Current change in the modal system of English

A case study of *must, have to* and *have got to*

Joanne Close & Bas Aarts*
University of Leeds & University College London

This paper takes the variation between *must*, *have to* and *have got to* as a window through which to view changes in the modal system in Present-Day British English (1960s–1990s). The results from this study show a dramatic decrease in frequency of the core modal *must* and a significant increase in frequency of the semi-modal *have to* in the *Diachronic Corpus of Present-Day Spoken English* (DCPSE). Changes in the modal system affect both epistemic and root uses of *must,* although *have to* is only an active rival to root *must;* epistemic instances of *have to* (and *have got to)* are rare in the corpus. We suggest that a tendency to avoid expressions of strong commitment is the cause of the decline in root *must.* This is supported by similar falls in the frequency of *should* and *ought* found by Leech et al. (2009).

1. Introduction

The class of English modal verbs have undergone a number of important changes in the history of English, and continue to change in Present-Day English (PDE). Studies on current change suggest that the modals are decreasing in frequency, becoming monosemous, and are facing possible competition from rival semi-modals which are increasing in frequency (see Krug 2000; Leech 2003; Smith 2003; Mair & Leech 2006; Mair 2006; Leech et al. 2009).

This paper investigates the variation between the core modal *must* and the semi-modals *have to* and *have got to* in the *Diachronic Corpus of Present-Day*

Spoken English (see Section 3.1).[1] We use the label "semi-modal" for *have to* and *have got to* in order to highlight that these forms share some of the properties common to *must*, a member of the class of core modals.[2]

The questions we address in this paper are: (i) is the core modal *must* decreasing in frequency? (ii) is there evidence for "competition" between *must* and the semi-modals *have to* and *have got to*? (iii) is *must* becoming monosemous?, and (iv) what is the motivation for the change(s)?

The variation between the three forms is a window onto the changing modal system and, while the results cannot necessarily be generalized to other modals, the study will offer an insight into areas of the modal system susceptible to change and identify possible motivating factors behind these changes.

This paper is organized as follows: Section 2 reviews some of the existing work on current change in the modal system; Section 3 presents the corpus and describes the data collection method, including decisions taken on which data to exclude from the current study; Section 4 discusses the main findings and investigates possible motivations for the changes that have occurred; Section 5 is the conclusion.

2.   Current change in the modal system of English

2.1   Observing current change

Throughout the history of English, the modal verbs have undergone a number of significant changes, resulting in a class of verbs characterized by a number of syntactic properties which are not shared by main verbs (see Traugott 1989; Lightfoot 1991; Warner 1993). More recently, there has been a growing interest in *current change* in the modal system of English (Krug 2000; Leech 2003; Smith 2003; Mair 2006; Leech et al. 2009), as well as more generally. It is hoped that research into this phenomenon will lead to a greater understanding of the trigger(s) and pathway(s) of language change.

We might expect instances of current change to be subtle in comparison to changes measured over a longer period. While recent changes in the language

---

1.   We do not include in this study the reduced form *gotta* or the form *got to* not preceded by an auxiliary. We exclude these from this paper because the phonologically reduced *gotta* does not appear in the corpus and *got to* occurs only once.

2.   For Quirk et al. (1985: 141–146) *have to* is a modal semi-auxiliary, while *have got to* is a modal idiom. These combinations are not identical in their syntactic behaviour: *have to* takes *do*-support, while inversion is not possible for most speakers. It can occur as a non-finite form and be preceded by an auxiliary verb. By contrast, *have got (to)* is always finite, and therefore cannot be preceded by auxiliary verbs. It can invert with a subject.

may not appear as dramatic as some of the earlier changes such as the levelling of the English inflectional system from Old to Middle English, the shift from OV to VO or the appearance of *do*-support, they could be equally significant. There exist instances of historical language changes which involve a change in the underlying grammar of the language (see Kroch 1989; Lightfoot 1979; Warner 1993; Pintzuk 1995, among others). Current change has been linked to grammar change only very recently in this way. According to Leech (2003: 223), however, "if we understand such [grammatical] changes to include changes of frequency, significant grammatical changes do take place within a generation."

The nature of current change is such that it is more difficult to observe than a change which has occurred over a longer period of time, e.g. several centuries. In the study of a long-term change, the end-point of the change has typically been reached, and it is often possible to track the change through the period in which it occurred. A change measured over a shorter period of thirty years or so, however, will not necessarily be completed in the time period, and it is impossible to know *how*, *when*, or even *if* a particular change will complete. The danger of attempting to predict these is illustrated by the subjunctive which was believed to be decreasing to the point of extinction (Fowler 1965), but has been shown more recently to be undergoing a revival (Övergaard 1995; Hundt 1998; Leech et al. 2009). It is, of course, possible to track a change as it occurs, but it will often be necessary to observe the phenomenon at a later period (see Bauer 2002; Mair 2008).

