An auxiliary in decline?

A corpus-based study of the development in the use of the modal auxiliary *ought* from a synchronic and diachronic perspective

by

Nina Jeanette Lailey

A Thesis presented to
the Department of Literature, Area Studies and European Languages
The University of Oslo
in partial fulfilment of the requirements
for the Masters Degree
Spring term 2008
ACKNOWLEDGEMENTS

A number of people have helped me in my writing of this thesis, and I would like to take the opportunity to thank some of them here.

First, I would like to express my sincere thanks to my supervisor, Hilde Hasselgård for all her help and advice throughout the writing process.

I would also like to thank Johan Elsness, Stig Johansson and Per Lysvåg for introducing me to corpus linguistics and for all their help and as well as willingness to answer questions whenever I have needed it.

To Helle Linné Eriksen, student advisor at ILOS, for her help and advice, throughout my time as a student at the university – including practical assistance far beyond the call of duty.

Professor Geoffrey Leech, whose work inspired me and provided the starting point for this investigation

Stine Hulleberg Johansen for very useful comments and suggestions on the draft version of this thesis, as well as for being on hand for discussions, inspiration, advice and encouragement.

My parents, Torild and Andrew Lailey as well as my mother and father-in-law, Edel and Stein Hagen for encouragement, support and practical help, particularly in the form of babysitting services, throughout my student days.

My husband, Alexander Hagen. And my children, Mathias and Hannah Lailey Hagen, for their love, encouragement and support throughout.

My family and friends for all their encouragement

The front page illustration: © Cartoonbank.com
# Table of Contents

1) Introduction.

1.1 The background for the investigation in this thesis. 7

1.2. Varieties of English 9

1.3. Language change. 12

1.4. Spoken and written language. 14

1.5. Politeness theory 17

2) Modal Auxiliaries in General, and Ought in Particular.

2.1. Defining modal auxiliaries 21

2.2. The meaning and etymology of the modal auxiliary ought. 22

3) Research Method

3.1. Qualitative v.s. quantitative method. 24

3.2. Corpus linguistics and the use of corpora in a contrastive study. 27

3.3. The corpora used for this study 29

3.4. BROWN 29

3.5. LOB 31

3.6. FROWN and FLOB. 31

3.7. ACE 32

3.8. WC 32

3.9. BNC 33

3.10. COLT 34

3.11. DCPSE 35

3.12. ENPC 36

3.13. BYUC 37

3.14. Summary of the corpora used for this study. 38

3.15. Wordsmith and FileMaker Pro 5.5. 39
4) THE INVESTIGATION IN THE LOB, FLOB, BROWN, FROWN, ACE AND WC CORPORAS – RESULTS AND ANALYSIS.

4.1. Overall distribution 41
4.2. The reduction in the use of ought in British English from 1961-1991 42
4.3 Distribution in text categories in BrE and AmE. 44
4.4. The distribution according to the ideational dimension is Systemic Functional Grammar (SFG) 47
4.5. Systemic Functional Grammar and the experiential metafunction 47
4.6. Processs types in LOB, FLOB, Brown, Frown, ACE and WC 52
4.7. Distribution according to semantic meaning in written language. 57
4.8. Subject distribution in LOB, FLOB, Brown and Frown. 62

5) OUGHT IN SPOKEN LANGUAGE ACCORDING TO THE INVESTIGATION IN DCPSE 69

5.1. Distribution of ought according to text categories in the DCPSE Corpus. 69
5.2. The distribution of ought according to sociolinguistic parameters in the DCPSE 71
5.3. Diachronic development of ought use according to the DCSPE 73
5.3. Distribution according to semantic meaning in spoken language. 75

6) THE USE OF OUGHT IN RECENT BRITISH ENGLISH AS SEEN FROM THE BNC INVESTIGATION. 78

7) OUGHT IN RECENT AMERICAN ENGLISH ACCORDING TO THE BYUC 83

8) SOME SPECIAL USES OF OUGHT 86

8.1. The negative – oughtn’t. 86
8.2. The to-less ought 89
8.3. Ought in tag questions 93

9) ALTERNATIVE CONSTRUCTIONS TO OUGHT. 95

9.1. ENPC Results. 95
9.2. Dictionaries 106
9.3. Native Speaker intuition. 108
9.4. The development of the use of alternative expressions to *ought*.

10) CONCLUDING REMARKS

11) SUGGESTIONS FOR FURTHER RESEARCH.

REFERENCES

LIST OF FIGURES.
FIG. 1.1 – An illustration of Kachru’s circles of English. ................................................. 10
FIG 3.1 – The structure of the English-Norwegian Parallel Corpus.............................. 36
FIG. 4.1 – The occurrence of *ought* in the four corpora from the Brown Family used in this investigation (Brown, Frown, LOB, FLOB, ACE and WC) normalized into instances per 1 million words. ................................................................. 41
FIG. 4.2 – The distribution of *ought* (relative frequency per 100 000 words) In LOB and FLOB ............................................................................................................. 45
FIG. 4.3 – The distribution of *ought* (relative frequency per 100,000 words) In Brown and Frown ........................................................................................................ 47
FIG. 4.4 – Metafunctions acting together in a clause .......................................................... 49
FIG 4.5 – Distribution according to process types in LOB .................................................. 54
FIG 4.6 – Distribution according to process types in FLOB .................................................. 54
FIG 4.7 – Distribution according to process types in Brown ................................................ 55
FIG 4.8 – Distribution according to process types in Frown ............................................... 55
FIG 4.9 – Distribution according to process types in ACE .................................................. 55
FIG 4.10 – Distribution according to process types in WC .................................................. 56
FIG. 4.11 – The distribution of *ought* according to semantic meaning in Brown, Frown, LOB and FLOB ............................................................. 59
FIG 4.12 – The distribution of semantic meaning in LOB given in % ................................ 60
FIG 4.13 – The distribution of semantic meaning in FLOB given in % ............................... 60
FIG 4.14 – The distribution of semantic meaning in Brown given in % .............................. 61
FIG 4.15 – The distribution of semantic meaning in Frown given in % .............................. 61
FIG 4.16 – The proportion of subject types in LOB ............................................................ 65
FIG 4.17 – The proportion of subject types in FLOB ........................................................... 65
FIG 4.18 – The proportion of subject types in Brown .......................................................... 67
FIG 4.19 – The proportion of subject types in Frown .......................................................... 67
FIG 5.1 – The distribution of *ought* according to text categories in the DPCSE corpus given in instances per 1000 words: ................................................................. 70
FIG 5.2 – Usage of *ought* according to speakers’ level of education ............................. 71
FIG 5.3 – Usage of *ought* according to speakers’ gender ................................................ 72
FIG 5.4 – Distribution of *ought* over time given in instances per 1000 words............. 73
FIG 5.5 – The distribution of *ought* in the DCSPE according to speaker age given in instances per 1000 words ................................................................. 74
FIG 5.6 – The distribution of *ought* according to semantic meaning in spoken discourse according to the DCPSE. ................................................................. 75
FIG. 6.1 – The frequency of the word *ought* per million words in spoken, compared with written English accorded to the British National Corpus ........................................... 78
FIG. 6.2 – The distribution of *ought* in spoken texts according to speech categories given in instances per one million words (figures rounded to the nearest whole number) ................................................................. 79
FIG. 6.3 – The distribution of *ought* in written texts according to text categories in the BNC (figures rounded to the nearest whole number) .................................................. 79
FIG 6.4 – Distribution of *you ought* in text categories in the BNC (written texts) ........... 80
FIG. 6.5 – Frequency of the use of *ought* in written texts according to the age group of the author according to the British National Corpus ........................................ 82
FIG. 7.1 – The development of the use of *ought* in recent AmE as seen in the BYUC given in instances per million words ................................................................. 83
FIG.7.2 – The distribution of *ought* in spoken and written material according to the BYUC given in instances per million words ......................................................... 84
FIG.7.3 – The distribution of the use of *ought* in written text categories AmE as seen in the BYUC given in instances per million words ............................................... 85
FIG. 9.1 – Translations into Norwegian of *ought* in the ENPC ................................... 95
FIG. 9.2 – *Burde* translated into English according to ENPC ................................. 99
FIG. 9.3 – Translations of *skulle* in 100 instances from ENPC ................................. 101
FIG. 9.4 – *Bar* translated into English according to ENPC ...................................... 102
FIG. 9.5 – *Må* translated into English according to ENPC ...................................... 104

LIST OF TABLES
Table 1.1 – The results from the original investigation by Leech ..................................... 7
Table 2.1 – Attributes of qualitative and quantitative paradigms .................................. 24
Table 2.2 – The text categories and number of texts in each category in the Brown Corpus ....................................................................................................................... 30
Table 2.3 – Text categories in the LOB corpus ............................................................... 31
Table 3.1 – The build-up of the ENPC corpus ................................................................. 36
Table 3.2 – Details and summary of the corpora used in this study ................................ 38
Table 4.1 – The occurrence of *ought* in the various text categories in the LOB and FLOB Corpora ................................................................. 44
Table 4.2 – The occurrence of *ought* in the various text categories in the Brown and Frown Corpora ................................................................. 46
Table 4.3 – The distribution according to process types given in percent of total occurrences in each corpus (raw frequencies in brackets) ........................................ 52
Table 4.4 – The general distribution of process types given in % according to Matthiessen (1999) ......................................................................................... 53
Table 4.5 – The distribution of semantic meanings given in % (Raw Frequencies) ...... 59
Table 4.6 The distribution of subject types in LOB, FLOB, Brown and Frown given in raw frequencies ................................................................. 64
Table 5.1 – The distribution of *ought* according to text categories in the DPCSE corpus .................................................................................................................. 69
Table 5.2 – Distribution of *ought* over time ................................................................. 73
Table 5.3 – The distribution of *ought* in the DCSPE according to speaker age .......... 74
Table 5.4 – The distribution according to semantic meaning given in % (Raw frequencies) ................................................................................................. 75
Table 7.1 – The development of the use of *ought* in recent AmE as seen in the BYUC ..................................................................................................................... 83
Table 7.2 – The distribution of the use of *ought* in written text categories AmE as seen in the BYUC ................................................................. 84
Table 9.1 – The development of the use of alternative constructions to *ought* in LOB, FLOB, Brown and Frown ................................................................. 109
1) Introduction.

1.1 The background for the investigation in this thesis.

This study is based on Geoffrey Leech’s claim that the frequency of modal auxiliary verbs has changed over a time span of 30 years (Leech 2002). Leech performed an investigation by comparing findings in four corpora (LOB, FLOB, BROWN and Frown). The results are given in table 1.1 below.

Table 1.1 –The results from the original investigation by Leech

<table>
<thead>
<tr>
<th>Verb</th>
<th>British</th>
<th>English</th>
<th>Log likhd</th>
<th>Diff (%)</th>
<th>American</th>
<th>English</th>
<th>Log likhd</th>
<th>Diff (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would</td>
<td>3,028</td>
<td>2,694</td>
<td>20.4</td>
<td>-11.0</td>
<td>3,053</td>
<td>2,686</td>
<td>5.6</td>
<td>-6.1</td>
</tr>
<tr>
<td>Will</td>
<td>2,798</td>
<td>2,723</td>
<td>1.2</td>
<td>-2.7</td>
<td>2,702</td>
<td>2,402</td>
<td>17.3</td>
<td>-11.1</td>
</tr>
<tr>
<td>Can</td>
<td>1,997</td>
<td>2,041</td>
<td>0.4</td>
<td>-2.2</td>
<td>2,193</td>
<td>2,160</td>
<td>0.2</td>
<td>-1.5</td>
</tr>
<tr>
<td>Could</td>
<td>1,740</td>
<td>1,782</td>
<td>2.4</td>
<td>-2.4</td>
<td>1,776</td>
<td>1,655</td>
<td>4.1</td>
<td>-6.8</td>
</tr>
<tr>
<td>May</td>
<td>1,333</td>
<td>1,101</td>
<td>22.8</td>
<td>-17.4</td>
<td>1,298</td>
<td>878</td>
<td>81.1</td>
<td>-32.4</td>
</tr>
<tr>
<td>Should</td>
<td>1,301</td>
<td>1,147</td>
<td>10.1</td>
<td>-11.8</td>
<td>910</td>
<td>787</td>
<td>8.8</td>
<td>-13.5</td>
</tr>
<tr>
<td>Must</td>
<td>1,147</td>
<td>814</td>
<td>57.7</td>
<td>-29.0</td>
<td>1,018</td>
<td>668</td>
<td>72.8</td>
<td>-34.4</td>
</tr>
<tr>
<td>Might</td>
<td>777</td>
<td>660</td>
<td>9.9</td>
<td>-15.1</td>
<td>635</td>
<td>635</td>
<td>0.7</td>
<td>-4.5</td>
</tr>
<tr>
<td>Shall</td>
<td>355</td>
<td>200</td>
<td>44.3</td>
<td>-43.7</td>
<td>267</td>
<td>150</td>
<td>33.3</td>
<td>-43.8</td>
</tr>
<tr>
<td>Ought (to)</td>
<td>104</td>
<td>58</td>
<td>13.4</td>
<td>-44.2</td>
<td>70</td>
<td>49</td>
<td>3.7</td>
<td>-30.0</td>
</tr>
</tbody>
</table>

As a study of this kind has to be limited in order to be presented and thoroughly investigated within the scope of a thesis, I have chosen to focus on just the particular auxiliary ought. I have chosen this word because it is one the modal auxiliary which, according to

In addition the diachronic development, Leech points to an interesting difference if we look at use of modal auxiliaries from a synchronic perspective. American English (AmE) tends to make far less use of these verbs than British English (BrE), and there is generally a sharper rate of decline in BrE. "It is as if BrE is following rather reluctantly in the wake of a change in AmE, with something of a generation gap" (Leech 2002:67).

In his article Leech points out that it is the rare auxiliaries which seem to experience a reduction in use, while more common modals tend to hold their own better, or even as is the case with will and can, show an increase:

Broadly, the most frequent modals decline least, and the least frequent modals decline most in percentage terms, the rare modals shall, ought to and need (+ bare infinitive) having become much rarer. Some middle-order modals (especially must and may) also show very significant falls in frequency. (Leech 2002:67)

As a study of this kind has to be limited in order to be presented and thoroughly investigated within the scope of a thesis, I have chosen to focus on just the particular auxiliary ought. I have chosen this word because it is one the modal auxiliary which, according to

1 For more information on corpora, see further section 3.3- 3.14.
Leech, has seen a large reduction in usage from 1961 to 1991. It is also one of the words which is significantly more common in use in BrE than in AmE, a distinction which may be in interesting if we consider that the reduction in BrE may be due to “Americanization” of the language. In connection with this it may also be interesting to see how the word is used in other parts of the world, so I have also looked very briefly at Australian English (AusE) and New Zealand English (NZE).

Leech suggests that the presumed rise in semi-modals, such as be going to and have to, may account for some of the reduction in the usage of modals, particularly because of what he calls ‘democratization’ of language (Leech 2003:237). By this he means a tendency for certain words to increasingly acquire face-threatening potential (see further section 1.5.), so that speakers start to prefer alternative expressions in order to avoid this problem. Leech suggests that there may be an increasing tendency for language users to “suppress or avoid overt claims to power and authority” (Leech, 2003: 237). I intend to compare the use of ought with different semantic meanings, in order to investigate the validity of this claim, and to see if the decrease can be, at least partly, explained with some kind of semantic shift when it comes to the word ought.

Leech also suggests that what he calls ‘colloquialization’ of the language may lead to less frequent use of these relatively rare modals (Leech 2002:63). By this he means “a tendency for written style to drift toward oral styles over time” (Leech 2002:78).

If we assume that language change originates in conversation, and then works its way into the written language over time, the word ought is likely to be less frequent in spoken than in written language. I therefore intend to compare spoken with written language in order to investigate if this is, in fact, the case. Gunnell Tottie also deals with the issue of ought usage (Tottie 2002: 157), stating that it is rare in both AmE and BrE, and that most language users prefer to use the alternative should, particularly in spoken language.² My investigation into the difference between spoken and written English may uncover if this is a valid claim. I will also check her claim that ought without the support of to is exclusive to BrE.

I am particularly interested in investigating the nature of the reduction in use. How exactly is the word being replaced? In what contexts and by whom is the word still being used today, and where are we experiencing a reduction? Has the implicature of the word changed in any way, thus reducing its area of use and perhaps restricting it to certain domains? What

² Interestingly, Table 1 shows that should is also on the decline in the written material from Leech’s investigation, so there is no evidence in this material to suggest that writers are increasingly opting for should as a replacement for ought.
conclusions can be drawn regarding the future of this word, and perhaps the development of the general use of auxiliaries, from this perspective?

I also intend to explore how much information on a phenomenon like linguistic change can be accessed by using a variety of computerised corpora. Linguists differ in their opinions regarding whether corpus linguistics can be employed usefully in linguistic research. I have therefore tried to use a range of such corpora in order to investigate this problem from different perspectives. Wherever possible I have therefore chosen corpus research as my main research method, rather than, or before resorting to, other methods. This involves using corpora which were not in all cases designed particularly for my purpose.

Finally, I want to look at the distribution of the usage of the word with regard to genre, speaker/writer age and any other factors which may indicate formality or user domain to see if any patterns emerge, which may give us any indication with regard to future use of this particular modal.

My primary research questions are as follows:

- Is there an ongoing reduction in the use of the modal auxiliary *ought*?
- Is the reduction in use seen in equal distribution across different types of language use?
- Is there any evidence of a change in the semantic meaning of *ought*.
- Can corpus linguistics contribute to research into change in language use?
- Can a range of corpora, even if they are not designed especially for comparative purposes, be useful in a comparative study such as this, to shed light on the mechanisms behind linguistic change?

### 1.2. Varieties of English

This study looks at the use of *ought* from a synchronic as well as a diachronic perspective. By looking at the synchronic perspective – differences across variants of English – I hope to uncover some cultural factors which may influence the use of *ought*. In this section I briefly outline what is meant by “varieties of English” and which varieties I have elected to study in this thesis.

400 years ago schoolmaster and linguist Richard Mulcaster made the claim that “The English tongue is of a small reache, stretching no further than this island of ours, nay not there over all.” (Melchers and Shaw, 2003: 6)
English has experienced an unprecedented distribution since the 18th century, and although the empire no longer exists, it is safe to say that even today the sun never sets on the English-speaking world. Estimates over how many speakers of English there are in the world vary enormously, depending on how the writers define English speakers, and how much competence of English is required in order to qualify as a speaker of English, or native speakers of English. In Mulcaster’s day “The British” would probably be a reasonable answer to the question “Who are the native speakers of English?”. But what is the situation today?

Around ¾ of people who have English as their first language today actually live in America (USA and Canada). In addition to this, English is the main language of Australia and New Zealand, and English is also the first language of groups of people in other countries, such as South Africa. The majority of these English speakers have English as their main or even only language, so there is no reason to class them as anything other than native speakers. Estimates vary according to how authors define native speakers. David Crystal (2003:61) estimates what he calls inner circle speakers of English (those from the traditional bases of English where English is the first language) to 320-370 million people, with outer circle speakers (from areas with other native languages but where English language has an important official function) to 300-500 million speakers. Melchers and Shaw (2003: 8) estimate the number of what they refer to as English first-language speakers (L1 Speakers) to be around 375 million, while and equally large group is said to consist of English second-language speakers (L2 Speakers). According to them, 750 million people have foreign language proficiency in English (EFL Speakers).

Jennifer Jenkins (2003) refers to Kachru’s classification of English as a world language, which divides the English-speaking world into inner, outer and expanding circles (Fig.1.1). The circles overlap somewhat, but the inner circle refers to the traditional bases of English, where the language first travelled to when leaving its island home, and where English is the primary language. The inner circle thus consists of the USA, the UK, Canada, Australia and New Zealand.

This paper concerns itself with inner circle varieties, primarily The UK and the USA, as the English spoken in the inner circle is said to be norm-providing (Jenkins, 2003:15-16). Fig.1.1 below is based on the numbers suggested by Melchers and Shaw.

**FIG. 1.1 – An illustration of Kachru’s circles of English.**
With the geographically widespread distribution, even of inner circle or native speaking areas of English, it is not surprising that the areas have developed distinctively different varieties of the language, given the variety of extralinguistic as well as intralinguistic factors governing language change, which I will go on to discuss in some detail in section 1.3. The most obvious differences are perhaps the widely reported phonetic differences. Even in written English, however, there are a number of striking differences between varieties of inner circle English. These differences include grammatical as well as lexical differences. In general, British and American English are thought to be the two varieties displaying the most diversity, with Australian and perhaps even more so, New Zealand English, tending to side with British English. (Melchers and Shaw 2003: 104-111).

Noah Webster’s work, which created an American standard as early as in 1789 is probably, at least in part, responsible for this. Australian and New Zealand varieties followed British standards until much more recent times, and the spelling standards have thus continued to follow the development of British English until fairly recently.

These differences, however, are not absolute. In most cases there is a tendency in each area to follow certain conventions, but there are exceptions in many cases. One example is the word theatre, which is often said to be spelled theatre in BrE and theater in AmE. A corpus search in FLOB and Frown, however, reveals that BrE in general adheres to the British standard, while the American Frown corpus reveals that in as many as 59 of the 99 instances,
the so-called British variant is used, while the “American” standard is only employed in 40 instances.

It must also be emphasized that a division into AmE, NZE, AusE and particularly BrE is a gross simplification, particularly when dealing with lexical differences. There are plenty of dialectal varieties within the regions as well. As Melchers and Shaw point out:

Linguistically [...] English English is closer to Australian, New Zealand and South African English than to Scottish or Irish English [...] A country-by-country approach is unproblematic in a case like New Zealand where a variety of English is used within a well-defined geographical area which is also a political unit, but we are very much aware that this is not always the case (Melchers and Shaw, 2003:42)

In varieties of English, such as we are dealing with here, however, the regional varieties are codified into a written standard, unlike with dialectal variants. This study, therefore, concerns itself with differences between the main native speaker varieties, paying little attention to dialect variation.

For this particular study, I am studying tendencies. Ought is not limited to one variety, but may be used to varying degrees and somewhat differently between the main varieties of English. By using corpora, I will quantify and attempt to interpret such tendencies and stylistic preferences in order to explain and perhaps even predict language change.

