Fehringer C, Corrigan KP.
"You've got to sort of eh hoy the Geordie out": Modals of obligation and necessity in 50 years of Tyneside English.

*English Language and Linguistics* 2015, 19 (Special Issue 2), 355-381.

Copyright:

This is the authors' accepted manuscript of an article published in its final definitive form by Cambridge University Press, 2015.

DOI link to published article:

[http://dx.doi.org/10.1017/S1360674315000131](http://dx.doi.org/10.1017/S1360674315000131)

Date deposited:

16/03/2015

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](http://creativecommons.org/licenses/by-nc/3.0/)

Newcastle University ePrints - eprint.ncl.ac.uk
“You’ve got to sort of eh hoy the Geordie out”: **Modals of obligation and necessity in 50 years of Tyneside English**¹

¹ We are grateful to the Centre for Research in Linguistics and Language Sciences at Newcastle University for an award to assist with the completion of this project from their SDF Fund for Research Collaboration and Infrastructure. We are also grateful to the two anonymous reviewers of the first draft of this paper for their very helpful feedback.
This paper examines the use of the semi-modals *have to, have got to* and *need to* in the Diachronic Electronic Corpus of Tyneside English (DECTE), a corpus of spoken north eastern English dating from the late 1960s to the present day. It will be shown that the semi-modals have, in many contexts, replaced the historically older *must* as markers of obligation and necessity in this variety. Moreover, the two most frequent variants in the corpus, *have to* and *have got to*, will be examined in the light of current theories of grammaticalisation. Internal and external constraints, which have been shown in the literature on root modality to have played an important role in the distribution of variants in other regional varieties of British and North American English, will be tested in DECTE. The paper will also examine the rise of *need to* in this North Eastern variety, as the most recent addition to the group of variants marking obligation and necessity.

**Key words:**

Diachronic Electronic Corpus of Tyneside English; Grammaticalisation; Modals of Obligation and Necessity; North Eastern English; Regional Variation; Diachronic Change
The subject of modal auxiliaries in English has generated much interest over the past decades with many detailed studies of the system from Old English to the present day showing how they have changed in their distribution and function over time (see e.g. Krug 2000, Leech 2003, Close and Aarts 2010). One particular area of focus has been on the expression of obligation and/or necessity, where the original modal *must* has been shown to be losing ground to newer periphrastic modals, such as *have to*, *have got to* and, more recently, *need to*, which have come to express similar functions (Biber et al. 1998: 205-6). Consider the following examples:²

(1)  

a. The mothers and fathers *must* take their children or they think they *must* take their children to school (*pause*) in a car (*pause*) (0708a/NECTE2/m)  

b. I mean you *have to* have O Levels (05b/PVC/f)  

c. They’re not going to offer you a job in the hand like that (*pause*) you’ve *got to* go out and look for it (06b/PVC/m)  

d. You *need to go* to college for that (1005a/NECTE2/f)  

The periphrastic forms have been described as ‘semi-modals’ (e.g. Biber et al. 1998, Mair and Leech 2006), ‘quasi-modals’ (Fischer 1994, Collins 2009), ‘emerging modals’ (Krug 2000) or ‘peripheral modals’ (Denison 1998) and have been shown to be

---
² These examples from DECTE (see section 2) are marked with a speaker identification number followed by the time period of the recording (Tyneside Linguistic Survey (TLS) = late 1960s to early 1970s, Phonological Variation and Change (PVC) = 1991-1994, Newcastle Electronic Corpus of Tyneside English 2 (NECTE2) = 2007-2010). Finally, *m/f* identifies the sex of the speaker. The data is freely available to download upon completion of an access request form at: [http://research.ncl.ac.uk/decte/corpus.htm](http://research.ncl.ac.uk/decte/corpus.htm).
outnumbering the original ‘central’ or ‘core’ modal must in large-scale corpus studies of modern English (see e.g. Biber et al. 1998, Krug 2000, Leech et al. 2009).

These semi-modals are interesting in that they are often considered to be classic examples of the process of grammaticalisation, in which sequences of elements, some lexical and some grammatical, combine together to express certain procedural functions in a novel way (see Hopper and Traugott 2003, Bybee et al. 1994). Grammaticalisation theory has played a very important role in informing analyses of syntactic and semantic change, particularly with respect to differences in the modal system of English across historical time periods or between speakers representing diverse social or regional spaces synchronically (Leech 2003: 236). If we take have to as an example, we can see how the lexical verb have, meaning ‘to possess’, has become more general in its meaning in the grammatical construction have to. Consider the following steps in the evolution of have to from a lexical to a grammatical morpheme (based on Krug’s 2000: 55 adaption of Heine’s 1993 model):

(2) Stage  I  I have a letter  [Possession Schema]
            II  I have a letter to mail  [Purpose Schema: Possession Schema + purpose/goal adjunct]
            III I have a letter to write  [the possessive meaning of have has undergone semantic bleaching]
            IV  I have to write a letter  [have to now functions as a unit lexeme expressing the modal notion of obligation]
            V   I have to write  [object complement can now be deleted]
Particularly crucial is the development from Stage III to Stage IV. At Stage III, the possessive meaning of *have* has undergone semantic bleaching (see Hopper and Traugott 2003: 32), as the letter is not yet written and therefore cannot be in the possession of the speaker. By Stage IV, the object has moved from the main clause *[I have a letter] [to write]* to the subordinate clause, resulting in *have* and *to* functioning as a cohesive unit *[I have+to write a letter]*. This development of contiguous *have to* is commonly agreed to be the most important step in the grammaticalisation of this construction (Fischer 1994). Furthermore, the fact that *have to* has become generalised to contexts in which no direct object is present and therefore no possessive reading is possible (as in Stage V, *I have to write*) shows that this construction has become a grammatical marker of obligation and/or necessity independent of its original lexical meaning of possession (Stage I).

A further indicator that a construction is becoming grammaticalised is its cohesiveness as a unit, evidenced by the absence of intervening elements: e.g. *I have really to go*, *I’ve got actually to pretend* (Krug 2000: 67) and the tendency to undergo phonological reduction: e.g. *I hafta go*, *I’ve gotta go* (see Bybee et al. 1994: 6 for discussion). *Need to* differs from the other semi-modals in that it has not (yet) undergone similar phonological reductive processes. This might have to do with the fact that *need to* is a relatively recent addition to the group of semi-modals (see section 3.4 below).

Phonological reduction often goes hand-in-hand with frequency of use, and an increase in frequency has also been shown to be an important part of the grammaticalisation process (Bybee et al. 1994: 8). Grammaticalised forms become more
frequent due to their increased semantic generality which influences their ability to be used in a wider range of contexts. Corpus studies of the discourse frequency of the semi-modals have consistently shown that have to and have got to have been rising in frequency over the centuries (see e.g. Biber et al. 1998, Krug 2000, Leech 2003, Smith 2003, Collins 2009). Furthermore, in more recent decades, need to is following suit (Smith 2003, Nokkonen 2006, Collins 2009).

Indeed, a number of researchers believe that this change is continuing, and that the semi-modals of obligation and necessity are still undergoing grammaticalisation (e.g. Krug 2000, Tagliamonte 2004). Recently, a number of important quantitative, variationist studies on the modals of obligation/necessity have emerged which not only measure rates of discourse frequency but also examine the internal and external factors which contribute to the variation of forms (see Corrigan 2000 for Northern Irish English; Tagliamonte 2004, 2013 and Tagliamonte and Smith 2006 for English in England, Northern Ireland and Scotland; Jankowski 2004 for British versus American English and Tagliamonte and D’Arcy 2007 for Canadian English). Such studies allow us to gain more insight into the process of ongoing grammaticalisation in that they not only reveal changes in frequency over time but also diversity in encoding: i.e. how variant forms can become specialised so that they take on particular functional roles such as be to which expressed root modality in the historical variety of South Armagh English reported on by Corrigan (2000).