2.2   Core modal frequency in Present-Day English

One of the observed changes in the modal system of Present Day English is a decrease in frequency of the core modals (see Krug 2000; Leech 2003; Smith 2003; Leech et al. 2009). Leech (2003) investigates modal usage in the Brown quartet of corpora. The term 'Brown quartet' refers to four comparable corpora: the Brown Corpus and Frown (Freiburg-Brown) Corpus from 1961 and 1991–2, respectively, and two matching corpora of British English, LOB (Lancaster-Oslo/Bergen) and FLOB (Freiburg-Lancaster-Oslo/Bergen), again from 1961 and 1991–2, respectively. These corpora each contain a million words in fifteen written text types. He observes that the decrease in modal usage is apparent in 13 out of 15 categories ranging from a decline of 31.2 percent (D, Religion) to a decline of 0.5 percent (H, Miscellaneous). In the remaining two categories, there is a small and non-significant increase; in J (Learned) of 2.4% and L (Mystery Fiction) of 8.4 percent. Furthermore, the British and American corpora show a decrease over time (from 1961 to 1991) in the case of every modal (individual modals differ in frequency), except *can* and *could* which show an increase of 2.2 percent and 2.4 percent, respectively, in British English. Leech's results also show that the infrequent modals

shall, *ought to*, and (in British English) *need* have decreased drastically, and a similar decline has taken place in the mid-frequency modals *may* and *must*, while *would* and *should* have decreased less drastically. Finally, *will*, *can*, *could* and *might* have not significantly changed. The results presented by Leech et al. (2009) also confirm that the modals which show the steepest decline are the less common modals.

2.3   Core modals vs. semi-modals

Mair & Leech (2006) point out that although many changes involve a competing construction or constructions, there is not necessarily a correlation between the form(s) losing ground and those gaining. On discovering a decrease in the frequency of the core modals in the Brown family of corpora, Leech (2003: 229) suggests that perhaps the semi-modals are "gradually usurping" the functions of the core modals, but concludes that this is not the case. The reason for his conclusion is that, although semi-modal usage is increasing overall, some semi-modals are actually declining (for example *be to*, *(had) better*, *(have) got to* and *be going to* in FLOB), and most semi-modals are much less frequent than the core modals (for example, *will* is about ten times more frequent than *be going to*). These results are supported by Krug (2000), and also by Smith (2003: 249) who states that "the rise of [*have to*] ... by no means makes up for the shortfall in *must*."

Mair (2006: 105) also investigates modal frequency using the *Santa Barbara Corpus of Spoken American English* and the spoken component of the British component of the *International Corpus of English* (ICE-GB). His results indicate that *have to* is the most common form in both varieties, followed by *have got to* and *must* in British English and *need to* and *must* in American English. On the possibility of *have to* taking over the functions of *must,* he comments that "epistemic *must* is secure in spoken British English, as *have to* in this function has a marginal status at best" (Mair 2006: 105).

It seems, then, that if there is competition between *must* and *have to/have got to* this is in the expression of obligation or necessity, i.e. root meaning. In their investigation of modality in British dialects, Tagliamonte & Smith (2006) find exactly this; their findings indicate that "*[m]ust* is obsolescent and there is an unanticipated resurgence of *have to* alongside pan-dialectal grammatical reorganization" (Tagliamonte & Smith 2006: 341).

2.4   Monosemy of modals

Leech (2003) carries out a semantic analysis of *should* and every third instance of *may* and *must* found in the Brown corpora to discover whether a decline in

frequency can be attributed to one particular sense of a modal. He provides some evidence that the decline in frequency of the modals is linked to the tendency for modals to become "more monosemous", but points out that this is not always the case. In fact, a different pattern can be observed for each of the three modals investigated. *Must,* the modal we are interested in here, shows a decline in both root (event modality; obligation) and epistemic (logical necessity) senses between the 1960s and the 1990s, particularly in the SEU-mini and ICE-GB-mini corpora of spoken English,[3] but the root sense has shown the largest decrease in frequency.[4] Leech warns that the results of the study as a whole are tentative because of uncertainties in sampling and coding procedures. It is also important to bear in mind that comparing the results of semantic analyses across studies is difficult, due to the challenging nature of the task of semantic coding.

3.   Corpus and data collection

3.1   Current change and the *Diachronic Corpus of Present-Day Spoken English*

There are a number of "pitfalls of anecdotal observation" as far as current change is concerned, most notably "[t]he spread of salient new uses is exaggerated, while the less salient persistence of older forms is not noted or [...] a diachronic trend is read into a situation which merely shows variable or fluctuating usage" (Mair 2008: 1111). For this reason, studies on current change are carried out more reliably using corpora.