1.3. Language change.

Language is, by nature, a living thing. It changes and adapts when coming into contact with users. Consider for example this line from Chaucer’s prologue to The Pardoner’s Tale, which appears strange and foreign to most modern readers:

I peyne me to han an hauteynspeche. (Burgess 1987:194)

It looks and sounds like a foreign language, but is in fact, 14th Century English. If a language, over time, changes as much as the above quote by Chaucer suggests – even within the country of its origin, it is unreasonable to believe that it should not change and be adapted in areas where it is exposed to a well of other languages, cultural influences, as well as technological and physical conditions (such as climate). Gunnel Tottie (2002: 91) distinguishes between intralinguistic (within the language) and extralinguistic (the contextual situation in which the language is used) forces behind language change.

One of the main arguments for a range of diverse varieties of English is that the political situation over the last three to four hundred years has lead to an unprecedented spread of English throughout the world – first through the British Empire, and later through
the economic, technical and financial dominance of Britain and later of the USA. Even within England, language is constantly changing.

In Middle English there was no standardisation of spelling, and scribes would tend to write as they spoke. As the level of literacy increased, it was considered that a common standard was needed. The result was dictionaries, as first published in England by Dr. Johnson in the 18th century. Although written standards also adapt and change over time, they are more permanent than spoken standards, and tend to control or slow down the process, ensuring that the language remains intelligible to all users.

An initiative by Noah Webster as early as 1789, set out to create a separate General American standard of English, as part of the establishment of American independence (Melchers and Shaw 2003:14). This created the foundations for a separate American English. Australian and New Zealand standards, however, were laid down much closer to our own time, and were generally based on the British variant. Language has a tendency to become simplified over time, and this can clearly be seen in the spelling differences between British and American words, like theatre/theater, programme/program and defence/defense.

American and British English have had more time to develop differently during centuries of limited contact between the varieties. In the past century or so, however, communications have been vastly improved, and with the development of a global economy and media society increasingly led by the USA, many comments have been made regarding the influence of American culture and language on the rest of the world, and in particular on the English-speaking world. Tottie (2002: 147) claims that British English is nowadays being influenced by American English and […] several grammatical features that used to be regarded as typical of American English are now no longer exclusively American.

Charles Barber (1964:141) goes as far as saying that “American English is an accelerator of change in BrE”.

In an attempt to discover whether these claims are accurate, Marianne Hundt (1997:147) performed a series of studies, concluding that:

AmE, with the occasional exception, is usually more advanced in ongoing morphological and syntactic changes. AmE thus turned out to be the ‘centre of gravity’ in most cases, even if the changes do not affect the general rules but a shift in stylistic preferences.

As this study is particularly concerned with such stylistic preferences, one interesting aspect is investigating whether or not AmE is ‘the centre of gravity’ in the case of ought as well.
1.4. Spoken and written language.

When using the terms text and discourse in this paper, I am simply applying them to pieces of language in use, and not limiting them to written text or discourse. There are a number of key differences between spoken and written language. The main difference is, of course, the transience of spoken language, compared with written texts.

I am using corpus material in order to conduct my research for this paper, and the spoken texts are transcribed. Although the encoding of the language which occurs in the process of transcription has the effect of converting it into written text, I have chosen to define it as spoken material. This is simply because it is the closest we are able to get to spoken material when working with corpora.

There are a number of challenges with the use of spoken language corpora. When encoding spoken material and converting it to written texts, the makers of the corpora have tried to stay as true to the material as possible and give an as complete as possible account of the exchanges. Despite of their diligence, however, we unavoidably lose many contextual features of the discourse. The transcriptions reveal that some of the material is too unclear in the recording to be heard (examples 1 and 2), or the sentence is broken off before it is completed as in example 3.

(1) Course I ought to be more <unclear syllable> <unclear word> uneven. (DCPSE DL-B17 0754)

(2) <voice quality: laughing> You ought to <unclear> <end of voice quality>! (BNC F7U 916)

(3) Yeah it ought to be worth it cos… (COLT text)

We also lose extralinguistic features like tone of voice, facial expressions and body language, as well as shared experience or knowledge between the participants regarding the subject matter. In the example 4 below, for instance, we have no way of knowing we have no way of what the subject (Comley) is supposed to have got in where. Nor is it clear what sort of contributions are referred to in example 5, but it is obvious that the speaker expects that the intended listener can imply this from the situational context or from the context of shared previous knowledge. Sometimes this is not essential information in order to classify an example, but it may be. These features may be semantically significant, and semantic classification of this spoken material will almost certainly be less reliable as a result of these challenges. E.g. In the two examples below (examples 4 and 5) it is difficult to determine whether to classify the semantic meaning of ought and determine whether it is used here with
epistemic meaning, some type of moral expectation or judgement or expectation of an event, either because the context is unclear, or because it is semantically dependent on the extralinguistic features which are present in the situational context or within the common knowledge of the participants, but unavailable to me when using the corpus.

(4) Yes— Uh. You see the other man, Comley, ought also to have got his in on time. (DCPSE DL-B01 0047)

(5)(X) John I think this, this is wh—, I think this ought to be the last contribution. <unclear> <unclear> (BNC F8R 987)

Without the full context, researchers are forced to rely on their own judgement, powers of reasoning and sometimes even guesswork to a much greater extent than the original participants of the exchange. That is certainly at times unreliable, and will probably result in some mistakes in the classification. Alas, these are some of the limitations which are imposed on linguists when working with corpora containing transcripts of spoken texts.

The problems associated with spoken language corpora are unavoidable. On the other hand, however, corpora containing spoken texts provide us with a broad range of spoken material which can be quickly processed on the computer. I believe the advantages of this far outweigh the limitations and justify the inclusion of such a lot of spoken material in the investigation.

Spoken language, of course, covers a range of linguistic contexts ranging from planned speech, which is written to be spoken and shares many characteristics with written language to more informal exchanges. The more informal register of conversation, however, is quite different in many respects. Consider the conversation below, taken from the BNC:

(6)
A: <laugh>
Cor!
B: <pause> It’s like looking at the mirror innit?
C: A receding hair line is a sign of intelligence.
B: Thanks Al!
A: <laugh> <unclear>
B: Obviously <pause> <unclear> <pause> obviously it’s a genetic throwback with Chris, it’s not worked with him!
C: Oh!
Ha ha!
<pause> Throw up, yeah!
B: Right
A: <laugh>
B: an <unclear> used to be called <pause> metermarism <pause> and I say, it’s not a word <pause> it’s not a term I’ve seen on <unclear> for many many years <pause> and you don’t seen it many <pause> erm <pause> modern <pause> organic texts. I think it’s still around <pause> and they’ve gotta call it something <pause> if erm <pause> if you think C: <unclear> word. B: Right now! <pause> Stereo-isomerism next <pause> exhausted <pause> icuma C: <voice quality: laughing> You ought to <unclear> <end of voice quality>! Not a lot! B: We’ve exhausted the structural side of isomerism, now we can look at the stereo isomerism. (BNC F7U)

Example 6 above is a fairly typical conversation between three people, but understanding it poses a problem for the reader. Biber et al (1999:1041) point to similar difficulties when interpreting a piece of conversation:

The shared background as well as the shared physical and temporal space required to fully understand this excerpt are considerable. In this respect, although the difficulty of making sense of it on the page may be an unfamiliar and disorienting experience for many readers, the extract is typical for conversation.

They go on to list several characteristics of conversation. For further information on the particular features of conversational language, I refer to chapter 14 in Biber et al (1999) in its entirety.

One feature typical of spoken discourse is its immediacy. While written texts are often changed and rewritten before being presented, spoken discourse can not be taken back and re-done. This is, of course, particularly true of spontaneous spoken discourse.

As a result of this immediacy, spoken texts typically display features like false starts, repetitions and comment clauses, which are generally not found in written material (see examples 7-10 below)

(7) But erm <pause> I th-- <pause> they are worried <pause> certainly, one or two of them are worried, but er <pause> some of this equipment is left on, and left on unnecessarily <pause> and we ought to be encouraging people to switch off what isn’t (BNC FLS 335)

(8) I ought to ask Malcolm if he c-- I I don’t know what day you'll be in? (BNC FXR 493)
(9) The <> the this this you ought to ought to hang on hang on to it (DCPSE DL-B22 0195)

(10) I mean, like you say, you ought you know <pause> you ought to buy her a talking alarm clock for Christmas. (BNC KFB 992)

As David Crystal (1992:126) points out, this does not mean that there is any truth to the claim that conversational language lacks grammar. Comment clauses, for example, are not inserted randomly. Their position in a sentence is governed by strict semantic considerations. These considerations, however, frequently differ from those which govern the grammar of written language. Because of its immediacy, however, spoken language is likely to be more susceptible to language change.

1.5. Politeness theory

One of Leech’s suggestions as to why ought is undergoing a reduction in use is to refer to politeness strategies. I will therefore briefly account for some of the main theories in the field of politeness in order to explain why this may be relevant here.

The theories originally put forward by Brown and Levinson in 1987 remain among the most influential in the field of politeness theory. Central to this theory is the concept of ‘face’, which is best understood as a person’s ‘self-worth’. According to Brown and Levinson, certain illocutionary acts or speech acts, can either threaten or enhance a recipient’s ‘self-worth’, and such acts may also threaten the speaker’s ‘self-worth’ if the response to the utterance is undesirable.

Face is something that is emotionally invested, and that can be lost, maintained, or enhanced and must be constantly attended to in interaction (Brown and Levinson 1987: 61)

Brown and Levinson distinguish between what they refer to as positive and negative face.

Negative face is the desire to be autonomous and not to infringe on or be infringed upon by the other person. Brown and Levinson define negative face as “The want of every ‘competent adult member’ that his actions be unimpeded by others.” (Brown and Levinson 1997:62). Positive face is the desire to be liked and appreciated, or as Brown and Levinson put it:” The want of every member that his wants be desirable to at least some others” (Brown and Levinson 1987:62).
Face Threatening Acts (FTAs) are speech acts which threaten the receiver’s positive or negative face. In order to minimise the effects of such FTAs, speakers tend to employ certain strategies in order not to cause offence.

Negative politeness has the function of minimising impediments on the other member of an exchange, while positive politeness has the function of meeting the face needs by performing an action like complimenting or showing concern for another person. Brown and Levinson outline four main types of politeness strategies language users utilise in order to improve face; bald on record, negative politeness, positive politeness, and off-record or indirect strategy.

Bold on record statements are direct statements where the speaker says exactly what they mean without any redress or attempt at reducing any threats to the recipient’s face. This is often used in advertising or, as in example 11 below taken from the Labour Party campaign material for the UK 2008 local elections, in slogans, where the aim is to get a message across with few words.

(11)

It may also be used in situations where one member of the exchange is able (or even expected to) exercise power over another, for example by a commanding officer in the armed forces (Example 12), a teacher in a classroom (example 13) or a judge giving a command in a courtroom (example 14).

(12) Forward march!

(13) Sit down quietly everyone and open your books on page 52.

(14) Take him down!

3 The examples in this section are not real exchanges taken from corpora. I have invented them for the purpose of demonstration.
These exchanges are certainly the most economical forms of communication, and well suited to get a message across clearly, with little room for misunderstanding or disagreement. They are, however, likely to be interpreted as commands, and not expected to be used unless the context allows for it. In most contexts speakers will use politeness strategies in order to minimise the face threatening potential of the utterance.

Negative politeness appeals to the recipient’s desire not to be impeded upon. It is thus oriented toward the recipient’s negative face and at least appears to give options as in example 15, or offering an exchange of goods or services which apparently benefits the hearer too (example 16).

(15) Would you mind sitting down and opening your books, please?

(16) If you sit down quietly and open your books on page 52 as quickly as you can, we can all go home a little earlier today.

The threat an utterance poses to the hearer’s negative face may also be reduced e.g. by dressing the request as a reminder (example 17) or as a suggestion or piece of advice (example 18)

(17) Have you remembered to do the washing up?

(18) You really ought to get the washing up done and out of the way, you know.

Negative face may also be addressed to by acknowledging the imposition and, for instance, promising a return of the favour at a late date, as in example 19 below:

(19) I know you’re busy, but if you would be so kind as to help me with my paper now, I promise that I’ll do anything you want me to afterwards.

The speaker may also try to minimise the size of the imposition, as in example 20.

(20) Could you just spare me one second and read through this? It shouldn’t take long.

Positive politeness orients itself toward the recipient’s positive face. Common strategies includes various ways of appealing to recipients’ desire to be liked or approved of (example 21), as well as reducing the feeling of social distance or unevenness in the balance of power between the participants, for instance by pointing to common benefits or goals (example 22).

(21) I would so appreciate it if you would do this for me.
Everyone in this company, regardless of rank, really needs to work hard in order for us to become the best in our field.

Utterances which reveal that the speaker is indifferent to, or oblivious to the hearer’s feelings are a potentially a threat to the hearer’s positive face.

Off-record or indirectness strategies use indirect language in order to not appear to be imposing (example 23).

(23) I would just love it if someone took me out for a meal next weekend.

Jenny Thomas states that speakers:” Tend to employ indirectness to a greater extent with people who have some power of authority over us than those who do not” (Thomas 1995: 124)

She also explains that this includes three types of power. Firstly, a person may have legitimate power where the status of the person by virtue of age, role or status gives them the right to prescribe or request something from the person the utterance is directed at. Secondly a person may have referent power in the sense that the speaker may admire them or wish to be liked by them. Thirdly, they may have expert power, where they have power by virtue of possessing some special knowledge or expertise that the other participant in the exchange relies upon. These separate categories of power are not mutually exclusive. In practice, of course, a combination of two or all of these is not uncommon. Naturally, politeness also tends to increase with the size of the imposition.

In the case of ought, the presumption that the reduction of use is tied to politeness implies that the potential for the utterance to be face-threatening is increasing. If that is indeed the case, speakers will try to find alternative constructions as a face-saving strategy.

If ought is indeed in the process of acquiring increased face-threatening potential, we would perhaps expect to see a shift in the semantic meaning of ought. If the semantic meaning changes, and ought is thought of as more face-threatening in modern texts, then a reduction in use is a likely consequence because speakers will prefer to only use it if they are in a position to impede upon the recipient as a result of being in a position of power or familiarity which allows for such impositions. In domains where this is not the case, use of ought will be restricted.
2) Modal auxiliaries in general, and ought in particular.

2.1. Defining modal auxiliaries

Modal verbs provide an important semantic function in the English language. Kennedy (2002: 79) points out that

Verbs constitute about 20 per cent of all the word tokens used in English, and in written texts, modal verbs typically constitute around 8 per cent of all verb forms.

**Modal auxiliaries** Biber et al (1999:73) or **Central Modals** (Quirk et al 1985: 137) are defined as “verbs used to build up complex verb phrases, and unlike lexical verbs they can not occur alone. Modal auxiliaries do not have non-finite forms” (Biber et al 1999:73). Both these works define nine central modal auxiliaries: *can, could, may, might, must, shall, should, will and would*.

Verbs like *dare (to), need (to), ought to* and *used to* are regarded as **marginal modals** by Biber et al (1999:73), and Quirk et al (1985: 136), because they sometimes behave like ordinary modals. The marginal modals are sometimes constructed with what Svartvik and Wright (1977:179) refer to as *do*-periphrases and a following *to*- infinitive as in the following sentence: *Do we need to go?* They also include *have (to)* among the marginal modals. Greenbaum (1996:80) includes *ought* amongst the modal auxiliaries while Leech (2002:66) includes both *ought* and *need* among the modal auxiliaries in his study. Keith Mitchell (2003) has performed an investigation into the constructions *had better* and *might as well*, concluding that:

What I hope to have demonstrated in this paper is that *had better* and *might as well* share the majority of the morphological, syntactic and semantic properties of the central modal auxiliary verbs of English and deserve to be counted among them (Mitchell, 2003:147)

The notion of modal auxiliaries therefore appears rather slippery and the definition far from clear-cut.

As my investigation is based on the previously mentioned study performed by Geoffrey Leech (2002), I have elected to adhere to his definition and therefore regard *ought* as a belonging to the category of modal auxiliaries, and have not differentiated between marginal modals and modal auxiliaries. I have rather chosen to regard them all as belonging to the same category. This is not because I do not recognize that a distinction between marginal and central modals, but rather that for this particular study the difference is of no great consequence as it is concerned with meaning and use rather than syntax.
2.2. The meaning and etymology of the modal auxiliary *ought*.

*The Oxford English Dictionary (OED)*\(^4\) defines the word *ought* as:

Expressing duty of any kind, originally used of moral obligation, but also in various more general senses, expressing what is proper, correct, advisable, befitting, or expected. Originally and chiefly in past tense form (indicative or subjunctive), which may be either past or present in meaning. (The only current use in standard English.)

The *OED* also specifies that the subject is properly the person bound by the obligation. As such the meaning is similar and may in some cases be used synonymously with other modals, such as *should* and the more emphatic *must*.

Hasselgård, Johansson and Lysvåg (2001:199) identify two functions of *ought*. The first is to indicate epistemic modality (where the speaker thinks X is the case), and root modality (indicating some kind of obligation). Merriam-Webster, however, give four separate functions of the word: express obligation *ought to pay our debts*, advisability *ought to take care of yourself*, natural expectation *ought to be here by now*, or logical consequence *the result ought to be infinity*. The first two categories largely belong in Hasselgård et al’s category of epistemic modality, while natural the expectation and logical consequence meanings of *ought* largely correspond to the notion of root modality.

According to Huddleson and Pollum (2002:186-187) *should* and *ought* are interchangeable and denote medium strength epistemic (usually subjective with a strength comparable to *probably*, but involving inference), or deontic indicating what the speaker thinks is ‘right’ either morally or as a matter of expediency) modality. They point out that the deontic use is more basic than the epistemic, as an epistemic reading is hardly possible with past time situations. According to Huddleston and Pollum one significant way in which *ought* differs from *should/shall* is that it can not be used is the same way to mark low-degree modality, by which they mean with little discernible modal meaning of its own.

Etymologically, *The Oxford English dictionary online* states that this word stems from the old Swedish *owe*. It occurs sporadically in a modal verb sense in English from the 16\(^{th}\) century.

*The Merriam-Webster online dictionary* includes the following etymological information:

Middle English *oughte* (1st & 3d singular present indicative), from *oughte*, 1st & 3d singular past indicative & subjunctive of *owen* to own, owe — more at *owe* Date:12 century

By 1961, Leech’s investigation based on the four corpora, LOB, FLOB, Brown and Frown, reveals that it is much more common in British than in American English. By 1991 the differences between the two varieties are much less notable. They both show a decrease, but the decrease is much more dramatic in British English. This variety appears to have become more similar to American English with regards to the use of \textit{ought}. (see table 1.1).
3) Research method

3.1. Qualitative v.s. quantitative method.

There is a long-standing argument about which is the best approach to linguistic research – a qualitative or quantitative approach. What is meant by this and are the two, in fact, mutually exclusive?

Reichardt and Cook (1979: 10) have devised this table (2.1) outlining the different attributes of qualitative and quantitative paradigms

<table>
<thead>
<tr>
<th>Qualitative paradigm</th>
<th>Quantitative paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Advocates the use of qualitative methods</td>
<td>● Advocates the use of quantitative methods</td>
</tr>
<tr>
<td>● Phenomenonologism and verstehen. ‘concerned with understanding human behaviour from the actor’s own frame of reference’</td>
<td>● Logical-Positivism. ‘seeks the facts and causes of social phenomena with little regard for the subjective states of individuals.</td>
</tr>
<tr>
<td>● Naturalistic and uncontrolled observation</td>
<td>● Obtrusive and controlled measurement</td>
</tr>
<tr>
<td>● Subjective</td>
<td>● Objective</td>
</tr>
<tr>
<td>● Close to the data; ‘the insider perspective’</td>
<td>● Removed from the data; ‘the outsider perspective’</td>
</tr>
<tr>
<td>● Grounded, discovery-oriented, exploratory, descriptive and inductive</td>
<td>● Ungrounded, verification-oriented, confirmatory, reductionist, inferential, and hypotheticodeductive.</td>
</tr>
<tr>
<td>● Process-oriented</td>
<td>● Outcome-oriented</td>
</tr>
<tr>
<td>● Valid ‘real’, ‘rich’ and ‘deep’ data</td>
<td>● Reliable ‘hard’ and replicable data</td>
</tr>
<tr>
<td>● Ungeneralizable. Single case studies.</td>
<td>● Generalizable; multiple case studies</td>
</tr>
<tr>
<td>● Holistic</td>
<td>● Particularistic</td>
</tr>
<tr>
<td>● Assumes a dynamic reality</td>
<td>● Assumes stable reality</td>
</tr>
</tbody>
</table>

(Table quoted from Larsen-Freeman and Long, 1991:12)
The use of corpora for research is clearly largely a quantitative approach, although some of the qualitative paradigms, like the emphasis on naturalistic and uncontrolled observation, are also present.

It allows us to study language empirically and count number of occurrences of the phenomena we are interested in. It has been criticized because it measures linguistic *performance*, rather than *competence*. Corpora are collections of utterances. This is by its very nature not a complete survey if the internalised knowledge possessed by native speakers of a language. The phrases in example 24 and 25 are not found at all in the 100 million word British National Corpus, yet claiming that they are impossible or incorrect utterances in English would seem quite ridiculous.

(24) I will work from home today.

(25) The best actor in the play.

In fact, most speakers of English would probably think of these utterances as quite ordinary, and probably even be able to think of instances where they might be used. They do not seem any less likely, if we use introspection, than examples 26 and 27, which are, in fact, found in the BNC.

(26) I’m going paddling today (BNC ACK 2855)

(27) The only innocent person to die in the play. (BNC AN4 1422)

Likewise, using corpora in order to uncover grammatical constructions has to be done by using qualitative/introspective criteria as well. The construction *the black cat* is found 18 times in the BNC. *The big cat* is found 20 times. We cannot deduct from this, assuming that we have no prior knowledge of the word “cat”, that any adjective may be inserted between the article and “cat”. *The square cat* and *the white cat* are both found three times in the BNC. This should not be taken to mean that they are both as likely as each other. Most native speakers would claim that the former would need a rather particular context in which to occur, while the latter is much more likely.