The present paper will contribute to ongoing research in this ‘complex and unsettled’ area (Tagliamonte 2006: 344) by investigating the development of the modals of obligation/necessity in Tyneside English, using DECTE, which captures the speech patterns of communities in the North East of England between the 1960s and 2010. The
paper will follow the same methodology as the quantitative variationist studies of Jankowski (2004), Tagliamonte (2004, 2013) and Tagliamonte and Smith (2006) in order to ascertain whether these modals are continuing to undergo grammaticalisation in this region to the same degree as has been reported for other dialects in the British Isles and North America. Thus, both the frequency of the competing variants and the internal and external constraints on their distribution will be examined. In common with the research just noted, this analysis is also based on spoken dialect data. However, it differs from it in one key respect, i.e. the approaches of Tagliamonte (2004, 2013) and Tagliamonte and Smith (2006) are synchronic and comparative across regional space, whereas the research described here concentrates on diachronic change within a single dialectal variety. It also differs from Jankowski’s (2004) study in that her data is taken from written rather than spoken sources. Furthermore, this paper will also include an analysis of the most recent semi-modal need to, which has had little consideration in the variationist literature heretofore.

In sum, the following questions will be addressed: i) Are semi-modals taking over from older modal forms as markers of obligation and necessity in our North East data? ii) Do changes in frequency and in the patterning of internal constraints on the system of modality over the three time periods investigated provide evidence of the ongoing grammaticalisation of the semi-modals in the North East? iii) Do sociolinguistic variables such as sex and education, which have been found to be significant in determining the direction of change for the semi-modals in some varieties of English, also play a role in the distribution of the semi-modals in the North East?

2. The Diachronic Electronic Corpus of Tyneside English (DECTE)
DECTE is a collection of text transcriptions and audio files of interviews with a wide variety of people from the North East of England, dating from the 1960s up to the present day. It is a diachronic corpus, not only as regards the span of time across which the interviews have been and continue to be collected, but also as a reflection of the even greater depth covered in terms of the lifetimes of the people who have been interviewed (i.e. the oldest speaker from the first sub-corpus was born in 1895).

In total, DECTE currently contains 99 interviews, recording 160 speakers in 804,266 words of text and 71 hours 45 minutes and 43 seconds of audio. The interviews come from three different research projects carried out at Newcastle University. The first and second of these are the *Tyneside Linguistic Survey* (TLS) of the 1960s-1970s and the *Phonological Variation and Change in Contemporary Spoken English* (PVC) project of the 1990s. The third constituent part, NECTE2, extends the corpus into 2010 with further sets of interviews that have been collected annually since 2007. (Full information on these sub-corpora can be found in Allen et al. 2007, Beal et al. 2014 and Corrigan et al. 2012 as well as Corrigan et al. 2014 and Mearns (in press)). Table 1 summarises the dimensions of DECTE that we draw on in this study.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews</td>
<td>37</td>
<td>18</td>
<td>44</td>
</tr>
<tr>
<td>Informants</td>
<td>37</td>
<td>35</td>
<td>88</td>
</tr>
</tbody>
</table>
In order to be able to account fully for the distribution and function of the modals of obligation/necessity in present-day English (PDE), it is necessary to understand their history and development. This section will give a brief synopsis of the origin and development of each form.

Table 1: The Time-frame for the DECTE Data in our analysis

<table>
<thead>
<tr>
<th>Female</th>
<th>Male</th>
<th>Age: 16-20</th>
<th>21-30</th>
<th>31-40</th>
<th>41-50</th>
<th>51-60</th>
<th>61-70</th>
<th>71-80</th>
<th>81-90</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>17</td>
<td>2</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>17</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>49</td>
<td>39</td>
<td>34</td>
<td>24</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>

3 THE MODALS OF OBLIGATION AND NECESSITY

It is clear from this table that the speakers listed here have a rather uneven profile in terms of the numbers of males and females as well as their age ranges. This is not unusual with corpora that consist of legacy data as DECTE does. The imbalance is not especially relevant for the qualitative aspects of the study reported on here and they have been taken into account when dealing with the quantitative analyses so that the sub-corpora can be compared with one another accountably.
3.1 Must

Must is the oldest of the modal auxiliaries denoting obligation and has been present in the language since the Old English (OE) period. Its OE form was mot which, at that time, was mainly used to express permission and possibility, e.g. Me mæig … gif hit mot gewiderian, mederan settan, linsed sawan ‘One can … if it may be fair weather, plant madder, sew linseed’ (Warner 1993: 164). Uses of mot expressing obligation can also be attested in OE, although it was in ME that this function began to dominate, as may gradually took over from mot as a marker of permission and possibility (Denison 1993: 303). Examples such as (3a) from Warner (1993: 165) and (3b) from Fischer (1992: 263) demonstrate that mot occurred most commonly as a marker of obligation in ME. Here, the newer forms most (also attested in ME as must) and moste are used, which were originally preterite forms:

(3) a. Þu most hebbe redi mitte / Twenti Marc ine þi slitte
   You must have ready at-hand twenty marks in your sleeve
   ‘You must have ready at hand twenty marks in your sleeve’

   b. And seyde he moste unto Itayle, ...
   And said he had [to go] to Italy

Studies of modal verbs have traditionally classified their meanings as either ‘root’ on the one hand or ‘epistemic’ on the other. For instance, Coates (1983: 21) states that the

---

4 Some scholars use the term ‘deontic’ as a synonym for ‘root’ (e.g. Tagliamonte and Smith 2006) while others (e.g. Palmer 2001) divide root modality (referred to by Palmer as ‘event modality’) into ‘deontic’ and ‘dynamic’, the former referring to contexts where the obligation is imposed on the speaker from an external source and the latter denoting internally imposed obligation. For the sake of simplicity and
'root meaning’ of a modal refers to the logic of obligation and permission, whereas ‘epistemic meaning’ is to do with knowledge and belief about possibilities and probabilities. Although must did start out with root meanings in OE and these continued into later stages of the language (see examples (3) above), epistemic uses are well attested by the fourteenth century, e.g. Yif preisyng make gentilesse, thanne mote they nedes ben gentil that been preyed. ‘If praising creates nobility, then they who are praised must necessarily be noble’ (Warner 1993: 175).

From the Early Modern English (EModE) period onwards, must is firmly established as a marker of both root and epistemic modality. Consider the following examples from the Oxford English Dictionary, where (4a) expresses root and (4b) expresses epistemic meaning:

(4)  a. The Thracians when they must pass ouer frozen streames, sende out their Wolues (S. Gosson To Sir R. Pipe in Schoole of Abuse f. 39, 1579).

b. He like an Asse because he hath a faire wife, thinks that per Consequens he must be a Cuckold (Tarltons Newes out of Purgatorie 23, 1590).

Indeed, in Tyneside English, must is considered to function primarily as an epistemic marker (Trousdale 2003:277 and Beal 2012: 67). In her 150,000 word spoken corpus of 1970s Tyneside English, McDonald (1981: 253) found that epistemic readings with must occurred 89.02% of the time compared to 52.91% in her non-Tyneside (i.e. mixed dialect) corpus. In DECTE, both root and epistemic occurrences of must are
present (see (5a) for the former and (5b) for the latter). However, epistemic readings dominate (see Table 3).