One of the strengths of the methodology we have adopted is that it uses a corpus of spoken English. This is particularly important for the study of the semi-modals as "these forms, being typically colloquial, are not likely to show up in their true colours in the written language" (Leech 2003: 230).

The *Diachronic Corpus of Present-Day Spoken English* (DCPSE) contains 421,362 words from the British component of the *International Corpus of English* (ICE-GB) collected in the early 1990s and 464,074 words from the *London-Lund*

---

3.   The SEU-mini corpus used by Leech (2003) contains 80,000 words (16 texts) from conversation, BBC discussions and news, sports and other commentaries, broadcast talks, etc., all dated between 1959–65. The ICE-GB-mini corpus contains 80,000 words from text categories chosen to match those of the SEU-mini corpus, the only difference being the date: texts are dated 1990–92.

4.   Leech's (2003) results show that the epistemic reading of *may* has increased in the written corpora, but all other meanings have shown a decline. With *should* the root sense of "weak obligation" is shown to be gaining at the expense of other senses, particularly the epistemic sense of "weak inference."

Corpus (LLC) collected between the late 1960s and early 1980s (for more information about the LLC and ICE-GB see Svartvik (1990) and Nelson et al. (2002), respectively). The period of time between LLC and ICE-GB is within the range of what is considered ideal for studying current change. In designing the Brown family of corpora a decision was made to have an interval of approximately thirty years between the Brown and LOB corpora (1960s) and the Frown and FLOB corpora (1990s) because "the interval of one generation or so […] is usually considered the minimum sufficient to clearly identify and document linguistic change in real time" (Mair 1998: 140).

With over 800,000 words of spoken English, DCPSE is the largest available single collection of parsed and checked orthographically transcribed spoken material. All of the sentences in DCPSE have been grammatically analysed and have been given detailed parse trees which contain *functional information* (subject, direct object, noun phrase head, etc.), *categorial information* (part-of-speech, phrase, clause, etc.), and optional additional *features* (intransitive (verb), common (noun), main (clause), etc.).

The corpus can be explored by using the *International Corpus of English Corpus Utility Program* (ICECUP) software developed at the Survey of English Usage. ICECUP offers a facility called *Fuzzy Tree Fragments* (FTFs; see Aarts et al. 1998 and Nelson et al. 2002).[5] An example FTF which searches for all verb phrases (VPs) in the corpus is shown in Figure 1.1. This FTF is a partial tree diagram, which can be constructed by users. FTFs are "fuzzy" in the sense that users can specify for themselves the level of detail that they are interested in. For example, if the user adds "intr" to the lower section of the FTF in Figure 1.1, the software will search for all intransitive verb phrases in the corpus. This is shown in Figure 1.2.

More complex FTFs can contain grammatical structures, words, lexical wild cards, etc. In the context of the present research, the FTF in Figure 2 searches the corpus for the string *got to* with the position for *have* left unspecified (indicated by the symbol '¤'). This FTF will find cases of *'ve/'d got to* as well as *have/has/had*
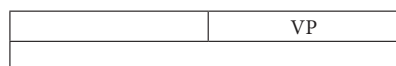
	VP

Figure 1.1. An FTF which instructs the search software to find all the verb phrases in the corpus
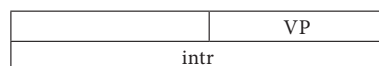
	VP
	intr

Figure 1.2. An FTF which instructs the search software to find all the intransitive verb phrases in the corpus

---

5. See also www.ucl.ac.uk/english-usage/resources/ftfs for further information on conducting experiments using FTFs.
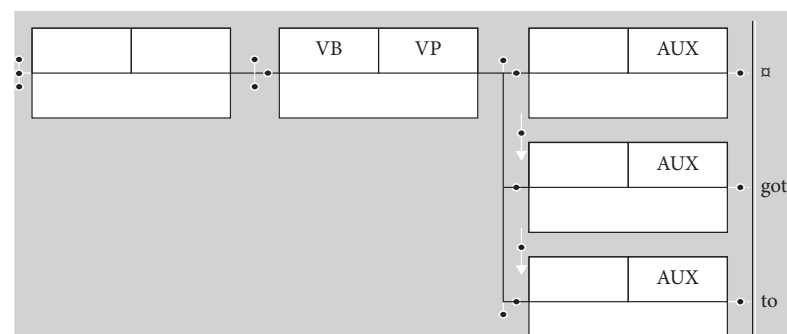
Figure 2. FTF for *(have) got to*

got to, and even cases where an adverb intervenes between *have* and *got*.[6] This is because in Figure 2 the arrows between the nodes are set so that something may intervene between the nodes. This is an example of user choice. Users can change this setting if they wish to specify that the nodes should follow each other immediately.