Thus, turning to corpora in hope of finding out all possible, or even likely utterances in English simply by counting number of occurrences would not be a particularly good approach. As McEnery and Wilson point out:

---

5 They are, in fact, not taken from anywhere in particular. I simply made them up.
6 In this case, “The square cat” turned out to be the name of a bistro.
The number of sentences in a natural language is not merely arbitrarily large. It is no use sitting around speculating about the number of sentences in a natural language. The number is uncountable – the number of sentences in a natural language is potentially infinite. (MacEnery and Wilson, 1997:7)

A corpus then, however large, only contains a selection of possible constructions. However, the corpus does have one big advantage. The data is observable and it allows us to measure the extent to which a phenomenon occurs in language. No other approach allows us to “weigh and measure” language in the same way. Working with corpora also makes it quite clear that frequency and change in frequency is not something native speakers are particularly good at recognizing. We may notice some things in the language more than others, and therefore assume that they are more frequent than they actually are. A quantitative approach like this is probably more reliable than introspection when measuring something like the reduction in use of a modal auxiliary.

Quantitative approaches like corpus research also has the advantage of being replicable. Any researcher using the same corpus looking for the same construction will find the same results. Using different corpora is one way of controlling whether chance, or indeed the compilation methods used in the various corpora, is in any way responsible for the findings. The more confirmation we gain from searches in different corpora, the more certain we can be that a phenomenon is really occurring. By looking at how the different corpora are compiled, we can investigate how and in which contexts the phenomenon occurs.

In the case of modal auxiliaries, we can measure and compare the use across different varieties of English, as well as over time. By checking occurrences in different corpora, while at the same time applying qualitative criteria and using introspection, we can say something about the extent to which a change is occurring as well as perhaps where and how.

McEnery and Wilson observe that: “a stage of qualitative research is often a precursor for quantitative research, since, before linguistic phenomena are classified and counted, the categories for classification must first be identified.” (McEnery and Wilson, 1997: 62). Thus, by using corpora, the two approaches may to a certain extent be united. “Qualitative analysis can provide greater richness and precision, whereas quantitative analysis can provide statistically reliable and generalisable results” (McEnery and Wilson, 1997: 63).

The present study is a combination of the two approaches. It is quantitative in the counting of frequencies to reveal systemic grammatical and semantic variation and change rather than being arbitrary. It is, however, qualitative in taking meaning and context into account.
3.2. Corpus linguistics and the use of corpora in a contrastive study.

Corpus linguistics has revolutionized the way in which language and the use of language may be researched. The term “corpus” comes from the Latin “body” and in this instance simply refers to a body of text. Any collection of text selected for analysis may become the corpus in a study. Corpus linguistics generally, however, tends to be taken to mean the use of computerized corpora which are available to modern linguists. These range from large, relatively generalized corpora like the BNC, Bank of English and the ‘Brown Family’ of corpora (I say relatively generalized because even these corpora are restricted to certain domains), which aim to represent a cross section of language in use, to smaller corpora, like COLT which are designed specifically to look at certain types of language in particular.

Generative linguists, like Noam Chomsky, have criticized the use of corpora in linguistic research. Chomsky is one of the harshest critics of corpus linguistics as a research methodology. He has suggested that the corpus could never be a useful tool for the linguist, as the linguist must seek to model language competence rather than performance, and when asked his opinion of modern corpus linguistics in an interview with Bas Aarts, he claimed that: “It doesn’t exist” (Aarts 2000:6).

He goes on to argue that language research should be conducted in a similar manner as is used in the natural sciences:

You don’t take a corpus, you ask questions. You do exactly what they do in the natural sciences. You do experimentation. A scientist doesn’t take photographs of what is happening in the world, and then try to reduce it to data (...) One of the big insights of the scientific revolution (...) is that arrangement of data isn’t going to get you anywhere (Aarts 2000:6)

Chomsky argues that linguistic performance – the actual utterances produced in an exchange, is a poor measure of linguistic competence, i.e. the internalised knowledge of the language that the language user possesses. Language research, according to Chomsky, should be concerned with rule-governed approaches, rather than description: "A corpus never tells you what is impossible. In fact, it doesn’t even tell you what is possible” (Aarts 2000:6).

Grammarians who are critical of the use of corpora argue that if we are unable to measure linguistic competence how can we know if what we are discovering are relevant linguistic phenomena? Michael Stubbs (2001:60) feels this is an over-simplifying polarization between competence and performance, and points to the way in which Hymes unites the two

---

7 This term is used by Leech and Smith (2005) and refers to “English language corpora modelled on the original Brown university corpus, such as LOB, FROWN, FLOB, Wellington etc.” (Leech and Smith, 2005: Abstract)
concepts. Hymes (1992:31-59) discusses if a sentence is formally possible, but also looks at whether it is psycholinguistically feasible as well as sociolinguistically appropriate.

As Charles Meyer (2002: preface) argues that corpus linguistics is not as much a separate discipline within linguistics as, for example, sociolinguistics or psycholinguistics, but rather a method by which one can study language. In other words, corpus linguistics does not supply the linguist with a theoretical framework for linguistic analysis. It does, however, constitute a tool which I would argue that linguists from most or, if not all, of the various disciplines may find it useful to apply in their research. It does not in itself provide us with answers, but if we have questions regarding the extent to which a phenomenon occurs in language, corpora allow us to quantify this. We can ‘weigh and measure’ language using corpora, and say something about not only to what extent phenomena do occur in language, but also say something about how language varies between domains (social context, geographical areas text types, oral/written language, native/learner language etc.). Depending on the corpora the linguist chooses to use, the possibilities are almost unlimited. That allows us to make predictions and comment on how a language tends to function when used in actual real-live situations.

Critics of corpus linguistics prefer to rely on native speaker intuition. However, all native speakers are part of a society. Class, background, nationality, gender, level of education etc. may influence the choices they make when using language. Corpora open up for researchers to use quantitative method in order to describe how speakers in general behave. As Wendy Anderson (2008:7) puts it: “Corpora allow us to detect and describe variation at every level from idiolects to national variation, and these corpora do not need to be explicitly designed for these purposes.”

Its focus is on description, rather than the universal (Granger 1998:3), and corpus linguistics is a descriptive, rather than a rule-governed approach to studying language.

As Filimore puts it:

I don’t think there can be any corpora, however large, that contain information about all of the areas of the English lexicon and grammar that I want to explore…[but] every corpus that I have had the chance to examine, however small, has taught me facts I couldn’t imagine finding out any other way. My conclusion is that the two types of linguistics need one another. Filimore (2002) quoted in Semino and Short (2004:8)

For this particular investigation, a corpus study is particularly well suited. When investigating language change, we are not particularly concerned with how it is possible to use language, or the correctness of utterances, but rather how language is actually used at various times, in different geographical areas, or in particular domains. Despite its limitations,
corpora do provide us with the opportunity to investigate this, and thus may shed some light on the extent to which there is a language change in progress, and also where and perhaps even what factors are instrumental in driving the change forward.

3.3. The corpora used for this study

Although, as I have mentioned previously, in theory any collection of texts may be referred to as a corpus, the term is today generally taken to refer to electronic, machine-readable corpora. For this study, I have used 11 different such corpora in order to gain insight from a varied selection in order to get a broad basis for comparison. I will present the corpora in order to clarify why I have chosen these particular corpora and also in order to say something about their strengths and limitations. I base my descriptions to a large extent on the manuals of the respective corpora, which are available online. Readers in need of more information would be well advised to view these websites for themselves. I will give the web addresses in a footnote for each section.

I am working mainly with static corpora in this study. The texts selected are permanently added to the corpora, and the corpora in this study are in a sense “frozen in time”, allowing scholars to investigate different phenomena using the same texts. Of course, in such static corpora, the language tends to become rather dated after a while, and most of the modern corpora in this investigation are made up of texts which are at least 15-20 years old at the time of writing this thesis. An alternative to this is using dynamic corpora, where new texts are routinely added. Such corpora include, for example, the Harper Collins COBUILD team and Birmingham University’s Bank of English.

The brand new Brigham Young University Corpus of American English (BYUC), which I have made use of in this study, is also a dynamic corpus, and will be updated twice a year. However, as it only became available to me in March 2008, I have only made use of the very first version, and for obvious reasons not the future updates. However, the first version of this particular corpus includes material from 1990-1997, so this corpus has given me the opportunity to study some very recent American English material.

3.4. BROWN

8 The manual for the Brown corpus may be accessed at:
http://www.hit.uib.no/manuals/brown/index.html
http://khnt.hit.uib.no/icame/manuals/brown/INDEX.htm
The Brown Corpus of Standard American English (Brown) was the first of the modern, computer readable, general corpora. It was compiled by W.N. Francis and H. Kucera at the Brown University, Providence, RI.

It consists of around one million words, and was, as such, a large corpus when it was compiled. Although it is not among the largest available corpora today, it is still widely used for linguistic research. It became the founder member of the so-called ‘Brown Family of Corpora’. ‘The Brown Family’ consists of six corpora, all compiled in a similar manner, and provide an excellent tool for comparison of different varieties of English. I have used all of them in this study and will describe each of them in the following sections.

The texts in Brown are American, written texts, which were published in 1961. There were 500 texts, from which 2000 words have been taken. The texts are sorted into 15 text categories in order to provide a broad basis for comparison between genres. The number of texts in each category varies (see table 2.2 below).

<table>
<thead>
<tr>
<th>Non-Fiction</th>
<th>Fiction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – Press - Reportage</td>
<td>K – General Fiction</td>
</tr>
<tr>
<td>44 texts</td>
<td>29 texts</td>
</tr>
<tr>
<td>B – Press - Editorial</td>
<td>L – Mystery</td>
</tr>
<tr>
<td>27 texts</td>
<td>24 texts</td>
</tr>
<tr>
<td>C – Press- Reviews</td>
<td>M – Science</td>
</tr>
<tr>
<td>17 texts</td>
<td>6 texts</td>
</tr>
<tr>
<td>D – Religion</td>
<td>N – Adventure</td>
</tr>
<tr>
<td>17 texts</td>
<td>29 texts</td>
</tr>
<tr>
<td>E – Skill and Hobbies</td>
<td>P – Romance</td>
</tr>
<tr>
<td>36 texts</td>
<td>29 texts</td>
</tr>
<tr>
<td>F – Popular Lore</td>
<td>R – Humour</td>
</tr>
<tr>
<td>48 texts</td>
<td>9 texts</td>
</tr>
<tr>
<td>G – Belles Lettres, Biography, Essays</td>
<td></td>
</tr>
<tr>
<td>75 texts</td>
<td></td>
</tr>
<tr>
<td>H – Miscellaneous; Government &amp; House</td>
<td></td>
</tr>
<tr>
<td>30 texts</td>
<td></td>
</tr>
<tr>
<td>documents</td>
<td></td>
</tr>
<tr>
<td>J – Learned</td>
<td></td>
</tr>
<tr>
<td>80 texts</td>
<td></td>
</tr>
</tbody>
</table>

Its size, varied compilation and similar structure to the other ‘Brown Family’ corpora are among the main strengths, and makes it ideally suited for a comparative study of this kind. There are, however, some weaknesses. Firstly, there is only written material available in this corpus. Also, there are only published texts, so no personal letters or private, more informal material. Nor is there any poetry or drama represented in this corpus. This entails some major limitations as far as being representative for language in use. However, it still gives a good variation of linguistic domains, and provides a very good starting point for a study such as this.
3.5. LOB

The Lancaster-Oslo-Bergen Corpus (LOB) was compiled as a British English counterpart to the Brown Corpus.

It is the result of cooperation between the University of Lancaster, the University of Oslo, and the Norwegian Computing Centre for the Humanities at Bergen. During 1970-1976 the project was conducted at the University of Lancaster, the Department of Linguistics and Modern English Language, under the direction of Geoffrey N. Leech, before being moved to Norway in 1977. There, it was directed by Stig Johansson, Department of English at The University of Oslo. The LOB Corpus was completed in 1978.

Table 2.3 – Text categories in the LOB corpus.

<table>
<thead>
<tr>
<th>Non-fiction</th>
<th>Fiction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – Press; Reportage</td>
<td>K – General Fiction</td>
</tr>
<tr>
<td>44 texts</td>
<td>29 texts</td>
</tr>
<tr>
<td>B – Press - Editorial</td>
<td>L – Mystery</td>
</tr>
<tr>
<td>27 texts</td>
<td>24 texts</td>
</tr>
<tr>
<td>C – Press- Reviews</td>
<td>M – Science</td>
</tr>
<tr>
<td>17 texts</td>
<td>6 texts</td>
</tr>
<tr>
<td>D – Religion</td>
<td>N – Adventure</td>
</tr>
<tr>
<td>17 texts</td>
<td>29 texts</td>
</tr>
<tr>
<td>E – Skills, Trades and</td>
<td>P – Romance</td>
</tr>
<tr>
<td>Hobbies</td>
<td>29 texts</td>
</tr>
<tr>
<td>38 texts</td>
<td></td>
</tr>
<tr>
<td>F – Popular Lore</td>
<td>R - Humour</td>
</tr>
<tr>
<td>44 texts</td>
<td>9 texts</td>
</tr>
<tr>
<td>G – Belles Lettres,</td>
<td></td>
</tr>
<tr>
<td>Biography, Essays</td>
<td></td>
</tr>
<tr>
<td>77 texts</td>
<td></td>
</tr>
<tr>
<td>H – Miscellaneous</td>
<td></td>
</tr>
<tr>
<td>30 texts</td>
<td></td>
</tr>
<tr>
<td>J – Learned</td>
<td></td>
</tr>
<tr>
<td>80 texts</td>
<td></td>
</tr>
</tbody>
</table>

It is largely compiled in the same way as the Brown Corpus, and thus shares the strengths and limitations pointed out in section 2.1.

3.6. FROWN\(^\text{10}\) and FLOB\(^\text{11}\).

The Freiburg-Brown Corpus of American English (Frown) was compiled on the initiative of Christian Mair at Germany’s Freiburg University. It was created to mirror the Brown Corpus, but with American English language from the early 1990s. Compilation and text categories resemble those from the Brown Corpus, with 500 texts from which 2000 word excerpts have been selected.

The Freiburg-LOB Corpus of British English (FLOB) was created at the Freiburg University in Germany, also on the initiative of Christian Mair, and parallel with the compilation of the Frown Corpus. It is designed as an update to the previously described LOB

\(^9\) The manual for the LOB corpus may be accessed at: http://khnt.hit.uib.no/icame/manuals/LOB/index.htm

\(^10\) The manual for the Frown corpus may be accessed at: http://khnt.hit.uib.no/icame/manuals/frown/INDEX.htm

\(^11\) The manual for the FLOB corpus may be accessed at: http://khnt.hit.uib.no/icame/manuals/FLOB/INDEX.htm
The Australian Corpus of English (ACE) was compiled at Australia’s Macquarie University in New South Wales by Pam Peters from 1986. ACE was the first systematically compiled heterogeneous corpus in Australia, and was designed to resemble the Brown/LOB corpora, in order to be suited for comparative research with Australian, British and American English. It is not, however comparable to its British and American cousins with regard to time of origin of the texts. With ACE, the desire to create a contemporary corpus, prompted the decision to compile published texts from 1986.

Although this is somewhat of a weakness when looking at the diachronic development of language, ACE nevertheless provides us with a very useful tool for studying language from a synchronic perspective. I have therefore used it to look at where Australian English finds itself with regard to whether it seems to be influenced mainly by the British or American variant. The text categories and sampling procedures closely resembles those of its American and British counterparts. Unfortunately copyright issues have led to only a partial version of the ACE corpus being available at the University of Oslo. The version available to me consists of 700,000 words, rather than one million. This has to be taken into account when analysing and comparing the findings from this corpus to the others.

3.8. WC

The manual for the WC corpus may be accessed at: http://khnt.hit.uib.no/icame/manuals/wc/INDEX.htm

---

12 The manual for the ACE corpus may be accessed at: http://khnt.hit.uib.no/icame/manuals/ace/INDEX.htm
13 The manual for the WC corpus may be accessed at: http://khnt.hit.uib.no/icame/manuals/wc/INDEX.htm
The Wellington Corpus of Written New Zealand English (WC) was developed under the supervision of Laurie Bauer at the Victoria University of Wellington between 1986 and 1992. It was modelled on the ongoing project of compiling the ACE corpus, so the texts were taken from publications from the year 1986. The text categories have been created to mirror those of LOB. The non-fictional sections, however, are not subdivided in WC.

I use ACE and WC merely for reference in this study.

3.9. BNC

I have used the British National Corpus (BNC) because it is much larger, and contains spoken as well as written material.

Unlike in the ‘Brown Family’ of corpora, the written material of the BNC also consists of some unpublished material. Published texts are generally less spontaneous than unpublished material like personal letters. Published texts usually undergo a longer process of planning, editing and adaptation to the norm of the genre than an unpublished text. Unpublished material may therefore be seen as closer to spontaneous language. The range of language, as well as the sheer size of the corpus therefore provides us with the opportunity to study BrE in use in great detail, and can be assumed to provide examples of a wide range of language uses, from spoken conversational language to formal texts. I have used the BNC quite extensively for this investigation, in order to investigate how ought is being used in BrE, both as a comparison with the findings from LOB and Frown, and as a source of new information with regard to distribution.

The BNC was created by a large number of participants, organisations and individuals and developed at the University of Zurich. It gives a great deal of information regarding the context of language use, for example the age and sex of the speaker, the sort of text or speech situation etc., which may be relevant in a study such as this. The BNC consists of around 100 million words and is:

Designed to characterize the state of contemporary British English in its various social and generic uses. (Aston and Burnard 1998:289)

90% of the material is written, with the remaining 10% being made up of spoken, transcribed discourse. The BNC World edition was made available online in 1991. Written texts are divided into the categories; Imaginative, Belief and thought, Social Science,

14 The manual for the BNC corpus may be accessed at: http://homepage.mac.com/bncweb/manual/bncwebman-home.htm
Commerce and Finance, Arts, World Affairs, Applied Science, Natural and Pure Sciences and Leisure. Spoken discourse is divided into domains, monologue and dialogue, and context. The findings can be sorted into a series of variables, including age and sex of author, types of audience etc. The BNC is an invaluable tool when studying real language in use. Its main limitation is perhaps that it largely contains British English.

Unfortunately, the BYU Corpus of American English only became available at a late stage in this study, so I have not been able to make as much use of that as I would have liked. Therefore, the focus is perhaps slightly more on the British variant, rather than a purely contrastive study of differences between varieties of English.

The BNC is also limited to language in the early 1990s. Thus, it provides us with valuable insight into one variety at one particular time, but is not suited for comparative studies on its own.

3.10. COLT

While most of the corpora used in this study are designed to give a balanced view of language in use, some corpora are best suited to study specific types of language.

I have used The Bergen Corpus of London Teenage Language (COLT). My reason for including this particular corpus is that teenagers are, in a sense, the future users and shapers of any language. If a word is used significantly less by the younger part of the population, we can to some extent predict that the usage of the word will continue to decline in the future. The proportion of teenagers in the BNC generally is also significantly lower than other age groups, so the inclusion of the COLT corpus may give more reliable results as far as this age group is concerned. COLT is also a corpus constructed by transcribing spoken language. Again, the trends arising in spoken conversation may give indications on future usage of a particular word in written, more formal situations. Language change tends to begin with changes in spoken language, as spoken language is, by nature, less durable and faster to adapt and change. COLT is, however, a relatively small corpus, and geographically restricted. In addition, the material is limited to spoken transcripts, so caution must be exercised when drawing conclusion about the language more generally on the basis of these findings. It is important to point out that COLT is, in fact a constituent of the BNC, so the material from COLT is also part of the BNC material. Using COLT, however, allows me to isolate the

15 The homepage for the COLT corpus may be accessed at: http://torvald.aksis.uib.no/colt/
findings from spoken London teenage language. I have used COLT purely as a reference corpus in this study.

3.11. DCPSE

The Diachronic Corpus of Present-day Spoken English (DCPSE), constructed at the Survey of Spoken English Usage at University College, London, is another corpus designed to investigate rather specific language use, as it contains only spontaneous spoken language. It is, however, much broader in its scope than COLT, and is naturally a larger art of this investigation. On the other hand, the fact that it is based on spontaneous speech means that it does not include the rehearsed and written-to-be-spoken material, which is quite prominent in the spoken section of the BNC. This gives us insight into an entirely different type of language.

This corpus contains around 800,000 words of transcribed speech. 400,000 words are taken from the spoken section of ICE-GB, collected in the early 1990s, while the remaining 400,000 words are taken from the London-Lund Corpus with language ranging from the 1960s until the 1980s.

The DCPSE has been grammatically annotated, thus providing the user in syntactic tree diagrams. This is perhaps not essential for a study of this type, where the use of ought is restricted to that of a modal auxiliary. If we consider the other modals, however, this may me very useful information indeed.

Of the total 1307 examples of the word will, far from all the instances are in a modal auxiliary sense. There are quite a few of these instances where will is actually a noun and head of a noun phrase, as in examples 28 and 29 below, and as such entirely irrelevant to a study concerned with modal auxiliaries.

(28) ...they are as instruments of international will upholding international law and supporting the United Nations. (DCPSE DI-D07 0035)

(29) She signed the will. (DCPSE DL-H01 0278)

The syntactic annotation in the DCPSE corpus therefore quickly enables us to disregard such irrelevant entries. The encoding is not entirely reliable, but at least gives us a pretty good indication of the likely functions of a particular word.

---

16 The manual for the DCPSE corpus may be accessed at: http://www.ucl.ac.uk/english-usage/projects/dcpse/
3.12. ENPC\textsuperscript{17}

The English-Norwegian Parallel Corpus (ENPC) contains original Norwegian and English texts and their translations. The ENPC originated as a research project at the Institute of British and American Studies (IBA) at the University of Oslo. The majority of the content is novels and general non-fiction books, with chunks of 10,000-15,000 words taken from 30 original fictional texts and 20 original non-fictional original texts in each language. The build-up of the corpus is illustrated in table 3.1 below, which is taken from section 1.2 in the ENPC manual.

Table 3.1 – The build-up of the ENPC corpus

<table>
<thead>
<tr>
<th></th>
<th>Original texts</th>
<th>Translated texts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English</td>
<td>Norwegian</td>
</tr>
<tr>
<td>Fiction</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Non-fiction</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Total texts</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Total number of words</td>
<td>671,700</td>
<td>629,900</td>
</tr>
</tbody>
</table>

All in all, the ENPC corpus contains slightly over 2.6 million words (2,662,500 words). Figure 3.1 below, which is also taken from section 1.2 in the ENPC manual, is a schematic illustration of the core corpus. The four boxes indicate the core parts of the ENPC, while the lines between them show the main types of studies that are made possible by this structure, namely contrastive studies of parallel original texts or on original texts and their translations. The ENPC is also well suited for different types of translation studies.