(5)  a. So you **must** always be sure that you’re always building on your experience (G23/TLSTM)

       b. If they knew that eh you know your best friend knows that you've won a quarter of a million pound … it **must** make a difference you know somehow (G02/TLSTM)

The decline in frequency of root **must** has been examined widely by scholars working on Standard English (British, American and Canadian: see e.g. Krug 2000, Leech et al. 2009, Collins 2009, Close and Aarts 2010, Smith 2003 and Dollinger 2008). Most striking is the difference between spoken and written English, as root **must** is becoming considerably less frequent in speech than it is in writing (Close and Aarts 2010: 177). This, according to Leech (2003) is resulting in a move towards monosemy for this variant: i.e. **must** is becoming specialised for a single meaning, namely epistemic.

A frequency count of DECTE data across real time reveals that occurrences of root **must** are already fairly rare in the TLS corpus of the 1960s-70s and they become even rarer in the subsequent two corpora. Table 2 shows the frequency of root **must** in relation to that of other root modals such as **have to, have got to** and **need to**.  

---

5 It is important to bear in mind that while comparison between these three sub-corpora can be considered to be indicative of real time change in the region, the data differs from that which is generated via panel studies, for example. The interview protocols between the TLS sub-corpus and the PVC and NECTE2 sub-corpora are not identical and this may have some impact on the findings that cannot be ruled out. 

6 Full frequency results for all root modals can be found in section 5 below.
If we consider epistemic meanings of *must*, however, we can see that they have actually increased over the three sub-corpora of DECTE:

<table>
<thead>
<tr>
<th></th>
<th>TLS (60s/70s)</th>
<th>PVC (1991-94)</th>
<th>NECTE2 (2007-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>root <em>must</em></td>
<td>11</td>
<td>9.3</td>
<td>5</td>
</tr>
<tr>
<td>other root modals</td>
<td>107</td>
<td>95.2</td>
<td>195</td>
</tr>
<tr>
<td>TOTAL</td>
<td>118</td>
<td>200</td>
<td>286</td>
</tr>
</tbody>
</table>

**Table 2: Frequency of root *must* in DECTE**

Our data indicate that, at least in DECTE, *must* is gradually becoming monosemous, which ties in with Leech’s (2003) observations on spoken standard English. Below are some examples of *must* in the NECTE2 sub-corpus, all of which have the epistemic meaning in question:

(6) a  We **must** be the only major city in the country that hasn't got a direct link with London (1003a/NECTE2/m)
b Well, what, we **must** be like one of only four or five English teams to have actually won a European trophy, like a major one, not the Inter Toto

(1007b/NECTE2/m)

c He's been there for like, **must** be about twenty five years (1007a/NECTE2/m)

In sum, root *must* does still exist in DECTE but it is very rare, which ties in with findings from other corpora of North Eastern English (e.g. McDonald 1981). Our analysis of this data thus contribute to the growing body of evidence that the modal system in North Eastern English as a whole is being reorganised, with *must* becoming increasingly monosemous.

In his investigation of modal verbs in American English, Myhill (1995) claims that the main reason for the decline of root *must* is essentially social. He points out that the older modal expressions generally involve a clear social order and absolute evaluations based on ostensibly universal principles (1995: 160). For example, *You must always tell the truth* reflects a general societal norm. By contrast, the use of the newer semi-modal verbs presupposes more or less equal power relationships (or at least the semblance thereof) and focus on interactive factors such as mutual co-operation, emotional appeals, advice and so on. Smith (2003: 259) agrees with Myhill that *must* is a ‘casualty of a changing society’ and argues that this is also the case in British English, where society has been moving towards greater democratisation. Jankowski (2004: 97) provides evidence from written corpora that British English and American English are undergoing the same change, although roughly fifty years apart, with American English leading the way. By contrast, Close and Aarts (2010) prefer to explain the demise of *must* as resulting from a general decline in forms expressing strong commitment. It is
not clear to us, however, whether this is actually the case or whether speakers are simply preferring to express commitment in other ways.

The idea expressed by Myhill (1995) and Smith (2003) that must is associated with formal language is supported in DECTE, in which many utterances with must are examples of instructions or rules and regulations. This is particularly the case in the PVC sub-corpus from the early 1990s, where the four examples of root must below reflect formal usage of different sorts, either in polite language (7a), external imperatives (7b), technical subjects (7c) or direct quotes (7e):

(7)  

a. Well I must eh you must excuse me (02b/PVC/m)  
b. And then we've got this what is it the Citizen's Charter or whatever it is … which says that standards must be achieved (11b/PVC/m)  
c. When you buy an upgrade you must make sure that it’s going to be compatible with the work that you’ve already done (11b/PVC/m)  
d. My father thought well he was a French polisher so … he thought that getting a trade was the main thing…. ‘You must get a trade’ he said (14b/PVC/m)

3.2 Have to

The question as to whether Old English habban + to-infinitive could have the sense ‘be obliged to V’ has been much discussed in the literature. One very early example is the following (from Denison 1993: 316), which illustrates the ambiguity between have … to in its original sense of possession (see (2) above) and have to as a unit expressing obligation:
In Middle English, however, a greater number of unambiguous examples appeared which showed *have to* functioning as root modal (Denison 1993: 317). At that time, *have to* began to compete with *must* for this function (see example (9) below from Brinton (1991), quoted in Tagliamonte and Smith 2006: 348):

(9) *I moot go thider as I haue to go*

‘I must go thither as I have to go’ (Chaucer, *Canterbury Tales*, 1386-1400)

An investigation of the frequency of *have to* in standard English shows that this form began to increase dramatically at the beginning of the modern English period. In fact, Krug (2000), on the basis of frequency data from written English corpora, argues that *have to* underwent a crucial period of grammaticalisation in the late nineteenth century. Specifically, he points out that the frequency of *have to* rose almost six fold in the second half of the nineteenth century compared to the previous fifty years (Krug 2000: 80).
Indeed, studies of modality in modern English have often focused on the role that the rise of semi-modals such as *have to* has played in the demise of core modals such as *must* (e.g. Biber et al. 1998, Leech 2003, Leech et al. 2009, Smith 2003, Close and Aarts 2010, Collins 2009). As we might expect of a corpus of spoken language, DECTE shows an extremely high rate of occurrence of *have to* as opposed to *must*:

<table>
<thead>
<tr>
<th></th>
<th>TLS (60s/70s)</th>
<th>PVC (1991-94)</th>
<th>NECTE2 (2007-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><em>have to</em></td>
<td>37</td>
<td>31.4</td>
<td>34</td>
</tr>
<tr>
<td><em>must</em></td>
<td>11</td>
<td>9.3</td>
<td>5</td>
</tr>
<tr>
<td>other root modals</td>
<td>70</td>
<td>59.3</td>
<td>161</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>118</td>
<td></td>
<td>200</td>
</tr>
</tbody>
</table>

Table 4: Frequency of *have to* versus *must* in DECTE

The following examples are taken from all three sub-corpora.9

(10) a. We have to study and that (G11/TLS/f)
    b. I have to go to the doctor’s I’ve got a bad back (08b/PVC/f)
    c. You could see her bum. You could see her boobs were falling out of it. You have to choose legs or boobs. You don’t do both (1018a/NECTE2/f)

3.3 *Have got to*

---

9 Note that all examples have root meaning and there are no epistemic examples of *have to* in DECTE.
As a semi-modal of obligation/necessity, *have got to* is first attested much later than *must* and *have to*, namely, not until the nineteenth century (Biber et al. 1998, Krug 2000). Krug (2000: 62-4) gives a number of examples from the work of Dickens. Consider the following extract from *Oliver Twist* (1837/38):

(11) ‘I’ve got to be in London to-night; and I know a ‘spectable old gentleman as lives there, wot’ll give you lodgings for nothink’ (VIII, 102).

