

Penry Williams, C., & Korhonen, M. (2020). A sociolinguistic perspective on the (quasi-)modals of obligation and necessity in Australian English. *English World-Wide*, 41 (3), 267–294.  
doi:10.1075/eww.00051.pen  
<https://benjamins.com/catalog/eww.00051.pen>  
Copyright held by John Benjamins

The publisher should be contacted for permission to re-use or reprint the material in any form.

## **A sociolinguistic perspective on the (quasi-)modals of obligation and necessity in Australian English**

Cara Penry Williams and Minna Korhonen

University of Derby/La Trobe University and Macquarie University

This article examines the distribution and sociolinguistic patterning of (quasi-)modals which express strong obligation/necessity, namely *must*, *have to*, *have got to*, *got to* and *need to*, in Australian English. Variationist studies in other varieties of English have had contrasting findings in terms of distributions of root forms, as well as their conditioning by social and linguistic factors. The corpus analysis suggests real-time increased use of *need to* and decrease in *have got to* through comparison to earlier findings. The variationist analysis shows quasi-modals *have to*, *have got to* and *got to* as sensitive to speaker age and sex, and a recent increase of *have to* via apparent time modelling. Linguistic conditioning relating to the type of obligation and subject form is also found. The study contributes to sociolinguistic understanding of this large-scale change in English and the place of Australian English amongst other varieties.

Keywords: Australian English; root modality; epistemic modality; quasi-modals; language change; age; sex; apparent time

### **1 Introduction**

#### **1.1 Context and focus**

Modal auxiliaries in English are undergoing long-term change in their uses and forms. Overall, it is argued that there is a decreasing use of modals potentially attributable to grammaticalization, colloquialisation or a combination of the two (Leech 2013). In addition, there is change in the forms being used to express modality: alongside true modal auxiliaries (e.g. *must*, *may* and *will*) there are periphrastic forms referred to as *quasi-* or *emergent modals* including *semi-modals/modal idioms* and *lexico-modals/semi-auxiliaries* (e.g. *have to*, *be able to* and *be going to*) (Krug 2000; Collins 2009a; Leech et al. 2009). These quasi-modal forms differ in some of their properties from true modals (see Collins 2009a) but function similarly. In fact,

quasi-modals are increasing in frequency across varieties of English and in some cases look to be replacing the older modal auxiliary forms (Leech 2013). Currently, quasi-modal use is considered more “informal” in comparison to true modal use and is far more common in spoken language than in writing (Leech et al. 2009; Collins 2014). In particular, it is in the domain of expressing necessity and obligation that there has been the most change (Mair 2014) and quasi-modals are used more frequently than the traditional modal *must* (Collins 2009a: 33), which is in marked decline (Baker 2011; Leech 2013; Hansen 2018).

This study investigates the current state of this change in spoken Australian English (AusE). It examines overall frequencies of forms expressing strong obligation and necessity, then utilises a variationist analysis to establish their distribution while accounting for the role of social and linguistic constraints. The methodology is closely modelled on research conducted in other contemporary varieties of English on the variation in root modality.

Section 1.2 introduces the linguistic forms, briefly outlining the pertinent part of the modal system. Section 2 details the most relevant previous research, first in AusE and then in variationist sociolinguistics. Section 3 includes the research questions, and describes the data and methods of analysis used in the study. The results of the study are reported on in Section 4, first in a corpus-style analysis of the overall distribution of forms, followed by a variationist analysis of social and linguistic constraints on the use of the (quasi-)modals investigated. Section 5 presents overall conclusions for AusE and the study of the variation in (quasi-)modals of obligation and necessity more generally. The paper provides a first sociolinguistic account of this variation in AusE, showing how it is participating in these large-scale, cross-variety changes.