The extension of the corpus, which has included more languages (German, Dutch and Portuguese), has also opened up for research into translation studies focusing on intralinguistic translation difficulties, compared with language-specific problems.

FIG 3.1 – The structure of the English-Norwegian Parallel Corpus

\textsuperscript{17} The manual for the ENPC corpus may be accessed at::
http://www.hf.uio.no/ilos/forskning/forskningsprosjekter/enpc/ENPCmanual.htm
My use of the ENPC is very limited in this study. By looking at how *ought* is translated into English, and then how these translations are translated from Norwegian into English, I hope to produce a list of alternative expressions in order to see if an increase in these alternative expressions can account for some of the decrease in the use of *ought*.

### 3.13. BYUC

The Brigham Young University Corpus of American English (BYUC) is the first large corpus of American English, and was compiled under the leadership of Mark Davies. This is a very large corpus. At present it contains more than 360 million words of American English, and it will be updated twice a year, so that it will continue to grow in the years to come. It has a rather sophisticated search engine that allows the user to perform quite detailed searches. It also consists of around 20% spoken language. This corpus will undoubtedly prove to be a valuable resource to researchers who want to study American English, and to anyone wishing to perform comparative investigations with British English, using a corpus like the BNC. It also contains more recent language than the other corpora I have used, and contains language produced as recently as 2007. It has therefore enabled me to investigate the development of the use of *ought* in American English almost up to the present day. It is freely available online.

Unfortunately, the BYUC only became available a few weeks before this investigation was completed. I have therefore not been able to make as much use of it as I perhaps would have liked.

---

18 The manual for the BYUC corpus, as well as the corpus itself may be accessed at [http://www.americancorpus.org](http://www.americancorpus.org)
An unfortunate limitation is the relatively narrow selection of text domains compared with, for instance, the BNC.

The spoken section contains only broadcast speech, mainly from television and radio. Even though the degree or formality, rehearsal and pre-planning probably varies, broadcast language is rarely, if ever, as spontaneous transcripts form real-life face-to-face or telephone conversation. It would have been interesting to see more informal spoken language in AmE, as we have access to this in BrE.

The written section also contains exclusively published material from three sources: Magazines, Newspapers and Academic prose. This provides information about the distribution of linguistic phenomena according to parameters like formality to a certain extent, but it would be interesting to see even less formal language, like personal letters or even e-mail language in such a modern corpus.

Nevertheless, the BYUC provides a well of information on the American variety of English. The search engine provides information on important parameters like the year the texts were produced, as well as the type of medium within the broad categories. The “chart” option in the search engine instantly provides information regarding distribution according to text categories as well as the development of a phenomenon over time. It also allows the user to perform comparative studies between categories. In addition, the search engine has a number of rather sophisticated search options, which time restrictions have not allowed me to explore fully for this thesis.

### 3.14. Summary of the corpora used for this study.

**Table 3.2 – Details and summary of the corpora used in this study.**

<table>
<thead>
<tr>
<th>Name of Corpus</th>
<th>American or British English</th>
<th>Date of data collected</th>
<th>Written or transcripts from spoken material</th>
<th>Size and design of corpora.</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOB Corpus</td>
<td>BrE</td>
<td>1961</td>
<td>Written</td>
<td>Each of these corpora contains around one million words. The words are taken from 500 text samples from up to 15 different genres. Because the corpora are built using the same design and sample method, they are particularly suited for a comparative study.</td>
</tr>
<tr>
<td>Brown Corpus</td>
<td>AmE</td>
<td>1961</td>
<td>Written</td>
<td></td>
</tr>
<tr>
<td>WC Corpus</td>
<td>NZE</td>
<td>1986</td>
<td>Written</td>
<td></td>
</tr>
<tr>
<td>FLOB Corpus</td>
<td>BrE</td>
<td>1991</td>
<td>Written</td>
<td></td>
</tr>
<tr>
<td>Frown Corpus</td>
<td>AmE</td>
<td>1992</td>
<td>Written</td>
<td></td>
</tr>
<tr>
<td>ACE Corpus</td>
<td>AusE</td>
<td>1986</td>
<td>Written</td>
<td>This Corpus is assembled in the same way as the rest of the ‘Brown...</td>
</tr>
<tr>
<td>Corpus</td>
<td>Language</td>
<td>Period</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>--------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>BNC</td>
<td>BrE</td>
<td>1991-1994</td>
<td>Written (about 90%) and Spoken (about 10%)</td>
<td>This corpus contains around 100 million words, selected from a wide variety of sources. It is designed with the intention of representing a large cross-section of British English language use.</td>
</tr>
<tr>
<td>DCPSE</td>
<td>BrE</td>
<td>1950-1990</td>
<td>Spoken</td>
<td>Contains around 800,000 words of spoken present-day English. Half are taken from the spoken section of ICE GB, while the other half are taken from the London-Lund Corpus.</td>
</tr>
<tr>
<td>COLT</td>
<td>BrE</td>
<td>1993-2003</td>
<td>Spoken</td>
<td>Around 500,000 words of transcriptions from around 55 hours of spoken, spontaneous language by 13-17-year-olds in London. The COLT corpus is a constituent of the BNC.</td>
</tr>
<tr>
<td>BYUC</td>
<td>AmE</td>
<td>1990-2007</td>
<td>Spoken (80%) and Written (20%)</td>
<td>More than 360 million words of written (80%) and spoken (20%) American English. Selected from a variety of published sources.</td>
</tr>
<tr>
<td>ENPC</td>
<td>Eng/NO.</td>
<td>1994-1997</td>
<td>Written</td>
<td>Around 2.6 million words of original texts and their translations (English to Norwegian and Norwegian to English)</td>
</tr>
</tbody>
</table>

**3.15. Wordsmith and FileMaker Pro 5.5.**

LOB, FLOB, Brown, Frown, ACE, WC and COLT were all accessed using the Wordsmith software. Using the programmes ‘concordance’ function, the programme allows the user to access the number of occurrences of a phenomenon in the entire corpus, along with the context in which it appears, as well as information regarding the distribution of the phenomenon I am interested in, with regard to text categories.

The search for *ought* LOB, FLOB, Brown, Frown, ACE, and WC yielded a relatively high number of instances. The Wordsmith software gives information regarding the number of tokens or instances in each corpus, and the distribution according to text categories. The similarities is size and composition of these corpora, as well as their suitability for diachronic as well as synchronic study, however, make them interesting for more detailed comparison. Uploading the instances from these corpora to FileMaker allows the researcher to create his or her own database and organize the material, sorting the instances according to desired criteria.
Irrelevant instances can easily removed, and the sorted material can easily be compared, revealing differences between the corpora or time periods, etc. This gives the opportunity to set ones own parameters when conducting research.
4) The investigation in the LOB, FLOB, Brown, Frown, ACE and WC corpora – results and analysis.

4.1. Overall distribution

Figure 4.1 below shows the occurrence of the modal auxiliary verb *ought* in the six corpora of the ‘Brown Family’: LOB, FLOB, Brown, Frown, ACE and WC. The table clearly shows a decrease in the use of *ought*, both in American and British English, but the decrease is more dramatic in the case of BrE. It also indicates that the word was more common in BrE in the 1961 LOB Corpus than the American Brown Corpus, but the difference between AmE and BrE is less marked in the two more recent corpora as seen in Figure 4.1.

Australian and New Zealand English are, of course, only seen in 1986, so we do not have the advantage of a diachronic perspective on those two varieties. In the case of Australian English, I also had to normalize the figure and round it to the nearest whole number, as the version of ACE I am using only contains around 700,000 words. The ACE figure is therefore not as accurate as the others, and the figures from ACE and WC, which show distribution in Australian and New Zealand English, are not entirely comparable. They are, however, perhaps expectedly, somewhat higher than the figure from FLOB, but significantly lower than the LOB figure. They are also only a little bit lower than the Brown figure. As the ACE/WC figures show language less only around 6 years older than Frown/FLOB figures and more than 20 years more recent language than the LOB/Brown corpora, this does perhaps indicate that New Zealand and Australian English follow a path which resembles BrE more closely than AmE when it comes to the usage of the modal auxiliary *ought*.

**FIG. 4.1 – The occurrence of ought in the four corpora from the Brown Family used in this investigation (Brown, Frown, LOB, FLOB, ACE and WC) normalized into instances per 1 million words.**
As the reduction is much most striking in BrE, and I have access to a greater variation of BrE and AmE corpora, I have decided to investigate these two varieties, and in particular BrE in more detail. I will, however, try to include comparisons with other varieties of English, and in particular American English, wherever possible and relevant.

4.2. The reduction in the use of ought in British English from 1961-1991

I will first make some general observations in this section, before going on to investigate some of the issues more closely in the following sections in this chapter.

The word is used in declarative sentences, as in examples 30-32 below, in the vast majority of instances, with only a tiny minority in interrogative sentences (5 in LOB and 3 in FLOB).

One striking thing is that the distribution with regard to person is extremely uneven (See further in section 4.8). In addition, there is a striking number of examples containing ought in the passive voice, often without an expressed agent.

(30) The two men ought to be held in restraint. (LOB A text)

(31) It is also a tax-measure that ought to be fully discussed. (LOB B text)

(32) They are also hugely energy-sapping and therefore ought to be separated to avoid too great a strain. (FLOB E text)

This is consistent with a tendency toward indirectness in order to conform to principles of politeness. A sentence beginning with “you ought…”, as in example 33 below may be interpreted as a command, criticism, or at least a strong recommendation, and unless
the speaker is in a position to command or pressurise the hearer in any way, or is criticizing, it
may be that he/she would rather use a less direct way of communicating these wishes.

A good way of doing this might for example be to use passive constructions or, as in
example 33 below, an unspecified, perhaps slightly vague, third person subject.

(33) …by some who know little, and others who ought to know better.
  (FLOB F text)

The implicature here is that something really should be known, and by the receiver of
the message or at least someone the receiver will be able to identify, but the speaker has
successfully avoided being quite so direct, while still conveying the message so that it leaves
no room for misunderstandings. *Ought* may also be used by speakers employing the tact
maxim described by Jenny Thomas (1995:160) where she talks of “mitigating the effect of a
request by offering optionality” (see example 34 below)

(34) Perhaps we ought to have some lunch?” (BNC H94 3952)

Giving options, or at least appearing to do so, is widely used in British, as indeed
elsewhere in Western culture, as commands are considered face-threatening acts, acts which
are described by Thomas (1995:169) as:

*Illocutionary acts which have the potential to damage a receivers positive face…or negative
  face…*

In other words: acts causing offence, by implying criticism, expressing disapproval or
cornering someone into making an offer. Such acts are really only expected in certain
contexts. In the case of this word “you ought to know this by now” would be considered more
face threatening than “This ought to be familiar to most people”, because it has a much
stronger implicature of criticism directed at the receiver than the second sentence, and the first
instance would be more likely to be used by someone who is very familiar with or in a
position of superiority in relation to the receiver. If the use of these constructions is reserved
for certain situations, and preferred in conjunction with different politeness strategies (like
using the third person and passive constructions) this will, in time, enhance its face-
threatening potential, and may lead to speakers increasingly preferring to avoid it altogether.
In this context a detailed look into the semantic meaning of *ought* might be useful. Only the
root meaning *ought* has face-threatening potential, so any reduction in *ought* in the epistemic
sense can not be accounted for by referring to politeness theory. See further discussion section
4.7.
It is often difficult to distinguish the epistemic use of *ought* from the use implying an obligation. Sometimes speakers appear to use *ought* apparently epistemically:

(35) The job ought to be finished on Friday. (LOB B text)

The context may, however, reveal that, in reality, the meaning is something like: “I expect this to be done by Friday, and you are responsible for making sure it is”.

It is notable that some of the other modals which were shown in Leech’s study to be used much less frequently in recent times, like *shall*, *need* and *must*, also have the same function of strongly directing someone, while the words which actually increase in usage, *can* and *could* seemingly give options.

### 4.3 Distribution in text categories in BrE and AmE.

Literary authors tend to employ techniques in order to ensure variation of vocabulary. This is probably one reason why they are slower to stop using a word on the decrease than non-fictional writers. Because the language in fiction tends to be directed at imaginary characters, there is not the real potential of causing embarrassment or offence, so it could be that writers feel somewhat freer to make use of *ought* there. This would also explain why it is seemingly easier to use a first person subject, so directing the implicature toward the speaker him/herself than toward the listener as a second person subject. A detailed look at text categories where *ought* is used, regardless of what type of subject it is directed at, reveals a tendency to use *ought* in fiction, rather than in non-fictional texts, with a particularly low frequency in press texts. In table 4.1 the raw figures from LOB and FLOB are given. As the categories vary in size, I have also calculated the relative frequencies in each category, given in instances per 100,000 words. This gives a more accurate picture of the development in BrE. Fig. 4.1 gives a graphic illustration based on the relative frequencies.

Some of the categories in which we find high frequencies are non-fictional. These include the section containing Belles Lettres, Biographies and Essays, as well as the section containing religious texts. This type of text can be said to be found in the centre ground between fiction and non-fiction, but they normally portray a subjective image of reality, and thus often contain subjective judgements which perhaps may account for some of this relatively high usage of *ought*.

Table 4.1 – The occurrence of *ought* in the various text categories in the LOB and FLOB Corpora.
<table>
<thead>
<tr>
<th>Category</th>
<th>LOB Raw Frequency</th>
<th>LOB Relative Frequency (per 100,000 words)</th>
<th>FLOB Raw Frequency</th>
<th>FLOB Relative Frequency (per 100,000 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – Press, Reportage</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B – Press, Editorial</td>
<td>7</td>
<td>11</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>C – Press, Reviews</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>D – Religion</td>
<td>8</td>
<td>20</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>E – Skills, Trade, Hobbies</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>F – Popular Lore</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>G – Belles Lettres, Biography, Essays</td>
<td>25</td>
<td>14</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>H - Miscellaneous</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>J – Learned, scientific writing</td>
<td>14</td>
<td>7</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>K – General Fiction</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>L – Fiction, Mystery</td>
<td>11</td>
<td>19</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>M – Fiction, science</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>N – Fiction, Adventure/Western</td>
<td>10</td>
<td>14</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>P - Fiction, Romance</td>
<td>8</td>
<td>11</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>R - Humour</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

FIG. 4.2 – The distribution of *ought* (relative frequency per 100 000 words) in LOB and FLOB.

The distribution clearly shows a reduction in the use of *ought* in most text categories. The reduction is, however, particularly marked in category G (Belles Lettres, Biography, Memoirs etc.), and also in the fictional categories. If a reduction is taking place for the reasons I suggested above, it may be that it occurs first in the informative and directly communicative categories. As the use of this word is gradually being reduced in the everyday language, it is also reduced in fiction. The language of fiction then seems to follow on,
although slightly “lagging behind” the tendencies from general language use. It is still apparent, however, that this word is used more in the fictional categories, if compared with non-fiction, especially if we consider that there are around three times the number of texts in the non-fiction categories (374) than in the fictional ones (126) in the corpora. Category D (Religion) also yields a high number of occurrences, despite the fact that it contains relatively few text samples (17). The distribution does, however, appear to be evening out somewhat in FLOB when we consider the relative frequency, compared with the distribution in LOB, but even here it is in the fictional categories we see the largest reduction. There are, however, relatively few instances in each category, so care must be exercised in order to not draw too general conclusions based on this alone.

The reduction of *ought* use has not been as dramatic in AmE. Again, however, Table 4.2 and Figure 4.3 reveal that the reduction is more marked in the fictional categories. The distribution on non-fictional categories remains fairly unchanged, or even, as in category G show an increase.

It must be borne in mind, however, that the lower number of instances in the AmE corpora means that there are even fewer instances in each category in this variant than in BrE. Therefore, the stylistic preferences of individual authors may have an exaggerated effect on the overall result.

**Table 4.2 – The occurrence of *ought* in the various text categories in the Brown and Frown Corpora.**

<table>
<thead>
<tr>
<th>Category</th>
<th>BROWN Raw Frequency</th>
<th>BROWN Relative Frequency (per 100,000 words)</th>
<th>FROWN Raw Frequency</th>
<th>FROWN Relative Frequency (per 100,000 words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A – Press, Reportage</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>B – Press, Editorial</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>C – Press, Reviews</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>D - Religion</td>
<td>4</td>
<td>9</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>E – Skills and Hobbies</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F – Popular Lore</td>
<td>8</td>
<td>6</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>G – Belles Lettres, Biography, Memoirs</td>
<td>6</td>
<td>3</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>H - Miscellaneous</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J – Learned, scientific writing</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>K – General Fiction</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>L –Mystery and Detective Fiction</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>M – Science Fiction</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Text Category</td>
<td>Brown</td>
<td>Frown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N – Adventure and Western Fiction</td>
<td>7</td>
<td>9</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>P - Romance and Love Story</td>
<td>14</td>
<td>18</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>R - Humour</td>
<td>2</td>
<td>0.08</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

FIG. 4.3 – The distribution of *ought* (relative frequency per 100,000 words) In Brown and Frown.

4.4. The distribution according to the ideational dimension is Systemic Functional Grammar (SFG)

Having uploaded the examples FileMaker, I categorized the instances according to one aspect of Michael Halliday’s System of Systemic Functional Grammar. Before revealing the results from this part of the investigation, I would very briefly like to outline the main ideas behind the SFG system, as well as the ideational dimension in particular (section 4.5). I will then go on to reveal and discuss my findings in section 4.6.

4.5. Systemic Functional Grammar and the experiential metafunction

Systemic Functional Grammar (SFG) is a grammatical model of language first developed by Michael Halliday in the 1960s. The term "systemic" is used to refer to the view of language as: "A network of systems, or interrelated sets of options for making meaning" (Halliday 1994:15).

The term "functional" is used to indicate that the approach is concerned with meaning. Systemic functional grammar is mainly concerned with the *choices* that are made available to speakers of a language by their grammatical systems. These choices are thought to be
meaningful and relate speakers' intentions to the concrete forms of a language. Traditionally the choices made by speakers (and writers) are viewed in terms of either the content or the structure of the language used. SFG seeks to present a view of language in terms of both the structure (grammar) and the words (or lexis). The term "lexicogrammar" is used to describe this combined approach. While traditional grammar concerns itself mainly with syntax, and to some extent morphology, SFG does not distinguish between lexis as grammar, but views them both as carriers of meaning.

Halliday’s main purpose for creating this new type of grammar was to:

Not to orient the grammar to any single defined area of application, but to provide a general grammar for purposes of text analysis or interpretation (Martin et al. 1997:2)

Geoff Thompson refers to SFG compared with traditional grammar as moving from:


SFG looks at language from three perspectives or metafunctions, to use Halliday’s more technical term, when analysing text. Below, I will give a brief outline of the metafunctions based on Halliday and Matthiessen (2004)

1. The **textual** perspective of the clause as message, which mainly concerns itself with the flow of information in a clause, dividing it into Rheme and Theme elements. This metafunction can be said to create discourse.

2. The **interpersonal** perspective of the clause as exchange. This perspective is primarily concerned with the MOOD system where language is thought to contain four basic functions: Offer, command, statement and question, matched by the appropriate responses accepting an offer, carrying out a command, acknowledging a statement and answering a question. This metafunction has the function of enacting human relationships.

3. The **experiential** or **ideational** perspective of the clause as representation, which construes human experience. Here, the focus is on process and participant analysis as a way of explaining how the exchange is a representation of the world.
All the three metafunctions act together in creating meaning in discourse, as illustrated in figure 4.4 above.

I will not go into SFG in more detail, because the complexity of the system would demand far more time and space than what is available to me, but Halliday and Matthiessen (2004) as well as Geoff Thompson (2004) can be consulted for further information.

The one metafunction which is most relevant for my study, however, is the experiential or ideational metafunction. By looking at the process types in which *ought*
appears I hope to discover what kind of contexts ought appears in, and by that perhaps discover any differences in the way ought is used.

I will therefore briefly outline the possible process types. Again, this is based on Thompson (2004):

1. **Material processes.** These are the processes which involve some kind of physical actions, like walking, running, waving etc. There is typically one participant responsible for bringing about the action – an actor. Example 36 below is a good example of a material process. Example 37 is also a material process, although it is a negative sentence, where the action is not brought about. It is also in the passive voice, so the actor is somewhat unclear.

   (36) I ought to get up and shut the window, she thought, and lay still (LOB_L text)

   (37) projects could bring results, and hence they ought not even be investigated empirically. (FROWN J text)

Some clauses, however, are much more difficult to classify. Consider example 38 and 39 below:

   (38) Whatever act could lead to this outcome is the one we ought morally to perform – it is where our moral duty. (FLOB J text)

   (39) …don’t know what penance ought to be assigned to them. (Frown G text)

Verbs like perform and assign certainly indicate material processes, but in these instances it is perhaps not something physically performed? In particular assigning penance in example (39) These particular acts (e.g. the actual assignment of penance) are perhaps verbal acts, or even mental? If the object is abstract the process becomes abstract. I have still opted to class these as material processes, but I realise that that is not necessarily the only option.

2. **Verbal Processes** are processes of saying. These are the processes generally brought about by a sayer and involve not just direct quotes, but all verbs describing speech,
such as argued, claimed, said, whispered etc. Examples from the material include the below examples 40-42:

(40) This idea behind it is presumably that the catechism ought to mention one ministerial action in the Church of England (LOB D text)
(41) Catholic scholars who believe that religious truth ought to be explained in a rational way if possible (ACE G text)
(42) The doctor ought to tell the patient. (WC G text)

Consider example 43 below, however.

(43) He has said that they ought not to discourage anyone. (Brown B text)

Discourage certainly might be a verbal process, but without much more context it is impossible to say. It could be a material process, depending on the means of discouragement and how it is delivered to the recipient.

3. Mental Processes are processes of cognition, so thinking as well as sensing and evaluating constitute mental processes. This includes processes signalled by verbs like think, believe, consider and understand. Examples from the material include examples 44-46 below.

(44) he said there was nothing in Back to Methuselah, - "ought to have known that", - and "I look at my bookshelves despair (Brown G text)

(45) he suggested the poor fool "ought not to believe everything you read in the papers .... (FROWN B text)

(46) The Pope "can and ought to reconcile and adjust himself with progress, liberalism and modern civilisation (ACE G text)

Behavioural Processes are processes of often involuntary or at least subconscious behaviour. These behaviours include things like breathing, coughing and flinching or, as in example 47, dyeing.

(47) He ought not to have died (BNC BLW 910)
The participants in behavioural processes are behavers as opposed to the more consciously proactive actors of material processes.