As this example illustrates, early uses of *have got to* are usually accompanied by features of non-standard speech, which suggest that the speaker is of low status. Thus, it seems likely that *have got to* may have entered the language via the colloquial spoken register. Later, however, *have got to* became more general and writers such as George Eliot, Lewis Carroll and Oscar Wilde regularly used the collocation in dialogue representing all classes of speaker. The construction was still largely restricted to dialogue, however, and even today it is rare in the written language.

Krug (2000: 89) states that, following *have to*, which rose dramatically in frequency in the second half of the nineteenth century, *have got to* underwent a similar development in the early twentieth century. Thus, he argues that both variants were subject to the same grammaticalisation process at different stages of history. However, there is a crucial difference between these forms: *have got to* differs from *have to* in that the former entered English directly as a marker of root modality and not first as a possessive marker (Krug 200: 73). Thus, *have got to* did not undergo the stages of

---

10 See Krug (2000: 62-63) and Visser (1969: 1479) for more detailed discussion of the use of *have got to* by nineteenth and early twentieth century authors.
reanalysis that *have to* went through (see (2) above) but was used synonymously with *have to* straight from the outset. Krug (2000: 64-5) argues that the use of *have got to* was motivated by discourse factors, for example, the need for greater expressivity as a result of the contraction of *have*.

In DECTE, *have got to* thrives, which perhaps is not surprising for a corpus based on colloquial spoken English:

(12) a. I never try to change for anybody. The only way I try to change is if I gan anywhere and they cannot understand you … so you’ve got to sort of eh hoy the Geordie out (G27/TLS/m)
b. They’ve got to make their bed … they’ve got to wash the dishes (18b/PVC/f)
c. She’s got to be in uni nine till five every single day (1011b/NECTE2/f)

Recent studies of standard British English have shown that *have got to*, even in colloquial spoken English where it had increased dramatically since the turn of the twentieth century, has actually started to decline (Tagliamonte and Smith (2006), Close and Aarts (2010)). Krug (2000: 63) had labelled *have got to* a ‘success story’, and that certainly appears to have been the case up to the middle of the twentieth century, but evidence from written and spoken corpora show that, towards the end of the twentieth century, *have got to* loses a considerable amount of ground to *have to*. Close and Aarts (2010: 176) state that it is impossible to be sure whether *have got to* simply reached its peak in the 1990s prior to levelling out in the future or whether the decrease in frequency merely represents fluctuating use in the corpus investigated in their research, and that it is therefore necessary to examine the frequency of this form at a later period.
Interestingly, frequency data from DECTE does reveal a similar pattern when comparing *have got to* with *have to*, although the time periods of the corpus are not exactly commensurate with those used in the study by Close and Aarts (2010). *Have got to* underwent an increase in frequency in the PVC 1990s sub-corpus of DECTE but then declined in the NECTE2 sub-corpus from 2007-10. As Close and Arts are not dealing with North Eastern English, we are not comparing like with like in this respect either, however the parallels are noteworthy:

<table>
<thead>
<tr>
<th></th>
<th>TLS (60s/70s)</th>
<th>PVC (1991-94)</th>
<th>NECTE2 (2007-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><em>have got to</em></td>
<td>69</td>
<td>58.5</td>
<td>149</td>
</tr>
<tr>
<td><em>have to</em></td>
<td>37</td>
<td>31.4</td>
<td>34</td>
</tr>
<tr>
<td>other root modals</td>
<td>12</td>
<td>10.1</td>
<td>17</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>118</td>
<td>200</td>
<td>286</td>
</tr>
</tbody>
</table>

Table 5. Frequency of *have got to* versus *have to* in DECTE

3.4 *Need to*

Of all the semi-modals of obligation/necessity, *need to* is the least researched variant (Nokkonen (2010: 46); Collins (2009: 57)). This might be due to the fact that it is a relatively new addition to the group, in that only in the last couple of decades has its frequency increased sufficiently in order to be able to place it on a par with *have to* and *have got to*. Leech (2003: 230) notes the ‘remarkable rise’ in its frequency in the late
twentieth century and argues that it is now gaining semi-modal status.\footnote{Need to is not to be confused with the modal need, which is mainly used in non-assertive contexts: e.g. You need not come.} Moreover, like the other semi-modals, need to has been shown to be more frequent in the spoken language than in writing (Collins (2009: 76), Nokkonen (2006: 51)).

Our DECTE data also reveal that need to has increased in frequency over the three time periods in relation to the other root modals:

<table>
<thead>
<tr>
<th></th>
<th>TLS (60s/70s)</th>
<th>PVC (1991-94)</th>
<th>NECTE2 (2007-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>need to</td>
<td>1</td>
<td>0.8</td>
<td>12</td>
</tr>
<tr>
<td>other root modals</td>
<td>117</td>
<td>99.2</td>
<td>188</td>
</tr>
<tr>
<td>TOTAL</td>
<td>118</td>
<td>99.2</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 6. Frequency of need to in DECTE

Note, however, that the relatively recent rise in frequency of need to does not mean that it is a new addition to the language. In fact, it appeared before the other semi-modals, being attested before 1400. Consider the following example from the OED:

(13) A good phisician needeth to loke wel a-boute and be ful ware.

‘A good physician needs to look well about and be fully aware’

(J. Trevisa tr. Bartholomaeus Anglicus De Proprietatibus Rerum (BL Add.) f. 102v, 1398)
This shows that a grammatical form can exist in the language for centuries before it becomes used productively (Biber et al. 1998: 206).

The primary function of need to is to express inherent necessity: i.e. obligation that is motivated internally by the speaker (i.e. subjective obligation) rather than being imposed externally (i.e. objective obligation) by, for example, rules and regulations (see Bybee et al. 1994: 177 for discussion). For instance, if a speaker says I need to go to confession, he/she is feeling a compulsion which originates within himself/herself (see Nokkonen 2010: 47). By contrast, if he/she says I have to start work at eight o’clock, the obligation is imposed externally and is therefore objective. Having said that, need to can also be used to express objective obligation (e.g. I need to start work at eight o’clock), and it has thus now come to compete with the productive semi-modal have to and have got to. Consider the following examples from Nokkonen (2010: 49, 56, 61) in which (a-b) illustrate subjective obligation and (c-d) objective:

(14) a. I need to be a bit skinnier and a bit taller

b. She’s worried as well about whether she’s going to get another job, cos er, I mean they need to have erm a second income, I should think

c. Well, you’ll need to write it down tomorrow

d. Cos you need to have a bath and bed

Smith (2003: 260-1) claims that need to is often used to suggest that the action in question is being recommended (rather than imposed) for the doer’s own sake. This allows the speaker to downplay his/her authority and lay down obligation in a more
indirect way, particularly if *need to* is combined with a first person plural subject or a passive construction: *we need to ...; it needs to be* (Leech et al. 2009: 110-1).