## 1.2 The forms

As shown by Tagliamonte and Smith (2006) and Tagliamonte and D’Arcy (2007), and research therein, the system for Old English originally contained only *mot* to express a meaning of obligation and necessity. *Mot* became *must* in Middle English and *have to* was also first attested in this period. In Modern English, there were further additions with *have got to*, *got to* and *need to* first appearing in the 19th century. In the 20th century, all these

forms continued to be used with the additional possibility of a reduced form of *got to* often represented in informal writing as *gotta* (note that *have to* can be rendered as *hafta* or *hasta* to show similar processes but is less commonly seen. Cf. *?needa*).<sup>1</sup> Leaving aside the issues of reduction and contraction, contemporary varieties of English may allow these as a set with closely equated meaning as exemplified in Figure 1.

We	<i>must</i>	<i>go for a drive up there. (G1_F15_O)</i>
You	<i>have to</i>	<i>express it in a way that sounds, a lot more professional. (DANIEL_M_Y)</i>
	<i>have got to</i>	
	<i>got to</i>	
You	<i>need to</i>	<i>really have a car and to be able to get around in a car. (G1_F09_O)</i>

**Figure 1.** The set of (quasi-)modal forms as roughly equivalent in the same utterances<sup>2</sup>

These examples from our data (described in Section 3.1) originally contained *must*, *got to* and *need to*, respectively, but illustrate the discourse equivalence of the forms (potential differentiation is discussed in Section 4.2). All are “strong” in terms of their usual modal strength compared to *should*, *ought to*, *be supposed to* and *had better*, which are typically categorised as “medium” (Collins 2009a: 33) or “weak” (Smith 2003) in their expression of obligation. The forms in Figure 1 can thus be understood as expressing necessity (Huddleston and Pullum 2002) and are semantically alike enough for a variationist analysis.<sup>3</sup> Their concurrent use for the same purposes is considered an example of layering, with new forms fulfilling the same function as the older ones (Tagliamonte and Smith 2006).

The set of forms investigated (*must*, *have to*, *have got to*, *got to* and *need to*) can encode different types of modality. We follow Coates (1983) in adopting a system with a two-way

<sup>1</sup> We refer interested readers to the extended discussions of historical developments in Tagliamonte and Smith (2006), Fehring and Corrigan (2015), and Hansen (2018, Ch. 6).

<sup>2</sup> In participant identifiers, \_M = male and \_F = female. Age group is added to the end: \_A = Adolescent, \_Y = Young adult, \_MA = Middle-aged, \_O = Older. The different naming systems for speakers is an artefact of the two data sources: Blayney (code) and Melbourne (pseudonym).

<sup>3</sup> Based on its interpretation as a medium or weak strength modal and that in most cases it is not included in the previous sociolinguistic investigations of modals of obligation and necessity, *should* is not included in this analysis. Readers can imagine *should* within the sentences in Figure 1 for a comparison regarding strength of obligation.

split between epistemic and root modality to be directly comparable to previous sociolinguistic studies although other systems and further distinctions are possible.<sup>4</sup>

Epistemic modality relates to beliefs and knowledge, and encodes likelihood of the truth in a proposition as exemplified in examples (1) and (2).

(1) But I haven't been there to see her for ooh it **must** be five years (G1\_F04\_O)

(2) playing backyard cricket definitely **has to** be number one (JACQUI\_F\_Y)

In (1) the speaker notes that when considering how long it has been since she has been to her sister's home, it is by deduction, a natural conclusion, that the time lapsed is necessarily five years. In (2) again the information is presented as a natural or certain outcome of reasoning, knowledge, here, regarding "typically Australian" activities.

Root modality, on the other hand, is understood as the primary or initial meaning of such constructions (Collins 2009a: 21). Use of this type of modality suggests a personal (group) obligation to act as shown in examples (3)–(5).