4. Relational and Existential Processes are process of “being” (example 48 below) or “having”. In the relational process, the process is the “action”.

(48) Dennis Vosper, Minister of State, Home Office, said there ought not to be a close association of alcohol and road transport (LOB A text)

Existential Processes are processes of existing (see example 48 below). These are often signalled by verbs like exist, but also appear and seem.

(49) Melville had, as did his age, a nai-umlautve sense of the relation he felt ought to exist between language and being. (FROWN G text)

These are closely related to, and often difficult to distinguish from relational processes. They are also quite rare. For this study I have chosen to class the two together.

4.6. Processs types in LOB, FLOB. Brown, Frown, ACE and WC

Having classified the clauses where ought appears, I have outlined a rough distribution of such clauses in each of the six corpora. Because the number of tokens, or instances, varies considerably between the corpora, I have decided to give the figures in percentages, in order for them to be comparable. I would also like to stress the difficulty in the classification at times, as I have pointed to in some of the examples in the previous section. The classifications are at times based on my instincts and may, as such, be somewhat unreliable.

I have also had to disregard some irrelevant instances, for instance where ought is not used in a sense of modal auxiliary, but rather mentioned as the subject of a sentence.

However, Table 4.3 below gives the distribution of ought according to process types in the six corpora.

Table 4.3 – The distribution according to process types given in percent of total occurrences in each corpus (raw frequencies in brackets)
Matthiessen (1999) performed a limited study in order to try to determine the general distribution of process types. His results are outlined in table 4.4 below.

**Table 4.4 – the general distribution of process types given in % according to Matthiessen (1999)**

<table>
<thead>
<tr>
<th>Process type</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material</td>
<td>51</td>
</tr>
<tr>
<td>Behavioural</td>
<td>5</td>
</tr>
<tr>
<td>Mental</td>
<td>9</td>
</tr>
<tr>
<td>Verbal</td>
<td>10</td>
</tr>
<tr>
<td>Relational</td>
<td>23</td>
</tr>
<tr>
<td>Existential</td>
<td>2</td>
</tr>
</tbody>
</table>

Although Matthiessen’s study was very limited, and an investigation concerned with process types in general must be expected to yield different results than one concerned exclusively with modal auxiliary clauses (for instance, behavioural clauses are naturally rare in clauses with auxiliaries denoting any kind of obligation or advice, as they refer to natural, involuntary or subconscious behaviour), Matthiessen’s results bear striking similarity to the results from LOB in particular, but also the results from Brown and Frown.

In order to get a more visual picture of the distribution according to the distribution of *ought* according to process types in the four corpora, I have shown the distribution in sector diagrams in the following part of this section. Figure 4.5 and figure 4.6 below show quite clearly that the distribution is very different in the British corpora with material from 1961, than in the 1992 material. While there is a heavy emphasis on material processes in LOB,
there is an even stronger emphasis on the relational/existential process type in FLOB. Also, there are very few verbal processes in the FLOB material.

**FIG 4.5 – Distribution according to process types in LOB**

If we go to the American corpora and look at Brown and Frown in figure 4.7 and 4.8. below, we see that the distribution in Brown is not entirely unlike that from LOB, where there is a clear emphasis on material processes.

Unlike in the British material, there does not appear to have been a shift in the distribution according to process types in the American material. It would seem that while LOB, Brown and Frown have a fairly similar distribution, FLOB is rather untypical with its great emphasis on relational/existential processes featuring* ought.*
However, when we take the other two varieties into account, FLOB does not appear to be as untypical as it would seem.

FIG 4.9 – Distribution according to process types in ACE
It is difficult to say if there is a pattern here at all. It could be that the use of *ought* in clauses dealing with a wider variety of process types in FLOB, ACE and WC may indicate a slight movement toward a less context-bound usage which may in turn slow down the reduction in the use of *ought* somewhat. It may also, however, be due to a variation of the topics in the four corpora. I would therefore exercise caution before drawing any firm conclusions on the basis of this part of my investigation. As this part of the investigation appears to give such inconclusive results, I have opted not to carry out a similar investigation on other corpora.

For the remainder of the study, I have also chosen to concentrate on the British and American varieties of English, as I have access to a wider variety of corpora in these two varieties than in AusE and NZE.
4.7. Distribution according to semantic meaning in written language.

In FileMaker, I also classified the examples according to semantics. I used the Merriam-Webster distinction between four different senses of *ought* as my starting point (See section 2.2). These are; *ought* in the sense of some kind of obligation on the part of the listener or reader, or, as in examples 50 and 51 below, the person who is being talked about:

(50) Palfrey thought the new owner *ought to pay* his people two months’ wages. (Brown F text)

(51) Haney didn’t want to encourage his company, but felt he *ought to buy* him a drink anyhow, to prevent possible trouble. (Brown L Text)

The second sense is that of offering advise, as in examples 52 and 53 below:

(52) Personally, I think we *ought to set up an immediate naval blockade of Cuba*. (Brown H text)

(53) I advised her that she should not go on and in my opinion she *ought to step down,*” (FLOB A text)

These first two senses are what might be called the root sense of *ought*. As with the root sense, the epistemic sense of the word has also been divided into two categories. The first is that of natural expectation, where the speaker anticipates that X is the case, as seen in examples 54 and 55 below:

(54) The Jazz festival *ought to prove* entertaining (Brown C text)

(55) That *ought to draw* a laugh, Nicolas reasoned, (Brown N text)

Finally, and closely related to the sense of natural expectation is the sense of logical consequence. Here, the speaker implies that as X is the case, it follows that Y is true, too. Examples 56 and 57 below are instances of *ought* in the sense of logical consequence:

(56) Fuel industries are 'basic industries', fuel *ought therefore to be cheap*, and the more that is consumed the better (LOB J text)

(57) As a woman under intense psychological siege. Edith (Angela Winkler) *ought to be the envy of her peers.* (WC C text)
I have chosen this the use of this classification because it gives more information than
simply dividing the instances into root and epistemic meaning. I also feel that the distinction
between an implication of obligation and the offer of advise is important if we try to explain
the reduction in the use of *ought* with reference to politeness theory. If my hypothesis is likely
to have some truth to it, I would expect a larger proportion of the more recent examples to
contain *ought* in the sense of implying obligation. If that proves to be the case, it may be that
*ought* is increasingly being perceived as carrying this implication of obligation. As that
function “takes over” as the dominant semantic meaning, adherence to politeness principles
may lead language users to increasingly avoid using it in contexts where this is not
appropriate. As the categories of obligation and advice both belong within the realm of root
meaning, a simple division into root and epistemic meaning will not suffice. Some of the
examples are unclear. Example 58 below, for instance, could imply both of the root senses of
one took it to mean “I expect you to ensure that it will lead to a more consistent picture of the
sources” (implying obligation) or e.g. “I suggest you make sure that it will lead to a more
consistent picture of the sources (because this will improve your grades) (advice). In fact, the
wider contexts reveals that *ought* in this case is used in the sense of logical consequence.

(58) for one thing, it ought to lead to a consistent picture of the
sources. (FLOB G text)

In some cases, the four-part distinction, simply does not include a good alternative to
cover the semantic meaning of an utterance. Consider, for example, example 59 below:

(59) There ought to be a word for us: obliviscents, people who
forget. (LOB R text)

For this part of the investigation I will only give the figures from LOB, FLOB, Brown
and Frown as they are directly comparable and give an indication of both the synchronic and
diachronic perspective. The results are given in table 4.4 below. Again, the variation in
number of instances between the corpora and my primary concern with distribution rather
than total number has led me to give the results as percentages. The figures therefore give
each semantic sense as the percentage of the total number of *ought* instances in each corpus. I
do, however, give the raw frequency in brackets in table 4.5.

The classifications are also based on my personal judgements, and some cases were
less clear-cut than others. The figures should therefore be taken as approximate. In some cases
semantic meaning was, in my opinion, so unclear that I was unable to determine the correct
meaning. I have indicated the number of such cases in a separate column.
Table 4.5 – The distribution of semantic meanings given in % (Raw Frequencies)

<table>
<thead>
<tr>
<th></th>
<th>Obligation</th>
<th>Advisability</th>
<th>Natural Expectation</th>
<th>Logical Consequence</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOB</td>
<td>44 % (45)</td>
<td>42 % (42)</td>
<td>6% (6)</td>
<td>6 % (6)</td>
<td>2% (2)</td>
</tr>
<tr>
<td>FLOB</td>
<td>57% (33)</td>
<td>28 % (16)</td>
<td>11% (6)</td>
<td>4 % (2)</td>
<td>0</td>
</tr>
<tr>
<td>Brown</td>
<td>34% (23)</td>
<td>44% (30)</td>
<td>21% (14)</td>
<td>0</td>
<td>1% (1)</td>
</tr>
<tr>
<td>Frown</td>
<td>38 % (18)</td>
<td>34% (16)</td>
<td>28% (13)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

In Figure 4.11 below, is a graphic illustration of the distribution of *ought* according to semantic meaning in the four corpora. Because the epistemic meaning is quite stable and not relevant to the issue of politeness, and to ease the reading of the diagram, I have added the figures from the natural expectation and logical consequence categories together and thus created a single category entitled epistemic meaning. As all the corpora are of the same size, I have used raw frequencies in Figure 4.11.

FIG. 4.11 – The distribution of *ought* according to semantic meaning in Brown, Frown, LOB and FLOB.

I would particularly like to draw attention to the very sharp decline in the ‘advice’ category from LOB to FLOB, as well as a relatively sharp decline from Brown to Frown. The reduction in the obligation category is marked, but not as sharp as in the case of advice. It is also sharper in BrE than in AmE, and the three categories appear almost equal in size in the most recent American material from Frown.
Figures 4.12 and 4.13 are graphic illustrations of the distribution of semantic meaning in the two BrE corpora LOB and FLOB. Clearly, the main emphasis is on obligation and advisability. Interestingly, it would seem that in the more modern material, there is quite a considerable increase from 44% to 57% in the proportion of instances indicating obligation, while the advisability function is reduced from 42% to 28%. This may indicate that *ought* is gradually being increasingly associated with imposing some kind of obligation, thus causing language users prefer other options according to the principles of politeness. The proportion of natural expectation, logical consequence and unclear remain relatively stable when taken together. The slight shift toward an increase in Natural expectation compared with logical consequence can probably be accounted for by two factors. Firstly, the numbers are low, so that individual examples have a disproportionate effect on the distribution. Secondly, they are often difficult to distinguish from one another and many cases are on the borderline between the two, so the classification may in some cases be unreliable.

**FIG 4.12 – The distribution of semantic meaning in LOB given in %**

![Pie chart showing distribution of semantic meaning in LOB](image)

**FIG 4.13 – The distribution of semantic meaning in FLOB given in %**

![Pie chart showing distribution of semantic meaning in FLOB](image)
In the American material, the graphical illustrations, shown in FIG 4.14 (Brown) and FIG.4.15 (Frown), are somewhat different. Here, natural expectation is a much more prominent semantic meaning, and becoming even more so. The number of cases is similar to the BrE natural expectation/logical consequence cases – only ranging from 8-14 cases in each of the million-word corpora. It is therefore the obligation and advisability functions which have a less prominent position in AmE. It also appears, particularly in the case of advisability, which has nearly halved the amount of instances from Brown to Frown, that the root semantic meanings are experiencing a further reduction.

FIG 4.14 – The distribution of semantic meaning in Brown given in %

FIG 4.15 – The distribution of semantic meaning in Frown given in %
This analysis has yielded some useful insights in my opinion. Although there are some uncertainties as far as the classification is concerned, the indications seem to be that there is a shift in the semantic meaning of *ought*, and that in this particular instance, BrE could well be “lagging behind” AmE as Leech suggests.

While the amount of instances where *ought* is used to indicate some kind of natural expectation or logical consequence remains reasonably stable, the advisability sense of the word is quite clearly experiencing a reduction in both varieties of the English language. This has resulted in an increase in the proportion (if not number due to the overall decrease of the word) of instances where *ought* is used to indicate some kind of obligation. Thus, the findings here seem to support the hypothesis that the semantic meaning of *ought*, at least in the root sense, is increasingly becoming associated with an implication of obligation. This may, in turn, lead speakers to avoid using it in contexts where they do not want to appear to be implying an obligation on the part of the receiver or target. The avoidance of *ought* in these ‘high-risk’ contexts may indeed go some way towards explaining the reduction in the use of *ought*


The classification of the material revealed that the distribution of subjects was very uneven. By subject, I do not necessarily refer to the *grammatical* subject, even though it often is, but rather the participant in the exchange who is referred to by the speaker as the one who *ought to* perform the act in question. To use the terminology from Halliday’s SFG system, the “subject” in this part of the investigation refers to the *actor, senser, sayer, behaver, token or*
the carrier of an attribute. The idea behind this part of the investigation is to investigate if speakers tend to avoid ought in situations where the risk of threatening “face” is high.

As I mentioned in the introduction to this section, there is a heavy emphasis on third person subjects. Often these third person subjects are rather vague and it is not necessarily entirely clear exactly who is referred to. In order to see if this tendency is increasing, or if it is more prominent in one or the other of the two main varieties (AmE and BrE), I sorted the instances according to subject in FileMaker. I then calculated the proportion of each of the type of subject occupied in each of the corpora.

The subject type which is most obviously has face threatening potential is the second person, as demonstrated in example 60 and 61 below:

(60) Of course, there was nothing you could do, but you still ought to be ashamed of yourself for letting it happen (Brown P text)

(61) Diana has lots of responsibilities here and you ought to help her out with some of them instead of sitting around here (LOB P text)

These are typically constructions might seem impolite if uttered by someone who is not is some kind of position of authority over the target. One way in which speakers may try to minimise this imposition is by using the more inclusive we as in example 62 below.

(62) I really think we ought to start. (LOB L text)

If this is uttered by someone like a manager of a company to his staff, the real implication may well be you. Likewise, if example 63 is an utterance produced by a police sergeant in charge of an investigation directed at his subordinates, there is every possibility that the actual work involved in searching for witnesses will be expected to be carried out by them.

(63) We ought to be able to find somebody who saw something out of place (Frown L text)

In addition to straightforward instances of I, you, we, they as well as third person references so someone specific, I found instances of more vague third person subjects often signalled by someone or, as in example 64 below, one.

(64) If one ought to have conscription for labour in peace, why not conscription for war? (LOB G text)

Quite a few examples, like example 65 below, refer to inanimate objects:
(65) The new federal program of help to economically depressed areas ought to provide some stimulus to growth. (Brown B text)

Still others had groups of people, for instance official bodies, teams or geographical areas (but where the implication is that the reference to people occupying or administering the territory in question, rather than the actual area itself). Se examples 66 and 67 below:

(66) The Government ought to proceed with the special meeting, while containing expectations (ACE B text)

(67) It is always the same. All Tysonians felt that the village ought to be as one. (LOB G text)

There were also some instances where the subject as a body part, rather than a person:

(68) aimed down where the stomach ought to be (LOB L text)

Some were very difficult to class. Consider for instance the two examples below:

(69) If the mind be immaterial, its functions ought to be unaffected by the condition of the body (LOB D text)

(70) My Self ought to exist, and might exist, wholly in the “I” – the Subject. (FLOB D text)

Is “the mind” a body part as “the brain” would be? Or is it a non-human thing? “My self” presents similar problems. Classification probably depends on the opinion of the researcher. As I do not want to go into any kind of theological or philosophical discussion about this, I have simply classed such instances as unclear.

Table 4.6 below gives the raw frequencies of occurrences of each type of subject in each of the corpora. The figures following below reveal the percentages for each corpus.

<table>
<thead>
<tr>
<th>Subject Type</th>
<th>LOB</th>
<th>FLOB</th>
<th>Brown</th>
<th>Frown</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>11</td>
<td>7</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>You</td>
<td>9</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Ordinary third person</td>
<td>16</td>
<td>12</td>
<td>29</td>
<td>12</td>
</tr>
<tr>
<td>We</td>
<td>9</td>
<td>7</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>They</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Vague third person</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Inanimate</td>
<td>36</td>
<td>25</td>
<td>19</td>
<td>16</td>
</tr>
</tbody>
</table>
Again, I have decided to use graphic illustrations to show the proportion of the different subject types in the various corpora. Fig. 4.15 shows the distribution in LOB, while Fig. 4.16 shows the FLOB distribution.

**FIG. 4.16 – The proportion of subject types in LOB**

**FIG. 4.17 – The proportion of subject types in FLOB**

<table>
<thead>
<tr>
<th>Subject Type</th>
<th>LOB</th>
<th>FLOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body part</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unclear</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Official body/team/geographical area</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Ordinary third person</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We</td>
<td></td>
<td></td>
</tr>
<tr>
<td>They</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vague third person</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inanimate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Official body/team/area/group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclear</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Official body/team/area/group</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>
The figures reveal what looks like an increasing tendency to use *ought* with inanimate subjects, as with examples 71 and 72 below:

(71) This idea behind it is presumably that the catechism ought to mention one ministerial action in the Church of England (LOB D text)

(72) A black badge of frayed respectability that ought never to have left his neck. (BROWN N text)

As the concept of *face* is restricted to humans (inanimate objects, of course, have no concept of self-worth), the use of *ought* with inanimate objects carries little risk of being considered face-threatening acts. Over 40% fall into this category in FLOB. This is followed by ordinary third person subjects. As third persons are probably not a part of the exchange, the speaker is unlikely to worry about seeming to give directions to them, and the face-threatening potential of the utterance is and then by the two equally sized categories of *I* and *We*. The use of *ought* along with *you*, as in example 73 has decreased quite sharply, perhaps indicating that speakers increasingly take precautions in order to not cause offence by using *ought* in situations where they do not want to appear to be directing the listener unduly.

(73) You ought to know that by this time. (LOB L)
If we turn to American English, figure 4.17 and 4.18 reveal similar tendencies.

**FIG. 4.18 – The proportion of subject types in Brown**

![Pie chart showing percentages of subject types in Brown]

- I: 1%
- You: 4%
- Ordinary third person: 8%
- We: 15%
- They: 36%
- Vague third person: 5%
- Inanimate: 3%
- Official body/team/area/group: 4%
- Body part: 24%

Again, it appears that the category referring to inanimate objects is on the increase, while the “you” category has decreased. First and ordinary third person subjects are common in American English as well.

**FIG. 4.19 – The proportion of subject types in Frown**

![Pie chart showing percentages of subject types in Frown]

- I: 4%
- You: 2%
- Ordinary third person: 6%
- We: 4%
- They: 24%
- Vague third person: 18%
- Inanimate: 4%
- Official body/team/area/group: 4%
- Body part: 4%

Again, it appears that the category referring to inanimate objects is on the increase, while the “you” category has decreased. First and ordinary third person subjects are common in American English as well.
These are relatively few instances, but there appears to be a consistent trend that writers are becoming increasingly reluctant to use *ought* with second person subjects.
5) *Ought* in spoken language according to the investigation in DCPSE

To see if oral language differs from written with regard to the use of *ought*, I initially turned to the DCPSE.

In this 800,000-word corpus, there are 164 instances of *ought*. Normalized into instances per million words in order to make it comparable to Brown, Frown, LOB and FLOB, the figure is 205. This is much higher than we would expect, and indicates a much higher frequency in spoken than in written language. This is surprising if we explain the reduction in modal usage by claiming it is due to colloquialization. In fact – this distribution is hard to explain at all, but perhaps the face-threatening potential of an utterance is easier to reduce in a directly communicative situation, like speech, by the use of paralinguistic features I have discussed in section 1.4. It may therefore be that it is easier to use this modal in a speech situation than in writing.

The DCPSE corpus uses the ICECUP software may be sorted according to a range of sociolinguistic features, which may shed some light on how and in which contexts and by whom *ought* is being used. It is also possible to see when the sample was collected. As the corpus stretches in time from 1956 up to 1992, it may also give an indication as to whether the reduction in use over time is taking place in oral language as well as in written.

5.1. Distribution of *ought* according to text categories in the DCPSE Corpus.

A good starting point in this type of investigation is to look at the distribution according to speech categories. The DCPSE has quite a detailed classification system. The categories vary quite extensively in size, however, so in addition to stating the number of instances in each category and the total number of words in each category, I have normalized the figures into instances per 1000 words. These results are given in table 5.1 below:

<table>
<thead>
<tr>
<th>Category</th>
<th>Hits</th>
<th>Total words</th>
<th>Per 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face-to face conversation</td>
<td>95</td>
<td>58 180</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formal</td>
<td>17</td>
<td>8792</td>
<td>1.9</td>
</tr>
<tr>
<td>Informal</td>
<td>74</td>
<td>50 388</td>
<td>1.5</td>
</tr>
<tr>
<td>Telephone conversation</td>
<td>9</td>
<td>6865</td>
<td>1.3</td>
</tr>
<tr>
<td>Broadcast Discussion</td>
<td>29</td>
<td>5819</td>
<td>5</td>
</tr>
<tr>
<td>Broadcast Interviews</td>
<td>7</td>
<td>2836</td>
<td>2.5</td>
</tr>
<tr>
<td>Spontaneous Commentary</td>
<td>2</td>
<td>8109</td>
<td>0.7</td>
</tr>
<tr>
<td>Parliamentary Language</td>
<td>3</td>
<td>976</td>
<td>3.1</td>
</tr>
<tr>
<td>Legal cross examination</td>
<td>1</td>
<td>885</td>
<td>1.1</td>
</tr>
<tr>
<td>assorted spontaneous</td>
<td>9</td>
<td>12679</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Some of these categories, in particular the ones containing language from legal cross-examination and the parliamentary language, are quite small. They therefore contain few instances and should perhaps be regarded as giving less reliable results than the larger categories like face-to-face and telephone conversation. Figure 5.1 below gives a graphic illustration of the distribution given in instances per 1000 words:

**FIG 5.1 – The distribution of *ought* according to text categories in the DCPSE corpus given in instances per 1000 words:**

<table>
<thead>
<tr>
<th>Speech Category</th>
<th>Instances per 1000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal face-to-face</td>
<td>1.9</td>
</tr>
<tr>
<td>Informal face-to-face</td>
<td>1.5</td>
</tr>
<tr>
<td>Telephone conversation</td>
<td>1.3</td>
</tr>
<tr>
<td>Broadcast discussion</td>
<td>5</td>
</tr>
<tr>
<td>Broadcast interview</td>
<td>2.5</td>
</tr>
<tr>
<td>Spontaneous commentary</td>
<td>0.7</td>
</tr>
<tr>
<td>Parliamentary language</td>
<td>3.1</td>
</tr>
<tr>
<td>Legal cross examination</td>
<td>1.1</td>
</tr>
<tr>
<td>Assorted spontaneous</td>
<td>0.7</td>
</tr>
<tr>
<td>Prepared speech</td>
<td>2.7</td>
</tr>
</tbody>
</table>

The distribution is clearly uneven in spoken language as it was in writing. Direct comparison is difficult as the categories do not correspond very well.