These uses of *need to* can also be found in DECTE. Hence, alongside the basic subjective obligation expressed in (15a), we find instances where the speaker downplays his/her authority by suggesting that the obligation is for the addressee's own sake (15b) and examples where passives are used to shift the focus away from the agents and onto the action that needs to be done (15c):

(15)  

   a. I need to feel like I've got enough money for a taxi (1020a/NECTE2/f)  
   b. I just sat down and said 'look, you need to be a little bit more mindful of what you talk about and how you say it' (1017a/NECTE2/m)  
   c. I don't know whether this is the sort of thing that needs to be talked about (14b/PVC/m)

In her sociolinguistic account of *need to*, using the demographic spoken component of the BNC from the early 1990s, Nokkonen (2010) argues that the use of *need to* is determined by social roles and functions: the participants in her study with high rates of *need to* are all professionals and share the obligation to direct politely (2010: 70). She argues that middle class people have a high degree of social awareness and this manifests itself in their attitude towards the addressee, which ties in with Smith's (2003) observations on downplaying authority and being cognisant of the addressee's feelings.

3.5 **Form and function**
In an attempt to account for variation in the domain of root modality, some scholars have attributed the distribution of competing forms to specific semantic and pragmatic functions. For instance, Myhill (1995: 163-7) argues that have to and have got to differ in that the former is typically associated with habitual obligations (e.g. He has to take the bus to work every day) while the latter expresses a more emotional appeal (e.g. I’ve got to talk to you). Must, on the other hand, is associated with social norms (e.g. ‘Don’t answer when they ring’ – ‘But I must’). Unfortunately, however, Myhill does not provide any statistical evidence to support his claims.

Moreover, an investigation of the meanings and pragmatic functions of the modals in DECTE does not support the argument that each form has a particular function. On the contrary, they appear to be in free variation with each other. Consider the following examples, all from the same speaker (09/TLS/f):

(16) a. I think you’ve got to weigh things up
    b. I think you have to more or less try
    c. And I think you must just weigh things up

Moreover, need to can also be added to the mix, since it freely competes with have to and have got to as (17) demonstrates:

(17) a. What I’m doing at the minute, doing child care, you need to go to college or university to get your Level 2 (05a/NECTE2/f)
b. The army wouldn’t take them now because I mean you have to have O Levels (05b/PVC/f)

c. They’re not going to offer you a job in the hand like that … you’ve got to go out and look for it (06b/PVC/m)

Thus, it appears that what we have in DECTE is a case of ‘layering’: a well-known principle of grammaticalisation in which new grammatical morphemes enter the language and co-exist alongside older morphemes expressing the same function (see Hopper 1991: 22). Must, the original core modal verb, was joined by the semi-modal have to which, in turn, was joined by have got to and these three variants all competed to mark root modality. This is apparent in the 1960s/70s sub-corpus. By the 1990s, need to had joined the group as the newest layer; however the oldest layer, must, had by that time largely been replaced by the semi-modals in the root function, so that by 2010 the three competing layers were have to, have got to and need to. Quantitative data charting the rise and fall of each form will be given in section 5 below.

4 DATA COLLECTION AND ANALYSIS

Initially, all occurrences of must, have to, have got to and need to were collected, and then, in order to ensure that the context in which these cases were occurring allowed for variation, exclusions were made (see 4.2.1. below). Thus, the analysis only focuses on contexts in which there is free competition between the variant forms.

4.1 Data collected and exclusions
As the present study is focusing on the modals of obligation/necessity in their root sense, all occurrences of forms with epistemic meanings were excluded. This led to the exclusion of many occurrences of must. Moreover, since have got to and must only occur in the present tense, only present forms of have to and need to were investigated.

Unlike Standard English, Tyneside English does not allow variation between don’t have to and haven’t got to when expressing the meaning ‘not necessary to’. Instead, haven’t got to is used as a synonym of mustn’t: e.g. I’ll send my dad for you. You know you haven’t got to play in there (04b/PVC/fl).12 Therefore, we also excluded negative utterances from our analysis. Similarly, interrogatives were not included as they did not provide enough of a variable context: have got to is used only once in interrogatives in the whole data-set, which is also the case with need to.

Finally, a few occurrences of must occurred in set phrases with say or admit. These tended to recur in different parts of the same speaker’s conversation, which indicated that they were prefabricated chunks. Therefore, following Tagliamonte and Smith (2006) inter alia, these were also excluded from the analysis.

4.2 Data analysis

After the initial filtering out of non-variable contexts, we are left with affirmative declarative utterances in the present tense as the basis for our analysis, which gives us, in total, 447 instances of root modal use.

12 See Beal (2004:127) for a discussion of haven’t got to in Tyneside and McDonald (1981:234-235) and Buchstaller and Corrigan (to appear) for more examples.
Each occurrence was categorised for the internal constraints discussed in 4.2 above: grammatical person of subject, definite vs. indefinite subject and subjective vs. objective obligation. In addition, each participant was categorised as male vs. female and as having secondary versus post-secondary education, following Tagliamonte and D’Arcy (2007). They were also classified according to the time period of their recording (TLS, PVC, NECTE2).

4.3 Statistical analysis

Although the variable rule program GoldVarb has a number of disadvantages, as articulated by Johnson (2009) inter alia, it has been the bedrock of the quantitative paradigm for some time now. As the comparison of the system of modality in DECTE with that articulated in previous research on other varieties like that of Tagliamonte and Smith (2006) is a key aim of this paper, our analyses will employ GoldVarb X, the same tool used to investigate these exact variables so that direct comparisons can be made between our findings and those of other researchers (Sankoff 1988 and Sankoff et al. 2005).

5 Frequency results

An investigation of the frequency of all the competing forms shows that, over the three time periods under investigation, must has declined in frequency and need to has risen sharply. Have to and have got to are the most frequently occurring variants at every

13 Secondary education includes Advanced Level and vocational qualifications, whereas post-secondary is reserved for participants who are studying at university or who already have a university degree.
14 We would like to express our thanks to Claire Childs for research assistance with the statistical analyses.
stage, with *have got to* dominating in the earlier two sub-corpora and then levelling off in the most recent sub-corpus, so that both variants now appear to occur with equal frequency (see Table 7 and Figure 1).

<table>
<thead>
<tr>
<th></th>
<th>TLS (60s/70s)</th>
<th>PVC (1991-94)</th>
<th>NECTE2 (2007-10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td><em>have to</em></td>
<td>37</td>
<td>31.4</td>
<td>34</td>
</tr>
<tr>
<td><em>have got to</em></td>
<td>69</td>
<td>58.5</td>
<td>149</td>
</tr>
<tr>
<td><em>need to</em></td>
<td>1</td>
<td>0.8</td>
<td>12</td>
</tr>
<tr>
<td><em>must</em></td>
<td>11</td>
<td>9.3</td>
<td>5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>118</td>
<td></td>
<td>200</td>
</tr>
</tbody>
</table>

Table 7. Frequency of root modal expressions in DECTE

![Figure 1. Bar chart showing frequencies of root modal expressions in DECTE](#)
6 Constraints conditioning variation

Tagliamonte and Smith (2006), Jankowski (2004) and Tagliamonte (2012, 2013), point out the difficulties that can arise when attempting a semantic and pragmatic investigation into the differences between variant forms. For instance, whether the difference is ‘emotional appeal versus habitual obligation’ (Myhill 1995) or ‘strong versus weak obligation’ (Coates 1983: 33), it is very difficult for investigators to categorise such functions impartially, as, inevitably, the analyst imposes his/her own subjective interpretations onto the data (Tagliamonte 2012:311). What we need are concrete grammatical correlates that will provide us with an objective measure of the functions in question. This section provides an outline of such correlates, which have been used by quantitative variationist studies such as Tagliamonte and Smith (2006) and Tagliamonte (2013), investigating a range of British dialects,¹⁵ Tagliamonte and D’Arcy (2007), focusing on Canadian English, and Jankowski (2004), comparing standard British and American data. The same correlates will form the basis of our analysis of DECTE in order to make comparisons as valid possible.