3.2   Data

In order to ensure reliability of results, it is important to identify contexts in which the forms being investigated may alternate. In the case of *have to*, *have got to*, and *must* there are a number of syntactic environments in which there is only one option, and these have been excluded from the study (see Tagliamonte & Smith 2006; Depraetere & Verhulst 2008). All exclusions are listed below, and are not included in the statistical tests in Section 4.

First, negated forms were excluded because the meanings are not equivalent:[7]

(1) a. You mustn't go (the modal is outside the scope of the negation: 'It is necessary for you not to go').
     b. You don't have to go (the modal is inside the scope of the negation: 'It is not necessary for you to go').

---

6. This FTF will only retrieve instances of *got to* which are preceded by an auxiliary verb (always *have*); cases of *got to (gotta)* will not be captured with this FTF. Cases in which *got* is a main verb followed by the preposition *to* would also not be retrieved using this FTF.

7. Depraetere and Verhulst (2008: 16) also exclude negative uses of *must* and *have to* based on the observation that "the scope of negation is different for *have to* (absence of necessity, 'not necessarily') and *must* (prohibition, 'necessarily not')."

c. You haven't got to go (the modal is inside the scope of the negation: 'It is not necessary for you to go').

Secondly, interrogative contexts were excluded because interrogative forms with *must* did not occur at all in the later period, and only one example occurred in the early period:

(2) **Must** Wordsworth speak for us in his intimations of immortality?

(DCPSE:DL-J01 #0129: 2:B)

Finally, also excluded were infinitival *have to* (including future forms) because equivalent forms of *must* or *have got to* do not exist; and the past form *had to* because *must* cannot refer to past time, so would not be an option here.

It is important to point out that tokens of *must have been* or *must have V-ed* have not been excluded. Although there are no tokens of *have (got) to have been* or *have (got) to have V-ed* in the corpus, a search of the web using WebCorp illustrates that both of these forms are possible.[8] An example of *have to have been* is given in (3):

(3) Just my tuppence worth, but 2007 **has to have been** the absolute worst X Factor *ever*.[9]

3.3   Semantic coding

A number of claims have been made in the literature about the frequency of the modals and semi-modals relating to their meaning(s). For instance, Coates (1983) states that *have to* and *have got to* are infrequent as epistemic modals and Leech (2003) claims that *must* is one of a few modals not becoming monosemous. In order to test these claims using DCPSE, each instance of *must*, *have got to* and *have to* was manually classified according to the type of modality that they express.

Modal meaning has been dealt with in a number of ways. The present study uses the coding system advocated by Leech & Coates (1980) and Coates (1983), in which modals are divided into two semantic classes, root and epistemic. In this approach modal meaning is dealt with in a gradience model where each class has both core and peripheral members. This can account for the fuzziness in modal meaning without the need for postulating additional classes.

---

8.   There are two plausible reasons why these forms are not present in DCPSE: (i) they are very infrequent, or (ii) they are a post 1992 innovation. It is impossible to determine which, if either, of these is correct.

9.   Example retrieved using WebCorp on 22/09/08. Web address: http://www.webcorp.org.uk/cgi-bin/view.nm?url=http://www.unrealitytv.co.uk/x-factor/the-worst-x-factor-ever/&term=has%20to%20have%20been&case=case.

According to Palmer (2001: 9) "with deontic modality the conditioning factors are external to the relevant individual, whereas with dynamic modality they are internal". As outlined below, we do not classify the source of the obligation for the instances of *must*, *have to* and *have got to*. For this reason, we prefer the term *root* over *deontic*, as this does not imply a contrast with *dynamic*. Our approach to semantic coding is described in more detail in Sections 3.3.1–3.3.4.

3.3.1   *Root modality*

Coates argues that root modality is fuzzy with a cline from strong obligation ('It is imperative/obligatory that x') to weak obligation ('It is important that x'), although the basic meaning of root *must* is 'It is necessary for ...'. Examples of root meaning are given in (4):

(4) a. My boss had said you **must** read those books.

(DCPSE:DI-B12 #0171:1:E)

b. I**'ve got to** go, I'm afraid, in an hour.   (DCPSE:DI-B37 #0216:2:B)

c. Actually you **have to** be much more honest about what you're doing.

(DCPSE:DI-A02 #0124:2:B)

3.3.2   *Epistemic modality*

Epistemic *must* "conveys the speaker's confidence in the truth of what he is saying, based on a logical process of deduction from facts known to him (which may or may not be specified)" (Coates 1983: 41). Coates divides epistemic *must* into core (confident inference) and peripheral (logical necessity) meanings:

i.   *confident inference*: 'I confidently infer that x'
ii. *logical necessity*: 'In light of what is known, it is necessarily the case that x'

Epistemic examples of *must*, *have got to* and *have to* are given in (5) below:

(5) a. I am coming to conclude that it **must** be because I get too much water in it or something.   (DCPSE:DI-B47 #0166:1:B)

b. Loose shirts over jeans **has got to** be a sort of temporary prejudice.