In spoken language, however, it is the more formal, even rehearsed language, that displays the highest frequency of *ought*. “Broadcast discussion” tends to be debates between people who are used to debating their particular issues (e.g. politicians or representatives from special interest groups) and have probably rehearsed their arguments extensively. This category is followed by parliamentary language, which is probably largely a more formal variant of the broadcast discussions, and again rehearsed extensively. This is again followed by prepared speech.

As spoken language displays a higher frequency of *ought* usage, it is perhaps a little unexpected that it is the speech categories which most closely resemble written language and which is probably often originally written to be spoken has the highest frequencies. However, these categories, or at least the first two, contain persuasive language and are designed to be
argumentative. It is therefore concerned with what ought to be said and done. This opens up for extensive use of modality in general, and particularly modals of obligation, as in examples 74 and 75 below:

(74) It ought to be a government monopoly (DCPSE DL D01 0323)

(75) I say now that it ought to be extended to a milder form (DL D03 0237)

Spontaneous language and face-to-face conversation, however, displays much less use of *ought*, and the majority of *ought* – phrases clearly occur in formal “official” language. Interestingly, telephone conversation has the lowest number of occurrences of the conversation categories.

5.2. The distribution of *ought* according to sociolinguistic parameters in the DCPSE

One of the sociolinguistic parameters that the DCPSE allows us to study is the level of education of the speakers. If *ought* is indeed used to a larger extent in formal language, one would expect people with a university education to use it more often than people who only have secondary school education. As Figure 5.2 shows, this is indeed the case.

The university educated speakers use *ought* more than twice as often as speakers who do not have a university education. However, the participants in the speech situations which contain the highest levels of *ought* usages probably tend to have a university education, so that probably accounts for at least some of this difference. In addition to being more inclined to use *ought* to begin with, people with higher education probably tend to find themselves in speech situations where *ought* is more likely to be used more often than people without a higher education.

FIG 5.2 –Usage of *ought* according to speakers’ level of education.
Another interesting sociolinguistic aspect is the distribution of *ought* usage according to the gender of speakers, shown in FIG 6.3 below.

**FIG 5.3 – Usage of *ought* according to speakers’ gender.**

According to the DCPSE, male speakers are nearly twice as likely as female speakers to use *ought*. Before drawing this rather sweeping conclusion, however, a few factors must be born in mind.

Men and women, even today, are not equally represented in high-profile jobs. Consider for instance the unequal representation in Parliament, which is where the language from one of the highest scoring speech categories in Figure 5.1 is taken from. This corpus contains language from as far back as 1958, when representation by women in formal speech situations was probably even lower. The broadcast genres are also likely to include a considerably larger than 50% proportion of men. I would therefore suggest that the difference between the genders when it comes to using *ought* may largely be accounted for by the fact that the people in language situations where *ought* is most likely to be used are largely men. In order to investigate whether this claim is valid, I searched on female language users in
formal face-to-face conversation, and found the higher figure of 2.1 instances per 1000 words, apparently confirming this suspicion.

5.3. Diachronic development of ought use according to the DCSPE

As the DCSPE contains language from the time span from 1958 to 1992 and gives information regarding when the material was collected, it is ideally suited for diachronic study. I divided the time span into four time periods of eight to nine years, and looked at the distribution in each period. Again, the categories vary greatly in size, so table 5.2 below shows the total amount of words, raw frequencies as well as normalized figures. Figure 5.4 gives a graphic illustration of the normalized figures.

Table 5.2 – Distribution of ought over time.

<table>
<thead>
<tr>
<th>Time period</th>
<th>Total words</th>
<th>Hits</th>
<th>Per 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1958-1966</td>
<td>11189</td>
<td>41</td>
<td>3.7</td>
</tr>
<tr>
<td>1967-1974</td>
<td>27077</td>
<td>70</td>
<td>2.6</td>
</tr>
<tr>
<td>1975-1982</td>
<td>15164</td>
<td>33</td>
<td>2.2</td>
</tr>
<tr>
<td>1983-1992</td>
<td>46885</td>
<td>43</td>
<td>0.9</td>
</tr>
</tbody>
</table>

FIG 5.4 – Distribution of ought over time given in instances per 1000 words

The distribution suggests a steady decrease over the time period which closely resembles the decrease suggested by LOB and FLOB. Thus, it appears that the decrease in spoken language parallels the decrease in written language. In a situation where a phenomenon is clearly on the decrease, as ought seems to be, I would expect the main users of the phenomenon to be older language users. The DCPSE also provides information on the age of the speaker, so I decided to investigate this theory as well. Table 5.3 and Figure 5.5 below give the results from this part of the investigation:
Table 5.3 – The distribution of *ought* in the DCSPE according to speaker age

<table>
<thead>
<tr>
<th>Speaker age</th>
<th>Total words</th>
<th>Hits</th>
<th>Per 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-25</td>
<td>19790</td>
<td>16</td>
<td>0.8</td>
</tr>
<tr>
<td>26-45</td>
<td>35763</td>
<td>54</td>
<td>1.5</td>
</tr>
<tr>
<td>46-65</td>
<td>22516</td>
<td>80</td>
<td>3.6</td>
</tr>
<tr>
<td>66+</td>
<td>3205</td>
<td>6</td>
<td>1.9</td>
</tr>
</tbody>
</table>

The results suggest that older speakers in general tend to use *ought* more than younger speakers. There is, however, one rather surprising exception. The speakers from the oldest group – those over 66 – use *ought* less than those aged 46-65, and only slightly more than those aged 25-45. However, there is far less language from this age group in the corpus, so this figure is probably less reliable. Another explanation could again be the sociolinguistic aspect. As seen in section 5.1, *ought* is more common in formal settings. Such settings are usually in the workplace in specific occupations. The over 66 age group is likely to consist of a large proportion of retired people, and it may be that a smaller proportion of the language users in this category find themselves in the speech situations where *ought* is most commonly used. On the other hand, the age group between 46-65 is likely to contain the highest proportion of language users who regularly find themselves in situations where formal language is called for. This may partly account for the uneven distribution between the age groups. The tendency is so clear, however, that if we consider the age groups from 18-65, older language users are more likely to use *ought* than younger speakers.

FIG 5.5 – The distribution of *ought* in the DCSPE according to speaker age given in instances per 1000 words.
As a comparison to the findings, I turned briefly to COLT. In the COLT corpus, there are only 22 occurrences of *ought*. As this corpus only contains slightly over 500,000 words, this roughly equals 0.044 occurrences per 1000 words. This fits in well with the general trend that the word is used less the younger the language users are, and places London teenagers even lower than where one would expect them to be, if we compare these figures with those from the DCPSE. However, as this is such a small corpus, and also geographically limited, and contains informal language, I do not regard the results as very representative with regard to the general usage of British English, even among teenagers. It does, however, seem to confirm my previous findings from other corpora, regarding the general frequency of *ought* by the younger group, and provides us with useful additional information concerning the use of *ought* by the age groups by which we have the least information from the other sources.

**5.3. Distribution according to semantic meaning in spoken language.**

A there were some interesting findings in LOB, FLOB, Brown and Frown regarding a shift in the most common semantic meanings of *ought* in more recent times, I decided to do a similar analysis with the examples from the DCPSE, in order to see if spoken language differs significantly from written language with respect to this. The results including the raw frequencies as well as the % of the total number of instances are given in table 5.4 below, while Figure 5.6 is a graphic illustration of this distribution.

**Table 5.4 – The distribution according to semantic meaning given in % (Raw frequencies).**

<table>
<thead>
<tr>
<th>Semantic Meaning</th>
<th>Total number of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obligation</td>
<td>63% (104)</td>
</tr>
<tr>
<td>Advice</td>
<td>26% (43)</td>
</tr>
<tr>
<td>Natural expectation</td>
<td>7% (11)</td>
</tr>
<tr>
<td>Logical consequence</td>
<td>1% (2)</td>
</tr>
<tr>
<td>Irrelevant</td>
<td>2% (4)</td>
</tr>
<tr>
<td>Unclear</td>
<td>1% (1)</td>
</tr>
</tbody>
</table>

**FIG 5.6 – The distribution of *ought* according to semantic meaning in spoken discourse according to the DCPSE.**
The figures reveal a clear dominance by the semantic meaning of obligation in the oral parts of the corpus. It must be borne in mind that classification is more difficult with transcripts of spoken texts, as I have said before. Many instances are quite unproblematic, like the obligative example 76 below and example 77 which clearly implies natural expectation:

(76) *I ought to know this.* (DCSPE DL A02 0090)

(77) *They ought to know something about that* (DCPSE DL B10 0349)

Others are more difficult to place in a distinct semantic category. Sometimes the context helps. Example 78 below could imply some kind of obligation that someone ought to be making “things” easier, but the context reveals a conversation regarding changing circumstances with the logical consequence that life should become easier.

(78) *Things ought to get much easier now really* DL B10 0789

In example 79 below, the implication could be some kind of natural expectation in the sense that “There should be a textbook here somewhere – I’m sure there is” The context, however, reveals a conversation implying that someone trying to persuade someone else to write a textbook. The implication is clearly that there really should be a textbook, but there isn’t. I have therefore classed it in the obligation category.

(79) *There ought to be a textbook.* (DCPSE DL–B10 0832)

Example 80 below:

(80) *You ought to go Louisa* (DCPSE DL–B32 0416)
looks at first glance as a direct order (so obligation) in the sense of “I want you to leave” or advice in the sense of “you will probably benefit from going”, depending on tone of voice and facial expression. As the Christian name is used, it looks rather emphatic or enthusiastic. However, the context reveals that neither participant in the conversation is called Louisa. In fact, they appear to be talking about Louisa, not to her. The use of *ought* appears to be advice, but the meaning appears to be more a sense of “You ought to say ‘Louisa…’ “

There are also four examples similar to 81 and 82 below:

(81) You ought to see what I got for your Christmas present (DCPSE DI B28 0256)

(82) You ought to hear what they shouted at me (DCPSE DI B32 0295)

Semantically, these do not fit in to any of the traditional categories. The meaning is more “It’s a shame you didn’t”, but really not advisable, or suggesting any obligation or expectations. There were four of these instances in the 800,000 word DPCSE corpus, but I only found two in the 4 million words which make up LOB, FLOB, Brown and Frown. Nor is this meaning of ought mentioned in Merriam-Webster's definition of the word. This semantic meaning therefore appears to be more common to spoken language.
6) The use of *ought* in recent British English as seen from the BNC investigation.

A BNC search yields 5838 instances of the word *ought* in 1624 different texts. This is a frequency of 59.9 pr million words. As the BNC represents BrE around 1990-94, it is comparable to FLOB. The figure from this corpus is so close to the figure from FLOB that it seems to confirm the previous findings.

As in the DCPSE corpus, The BNC search reveals that the frequency is much higher in the case of spoken texts, with 123.58 instances per million words (see figure 6.1). The frequency in written texts is 52.24 per million words. This may be partly explained by the fact that sentences generally tend to be shorter in spoken than in written language. Another common feature of spoken language is false starts and incomplete sentences. This means that there will be a larger number of clauses, and thus allows for more opportunities to use modal auxiliaries in spoken language.

These false starts, phatic pieces of language and incomplete sentences mean that sentences in spoken discourse are indeed shorter, but they cannot fully account for such a major bias in the distribution of modals between the two parts of the corpus, and we have to conclude that there is a clear difference in the favour of spoken language.

**FIG. 6.1 – The frequency of the word *ought* per million words in spoken, compared with written English accorded to the British National Corpus**

If we look at the categories for spoken language in the BNC (Fig 7.2), the figures indicate that the word is more frequent in formal settings. When divided into user domains, *ought* is used more than three times as often in a public/institutional setting than in a less formal setting, such as leisure, so there may be a trend toward democratisation, despite the
more frequent use in oral language. This again confirms the previously reported findings from the DCPSE corpus.

**FIG. 6.2** – The distribution of *ought* in spoken texts according to speech categories given in instances per one million words (figures rounded to the nearest whole number).

A look at text domains in written language in the BNC (Fig.6.3) largely confirms the distribution as seen in LOB and FLOB.

**FIG. 6.3** – The distribution of *ought* in written texts according to text categories in the BNC (figures rounded to the nearest whole number)

Again, we find that fiction texts (here labelled “imaginative”) or religion (here, “belief and thought”) use this word much more frequently than news and scientific texts.
We also see the same tendency as in LOB and FLOB, where first- and second-person subjects are less frequent than third-person subjects. A search for “I ought” yields 5.59 tokens per million words, “we ought” yields 6.97 tokens per million words, while “you ought”, which of course contains both the singular and plural second person reference, yields only 6.04 instances per million words. Also, the bias toward being used in fiction is even more striking in the case of second person usage. As shown in diagram 6.4 (below), of the total 373 instances of you ought in written language, as many as 266 are found in imaginative literature. This is massively above the frequency shown in any of the factual text categories. While you ought shows a frequency of 16.23 per million words in imaginative literature, it only shows a frequency of 2.99 instances per million words in the second most frequent category, belief and thought.

FIG 6.4 – Distribution of you ought in text categories in the BNC (written texts)

Incidentally, as many as six of the nine instances in this second category features the construction you ought inside a quote, marking it as something said by someone other than the author often exemplifying something someone may say to the reader, but frequently generally attributing statements to someone else, as in examples 83-85:

(83) …one can go so far as to say “You ought to go to sleep” (BNC CBI 169)

(84) “You ought to know that”, “You should have had more sense”, “You had better check that” (BNC CBI 12)
(85) ...it would be to the point to say "You should have known better" or "You ought to have had more sense", (BNC CBI 81)

This is also true of at least 11 of the 28 instances found in the Leisure category, for instance examples 86-88:

(86) ..a passer-by who had added, to the astonished Shanti, "You ought to be ashamed of yourself, frightening someone like that." (BNC BIY 741)

(87) The girls in the office would say in surprise, "You ought to be shopping!" (BNC BIY 1172)

(88) Sam Peckinpah takes a look at me and he says, "Don’t you think you ought to have something to eat?" (BNC CDG 1199)

This kind of usage certainly removes the implication of obligative direction from the author, and perhaps moves the usage of you ought closer to fiction in the sense that someone might say this to the reader and expect him or her to comply. In addition to this you does not refer to the addressee of the text, thus removing any face threatening potential.

As with the texts from the DCPSE, and perhaps indicative of future development, is the fact that the frequency generally decreases with the age of the writer (in written texts) as seen in figure 6.5. While the 60+ age group display 84.39 instances pr. million words, the frequency in the age group 25-34 is only 55.1 pr. million. Frequency does, however, increase slightly in the 15-24 age group, before falling again. There are, however, only a total of around 500,000 words spoken by this age group in the entire BNC corpus. This is much less than in the older age groups, and the mere 36 instances in total may give an unreliably high rate per million words. Even more so, the observation for the youngest group is based on just a single occurrence. There are very few speakers in the 0-14 age group as well, so this is not a reliable indication of frequency for these two groups of language users. These findings do, however correspond well with the findings from DCPSE, confirming the tendency for usage of ought to decrease with the age of the producer of discourse. As the correspondence between the two corpora is as high as it is, I consider this a fairly reliable indication of a real tendency. Unlike the DCPSE material, the over 66 age group does not show a lower frequency than the one before, although the decrease from the oldest age group to the next is only very slight. This difference from the findings from the DCPSE may be a result of differences in the criteria for the compilation of material.
FIG. 6.5 – Frequency of the use of *ought* in written texts according to the age group of the author according to the British National Corpus.
7) Ought in recent American English according to the BYUC

The BYUC contains more recent language than the other corpora in this investigation. It presently has a shorter time span than the other AmE corpora, Brown and Frown, but contains language from an 18-year period from 1990-2007. It is also much larger. I therefore performed a search to see if there is any evidence to suggest that the decline of ought indicated by the results from the other corpora continues in more recent American English.

Table 7.1 – The development of the use of ought in recent AmE as seen in the BYUC

<table>
<thead>
<tr>
<th>Time period</th>
<th>Total no. of words given in millions</th>
<th>Total instances</th>
<th>Frequency per million</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990-1994</td>
<td>103.1</td>
<td>7694</td>
<td>74.6</td>
</tr>
<tr>
<td>1995-1999</td>
<td>102.3</td>
<td>5850</td>
<td>57.2</td>
</tr>
<tr>
<td>2000-2004</td>
<td>102.9</td>
<td>4470</td>
<td>43.5</td>
</tr>
<tr>
<td>2005-2007</td>
<td>64.4</td>
<td>2329</td>
<td>37.3</td>
</tr>
</tbody>
</table>

As the most recent category is presently much smaller than the three pervious categories of five-year periods, Table 7.1 gives the total number of words in each category, total number of instances as well at the figure from each category normalized into instances per million words. Figure 7.1 below gives a graphic illustration of the normalized figures.

FIG. 7.1 – The development of the use of ought in recent AmE as seen in the BYUC given in instances per million words.
Again, the results quite clearly indicate a steady reduction in the use of *ought* over time. It would therefore appear that this downward trend continues in recent years, at least in American English. In fact, the normalized figures indicate a 50% decline in the use of *ought* from the early 1990s to the 2005-2007 time period.

The BYUC also allows us to consider the difference between spoken and written language in American English. The results are shown in FIG 7.2 below, and again reveal a preference for the use of *ought* in spoken rather than written language in American English, as we have seen in the British variant.

**FIG. 7.2** – The distribution of *ought* in spoken and written material according to the BYUC given in instances per million words.

![Bar chart showing instances per million words for spoken and written language.](chart.png)

The written language in the BYUC is divided into four rather broad categories, which nevertheless indicate degree of formality quite well, as well as differences between fictional and non-fictional literature. The distribution of *ought* between these written text categories is given in Table 7.2 below. The spoken discourse in the BYUC consists exclusively of broadcast material. This is precisely the type of language in the categories which displayed the highest frequencies of *ought* in the DCPSE (see further section 5.1). This may go some way to explain the apparently greater difference between the spoken and written language in the American material than we find in the British material.

**Table 7.2** – The distribution of the use of *ought* in written text categories AmE as seen in the BYUC

<table>
<thead>
<tr>
<th>Category</th>
<th>Total no. of words</th>
<th>Total instances</th>
<th>Frequency per</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td></td>
<td>132,7</td>
<td></td>
</tr>
<tr>
<td>Written</td>
<td></td>
<td>34,9</td>
<td></td>
</tr>
</tbody>
</table>
### FIG. 7.3 – The distribution of the use of *ought* in written text categories AmE as seen in the BYUC given in instances per million words.

<table>
<thead>
<tr>
<th>Text category</th>
<th>instances per million words</th>
<th>given in millions</th>
<th>million</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fiction</td>
<td>69.6</td>
<td>3645</td>
<td>52.4</td>
</tr>
<tr>
<td>Magazine</td>
<td>78.1</td>
<td>2183</td>
<td>28</td>
</tr>
<tr>
<td>Newspaper</td>
<td>73.4</td>
<td>1934</td>
<td>26.3</td>
</tr>
<tr>
<td>Academic Prose</td>
<td>73</td>
<td>2443</td>
<td>33.3</td>
</tr>
</tbody>
</table>

Again, as we have seen in the material from Brown, Frown, LOB and FLOB as well as the BNC, the fictional texts display a much higher frequency when it comes to the use of *ought* than non-fictional texts. Also, the fact that academic prose displays a slightly higher frequency than the other categories seems to fit with the previous findings which indicate that there is a higher frequency of *ought* in formal language, as academic prose is generally thought of as quite formal.
8) Some special uses of *Ought*

8.1. The negative – *oughtn’t*.

The evidence suggests that the negative contraction *oughtn’t* is rare in any variety, but perhaps distinctly British. I found only a single instance, in the ‘Brown Family’ corpora, and that was found in LOB. But of course, the single instance, shown in example 89 below, is not sufficient to draw sweeping conclusions on the basis of.

(89) Even if Mary had given a wink meaning "Sam’s expecting me, they oughtn’t to have fallen for that flattering assumption of sophistication. (LOB N text)

In BNC, however, I found 34 instances of *oughtn’t*. As BNC consists of 100 million words, this figure may be normalised to 0.34 instances per million words and thus supports the findings from the Brown corpora that although this is rare, it is occasionally used, at least in BrE. Typically, the instances of *oughtn’t* were followed by *to*, as in the LOB example.

(90) She oughtn’t to learn it from the police or the television news. (BNC CJF 1837)

(91) Foxy was saying how they were all standing on the balcony saying that she really oughtn’t to wear those shorts cos she’s got such thunder thighs. (BNC KP6 691)

There were, however, instances of *oughtn’t* which were not immediately followed by *to* in the BNC. 11 were interrogatives where *oughtn’t* was followed by the subject of the sentence, and then *to*:

(92) *Oughtn’t* we *to* ring up the police -- about the burglary? (BNC J54 1883)

(93) *Oughtn’t* we *to* give him just five minutes? (BNC AEB 3422)

(94) *Oughtn’t* you *to* be fetching Lisa -- she’s been there since 3 and if they’re having a party tonight they’d probably like to get her off their hands (BNC H9G 1451)

There were five instances of *oughtn’t* which were not followed by *to* at all. Of these, two were tag questions:

(95) I mean, that’s what it ought to feel like, oughtn’t it? (BNC EDJ 2247)

86
(96) After all, a general ought to have a wife, a family, oughtn’t he? (BNC GUG 2045)

But the remaining three were declarative sentences.

(97) They did what they oughtn’t. (BNC A69 1350)

(98) I oughtn’t snap, I think. (BNC HGL 1243)

(99) The interview ended shortly and our Elt said something about how perhaps he oughtn’t say that sort of thing on the Northwestern TV station (BNC J1J 600)

Previously, there have been no comparable corpora in size to the BNC that give us a similarly detailed investigation into AmE as in BrE, so Gunnel Totties claim that oughtn’t to is unique to BrE until recently seemed reasonable until very recently. The investigation using Brown and Frown, LOB and FLOB simply did not present any counter-evidence, and the diachronic investigation was restricted to ‘Brown Family’ corpora. With such uncommon constructions as oughtn’t that is probably rather unreliable.