(3) you **must** try and keep your uh young people active (G1\_M05\_O)

(4) And if you **have to** go to Sydney it's only three hours away (G2\_F11\_MA)

(5) I wanna get a ferret so now I've **got to** make him a cage so that'll be some good welding practice (G3\_M06\_A)

In (3), although the *you* is generic and does not refer to specific people and so the utterance functions more as a statement of convention, the root meaning comes via the idea that there

---

<sup>4</sup> Most previous sociolinguistic work uses the term *deontic* modality with the same referent (e.g. Tagliamonte and Smith 2006; Walker and Hoffman 2016). The use of *root* is preferable here in that previous work on AusE makes a further distinction between deontic and dynamic meanings (within non-epistemic meanings), following Palmer (1990), and also used by Huddleston and Pullum (2002). Collins (e.g. 2009a) analyses deontic modality, which he defines as "concerned with conditions relating to the completion of an action deriving from an external source" alongside dynamic modality which he describes as "concerned typically with an individual's ability or volition" (2009a: 73). As Fehringer and Corrigan (2015) note, using the term *root modality* avoids having two different meanings of *deontic modality*. Depraetere and Reed (2006) offer a helpful overview of different systems of division.

is a social obligation to keep young people active: it is a necessary course of action. In (4) again there is a requirement to go to Sydney rather than a statement of belief or knowledge and in (5) the speaker is under obligation to act. This type of meaning is further exemplified in examples (6)–(7) with *got to* and *need to*.

(6) I suppose there's lots of things you could want but you **gotta** live within your means  
(G2\_M07\_MA)

(7) I think that we **need to** somehow address this issue in Australia, where no one thinks they're a racist (KELLY\_F\_YA)

While the primary focus in this analysis is root uses, epistemic uses of the same forms are also described for completeness, following convention in much of the literature.

## 2 Previous research

In order to set the background for our study, we discuss two sets of relevant previous findings: those that relate to AusE (Section 2.1), and studies of other varieties of English which investigate use of these (quasi-)modals via variationist analyses (Section 2.2).

### 2.1 In Australian English

Collins' corpus studies (e.g. Collins 2005, 2009a, 2009b, 2014) extensively discuss the overall distribution of (quasi-)modals in AusE and make comparisons of their use to other varieties of English, showing that AusE usage falls between the English of the United States (USE), which is most advanced in the uptake of newer forms, and British English (BrE). For modals and quasi-modals of obligation and necessity specifically, Collins' (2009a) evidence from the synchronic *International Corpus of English - Australia* (ICE-AUS) shows, for example, that (quasi-)modals are sensitive to mode in that *must* is rare in speech, while quasi-modal alternatives are commonly used in spoken AusE.<sup>5</sup>

---

<sup>5</sup> See <https://www.ice-corpora.uzh.ch/en.html> for more information on ICE and ICE-AUS.

In order to provide a baseline for the current study, the findings pertaining to modals of obligation and necessity in AusE from Collins (2009a) are presented in Table 1. To make Collins' findings comparable to those reported in the current study, we have collapsed the categories of deontic and dynamic modality into *root* modality and only included "strong" forms. Furthermore, as our interests are in spoken language interactions, only results from the dialogue section of ICE are included.

**Table 1.** ICE-AUS spoken dialogue frequency of forms (adapted from the appendix in Collins 2009b)

	<i>must</i>		<i>have to</i>		<i>(have) got to</i>		<i>need to</i>		<b>Total</b>
	N	%	N	%	N	%	N	%	<b>N</b>
<b>Root</b>	65	6.1	649	60.7	258	24.1	98	9.2	<b>1070</b>
<b>Epistemic</b>	114	94.2	2	1.7	4	3.3	1	0.8	<b>121</b>

Although the results in Table 1 are not completely comparable with the current study due to the differences in data type and the definition of the variable *(have) got to* as one form, they are of interest as they suggest likely relative frequencies of the forms in AusE across a range of interaction types to which the results of the current study can be compared. Of particular interest is the fact that *must*, the only true modal, is least commonly used for root modality at 6% but it is not altogether rare in speech in that it accounts for over 94% of epistemic use of the forms in focus in the present study. Thus, it seems that *must* is specialising to epistemic use in AusE (see Trousdale 2003). In terms of the root forms, *have to* accounts for over 60% of instances, *(have) got to* 24% and *need to* just over 9%, indicating a strong preference for *have to* in spoken AusE.