6.1 Internal constraints

6.1.1 Type of subject

Coates (1983: 37) observes that root modals in clauses with second person subjects usually incorporate a stronger sense of obligation than those with first person, while

¹⁵The dialectal varieties found in DECTE were not amongst those investigated in previous research.
clauses with first persons tend, in turn, to have a stronger obligation interpretation than those with third person subjects. Examples of these different subject types incorporating a range of obligation interpretations from stronger to weaker depending on personhood can also be attested in DECTE:

(18) a. You’ve got to ask my dad’s - my mam’s permission (04b/PVC/f)
    b. I’ve got to gan away on Monday morning (04b/PVC/f)
    c. So I think schools have got to compensate for the weaknesses and ehm problems of parents (23/TLS/m)

One must take into account, however, that some subjects can be indefinite: particularly you, meaning ‘one’.

(19) a. It was just a harsh blow for everyone, you’ve just got to get used to it (02b/NECTE2/m)
    b. You’ve got to have your hair tied back (13b/PVC/f)
    c. I mean, you’ve just got to see Gateshead (04/TLS/m)

As only definite subjects have the possibility of encoding a strong reading, indefinite subjects must be kept separate in the analysis in order to test adequately whether, for example, second person does indeed embody the strongest possible sense of the hearer of the root modal utterance feeling obliged (Tagliamonte and Smith 2006: 359).

16 They can also be indefinite, e.g. They have to keep up with the Joneses now (Tagliamonte and Smith 2006:359), but this was rare in the DECTE data (a mere 8 tokens).
Our DECTE results clearly demonstrate that, in all three sub-corpora, grammatical person has no significant effect on the choice of *have to* versus *have got to* (see Tables 8-10 and, for the factor weights, Appendix, Table 16). Similarly, the definiteness of the subject plays no significant role (see Table 11). This contrasts with Tagliamonte and Smith’s (2006) findings, which show that, for definite subjects, *have to* is the most frequently used form in first person and third person contexts, whereas *have got to* strongly correlates with indefinite subjects (primarily you meaning ‘one’).

---

17 We included the frequencies for *must* and *need to*, although these numbers were too small to factor into the multivariate analysis.

18 This is not the case in Jankowski’s corpora, which might be due to the fact that *have got to* is less common in standardised written English than it is in spoken dialect data.
### Table 8. Grammatical person in TLS

<table>
<thead>
<tr>
<th></th>
<th>have to</th>
<th>have got to</th>
<th>need to</th>
<th>must</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st person</td>
<td>8</td>
<td>40.0</td>
<td>10</td>
<td>50.0</td>
<td>0</td>
</tr>
<tr>
<td>2nd person</td>
<td>1</td>
<td>16.7</td>
<td>4</td>
<td>66.7</td>
<td>0</td>
</tr>
<tr>
<td>3rd person</td>
<td>6</td>
<td>27.3</td>
<td>13</td>
<td>59.1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>31.2</td>
<td>27</td>
<td>56.2</td>
<td>0</td>
</tr>
</tbody>
</table>

Have to vs have got to: Chi-square = 1.28, d.f. = 2, p = 0.528

### Table 9. Grammatical person in PVC

<table>
<thead>
<tr>
<th></th>
<th>have to</th>
<th>have got to</th>
<th>need to</th>
<th>must</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st person</td>
<td>13</td>
<td>18.3</td>
<td>51</td>
<td>71.8</td>
<td>7</td>
</tr>
<tr>
<td>2nd person</td>
<td>1</td>
<td>7.7</td>
<td>10</td>
<td>76.9</td>
<td>1</td>
</tr>
<tr>
<td>3rd person</td>
<td>5</td>
<td>12.5</td>
<td>33</td>
<td>82.5</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>15.3</td>
<td>94</td>
<td>75.8</td>
<td>10</td>
</tr>
</tbody>
</table>

Have vs have got: Chi-square = 1.39, d.f. = 2, p = 0.499
<table>
<thead>
<tr>
<th></th>
<th>have to</th>
<th>have got to</th>
<th>need to</th>
<th>must</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; person</td>
<td>N 36</td>
<td>33</td>
<td>23</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>% 39.13</td>
<td>35.9</td>
<td>25.0</td>
<td>0.0</td>
<td>65.2</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt; person</td>
<td>N 6</td>
<td>13</td>
<td>5</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>% 24.0</td>
<td>52.0</td>
<td>20.0</td>
<td>4.0</td>
<td>17.7</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt; person</td>
<td>N 10</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>% 41.7</td>
<td>41.7</td>
<td>16.7</td>
<td>0.0</td>
<td>17.0</td>
</tr>
<tr>
<td>Total</td>
<td>N 52</td>
<td>56</td>
<td>32</td>
<td>1</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>% 36.9</td>
<td>39.7</td>
<td>22.7</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>

*Have to vs have got to: Chi-square = 2.56, d.f. = 2, p = 0.277*

Table 10. Grammatical person in NECTE2

<table>
<thead>
<tr>
<th></th>
<th>TLS</th>
<th>PVC</th>
<th>NECTE2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>specific</td>
<td>non-specific</td>
<td>specific</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td><strong>have to</strong></td>
<td>15</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>35.7</td>
<td>34.4</td>
<td>16.8</td>
</tr>
<tr>
<td><strong>have got to</strong></td>
<td>27</td>
<td>42</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>64.3</td>
<td>65.6</td>
<td>83.2</td>
</tr>
</tbody>
</table>

TLS Chi-square = 0.02, d.f. = 1, p = 0.888

PVC Chi-square = 0.61, d.f. = 1, p = 0.435

NECTE2 Chi-square = 0.88, d.f. = 1, p = 0.348
6.1.2 Subjective versus objective obligation

The pragmatic difference between subjective and objective obligation can also be tested quantitatively and can give further insight into the nuances expressed in the root modal system. Coates (1993: 32) sees subjective obligation as a prototypical function of root modality: i.e. the speaker is imposing authority on himself/herself or on other people:

(20) a. I love Waterloo Road. I have to watch it every Wednesday (05b/NECTE2/f)
    b. I think we need to go back to Tenerife in August (18b/NECTE2/f)
    c. He’s like ‘you’ve got to stick your hand through the window’ (16b/NECTE2/f)

By contrast, objective obligation comes from a source external to the speaker. For example, from the imposition of rules and regulations:

(21) a. I work in a pharmacy, just basically people come in with problems and I have to give them creams for eh things like that (04b/NECTE2/m)
    b. Yeah, ‘cause you have to do medicine if you want to be a pathologist
       (23b/NECTE2/f)
    c. She’s got to be in uni nine till five every single day (11b/NECTE2/f)

Tagliamonte and Smith (2006) demonstrate that, in subjective contexts, extensive layering of forms occurs. While must has traditionally had the main function of
subjective root modality, it now has ‘stiff competition’ (2006: 362) from have to and have got to, which is particularly interesting, as this has often been considered to be the core area of root meaning. Indeed, in her later analysis of a wider range of dialects, Tagliamonte (2013: 143) found that have got to was actually winning out in subjective contexts. By contrast, objective contexts favour have to (Tagliamonte and Smith 2006, Tagliamonte 2013), unless they have indefinite second person subjects which, as stated in 2.3.1 above, strongly favour have got to.

Our results from DECTE, however, reveal that subjective versus objective obligation has no significant effect on the choice of have to versus have got to (see Table 12). However, in the most recent NECTE2 sub-corpus, where need to has gained ground, we see that this variant is significantly favoured in utterances expressing subjective obligation (see Table 13).