However, in an elicitation experiment performed by Stig Johanson (1978:211), the sentence in example 100 below was presented to a number of native speakers of British and American English.

(100) You oughtn’t to waste our money on smoking (S. Johanson 1979:211)

The sentence received a slightly higher score on acceptability from British than from American university students. On a scale from 1-5 where 1 indicates unacceptability and 5 the highest rate of approval, the mean score from the British speakers was 3.96, while the mean score from the American speakers was 3.15. Despite the relatively high acceptability ratings, however, the corpus evidence suggests that speakers and writers rarely use these constructions when using the language.

With the release of the BYUC, however, the claim that oughtn’t is restricted to BrE could be put to the test. As Johansson’s elicitation test also indicated, it would seem that it is not quite so. Again, it is clear that this is indeed a rare construction. With a total of 51 instances in the BYU Corpus, which can be normalised to 0.0014 instances per million words, the evidence suggests overwhelmingly that it is much less common in the American variety of the English Language than in BrE. However, the 51 instances found in the BYUC prove that it is not exclusive to BrE.
13 of the instances were found in spoken language. As almost a fifth of the corpus contains spoken language, result showing that around a quarter of the instances were from spoken language are not surprising, as this bias toward spoken texts has been found throughout the investigation.

There are also some notable differences between the instances found in the BNC compared with these American instances found in BYUC. First and perhaps surprisingly, there does not seem to be the same strong correlation between oughtn’t and to. In the American corpus, only 25 – or just under half of the instances, were immediately followed by to as was far more common in the British variety.

(101) A party system is a good thing and you oughtn’t to wreck the party system with Ross Perot. (BYUC SPOKEN text 1992 CBS Facenation)

(102) “Hon, you oughtn’t to litter,” Rosie admonished. (BUYC 1996 FIC BkJuv:Eveningstar)

Only two instances (Example 103 and 104 below) were interrogatives where to appeared immediately following the subject

(103) Anyway, I think you owe it to me. Oughtn’t you to alter yourself as drastically as you altered me? (BYUC1994 FIC Fantasyscifi)

(104) Wasn’t he a pack-and-a-half—a-day man? Oughtn’t that to be absorbing him at present? (BYUC 2001 Fiction BkGen: Livingtotell)

The remaining 24 instances, however, were examples of oughtn’t being used as a modal auxiliary in the strictest definition of the word, without being followed by to at all, as in examples 105 and 106 below.

(105) People oughtn’t be penalized because they are successful. (BYUC 1990 Spoken CNN_Crossfire)

(106) She eased her body down. I oughtn’t be this tired. (BYUC 1996 Fiction KenyonRev)

Although with 24 instances in a 360 million word corpus this can by no stretch of the imagination be said to be a common feature in American English, or even a more common feature in AmE than in BrE, it does appear, at least, that when oughtn’t occurs at all, it is more likely to occur without being followed by to in American English than in British English.
The instances of *oughtn’t* seem to be distributed fairly evenly between the spoken, fictional and magazine genres. No instances are found, however, in newspapers or academic prose. This may, however be partly due to the rarity of the phenomenon. There is nothing in this material that seems to indicate that *oughtn’t* is tied to any particular genre. From the corpus evidence, some conclusions may be drawn regarding the use of *oughtn’t*:

1) It is a very rare phenomenon, and even less common in American English than in British English.

2) It is, however not, as Gunnel Tottie claims, exclusive to British English.

3) When it does occur, users of American English are more likely to use this negative form of *ought* as a true modal auxiliary without immediately following it with *to*.

### 8.2. The to-less *ought*

Interestingly, an investigation into the to-less *ought* (*ought* which is not followed by *to*) gave some striking results in the ‘Brown Family’ of corpora. In a more narrow definition of modal auxiliaries than the one I have used, this would be considered the only true instances of *ought* as a modal auxiliary. *Ought to* would by many be considered to be a semi-modal because it is quite clearly commonly followed by *to*. The Longman dictionary of the English language, in fact, goes as far as claiming that:

> Ought should be followed by to, a fact which is sometimes forgotten in such combinations as "ha ought and could have told me". Young people are beginning to omit to "you ought go" but this construction should be avoided in formal writing. (Longman 1984; 1041)

A corpus investigation confirms my suspicion that researching only to to-less ought would give so few instances that it would render any conclusion useless. I will, however, comment briefly on the findings from this part of the investigation, too.

I asked the Wordsmith programme to exclude all instances of *ought* which were immediately followed by *to* in the ‘Brown Family’ of corpora. Some results were irrelevant, like where the coding of the programme had added numbers after the instance in between the *ought* ant the *to*. There were also three instances discussing *ought*, which were, of course, disregarded:

(107) But like many another gap that appears in philosophy (here readers will be reminded of the familiar gap with which moral philosophers are plagued between the 'is' and the 'ought'. (LOB J text)
(108) Within the horizon of an anthropological ethic of transcendence, individuals may yet find it possible to exercise the freedom to call society into question and initiate a social transformation—one that brings the Is under the judgment of the Ought. (Frown D text)

(109) But its demonic propensity is to create an absolute or "closed" order (in which Is = Ought) that prevents the continuing transformation of self and society. (Frown G text)

Two of the remaining instances featured a to following a negation (ought not to)

(110) On review the Supreme Court, via Mr& Justice Frankfurter, found southern racial problems "a sensitive area of social policy on which the federal courts ought not to enter unless no alternative to adjudication is open". (Brown J Text)

(111) As Bishop Talbot had told me that I ought not to spend many years in Tatsfield, we held great family consultations. (LOB D text)

Four instances in the British corpora (Three from LOB and one from FLOB) and one instance in WC had a to following some kind of comment adverbial or modifier:

(112) The passages dealing with the five "other Ministries of Grace (confirmation, holy order, holy matrimony, the ministry of absolution, and the ministry of healing) ought, we suggest, to be dropped entirely. (LOB D text)

(113) If, as I suspect, there are many missing it for one reason or another who ought in their own, and in the public interest, to be having full-time education after the age of 18; and if we are determined, as we ought to be, that they shall have a more adequate opportunity, we could add the extra numbers to the universities and to the specialist colleges in a proportion similar to that already existing between them. (LOB G text)

(114) Defenders of classical culture are apt to argue that science and technology, which are concerned with means, ought properly always to be subordinate to the arts, the humanities and religion, which are concerned with ends. (LOB G text)

(115) The petitioner would then request that the house overrule the injunction or, alternatively, make a clear determination on where the suit ought best to be tried. (FLOB J text)
(116) Maori language **ought** at least to be available to students (WC B text)

The remaining seven instances, however, were examples of what I consider genuine to-less **ought** — i.e. where there is no to in the clause at all:

(117) Ought not an edifying Trial have made every effort to demonstrate this once and for all by showing how representative types of “mere” anti-Semites were drawn step by step into the program of skull-bashings and gassings? (Brown F text)

(118) But we ought not overlook the generally youthful character of Bishop’s experiments with surrealism. (Frown G text)

(119) there is not, or for him ought not be, any difference: there is only a singular, compact identity. (Frown G text)

(120) But then Wittgenstein is simply making a bet on the future course of science or else he is engaged in a priori anti-science, denying a priori that certain projects could bring results, and hence they ought not even be investigated empirically. (Frown J text)

(121) Ought he to smoke or drink? (FLOB J text)

(122) Carefully, he made sure the labels on the specimens for the laboratory and the clinical chemistry department read as they ought, and placed them aside for despatch the following morning. (FLOB L text)

(123) “I work harder than usual, smoke and drink more than I ought, but I still find myself —” He broke off. (FLOB P text)

Four of these seven instances were taken from the American corpora, one from Brown and three from Frown. These were all negations immediately followed by **not**. (examples 117-120)

The final three (121-123) were all found in FLOB. Example 121 is an interrogative, but unlike example 117 from the American selection, it is not a rhetorical question followed by **not**. Semantically, it may be rhetorical, as the answer is, perhaps obvious, but it at least appears to be a genuine interrogative. The other two are declarative sentences. Notably, the declarative sentences are subordinate clauses with ellipsis of the main verb, so perhaps there
is a pattern that if the verb is omitted after the auxiliary *ought* the infinitive marker may disappear as well.

These findings suggest that to-less ought is found in AmE as well as BrE. But the absence of the construction in WC or ACE suggests that it is, perhaps, largely restricted to AmE and BrE. In AmE, however, the to-less *ought* seems to be restricted to negations where *ought* is immediately followed by *not*.

In BrE the use of to-less *ought* is quite rare, but it is sometimes used in contexts where AmE would require a *to*. The fact that there are no findings in LOB, but three in FLOB suggests that this use of *ought* may be a rather recent development.

BrE (and to a lesser extent NZE) seems to allow for adverbials, modifiers and other comment clauses between *ought* and *to*. This is not found in the American or the Australian material. It is, however, found in LOB as well as FLOB and WC. Perhaps then, the extended use in time and limited spread to NZE suggests that the use of the to-less *ought* in BrE could have its origins in constructions where a comment clause is inserted between the *ought* and the *to*. The exclusion of *to* altogether may then be a more recent result of the continuing process of allowing to-less *ought* to be used in BrE.

This, at least partially, supports the findings by Svartvik and Wright (1977:185-201), who conducted an elicitation experiment using British teenagers as informants, which showed that:

The to-less *ought* construction is not only extremely popular, but is, in fact the preferred form in the negative and negative-interrogative sentences. (Svartvik and Wright (1977:191)

However, they also observed that:

In assertive contexts there is, however, a safe majority for the *ought to* construction...at least three-quarters of the subjects rejected all the non-assertive constructions with *ought*. (Svartvik and Wright 1977:192)

Rationalizations for the rejections largely consisted of responses from informants that the construction seemed ‘old-fashioned’, ‘Victorian’ and, in some cases, ‘sounded middle-class’. Unlike in Svartvik and Wright’s experiment the findings from the ‘Brown Family’ material suggest that the use of *ought* in negative sentences with the uncontracted form *ought not* seems to be tied to American usage to a large extent. BrE, on the other hand, allows for the to-less *ought* to be used in other contexts. However, the Svartvik and Wright experiment does date back to the 1970s. This must be borne in mind since the evidence suggests some change in the use of *ought since then

The contracted form *oughtn’t*, however, is used occasionally in BrE, but no evidence has been found of this in other varieties. See for instance {examples 97-99 in section 8.1}
In the material from COLT, which is comparable to the language use by the informants in Svartvik and Wright’s material, no occurrences of to-less ought were found except for an irrelevant instance (example 124 below).

(124)  Who is nothing? Oh bloke he says he’s nought taken ought. ... I told you. Yeah yeah Alex Alex. Tell me the answers quick tell me the answers.

This illustrates the problems with native speaker intuition, that elicitation tests rely on. Informants answer questions regarding what they think is how they and other language users express themselves. However, corpus studies often reveal that how language is used in practice quite often differs from what informants think.

These are, however, few instances of the to-less ought. Although the distribution is, in my view, striking, care must be exercised when generalizing on the basis of a small handful of instances.

8.3. Ought in tag questions

As all the results from the corpus investigations indicate that ought occurs more frequently in spoken texts, I wanted to investigate whether this pattern could be attributed to it being more likely to be used in tag questions in spoken than in written language.

I started out with the examples I had uploaded to FileMaker, which came from LOB, FLOB, Brown and Frown. Although the material contained a total of 12 interrogatives – 7 from LOB, 3 from FLOB, and one each form Brown and Frown, none of these were tag questions.

However, when I turned to the spoken corpus DCPSE, I only found a single instance there (example 125 below)

(125)  They ought to be outside anyway Oughtn’t they? (DCPSE DL-B28 0128)

It is interesting that to is ellipted from these examples, too. This perhaps supports my hypothesis that this is tied to an omission of the main verb (see further section 8.2).

This is clearly too small an example for any firm conclusions to be drawn. I therefore turned to the much larger BNC.

The BNC only displayed four instances of ought appearing in tag questions. The two examples below (126 and 127) featured ought in a positive tag:

(126)  He oughtn’t to have seen that, ought he? (BNC J10 1062)
(127) Klaus, isn’t it? and Klara -- Oh, of course -- “corrected -- “Erika, Fritz, Rosa ... we really ought to see you more often, ought we not? (BNC A7A 85)

The other two, shown in example 128 and 129, featured the negative variant *oughtn’t* in the tag:

(128) I mean, that’s what it ought to feel like, oughtn’t it? (BNC EDJ 2247)

(129) After all, a general ought to have a wife, a family, oughtn’t he? (GUG 2045)

Common to all of these examples was not, as I had expected, that they appeared in spoken texts. In fact, they all appeared in written, fictional texts. They did, however, appear as part of a fictional dialogue.

This is clearly very few instances, so I hesitate to draw any firm conclusions. However the fact that this construction is occasionally used by authors when simulating conversation but not in other contexts may indicate that the phenomenon is thought of as a purely spoken language phenomenon. The difficulties in finding this construction in genuine spoken language except for in the single instance from the DCPSE may also indicate that it is, in fact, less common than authors perhaps think.

I also performed a similar search in the BYUC in order to see how *ought* is used in tag questions in American English, but that search proved fruitless as I could not find a single instance of *ought* or *oughtn’t* in tag questions there.

It seems, therefore, that this is a purely British English phenomenon and a rare one, particularly in genuine non-fiction language, at that.
9) Alternative constructions to *ought*.

Finding a complete list of alternative constructions to a linguistic phenomenon is notoriously difficult, and I am not in any way implying that my list is by any means an exhaustive list of alternative constructions to *ought*. How to best go about finding these constructions is debatable. There are a few available resources, and I have considered some. An obvious resource is, of course, a dictionary. I also have a number of British English native speakers available to me, so native speaker intuition is another good source.

However, as one of my main intentions in writing this thesis is to demonstrate how corpora may be used in order to conduct research, and to explore some of the ways in which corpora may provide information in order to answer a series of research questions, I have decided to consult corpora first in this question as well.

Among the corpora available to me is the ENPC, which was presented in section 3.11. I have looked at how *ought* tends to be translated into Norwegian. I have then gone on to see how these Norwegian expressions have been translated into English. In this way I hope to come up with a list of alternative expressions. I then turned to dictionaries and native speakers in order to compare the results from the ENPC investigation with the list gathered from these sources. By doing this, I have attempted to collect a reasonable list of alternative expressions, which in turn could be searched for in LOB, FLOB, Brown and Frown in order to see if the alternative expressions increase in use while *ought* decreases over time.

9.1. ENPC Results.

The ENPC results seem to confirm my suspicions that *ought* is most commonly translated as *burde*. As in my results from previous corpus investigation, *ought* is much more commonly used in fictional texts than in non-fictional ones. Of the 52 instances in total in the ENPC English original material, only seven are found in the non-fictional section. I have, however, added the two sections together when looking at then preferred translations.

As table 9.1 below shows, *burde* is by far the preferred translation with 31 instances.

**FIG. 9.1 – Translations into Norwegian of *ought* in the ENPC**
In one of the instances from the non-fictional section, example 130-131 below, *ought* appears in a quote which is not translated at all:

(130) Mobs roamed the streets chanting "Spain, Spain, you *ought* to be ashamed", and with the discernment of mobs everywhere broke the windows of shops owned by local Jews. (RF1)

Mobben ravet i gatene og sang "Spain, Spain, you *ought* to be ashamed", og med samme sans for sammenhenger som mobber alltid har hatt, knuste de vinduene i de lokale jødiske butikkene. (RF1T)

In as many as seven of the instances, the translators have opted to re-write sentences in a way that has led to an omission of the modal auxiliary, for instance in example 133 below where the translator has opted for an object with complement, rather than the modal auxiliary, or example 135 where the translator has opted for a heller/enn contrasting sentence (perhaps best back-translated as rather/than construction) rather that the seemingly slightly less directive ‘more than’ construction in the original.

(132) “You mean you don’t like them, or you don’t think you *ought* to have them? (PDJ3)

(133) ”Mener du at du ikke liker dem, eller at du ikke *synes det er noe for deg?” (PDJ3T)

(134) We know that Jeremiah came of a family very highly connected with political and religious authority in the reign of Josiah the
King, but I suggest that we *ought* to be more interested in Jeremiah's extraordinary psyche than in his politico-spiritual orientation – though doubtless the two matters were closely related. (HB1)

(135) Vi vet at Jeremia kom fra en familie som hadde svært gode kontakter hos de politiske og religiøse myndighetene under kong Josias. Jeg foreslår likevel at vi **heller** interesserer oss for Jeremias usedvanlige psyke enn for hans politisk-åndelige orientering, selv om disse områdene utvilsomt var nært forbundet. (HB1T)

Such examples have to be disregarded, because the options for translators to re-write sentences are far too numerous to include in a list of alternative expressions, but perhaps the relatively high number of such rewritten sentences points to the fact that modality is a elusive type of meaning.

**Burde** is by far the most common translation of *ought*, with 31 of the 52 instances. This translation is used in when *ought* has a meaning of moral obligation attached to it, such as in example 137 below, as well as when the semantic meaning is more of a natural expectation as in example 139, and advice as in example 141, so it seems to cover most of the range of semantic meanings attached to *ought*. Figure 9.1. shows a much higher frequency than *ought*, indicating that *burde* perhaps has a broader range of modal meanings.

(136) Harry, she felt, *ought* to stand between her and these embarrassments. (FW1)

(137) Hun syntes Harry **burde** demme opp mellom henne og disse kjedelighetene. (FW1T)

(138) It seemed his tongue and mouth were drier, and more blunted, than they *ought* to be, though whether this was caused by laurel sap or by the juice of orange he could not tell. (JC1)

(139) Det virket som tunga og munnen hans var tørrere og mer numne enn de **burde** være, skjønt om dette skyldtes laurbærsøvje eller appelsinsaft kunne han ikke si. (JC1T)

(140) "I think he *ought* to be taking Sanatogen." (NG1)

(141) "Han **burde** ta Sanatogen." (NG1T)
However, I could not find any examples in the ENPC material where *ought* in the sense of prediction is translated as *burde*. The predictive *ought* seems generally to be translated into *skulle*, as in example 142. In fact, half of the instances where *ought* is translated into *skulle* are when *ought* has been used in the predictive sense.

(142) Hvis du har regnet riktig *skulle* tallet være fire tusen tre hundre og tre pund og femti pence. *(RD1)*

(143) If you ’ve done it right it *ought* to be four thousand three hundred and three pounds and fifty pence. *(RD1)*

The remaining examples where *ought* is translated into *skulle* are when it is used in the sense of advise such as in example 145 or obligation, as in example 147.

(144) I don’t think you *ought* to be sharing it with strangers." *(TH1)*

(145) Jeg synes ikke du *skulle* dele det med fremmede." *(TH1)*

(146) I think in the morning we *ought* to tell Daddy what she ’s doing." *(TH1)*

(147) Jeg synes vi *skulle* fortelle pappa om det i morgen." *(TH1)*

So although the semantic meanings of *burde* and *skulle* are quite similar in many instances, they do not completely overlap in this material. Other translations include two instances where *ought* is translated as *bør* (149 and 151). These carry with them the implication of perhaps moral expectation, or at least a strong sense of someone being expected or even required to do or know something.

(148) Doctors are deluged with it, but a lot of the advertising does n’t tell a physician what he *ought* to know – especially about side effects of drugs, including dangerous ones. *(AH1)*

(149) Legene blir dynget ned, men en stor del av stoffet sier ingenting om hva vi *bør* vite, f. eks. om bivirkninger, mer og mindre farlige. *(AH1)*

(150) “She ’s no better than she *ought* to be," *(DL1)*

(151) ”Hun er ikke bedre enn hun *bør* være,” *(DL1)*
Translations also include two examples of *ought* carrying an even stronger implicature of requirement (example 153) which have been translated into *må*, and one instance where the implicature is more of an unfulfilled wish where the translator has opted for *ønsket seg* in example 155.

(152) "Do you think that all children's books *ought* to have funny bits in them?" Miss Honey asked.  
(RD1)

(153) Synes du at det *må* være noe morsomt i alle barnebøkene? spurte frøken Honey.  (RD1T)

(154) He thought he *ought* to have some guidance towards those that were important, which might feed the yearning for whose satisfaction he was not sure what was really needed.  (NG1)

(155) Han *ønsket seg* veiledning i hvilke bøker som var viktige, hvilke bøker som kunne nærre den lengselen som han ikke med sikkerhet visste hva gjaldt.  (NG1T)

I then looked at the translations of *ought*, i.e. *burde, skulde, bør*, and *må*, and how these tended to be translated into English in order to see which alternative expressions to *ought* would be uncovered in that way. Figure 9.2 below shows how the most common translation, *burde* was translated into English according to the ENPC:

**FIG. 9.2 – *Burde* translated into English according to ENPC**

![Diagram showing the translation of "ought" into English according to ENPC](image)

---

19 This is just the word form *burde*, I have chosen not to include the lemma form *bør* but will deal with that separately later in this section.
As seen in figure 9.1, *ought* was by far the most popular translation, perhaps even showing a certain overuse by translators. Such an overuse fits in with claims made by Jenkins (2003: 16) that while inner-circle English is norm-providing, English used in the outer circle and expanding circle is norm. One would therefore expect ‘Expanding Circle English’, such as English in Norway, to be ‘lagging behind’ when it comes to a development of linguistic change. Most translators in the of texts in the ENPCE are native speakers of English, so they are perhaps EFL speakers (Melchers and Shaw 2003:8) but many of them probably live in Norway, and as such be slightly detached from the English native speaking community. This may lead to them being slightly behind language users living in an inner circle setting when it comes to picking up on tendencies when it comes to language change. The overuse of *ought* may also be explained by referring to the concept of ‘translationese’ – an automatic translation of *burde*, which leads to overuse in the translation (Gellerstam, 1986).

*Ought* was, perhaps not surprisingly, followed by *Should*. Or the negative *shouldn’t*. Other translations included the more emphatic *Need* (ed), as in examples 157 and 159 below

(156) Det medførte at han oppdaget at garasjen *burde* males, at takstein *burde* skiftes, og at garasjeporten hadde kilt seg fast. *(OEL1)*

(157) He noticed that the garage *needed* painting, that slates *needed* changing, and that the garage door had got stuck. *(OEL1T)*

(158) Det høye gresset *burde* vært klippet, men hun hadde ikke hatt ork til det. *(EG1)*

(159) The grass badly *needed* cutting, but she didn’t feel up to it. *(EG1T)*

. There were also single instances of *must* (example 161) *will have to* (example 163) and *Might be better* (example 165), which would all seem to be viable replacements for *ought* in this context.