(DCPSE:DI-B44 #0161:1:B)

c. I mean good guidebooks really are full of impressions, aren't they, whereas imaginative works or works of fiction often do blend facts in a distorted form. If you think of Hardy's use of landscape as a curious blend of the identifiable, which people who research in that find interesting, and something which **has to** be an impression there, wouldn't you agree?   (DCPSE:DL-A06 #0310:1:E)

3.3.3 *Performative modality*

Performative or speech act modals occur where the speaker is carrying out the action denoted by the verb. These occur frequently with verbs like *admit*, *say*, *confess*, as shown in (6).

(6) a. There's a piece here called Spring Fire which takes my interest, I **must say**.   (DCPSE:DI-D12 #0141:1:A)
 b. When you switched to the emphasis being on Architecture <,,> did you initially think that you wanted to go into that as a career or were you doing it just as a degree because you enjoyed the subject
                    (DCPSE:DI-A07 #0093:1:A)
   Pretty much yeah I'**ve got to admit**.   (DCPSE:DI-A07 #0094:1:B)
 c. I **have to confess** an often irking thought of am I really really two pounds less than Kate Hamilton.   (DCPSE:DI-B07 #0224:2:B)

Although performative modality is considered to be a type of root modality, these examples were coded and counted separately because of their specialized meaning.[10]

3.3.4 *Ambiguity*

Coates (1983: 47) asserts that "there is no overlap between the two fuzzy sets representing Root and Epistemic *must*" and "[c]ases where it is not possible to decide which meaning is intended are therefore ambiguous", as illustrated in (7) taken from Coates (1983: 47, her example (35)):[11]

(7) And anyway, I think mental health is a very relative thing. I mean, mental health **must** be related to the sort of general mentality, or whatever other word you use of the community you're living in.   (S.2.14.73)

The root meaning here is 'it's essential that mental health is related to ...', while the epistemic reading is 'it's inevitably the case that mental health is related to ...'.

It might be expected that ambiguous cases would be relatively frequent. However, from the Lancaster and Survey corpora, Coates classes only 11 examples out of a total of 436 as ambiguous.[12] In the ambiguous cases "syntactic features which could disambiguate such [ambiguous] utterances are either not present or are themselves ambiguous" (Coates 1983: 47).

---

10. Tagliamonte & Smith (2006) exclude these cases from their data based on the fact that they occurred exclusively with *must*, but in DCPSE *have got to* and *have to* can have performative meaning.

11. Coates's original mark-up has been removed from the example as it is not relevant here.

12. This could be viewed as indirect support for the coding system chosen by Coates; we might expect that the better the coding system is, the fewer indeterminate examples there will be.

In coding the data from DCPSE, it was assumed that the speaker had an intended meaning when producing the utterance, and an attempt was made to classify each example, keeping ambiguous cases to a minimum. Despite this, there exist a number of examples which are impossible to disambiguate for two main reasons: (i) the utterance is compatible with both root and epistemic meanings and/or (ii) the modal is followed by ellipsis or unclear words. The numbers for ambiguous cases are shown in Table 2.

4. Results and discussion

4.1 Overall frequencies

In order to examine a possible trend for declining core modal usage and to investigate the theory of competition between core modals and semi-modals, the overall frequency of core modal *must* is compared with the frequencies of the semi-modal forms in DCPSE. As shown in Table 1, during the thirty-year period the frequency of *must* has declined by over 55% while the frequency of *have to* has significantly increased by nearly 32%.

Table 1. Overall frequencies of *must, have got to* and *have to* in DCPSE (figures in bold are significant at $p < 0.01$)

(Semi-) modal	LLC frequency		ICE-GB frequency		Change in frequency	
	raw	per 100,000 words	raw	per 100,000 words	%	χ^2 score
must	427	92.01	172	40.82	-55.64	**36.29**
have got to	187	40.30	156	37.02	-8.14	3.10
have to	188	40.51	225	53.40	+31.82	**31.94**
Total	802	172.82	553	131.24	-24.06	**71.32**

Table 2. Distribution of semantic types of *must* in DCPSE

Source corpus	Epistemic		Root		Performative		Ambiguous		Total
	N	%	N	%	N	%	N	%	N
LLC	43.96	47.78	36.63	39.81	9.48	10.30	1.94	2.11	92.01
ICE-GB	21.12	51.74	16.14	39.53	2.37	5.81	1.19	2.91	40.82
Total	65.08	48.99	52.77	39.73	11.85	8.92	3.13	2.35	132.83

Somewhat surprisingly, *have got to* decreases in frequency by 8%. It is impossible for us to be sure whether *have got to* has reached its peak and is showing a decline in the 1990s data prior to "levelling out" in the future, or whether the decrease is simply fluctuating usage in the corpus. To be certain, it will be necessary to measure the frequency of *have got to* at a point later than 1992. Unfortunately, the lack of available corpora makes this impossible at this point in time.