<table>
<thead>
<tr>
<th></th>
<th>TLS</th>
<th>PVC</th>
<th>NECTE2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>subj.</td>
<td>obj.</td>
<td>subj.</td>
</tr>
<tr>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>have to</td>
<td>9</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>33.3</td>
<td>35.4</td>
<td>16.9</td>
</tr>
<tr>
<td>have got to</td>
<td>18</td>
<td>51</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>66.7</td>
<td>64.6</td>
<td>83.1</td>
</tr>
</tbody>
</table>

TLS Chi-square = 0.039, d.f. = 1, p = 0.843

PVC Chi-square = 0.29, d.f. = 1, p = 0.588

NECTE2 Chi-square = 1.99, d.f. = 1, p = 0.158
Table 12. Subjective vs objective obligation in DECTE

<table>
<thead>
<tr>
<th></th>
<th>Subjective</th>
<th></th>
<th>Objective</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>have to</td>
<td>61</td>
<td>43.0</td>
<td>61</td>
<td>42.4</td>
</tr>
<tr>
<td>have got to</td>
<td>47</td>
<td>33.1</td>
<td>68</td>
<td>47.2</td>
</tr>
<tr>
<td>need to</td>
<td>32</td>
<td>22.5</td>
<td>15</td>
<td>10.4</td>
</tr>
<tr>
<td>must</td>
<td>2</td>
<td>1.4</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>142</strong></td>
<td></td>
<td><strong>144</strong></td>
<td></td>
</tr>
</tbody>
</table>

Need to Chi-square = 11.97, d.f. = 3, p = 0.007

Table 13. Subjective vs objective obligation in NECTE2: need to

6.2 Sociolinguistic constraints

In addition to the internal linguistic constraints outlined in 3.2 above, it might also be the case that sociolinguistic factors play a role in determining the use of the modal variants. Age and sex have been repeatedly shown to be determining factors (Tagliamonte and Smith 2006, Tagliamonte 2013) though less work has been done on the interaction between different types of modal expression and social class (see Tagliamonte and D’Arcy 2007, Nokkonen 2010).19

---
19 Synchronic analyses of variation and change incorporate age as a dependent variable so as to track changes in apparent time. Hence, Nokkonen (2010:52) shows that younger speakers used higher frequencies of need to than her older participants did, demonstrating the increase in the use of the construction over time. Although Cukor-Avila and Bailey (2013) establish the pitfalls of using diachronic data like DECTE to uncover similar change in real time, our analysis thus far has already demonstrated that, these issues notwithstanding, there may well be changes in progress with respect to the expression of root modality in North Eastern English.
6.2.1 Sex

It is well known in sociolinguistic research that women tend to favour forms that are closer to the standard language and have more prestige whereas men often favour those which are non-standard (Labov 2001). Indeed, Tagliamonte and Smith (2006: 369) show that *have got to*, which is commonly considered more colloquial than the more neutral *have to*, is often avoided by women. This effect has also been demonstrated for Canadian English, with *have to* being favoured by females and disfavoured by their male peers (Tagliamonte and D’Arcy (2007: 78).

Our DECTE results reveal a slightly different picture. In the earlier sub-corpora, *have to* is significantly favoured by females. However, in the most recent sub-corpus, the use of *have to* amongst males has increased considerably and sex no longer appears to be marked by the choice of one modal variant over another.
### Table 14. The effect of sex on choice of modal variant in DECTE.

<table>
<thead>
<tr>
<th></th>
<th>TLS</th>
<th></th>
<th></th>
<th>PVC</th>
<th></th>
<th></th>
<th>NECTE2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
<td>m</td>
<td>f</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>have to</td>
<td>7</td>
<td>30</td>
<td>7</td>
<td>27</td>
<td>56</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.9</td>
<td>42.3</td>
<td>8.9</td>
<td>22.3</td>
<td>45.9</td>
<td>40.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>have got to</td>
<td>38</td>
<td>31</td>
<td>62</td>
<td>87</td>
<td>48</td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>80.9</td>
<td>43.7</td>
<td>78.5</td>
<td>71.9</td>
<td>39.3</td>
<td>40.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>need to</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>17</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>1.4</td>
<td>7.6</td>
<td>5.0</td>
<td>13.9</td>
<td>18.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>must</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.3</td>
<td>12.7</td>
<td>5.1</td>
<td>0.8</td>
<td>0.8</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>47</td>
<td>71</td>
<td>79</td>
<td>121</td>
<td>122</td>
<td>164</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TLS Chi-square = 12.89, d.f. = 1, p = 0.0003

PVC Chi-square = 5.21, d.f. = 1, p = 0.0225

NECTE2 Chi-square = 0.42, d.f. = 1, p = 0.519

6.2.2 Social class and education

Tagliamonte and D’Arcy (2007: 79) find that, in Canadian English, *have* to is consistently favoured by speakers with post-secondary education, whereas those with only secondary education prefer (*have*) got to. Less quantitative research has been carried out on social class in British English, as already noted, particularly as regards
have (got) to. As for need to, Nokkonen (2010: 64) demonstrates that it is upper middle class speakers who use this form most frequently, and she argues that this is a function of their professional lives (see 3.4 above).

In our analysis of DECTE only the most recent sub-corpus was tested for this variable, owing to the imbalance in levels of education in the first two sub-corpora (i.e. the small number of speakers with post-secondary education (N=8). Interestingly, our results for a different variety of British English reveal that education appears to have no significant effect on the use of the modal variants.²⁰

<table>
<thead>
<tr>
<th></th>
<th>Secondary</th>
<th></th>
<th>Post-Secondary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>have to</td>
<td>60</td>
<td>47.2</td>
<td>62</td>
<td>39.0</td>
</tr>
<tr>
<td>have got to</td>
<td>49</td>
<td>38.6</td>
<td>66</td>
<td>41.5</td>
</tr>
<tr>
<td>need to</td>
<td>16</td>
<td>12.6</td>
<td>31</td>
<td>19.5</td>
</tr>
<tr>
<td>must</td>
<td>2</td>
<td>1.6</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>127</td>
<td></td>
<td>159</td>
<td></td>
</tr>
</tbody>
</table>

Chi-square: 5.83, d.f. = 3, p = 0.1204

Table 15. The effect of education on choice of modal variant in NECTE2

7 Discussion of results

²⁰ See also Appendix, Table 16, for the factor weights.
In answer to the question posed above: are semi-modals taking over from older modal forms as markers of obligation and necessity?, our findings show that, in DECTE, semi-modals dominate, relegating *must* to a minor player in the root modality category. From a low starting point in the 1960s/70s *must* declined even further in the subsequent sub-corpora and is now virtually non-existent as a root modal. This pattern also mirrors that of other varieties of spoken English, in which *must* has almost become monosemous as a marker of epistemic modality, while *have to* and *have got to* have taken over as root modals (see e.g. Tagliamonte 2013).

As for changes in frequency within the category of semi-modals, our results charting the rise and fall of *have got to* largely mirror the findings of other similar studies of colloquial spoken English, such as those of Tagliamonte and Smith (2006) and Close and Aarts (2010). The popularity of *have got to* in our 1960/70s data, which increased even further in the 1990s, shows a decline in our 2007-10 sub-corpus, where *have to* now competes on an equal footing with *have got to*. This fluctuation in frequency over time can be seen as evidence for the on-going grammaticalisation of the competing variants. Interestingly, the rise of *have to* appears to have occurred a decade later in DECTE than in the corpora used by Tagliamonte and Smith (2006) and Close and Aarts (2010), which only extended up to the end of the 1990s. In our 1990s sub-corpus of DECTE, *have got to* almost achieved categorical status.