(160) Og fremfor alt,” tilføyde han med et smil, ”skal vi ikke krydre den ut fra våre egne formeninger om hvordan den *burde* vært. *(FC1)*

(161) And above all, he added with a little smile, we *mustn’t* season it according to our own preconceptions. *(FC1T)*

(162) En av dem *burde* pusse nesen hans snart. *(BV1)*
(163) One of them will have to go and blow his nose in a minute. *(BV1T)*

(164) Jeg tenkte at det kanskje burde stå under transportbyråer i stedet. *(LSC2)*

(165) I thought it might be better listed under moving and storage companies instead. *(LSC2T)*

Three instances had to be disregarded because the whole sentence containing *burde* had been left out in the translation, and three because, as in examples 167 and 169, the translator had opted for something entirely different than the use of a modal auxiliary.

(166) Jeg *burde* hatt alle betingelser. *(EHA1)*

(167) I would have met all the conditions. *(EHA1T)*

(168) *Er det likevel noe som burde interessere alle?* *(JG1)*

(169) *Is there nothing that interests us all?* *(JG1T)*

*Skulle* produces as many as 896 hits in the ENPC. As I am looking for tendencies and preferences in this study, and it is limited in time as well as space, it is beyond the scope of the present study to investigate such a vast number of instances in detail. I therefore decided to limit my search to the first 100 instances where there were more than 100 hits in the ENPC corpus. Of these 100 instances, the distribution of translations is as shown in Figure 9.3 below:

**FIG. 9.3 – Translations of *skulle* in 100 instances from ENPC.**
The findings reveal a rather more widespread distribution of translations in the case of *skulle* if compared to *burde*. Most of the instances are not relevant as an alternative to *ought*. *Should* is, of course, an option, and has been mentioned previously. *Need to* and *must* are, as mentioned previously, more emphatic, but could perhaps replace *ought* in some contexts. *BE to* and *BE supposed to* indicates more of a demand than the advisive function of *ought*, but in some contexts the advice in ough is so strong that the speaker expects or even demands that the advice is followed, so perhaps these expressions may occasionally overlap to some extent. I will therefore include them in the list of possible alternative expressions, even though the relevance is fairly restricted and highly context-dependent.

*Bør* is another one of the translations of *ought* found in the ENPC. It is found 34 times among the Norwegian originals, and the distribution of translations is shown in figure 9.4 below:

**FIG. 9.4 – Bør translated into English according to ENPC**
Again, *Shouldn’t* is the dominating translation, followed by *ought*. There were four instances featuring the homonymous use of *bør*, where it is used as a noun in the sense of *burden*, as in example 170 below:

(170) Det var en tung bør å løfte, og gulvet var myret under meg. (GS1)

(171) It was a heavy *burden* to lift and the floor felt like a quagmire underneath me. (GS1T)

These were, of course, disregarded, but the three instances of has/had to shown in examples 173, 175 and 177 below, although more emphatic and seemingly giving less optionality than is generally associated with the use of *ought*, could be thought of as advice because the context lays out something which would be of benefit to the listener or the subject. Therefore, I will regard them as possible replacements for *ought*.

(172) Når man som jeg hører hjemme på det trivielle plan, bør man vokte seg for å strebe etter det tragiske. (EHA1)

(173) Someone like me, at home with the trivial, *has to* watch herself so as not to aim for the tragic. (EHA1T)

(174) Når den ene snur seg i posen, må den andre snu seg, når den første puster ut, bør den andre helst puste inn. (KH1)
Each time one of them turned over in the sleeping-bag, the other had to turn too; when one breathed out, the other should be breathing in. (KH1T)

Når den ene sover, bør den andre sove, når den ene taler — drevet til det av en kjedsommelighet som kan ta livet av en mann — bør den andre tvinge seg til å lytte enten han liker det eller ikke. (KH1)

When one of them slept, the other had to sleep; when one felt the need to talk — driven to it by fatal boredom — the other had to force himself to listen whether he liked it or not. (KH1T)

The final translation of ought is the more emphatic må. Må is used 589 times in the Norwegian original material. As with skulle, I will limit my investigation to looking at the first 100 instances. The distribution of translations is as shown in Fig. 9.5 below.

FIG. 9.5 – Må translated into English according to ENPC

Må is clearly more emphatic than the common meanings of ought and must is by far the most common translation. Ought is not found at all in these 100 instances. Insist that and are forced to give no room for options, either, so they are not alternative expressions to ought. Might as is used here in example 179 below does not suggest the modality of ought and is thus disregarded, even though it is possible to think of situations where the two might be used interchangeably.

Så blir vi gående side om side på en taus, halvhjertet måte hvor hver bevegelse bærer så mye forlegenhet i seg, at det er som
selve situasjonen må sprekke og falle fra hverandre hvert øyeblikk. (KF1)

(179) We walk side by side in a silent, half-heartedly way, each movement betraying so much embarrassment that I feel the situation might blow up and fall apart any moment. (KF1T)

*Need*, however, as in example 181 could be replaced with *ought* as it has an element of advice about something which would be of benefit to the listener.

(180) “Tror nesten du må begynne med hårnett, jeg!” ler hun. (LSC1)

(181) “I almost think you need to wear a hair net!” she laughs. (LSC1T)

*Will have to*, however, when used in the sense of *would be a good way of achieving one’s aims* as in example 183 below may perhaps be possible to exchange for *ought*,

(182) Kanskje han må tegne Afrika, eller hele Eidsvoldsbygningen med flagget til topps, eller kanskje han må henge ut av vinduet etter venstre øre? (LSC1)

(183) Perhaps he has to draw Africa, or the whole Eidsvold building with the flag flying high, or maybe he will have to hang out of the window by the left ear? (LSC1T)

Likewise, *will* as used in example 185 below, where the emphasis is on “for his own good it would me advisable to...” there is the option of replacing it with *ought*.

(184) Og han må tenke på bladet han svelget, det var noe av det merkeligste, han må tenke seg godt om før han gjør noe lignende. (LSC1)

(185) And he has to think about the leaf he swallowed. It was amazing. He will think better of it before he does something like that again. (LSC1T)

So, having looked at the way *ought* has been translated into Norwegian, and then looking at the way in which these expressions have been translated into English in the ENPC, I am left with a list of expressions which to a variable degree could be replacements for ought. The list is as follows:
- Should.
- Need
- Must
- Might.
- Will have to
- Will
- Could
- BE to
- BE supposed to
- HAVE to

9.2. Dictionaries

I have looked in several dictionaries in order to find alternatives to *ought*. My list includes English-English dictionaries, Learners’ dictionaries, bilingual (English-Norwegian and Norwegian-English) dictionaries as ell as synonym dictionaries.

In the online dictionary resource [www.dictionary.com](http://www.dictionary.com), the synonyms *must* and *should* are listed. The Collins paperback dictionary does not list any alternatives at all. *The Longman Dictionary of the English language*, however, states that *ought* implies *can*.

*Macmillan Dictionary for Advanced Learners* states *should* as a more common alternative to *ought*. This is supported by another Learners’ dictionary – the *Oxford Advanced Learner’s Dictionary*.

I have used three synonym dictionaries for this investigation. Penguin’s two works *Dictionary of English Synonyms and Antonyms* and *Modern Guide to Synonyms and Related Worss* do not contain any reference to *ought* or even *should* at all. Nor does the much larger *Longman Synonym Dictionary*, so I was unable to find alternatives in the synonym dictionaries at all.

As bilingual dictionaries, I have used Cappelens *Engelsk-Norsk, Norsk-Engelsk Ordbok* and *Kunnskapsforlagets Engelsk-Norsk stor ordbok* and the corresponding *Norsk-Engelsk Stor ordbok*. 
In Cappelens dictionary, *ought* is translated as *burde* and *bør*. *Burde* is in the Norwegian-English section translated as *ought to* and *should*, while *bør* is only translated as a noun into *burden, weight, load* and *medvind*, while *bør* in the verbal sense is left out entirely of the dictionary.

Kunnskapsforlagets dictionaries, however, are much larger and more detailed. In the English-Norwegian dictionary, *ought* is translated into *bør, burde* and *skulle* the dictionary also specifies that these translations may be used when *ought* is used in a sense of advisability, obligation, probability and in the case of inadequacies or weaknesses, as in example 186 below, which is taken from the dictionary

(186) It ought to have been done last year.

*Bør* is not translated in the corresponding Norwegian-English dictionary, but there is a reference to *burde*. *Burde* is translated as *should* and *ought to*. *Skulle*, however, has a range of translations. When pointing to future events, *be going to, will and shall* are suggested. This sense in non-modal, though, and is thus disregarded. The same applies to the suggested sense of a promise, where *will* is suggested.

However, in the sense more relevant sense of obligation, *have to* and *must* are suggested. Also, *skulle* in the sense of possibility is suggested translated as *be supposed to* and *be said to*. The first can clearly be substituted with *ought*, but the second is in a different sense altogether. Another sense of *skulle* is in the sense of doubt, where the suggested translations are *would, is* or *could*. These are more openly interrogative than *ought* usually is thought to be, and do not seemingly have the advising functions associated with *ought*. However as seen in the corpus material, there are occasions where at least *would* and perhaps then *could* may replace *ought* in order to reduce the face-threatening potential of an utterance, so these will be included in my list. The list of possible replacements for *ought* resulting from the dictionary investigation then is as follows:

- Must
- Should
- Can
- BE supposed to
- Would
- could
Thus the list from the dictionary investigation is much shorter than that from the corpus investigation, and only would did not crop up in the corpus investigation. From this I conclude that use of parallel corpora is a more reliable method when trying to find alternatives to ought than looking in various dictionaries.

9.3. Native Speaker intuition.

In order to supplement my list of alternative expressions to ought, I asked a group of native speakers what they might use to replace it with in some of the sentences taken from the various corpora, taking care to include the four different meanings of ought (advice, obligation, natural expectation and logical consequence). The informant group was not selected by random means or with any great thought of representation with regard to geography, socioeconomic background, age, etc. This is purely meant to be a supplement to the corpus investigation using native speaker intuition. The informants came up with a number of alternatives, and also had some interesting views on the use of ought, some of which are reported below.\(^\text{20}\)

A more thorough way of going about this would perhaps be to devise an elicitation test based on a substantial number of the examples from the corpora, taking care to incorporate the various meanings of ought and distribute it to a larger number of native speakers taking care to record variables such as age, gender, social and regional background, level of education, etc. This would, however, be time-consuming and far beyond the scope of the present study.

All of the informants I used pointed to should as an alternative. This was by far the most popular substitute expression. One person wrote that:

I generally would use SHOULD if I was writing but during speach I think I would use OUGHT about 30/40% of the time and SHOULD the remaining (Female, 27, Buckinghamshire)

This informant thus confirmed the bias toward using ought in spoken discourse, rather than written.

Another informant reported that:

I would use 'Should' in every one of those sentences. I have NEVER heard anyone using the word 'ought' in the North of England. The word 'ought' up in working class northern england is though of as very.. outdated and a little snooty. (Not that it is - i'm just reflecting opinion up there!) I have only every heard anyone use 'ought' in those BBC period dramas. I wasn't aware that it was still in use! (Female, 29, North Yorkshire)

\(^{20}\) All comments bade by informants are given in their original form
While the two informants quoted below also agreed it seemed old-fashioned and “upper-class” they admitted using it occasionally.

I find the upper class use the word ought and also the people trying to be in with them. I’ve used the word in a sentence like "you didn’t oughter do that" but I’m a bucks country boy and have had many old relatives what I’ve got these sayings from. (Male, 34, Buckinghamshire)

I would like to stick up for the word ‘ought’. In fact I used it this morning when discussing a matter with my colleague. I couldn’t remember some information and said “I ought to know this”. (Female, 31, Buckinghamshire).

There seems to be a consensus among my native speaker informants that ought does indeed sound formal, “upper class”, and old-fashioned. The truth value of this view is supported by the corpus investigations in this study. Though it is beyond the scope of the present study, however, it might be interesting to investigate dialect variation in order to see if the claim that this is more common in southern accents (and in particular the Buckinghamshire region) than in other dialects. The findings from COLT seem somewhat contradictory to this claim, as the use of ought seemed to be much less common in London teenage language than in the more general BrE material from ENPCE. This, however may be largely due to the fact that COLT contains teenage as well as informal language.

Alternatives suggested by the informants are as listed below:

- Should
- Would be better off
- Had better
- BE supposed to
- HAVE to

9.4. The development of the use of alternative expressions to ought.

The list of alternative expressions gathered from the investigation in the three previous sections were investigated in LOB, FLOB, Brown and Frown to see if reductions in the use of ought was in any way mirrored by an increase in the alternative expressions. I had gathered 13 possible alternatives, and then performed the corpus investigations. The results are given in table 9.1 below.

Table 9.1 – The development of the use of alternative constructions to ought in LOB, FLOB, Brown and Frown.
** Of these 312 in LOB, 293 in FLOB, 251 in Brown and 213 in Frown were “is to” which is often used in irrelevant constructions like:

“the main purpose is to act as a buffer” (LOB LOB A text)

or

“The patient, a single woman, aged about 30, whose home is in Cornwall, is to be released from Rampton in a few days.” (LOB A text)

Although some alternatives, like the modal auxiliary need(s) show an increase, the figures do not suggest that there is any reason to suppose that alternative expressions in general are increasing as ought decreases. Most, in fact, seem to be decreasing, and the one replacement which is most strongly suggested by all my three sources as being the most likely replacement for ought, namely should, is amongst those decreasing expressions on American as well as British English.

The increase of need to, among a few other alternatives to ought was also found by Leech (2003: 229) when he searched for what he called ‘selected semi-modals’ (BE going to/Gonna, BE to, (Had) better, (Have) got to/Gotta, HAVE to, NEED to, WANT to/Wanna and used to) in LOB, FLOB, Brown and Frown. Not all semi-modals were included in his investigation, but Leech too concludes that:

The findings […] do not in general support the view that true modals are declining because they are being supplanted by semi-modals.(Leech 2003:229)

The concept of alternative expressions is a very slippery one and it is difficult to think of a good method of finding such alternatives.

One alternative approach to the ones I have resorted to may be to conduct an elicitation experiment, asking a large number of native speakers using sentences from the corpora making and sure to include sentences implying all the different meanings of ought, as well as spoken and written material. This approach may give some fairly reliable results,
particularly if participants were selected broadly to represent a range of age groups, geographical areas and from varying socioeconomic and educational backgrounds. However, this approach would be very time consuming. Nor would the above mentioned approach go far in solving what I see as the main problem with this part of the investigation, namely the elusiveness of modal meanings which leads to situations where there are few exact alternatives. Thus finding a large number of possible alternatives to *ought* in not the major obstacle. The problem is that most, if not all, of the alternatives can in some contexts replace *ought*. However, most are often, some even mostly, used in contexts where *ought* would at the very least be an unlikely option, if not impossible. Some corpora are tagged for word class and can be asked to only find the relevant instances in a corpus. However, as I want to compare LOB, FLOB, Brown and Frown, and tagging is not available in all four, this is not an option. The tagging of the corpora where it is available is also quite unreliable, so for accuracy a manual check is required.

One would therefore have to manually go through all the instances of each alternative construction in each of the corpora, sorting them for relevance in order to get accurate results in this part of the investigation. That would be a considerable task and far beyond the scope of this particular investigation. I have therefore had to assume that the distribution between relevant and irrelevant instances of alternatives have remained fairly constant over the years.

Because of the difficulties and uncertainties about the methodology and the conflicting results, I consider the findings from this part of the investigation to be inconclusive. However, although some alternative expressions to *ought* seem to be increasing, these findings do not point to any general tendency whereby semi-modals or alternative expressions are increasing and replacing *ought*. 
10) Concluding remarks

Firstly, all the evidence suggests that the use of *ought* is decreasing, and that this is an ongoing development. The reduction in the use of *ought* appears set to continue, both because of the decrease in use by younger users and the restriction in domain. The figures from the BYUC also confirm that it has decreased quite sharply in American English in the last 18 years.

The reduction in the use of *ought* is marked in BrE, and BrE is in this instance becoming more similar to AmE. This supports Leech’s theory on Americanization, but it is hard to measure the effect of Americanization of language and to pin it down as a direct cause of language change. There is a reduction in the use of *ought* in AmE, too, but the frequency was always lower here, so the decrease is less marked. However, it is clearly happening in American English, too, and the findings from the BYUC reveal that the process is ongoing even in very recent language.

The more frequent use in oral language is surprising. It seems to contradict the belief that colloquialization is a force behind the reduction in *ought* usage. It also seems to contradict the claims made by Gunnel Tottie that there is a stronger tendency for language users to opt for *should* in spoken than in written language.

In the light of my findings, then, I believe that perhaps most weight should be attributed to the idea of democratization of language as the most influential contributor to the development. As the use in speech also seems to be largely restricted to more formal settings, there may be a tendency toward ‘democratization’ in spoken as well as in written material. Leech’s findings in the case of increased use of semi-modals also indicate that ‘democratization’ may be a factor. The statements put forward by native speaker informants seem to support this, as there appears to be a general consensus that *ought* is formal and somewhat ‘old-fashioned’ *Ought* appears to be acquiring a more face-threatening potential as an increasing proportion of the instances of *ought* implies some kind of obligation. My assumption then is that producers of discourse will avoid using it, particularly where there is a potential for the utterance to be misunderstood in contexts where recipients may infer a demand on the part of the speaker or writer, or in the instances where language users are uncomfortable with impeding in recipients.
This could be a self-enhancing process, where the word is gradually being used less and restricted to contexts where it appears to be face-threatening. It then tends to be regarded as a potentially face-threatening utterance by language users, and thus avoided in contexts where this effect is not desired. As the use of this word becomes more restricted and is seen and heard less often, it also acquires a kind of old-fashioned and specialised meaning, associated with older people, so it decreases in use. In the next instance, the less people see it in use the less likely they are to use it themselves, again adding to the specialisation in use.

As the three mechanisms outlined above tend to be at work at the same time, they all pull in the same direction and probably enhance each other’s effect. The result of this is a continued decrease in the use of *ought*. This theory is in my view particularly supported by the fact that the different semantic meanings of *ought* seem to be decreasing at a different pace. While epistemic meaning where there is no face threatening potential remains relatively stable, the root meaning is where the major change is taking place. It is interesting that this happens while, at the same time, the semantic meaning of obligation is becoming more dominant within the root sense, thus increasing the potential for it to be perceived as a potentially face threatening utterance.

Finally, I turn to the question of whether corpus linguistics can provide valuable information in the study of language change.

Hopefully, this thesis has presented a wide variety of issues which can be investigated in this way, and I hope to have demonstrated how information on actual language use may be accessed using corpora. I believe this enables researchers to describe and observe linguistic phenomena.

When used to find alternative constructions to *ought*, the corpus investigation proved the most fruitful if compared with my limited use of native speakers, as well as my rather extensive investigation into dictionaries.

I would also like to point out one interesting observation with regard to my corpus investigation when compared to native speaker intuition. Although originally Swedish, Gunnell Tottie is living in California with an American husband, She is also a linguist. Despite this, her claims that I referred to in the opening chapter of this thesis, which were presumably based on her instincts as a native or near-native speaker, that most speakers would prefer to use *should* rather than *ought*, particularly in spoken language, is at only partially correct.
Should is definitely the preferred construction, but all the corpus evidence reveals that, contrary to this claim, ought is much more likely to occur in spoken than in written language in American as well as in British English.

Corpus studies, like most research methodologies have their limitations. However, in my view, their ability to enable the researcher to quantify linguistic phenomena and to observe real language in use, provide us with insight difficult to gain from other sources.

For a study such as this, which is concerned with preferences and tendencies rather than correctness, I believe that a corpus investigation is particularly well suited.
11) Suggestions for further research.

There are many insecurities regarding the development of the use of *ought* in particular as well as modal auxiliaries in general, perhaps particularly the rare modal auxiliaries. The recent release of the BYUC and the forthcoming American National Corpus, which will resemble the BNC in composition, will allow researchers to carry out more in-depth studies concerning the differences between British and American English.

In an interview with Paloma N. Pertejo (2006:152), Geoffrey Leech presents a new corpus, currently being compiled. This new corpus is compiled to be comparable with the other ‘Brown Family’ corpora, but uses texts from an earlier period, around 1931. Leech presents some preliminary findings from this corpus and states that:

Certain of the changes we observed in the 1961-1991 corpus comparison were also found in the earlier periods, so, for example, modal auxiliaries declined between 1931 and 1961, and between 1961 and 1991. (Pertejo, 2006:153)

It would, of course, be interesting to study the occurrences of *ought* from the 1931 corpus in detail in order to see if the patterns I have observed in LOB and FLOB form part of a continuum from the earlier period.

Another very difficult task which would nevertheless provide valuable insight is to carry out a more thorough search of what is replacing *ought* as it decreases in use.

Investigating other modal auxiliaries to see if some of the findings from the investigation into *ought* are similar to findings in other modals might prove to be an interesting aspect, and in particular the rare modals which are also experiencing a decrease.
References


Johansson, Stig, 1979, ‘American and British English grammar: An elicitation experiment’, in *English Studies, a journal of English Language and Literature* 60


**Corpora**

The Brown Corpus of British English (BROWN).

The Lancaster – Oslo/Bergen Corpus (LOB).

The Freiburg-Brown Corpus of American English (FROWN).

The Freiburg-Lancaster Oslo – Bergen Corpus (FLOB).

The Bergen Corpus of London Teenage Language. (COLT)

The Australian Corpus of English (ACE)
The Wellington Corpus of Written New Zealand English (WC)
The British National Corpus (BNC)
The Diachronic Corpus of Present-Day Spoken English (DCPSE)
The Brigham-Young University Corpus of American English (BYUC)
The English-Norwegian Parallel Corpus (ENPC)

**Dictionaries:**

Cappelens Engelsk-Norsk, Norsk-Engelsk Ordbok, 2007, Oslo: Kunnskapsforlaget


Kunnskapsforlagets Engelsk-Norsk Stor Ordbok, 2002, Oslo: Kunnskapsforlaget

Kunnskapsforlagets Norsk-Engelsk Stor ordbok, 2002, Oslo, Kunnskapsforlaget


**Other online resources**

- *Oxford English Dictionary Online*
- *Merriam-Webster online dictionary.*
- The BYUC homepage http://www.americancorpus.org (02.05.2008)