4.2   Root and epistemic modality

One of the aims of this study is to determine whether *must* is becoming monosemous. Using the semantic coding carried out on the data (see Section 3.3 above for details), the frequencies of epistemic, root and performative uses were calculated. The results are presented in Table 2.

   To discover if the decline in root and/or epistemic *must* is related to the use of the semi-modals, root and epistemic uses of the three forms were compared. The results for root and epistemic uses are presented in Tables 3 and 4, respectively.

   Table 3 illustrates that as expressions of root obligation, *must*, *have got to* and *have to* combined have decreased by 12.98%. Again, the decrease in *must* and the rise in *have to* are found to be significant, and *have got to* shows a decline.

Table 3. Frequencies of root *must*, *have got to* and *have to* in DCPSE (figures in bold are significant at p < 0.01)

(Semi-) modal	LLC frequency		ICE-GB frequency		Change in frequency	
	raw	per 100,000 words	raw	per 100,000 words	%	χ^2 score
must	174	37.49	70	16.61	–55.69	**23.61**
have got to	184	39.65	151	35.84	–9.61	0.12
have to	185	39.86	208	49.36	+23.83	**12.32**
Total	543	117	429	101.81	–12.98	**36.05**

Results for epistemic use in Table 4 show that, although epistemic *must* has decreased by over 50% *must* is the still the preferred form for expressing epistemic necessity; epistemic *have got to* and *have to* are very rare in DCPSE.[13]

   The general view in the literature is that the decline of the core modals cannot be attributed directly to an increase in the semi-modals (Krug 2000; Leech 2003;

Table 4. Frequencies of epistemic *must*, *have got to* and *have to* in DCPSE

(Semi-) modal	LLC frequency		ICE-GB frequency		Change in frequency	
	raw	per 100,000 words	raw	per 100,000 words	%	χ^2 score
must	206	44.39	91	21.60	–51.34	0.19
have got to	2	0.43	4	0.95	+120.93	3.36
have to	2	0.43	3	0.71	+65.12	1.83
Total	210	45.25	98	23.26	–48.60	5.38

Smith 2003; Mair & Leech 2006). This conclusion was reached by studying frequencies in written language where, on the whole, core modals are more frequent than semi-modals. The situation with respect to the spoken data is quite different. In root contexts (see Table 3), *have to* is already more frequent than *must* in the 1960s (frequencies are 39.86 per 100,000 words for *have to* and 37.49 for *must*). By the 1990s, the frequency of *have to* is almost three times that of *must* (the frequencies are 16.61 and 49.36, respectively). This suggests that there may be a link between core modal decline and semi-modal increase.

4.3   Motivating factors

The frequency shift in the expression of root modality suggests a possible correlation between the decline of *must* and the rise of *have to*. The question, of course, is what is the motivation for the shift away from *must* and towards *have to* in root meaning? A number of suggestions to explain the decline in *must* have been made in the literature, many of which refer to the "authoritarian" nature of *must*. Myhill (1995: 202), for instance, attributes the decline in *must* (= 'obligation') and the rise of *should* (= 'weak obligation') around the time of the American Civil War to a growing tendency to avoid overt claims to authority by the speaker/writer. He calls this "democratisation". He claims that the "old" modals were linked with people controlling the actions of other people while the "new" modals are more personal, and are used to give advice to an equal, and the like. Smith (2003) shares a similar view, and makes the following remark:

> It seems probable that MUST is a casualty of a changing society where increasing emphasis is being placed on equality of power, or at least the appearance of equality of power, and the informality of discourse found in private conversation is becoming more acceptable, even usual, in official types of discourse. Just as these conditions are likely to disfavour the use of MUST, they should correspondingly favour other forms which express obligation less directly (Smith 2003: 259).

---

13.   The frequencies for epistemic *have to* and *have got to* are low which may lead to questions about the reliability of the chi-square test here, but log likelihood calculations confirm low significance scores: *must* 0.19, *have got to* 3.05, *have to* 1.67 (total 4.92).

The problem with the suppositions by Myhill and Smith, however, is that epistemic *must* has also decreased, and this verb is not related to power and authority. Leech et al. (2009: 88) suggest that the "partial decline [of epistemic *must*] could be due to contamination by the dramatic fall of deontic *must*". This is perhaps a questionable explanation because the root and epistemic meanings of *must* do not intersect, so it is not clear how a fall in one would cause a fall in the other (cf. Coates's (1983: 170) treatment of the meanings of *will* which do intersect). In their data, root *must* declines more drastically than epistemic *must*, which Leech et al. suggest is because "neither *have to* nor any other form has become widely adopted as an alternative expression of strong epistemic necessity". In the spoken data from DCPSE, however, the fall of epistemic *must* is almost equal to that of root *must,* although it is true that there is no rise in epistemic *have to* or *have got to.* It is, of course, possible that there are independent explanations for the fall of root *must* on the one hand, and epistemic *must* on the other. However, we suggest that the cause is the same.