With regard to internal constraints on the distribution of variants, our results demonstrate that, unlike in some other varieties of English, factors such as the grammatical person of the subject or the nature of the obligation (subjective versus objective) play no role in determining the distribution of *have to* and *have got to*. If we examine our most recent sub-corpus from 2007-2010, we see that *have to* and *have got
to enjoy equal status as modal markers and they appear to occur in free variation. The fact that there are no significant internal constraints on the distribution of these variants indicate that they are not restricted to particular grammatical environments. Indeed, they appear to be at a very advanced stage of grammaticalisation in the North East of England.

Similarly, there appear to be no external constraints operating on the distribution of the two most frequent semi-modals have to and have got to. Both forms are used equally by males and females and by educated and less educated speakers. The motivation for the rise of have to in DECTE does not accord with Tagliamonte and Smith’s (2006) and Tagliamonte’s (2013) accounts that females are leading the way in the change towards the more prestigious have to. On the contrary, in our earlier sub-corpora, females significantly preferred have to more than males did. However, by 2007-10, males had increased their use of have to so that, now, there is no significant difference in male and female usage in this region.

Moreover, there is no evidence in our data either for the assertion that have to is more prestigious than have got to. If this were the case, we would expect to see more educated speakers preferring have to. Yet results from the most recent NECTE2 sub-corpus reveal no significant difference between speakers who have diverse educational backgrounds as regards the choice of variant. This outcome differs from the findings of Tagliamonte and D’Arcy (2007) that educated speakers significantly favour have to in Canadian English. Indeed, there has been a long tradition of stigmatising got in North American varieties, whether this be got as a stative possessive (e.g. I’ve got a cat) or in the modal sense of have got to (see Tagliamonte et al. (2010: 161-162) for discussion). However, in British English, there is no similar evidence for the stigmatisation of have
Here, at least in the spoken language, *have got to* appears to have become socially acceptable relatively early on, as the dialogues of Eliot, Carroll and Wilde demonstrate (see section 3.3 above) and this situation remains the same today. Thus, it is not possible to explain the rise of *have to* in DECTE by appealing to extralinguistic factors.

Finally, *need to* has risen sharply in frequency over the last two decades. In our data, it emerges in the 1990s corpus and increases considerably in the most recent sub-corpus. This is particularly the case in contexts where the obligation is subjective (i.e. imposed internally by the speaker himself/herself). This latter constraint accords very well with Bybee et al.’s (1994: 177) interpretation of its primary function. Unlike Nokkonen (2010), we did not find any effect of sex or education on the use of *need to*, since, as an emerging marker in our data, the number of tokens is still far too low to quantify with sufficient rigour.

8 CONCLUSION

The semi-modals *have to* and *have got to* have taken over from *must* as the two major markers of root modality in DECTE. Over the past fifty years they have come to be on a par with each other and are not constrained either by internal or by external factors in this corpus of North Eastern English. In particular, the tendency for women to favour *have to*, which has been reported for other varieties of British and North American English, is not evidenced in our most recent data from this locality. It is not clear what is determining the rise of *have to*, but it is unlikely to be considerations of prestige, as

---

21 Fehringer and Corrigan (in press) demonstrate that sex and education play no significant role in determining the distribution of stative possessive *have got* versus *have* in DECTE either.
there is no evidence in DECTE of *have to* being seen as a stronger marker of status than *have got to* is. Indeed, the increase in the use of *have to* simply demonstrates that competing variants may rise and fall in frequency over time, which is evidence for the on-going grammaticalisation of the forms in question.

For future research, it will be of interest to ascertain: i) whether the use of *have to* will completely overtake that of *have got to*, as has happened with *must*; ii) whether internal constraints will develop over time which determine the distribution of the variants (e.g. the specialisation of *have got to* with second person indefinite subjects, as shown by Tagliamonte and Smith (2006) for other varieties; or iii) whether the free variation apparent in this variety of English will continue unconstrained. Moreover, the newly-emerging *need to* modal also requires further research. At present, it appears to be occurring mostly in contexts of subjective obligation. However, the fact that it can also compete with *have to* and *have got to* in utterances expressing objective obligation indicates that it could, in time, become a serious competitor to the two main variants.
APPENDIX

Table 16. Multivariate analysis: *have got to* in all three subcorpora

<table>
<thead>
<tr>
<th></th>
<th>TLS</th>
<th>PVC</th>
<th>NECTE2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>0.676</td>
<td>0.836</td>
<td>0.485</td>
</tr>
<tr>
<td>Total N</td>
<td>106</td>
<td>183</td>
<td>237</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-61.724</td>
<td>-81.156</td>
<td>-164.172</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Factor weight</th>
<th>%</th>
<th>N</th>
<th>Factor weight</th>
<th>%</th>
<th>N</th>
<th>Factor weight</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subject type</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generic (2&lt;sup&gt;nd&lt;/sup&gt; person) objective</td>
<td>[.515]</td>
<td>62.7</td>
<td>51</td>
<td>[.451]</td>
<td>80.9</td>
<td>47</td>
<td>[.544]</td>
<td>55.4</td>
<td>74</td>
</tr>
<tr>
<td>Definite (1&lt;sup&gt;st&lt;/sup&gt;/3&lt;sup&gt;rd&lt;/sup&gt; person) objective</td>
<td>[.488]</td>
<td>64.7</td>
<td>17</td>
<td>[.520]</td>
<td>84.2</td>
<td>38</td>
<td>[.519]</td>
<td>53.2</td>
<td>47</td>
</tr>
<tr>
<td>Definite (1&lt;sup&gt;st&lt;/sup&gt;/3&lt;sup&gt;rd&lt;/sup&gt; person) subjective</td>
<td>[.472]</td>
<td>60</td>
<td>20</td>
<td>[.524]</td>
<td>81.2</td>
<td>64</td>
<td>[.402]</td>
<td>42.9</td>
<td>42</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Speaker sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>.722</td>
<td>84.4</td>
<td>45</td>
<td>.664</td>
<td>89.9</td>
<td>69</td>
<td>[.474]</td>
<td>46.2</td>
<td>104</td>
</tr>
<tr>
<td>Female</td>
<td>.331</td>
<td>50.8</td>
<td>61</td>
<td>.398</td>
<td>76.3</td>
<td>114</td>
<td>[.520]</td>
<td>50.4</td>
<td>133</td>
</tr>
<tr>
<td><strong>Range</strong></td>
<td>39.1</td>
<td></td>
<td></td>
<td>26.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scores closest to 1 show a favouring effect while those closest to 0 show a disfavouring effect. Brackets indicate that a score is not statistically-significant. Note that the disparity between the percentages and the factor weights in the 'subject type' group is most likely due to the fact that the factors in that group all have very similar weights/percentages and the factor group is not significant overall, in any time period.
Carol Fehringer

School of Modern Languages,

Old Library Building,

Newcastle University, Newcastle upon Tyne NE17RU, UK.

E-mail: carol.fehringer@ncl.ac.uk

Karen P. Corrigan,

School of English Literature, Language and Linguistics,

Percy Building, Newcastle University, Newcastle upon Tyne NE17RU, UK.

E-mail: k.p.corrigan@ncl.ac.uk


Nokkonen, Soli. 2010. ‘How many taxis there needs to be?’ The sociolinguistic variation of NEED TO in spoken British English. *Corpora* 5(1), 45-74.


