cambridge: Cambridge University Press.

Bauer, Laurie. 1998. *Vocabulary*. Florence, KY: Routledge.

Beal, Joan C. 2012. "'À la Mode de Paris': Linguistic Patriotism and Francophobia in 18th-century Britain". In Percy, Carol and Mary Catherine Davidson (eds.).

Campbell, Lyle. 1998. *Historical Linguistics: an introduction*. Edinburgh: Edinburgh University Press.

Carter, Ronald. 1987. *Vocabulary: Applied Linguistic Perspectives*. London: Allen & Unwin (Publishers) Ltd.

Cowan, Brian. 2005: *The Social Life of Coffee*. Newhaven: Yale University Press.

De Smet, Hendrik. 2004: "A Corpus of Late Modern English Texts". *ICAME Journal* 29: 69-82.

Deignan, Alice. 1999. "Linguistic Metaphors and Collocation in Nonliterary Corpus Data". *In Metaphor and Symbol* 14 (1): 19-36.

Denison, David. 1994. "A corpus of late Modern English prose". In Merja Kytö, Matti Rissanen & Susan Wright (eds.). *Corpora across the centuries*. Amsterdam and Atlanta GA: Rodopi.

Diller, H., H. De Smet, J. Tyrkkö. 2011. "A European database of descriptors of English electronic texts". In *The European English Messenger* 19, 21-35. Available: https://perswww.kuleuven.be/~u0044428/clmet3_0.htm [27 February, 2014]

Fromer, Julie E. 2008: *A Necessary Luxury*. Athens: Ohio University Press

Griffiths, John. 2007. *Tea: A History of the Drink that Changed the World*. London: Carlton Publishing Group.

Halliday, M. A. K, Wolfgang Teubert, Colin Yallop and Anna Čermáková. 2004. *Lexicology and Corpus Linguistics – An Introduction*. London: Continuum.

Hickey, Raymond. 2010. "Attitudes and concerns in eighteenth-century English". In Hickey, Raymond (ed). 2010: 1-20

Hickey, Raymond (ed). 2010. *Eighteenth-Century English - Ideology and Change*. Cambridge: Cambridge University Press.

Hines Jr., Philip. 1994. *Newdigate Newsletters*. Available: <http://khnt.hit.uib.no/icame/manuals/NEWDIGAT/INDEX.HTM> [21 February, 2014]

Horobin, Simon. 2015: "From teaspoons to tea-sots: the language of tea". Available: <http://blog.oxforddictionaries.com/2015/08/teaspoons-tea-language-tea/> [31 October, 2016]

Johnson, Samuel. 1785 (1) *A dictionary of the English language: in which the words are deduced from their originals, and illustrated in their different significations by examples from the best writers. Volume 1*. The sixth edition. London. Available: Eighteenth Century Collections Online. Gale. Abo Akademi University. 2 Feb. 2017
<http://find.galegroup.com.ezproxy.vasa.abo.fi/ecco/infomark.do?&source=gale&prodId=ECCO&userGroupName=aboacad&tabID=T001&docId=CW113728360&type=multipage&contentSet=ECCOArticles&version=1.0&docLevel=FASCIMILE>.

Johnson, Samuel. 1785 (2). *A dictionary of the English language: in which the words are deduced from their originals, and illustrated in their different significations by examples from*

the best writers. Volume 2. The sixth edition. London. Eighteenth Century Collections Online. Gale. Abo Akademi University. 2 Feb. 2017.
<http://find.galegroup.com.ezproxy.vasa.abo.fi/ecco/infomark.do?&source=gale&prodId=ECCO&userGroupName=aboacad&tabID=T001&docId=CW113729474&type=multipage&contentSet=ECCOArticles&version=1.0&docLevel=FASCIMILE>.

Kennedy, Graeme. 1998. *An Introduction to Corpus Linguistics*. London: Longman.

Kytö, Merja. 1996: *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts*. Available:
<http://khnt.hit.uib.no/icame/manuals/HC/INDEX.HTM> [26 February 2014]

Lee, David. 2001. *Cognitive Linguistics - An Introduction*. Oxford: Oxford University Press.

Markman, Ellis (ed). 2006 (1). *Eighteenth-Century Coffee-House Culture Volume 1: Restoration Satire*. London: Pickering & Chattoo.

Markman, Ellis (ed). 2006 (2) *Eighteenth-Century Coffee-House Culture Volume 2: The Eighteenth-Century Satire*. London: Pickering & Chattoo.

Markman, Ellis (ed). 2006 (3) *Eighteenth-Century Coffee-House Culture Volume 3: Drama*. London: Pickering & Chattoo.

Markman, Ellis (ed). 2006 (4) *Eighteenth-Century Coffee-House Culture Volume 4: Science and History Writings*. London: Pickering & Chattoo.

Milic, Louis T. 1990 (1). "A New Historical Corpus". *ICAME Journal* 14: 26-39.

Milic, Louis T. 1990 (2). "The Century of Prose Corpus". *Literary and Linguistic Computing* 5, 3: 203-208.

Moss, Sarah, and Alexander Badenoch. 2009. *Chocolate - A Global History*. London: Reaktion Books Ltd.

Mugglestone, Lynda. 2010. "Registering the Language - dictionaries, diction and the art of elocution" In Hickey, Raymond (ed) 2010: 309-338

Mugglestone, Lynda. 2012. "Patriotism, Empire and Cultural Prescriptivism: Anglicity in the *OED*". In Percy, Carol and Mary Catherine Davidson (eds.).

Nevalainen, Terttu, and Helena Raumolin-Brunberg. 2003: *Historical Sociolinguistics*. Essex: Pearson Education Limited.

OED: Oxford English Dictionary Online. 2010. Oxford University Press. Available:
<http://www.oed.com> [28 October, 2010]

Percy, Carol and Mary Catherine Davidson (eds.). 2012. *The Languages of Nation. Attitudes and norms*. Bristol: Multilingual Matters.

Porter, Roy and Marie Mulvey Roberts (eds.). 1996. *Pleasure in the Eighteenth Century*.

London: Macmillan Press LTD.

Roberts, Marie Mulvey. 1996. "Pleasures Engendered by Gender: Homosociality and the Club". In Porter, Roy and Marie Mulvey Roberts (eds.).

Romaine, Suzanne. 1982. *Socio-historical Linguistics - its status and methodology*. Cambridge: Cambridge University Press.

Saberi, Helen. 2010. *Tea - A Global History*. London: Reaktion Books Ltd.

Scott, Mike. 2004. WordSmith Tools. Available:
<http://www.lexically.net/wordsmith/version4/> [12 February, 2014]

Stockwell, Robert P. and Donka Minkova. 2001. *English Words: History and Structure*. Cambridge: Cambridge University Press.

Taylor, Ann and Arja Nurmi. 2006. *The Parsed Corpus of Early English Correspondence*. Available:
http://www-users.york.ac.uk/~lang22/PCEEC-manual/corpus_description/index.htm [21 November, 2012]

Teubert, Wolfgang. 2004. "Language and Corpus Linguistics". In Halliday, M.A.K. et al: 73-112.

Ungerer, Friedrich and Hans-Jörg Schmid. 1996. *An introduction to Cognitive Linguistics*. Harlow: Addison Wesley Longman Limited.

Vary, Simon. 1996. "The Pleasure of the Table". In Porter, Roy and Marie Mulvey Roberts (eds.).

Winterbottom, Anna. 2013. "Self-fashioning and auto-ethnography: Samuel Baron's description of Tonqueen (1686)". *Journeys* 85