The data from DCPSE cannot tell us which occurred first, the fall of root *must* or the rise of root *have to.* The figures for epistemic *must* suggest the former, otherwise it is difficult to explain why epistemic *must* has declined, but epistemic *have (got) to* has not increased.

As an explanation for the decrease in *must,* we appeal to what Huddleston & Pullum et al. (2002: 175) refer to as "strength of commitment" where "necessity involves a strong commitment". We suggest that the decline in *must* is a result of a decline in forms expressing strong commitment. In the root sense, *have to* is favoured over *must.* The difference between the two forms is illustrated in (8) where the conjoined clause is not possible with *must*, because the commitment is stronger and does not permit resistance (example from Sweetser (1988: 54), quoted in Hopper & Traugott (1993: 79)):

(8)   I **have to/??must** get this paper in, but I guess I'll go to the movies instead.

This explanation can be extended to epistemic *must* because this also expresses necessity (and hence involves strong commitment), although, as pointed out in Section 4.2, *have to* is not a rival to epistemic *must.*

If there is a move away from forms which express a stronger commitment, then we would also expect other forms which express necessity to show a decrease. This is in fact the case: *should* and *ought to,* both classified by Coates (1983) as modals of necessity and obligation, show a decrease in frequency in DCPSE and in written corpora (see Leech (2003) and Leech et al. (2009) for results from the Brown family of corpora).

5.   Conclusion

This study has investigated variation between the core modal *must* and the semi-modals *have to* and *have got to* using the *Diachronic Corpus of Present-Day Spoken English*. By considering only syntactic environments in which the forms are variable, any possible competition between *must* and the "rival" semi-modal alternates can be reliably observed. The results from the study show that even in the 1960s data, the frequency of root *have to* has surpassed that of root *must.* The continued decline of *must* and the rise in *have to* in the 1990s data suggests that *have to* is being used in contexts where previously we would have found *must.* Surprisingly, *have got to*, the most recent marker of root modality, has decreased. This does not appear to be a reflection of the DCPSE sampling, as Tagliamonte (2006) and Tagliamonte & Smith (2006) found that this was also the case in a number of English dialects.

The semantic coding carried out on *must, have to* and *have got to* in DCPSE has led to a number of interesting results. Firstly, although both epistemic and root senses of *must* have declined, as proportions of the total number of instances of *must* they have remained fairly constant. This supports Leech's (2003) claim that *must* is not becoming monosemous. Further study of the semantic senses of other modals in DCPSE is necessary in order to determine if this fact is peculiar to *must.* Secondly, *have to* is the most frequent marker of root modality, which is remarkable because we might expect the newer form *have got to* to be more frequent. And finally, the use of epistemic *must* has decreased, but the use of epistemic *have to* has not increased. This highlights the importance of using corpora to investigate language change, as simple "anecdotal observation" might have led to a prediction that this form is on the increase.

References

Aarts, Bas & Sean A. Wallis. 2006. *The Diachronic Corpus of Present-Day Spoken English* (DCPSE). [Software]

Aarts, Bas, Gerald Nelson & Sean A. Wallis. 1998. "Using Fuzzy Tree Fragments to Explore English Grammar". *English Today* 14:3.52–56.

Bauer, Laurie. 2002. "Hitting a Moving Target: An Interim Report on Language Change across Generations". *English Today* 18:4.55–59.

Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use.* Cambridge: Cambridge University Press.

Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries.* London: Croom Helm.

Depraetere, Ilse & An Verhulst. 2008. "Source of Modality: A Reassessment". *English Language and Linguistics* 12:1.1–25.

Fowler, Henry Watson. 1965. *A Dictionary of Modern English Usage*. Second edition. Revised by Ernest Gowers. Oxford: Clarendon Press.

Hopper, Paul J. & Elizabeth Closs Traugott. 1993. *Grammaticalization*. (Cambridge textbooks in linguistics). Cambridge: Cambridge University Press.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Hundt, Marianne. 1998. "It is Important that this Study (should) be Based on the Analysis of Parallel Corpora: On the Use of the Mandative Subjunctive in four Major Varieties of English." *The Major Varieties of English* (Papers from MAVEN 97) ed. by Hans Lindquist, Staffan Klintborg, Magnus Levin & Maria Estling, 159–175. Växjö: Acta Wexionensia.

Kroch, Anthony S. 1989. "Reflexes of Grammar in Patterns of Language Change". *Language Variation and Change* 1.199–244.

Krug, Manfred G. 1998. "*Gotta* – The tenth Central Modal in English? Social, Stylistic and Regional Variation in the British National Corpus as Evidence of Ongoing Grammaticalization". *The Major Varieties of English* (Papers from MAVEN 97) ed. by Hans Lindquist, Staffan Klintborg, Magnus Levin & Maria Estling, 177–191. Växjö: Acta Wexionensia.

Krug, Manfred G. 2000. *Emerging English Modals: A Corpus-based Study of Grammaticalization.* Berlin & New York: Mouton de Gruyter.

Leech, Geoffrey. 2003. "Modality on the Move: The English Modal Auxiliaries 1961 1992." *Modality in Contemporary English* ed. by Roberta Facchinetti, Manfred Krug & Frank Palmer, 223–240. Berlin & New York: Mouton de Gruyter.

Leech, Geoffrey & Jennifer Coates. 1980. "Semantic Indeterminacy and the Modals." *Studies in English Linguistics* ed. by Sidney Greenbaum, Geoffrey Leech & Jan Svartvik, 79–90. London: Longman.

Leech, Geoffrey, Marianne Hundt, Christian Mair, & Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study.* Cambridge: Cambridge University Press.

Lightfoot, David W. 1979. *Principles of Diachronic Syntax.* Cambridge: Cambridge University Press.

Mair, Christian. 1997. "Parallel Corpora: A Real-time Approach to the Study of Language Change in Progress." *Corpus-based Studies in English. Papers from the seventeenth International Conference on English Language Research on Computerized Corpora (ICAME 17)*, Stockholm, May 15–19 1996, ed. by Magnus Ljung, 195–209. Amsterdam: Rodopi.

Mair, Christian. 1998. "Corpora and the Study of the Major Varieties of English: Issues and Results." *The Major Varieties of English* (Papers from MAVEN 97) ed. by Hans Lindquist, Staffan Klintborg, Magnus Levin & Maria Estling, 139–157. Växjö: Acta Wexionensia.

Mair, Christian. 2006. *Twentieth-Century English: History, Variation and Standardization*. (= *Studies in English Language*) Cambridge: Cambridge University Press.

Mair, Christian. 2008. "Corpora and the Study of Recent Change". *Corpus Linguistics: An International Handbook* ed. by Anke Lüdeling & Merja Kytö, 1109–1125. Berlin & New York: Mouton de Gruyter.

Mair, Christian & Geoffrey Leech. 2006. "Current Changes in English Syntax". *The Handbook of English Linguistics* ed. by Bas Aarts & April McMahon, 318–342. Oxford: Blackwell.

Myhill, John. 1995. "Change and continuity in the functions of theAmerican English modals." *Linguistics* 33.157–211.

Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam & Philadelphia: John Benjamins.

Övergaard, Gerd. 1995. *The Mandative Subjunctive in American and British English in the 20th Century* (= *Studia Anglistica Upsaliensia*, 94). Uppsala: Acta Universitatis Upsaliensis. Stockholm: Almqvist and Wiksell International.

Palmer, Frank R. 2001. *Mood and Modality.* Second edition. Cambridge: Cambridge University Press.

Pintzuk, Susan. 1995. "Variation and Change in Old English Clause Structure." *Language Variation and Change* 7.229–260.

Smith, Nicholas. 2003. "Changes in Modals and Semi-modals of Strong Obligation and Epistemic Necessity in Recent British English". *Modality in Contemporary English* ed. by Roberta Facchinetti, Manfred Krug & Frank Palmer, 241–266. Berlin & New York: Mouton de Gruyter.

Svartvik, Jan, ed. 1990. *The London-Lund Corpus of Spoken English: Description and Research* (= *Lund Studies in English*, 82). Lund: Lund University Press.

Sweetser, Eve E. 1988. "Grammaticalization and Semantic Bleaching". *Berkeley Linguistics Society: General Session and Parasession on Grammaticalization* ed. by Shelley Axmaker, Annie Jaisser & Helen Singmaster.

Tagliamonte, Sali. 2006. "Historical Change in Synchronic Perspective: The Legacy of British Dialects". *The Handbook of the History of English* ed. by Ans van Kemenade & Bettelou Los, 477–506. Oxford: Blackwell.

Tagliamonte, Sali & Jennifer Smith. 2006. "Layering, Competition and a Twist of Fate: Deontic Modality in Dialects of English". *Diachronica* 23:2.341–380.

Traugott, Elizabeth Closs. 1989. "On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change". *Language* 66:1.31–55.

Wallis, Sean & Gerald Nelson. 2000. "Exploiting Fuzzy Tree Fragments in the Investigation of Parsed Corpora". *Literary and Linguistic Computing* 15:3.339–361.

Warner, Anthony R. 1993. *English Auxiliaries: Structure and History.* (= *Cambridge Studies in Linguistics*, 66). Cambridge: Cambridge University Press.

2nd proofs