In addition to his synchronic corpus studies, Collins (2014) provides a diachronic view on the (quasi-)modal use in Australian fiction from 1788–1999, based on the *Corpus of Oz Early English* and *AusCorp*, showing increasing use of *have to* and *have got to* and decrease of *must*. Collins (2014: 12–13) discusses various phenomena as possible reasons for the increase of *have to* in AusE, including the flexibility of *have to* and the influence of *democratisation*, that is, the creating of more equal power relationships between interlocutors.

There have also been some survey investigations into the use of (quasi-)modals of obligation and necessity in AusE which reveal age-based patterns in their preferred use. Collins (2007) reports that while the younger respondents in his *Australian Style*<sup>6</sup> survey strongly favoured *have to* in casual conversation (78% and 80%), they shifted toward *must* in writing (69% and 71%). In contrast, the older respondents did not change usage based on context, with use of *have to* at around 40% for both contexts and genres. *Have to* was preferred over *have got to* in both contexts by all age groups, but the two youngest groups shifted toward *have got to* for written communication and the youngest suggested it most for speaking. In addition to differences in usage across age groups (linked to language change), these findings also suggest differences in ideas surrounding formality (via social distance and mode/genre).

While the studies discussed above give a brief description of the use of the (quasi-)modals in AusE and suggest some age-based differences in preferences and usage, to-date there is no published sociolinguistic analysis of root modality use akin to those completed for other varieties of English. The main findings of such studies will be discussed next.

## 2.2 Sociolinguistic work in other varieties of English

The distribution of different sets of (quasi-)modals has been investigated using variationist methods in multiple regional varieties of BrE as well as Canadian English (CanE).<sup>7</sup>

According to these studies, root modal *must* is decreasing in relative frequency across the varieties (see also Section 2.1). This is sometimes attributed to change in social values with speakers increasingly avoiding bold assertions concerning what others are obliged to do, instead preferring to express obligation in forms less associated with authoritative discourse, especially in informal interview data (e.g. Tagliamonte 2004). However, the increase of particular quasi-modal forms has been found to differ by variety. For example, although in some BrE communities *have to* is used most (e.g. Northern Ireland, Scotland and Northern England), in others *got to* and *have got to* are more frequent (Tiverton, Devon) (Tagliamonte and Smith 2006). Based on interpreting the latter varieties as more innovative, Tagliamonte

---

<sup>6</sup> *Australian Style* is the newsletter of the Dictionary Research Centre at Macquarie University.

<sup>7</sup> Jankowski (2004) also studies USE with comparisons to BrE but the use of playscripts as data means results are not directly comparable to other studies, including this one.

(2012) argues for increasing use of *have got to* in BrE varieties, as part of a large-scale change and interprets majority use of *have to* as conservative in BrE. However, using the *Diachronic Electronic Corpus of Tyneside English* (DECTE), collected at three data points between the 1960s and 2000s, Fehringer and Corrigan (2015) found *have to* on the increase (% *have to* > (*have*) *got to* > *need to* > *must*). Furthermore, the real-time data display a more recent lowering use of (*have*) *got to* (59%, 75%, 40%, from oldest to newest data). In CanE, *have to* dominates and (*have*) *got to* is declining in both apparent-time (Tagliamonte and D'Arcy 2007; Denis and D'Arcy 2019) and real-time analyses (Denis and D'Arcy 2019).

In sum, depending on the interpretation and allowing for more local variation, there is either (1) a divergence between BrE and CanE (following Tagliamonte 2012) in that the former is increasingly using *have got to* and the latter *have to* or (2) an overall decreasing use of *have got to* and increasing use of *have to* across these varieties (following Fehringer and Corrigan's 2015 argumentation), contradicting earlier predictions of the continuing rise of *have got to* (e.g. Krug 2000).

In addition to regional variation, previous research has found that the root (quasi-)modals are sensitive to social (or external) factors (Tagliamonte and Smith 2006; Tagliamonte and D'Arcy 2007; Hoffman and Walker 2014; Walker and Hoffman 2016). In particular, Tagliamonte and colleagues find them to be conditioned by age and sex (and level of education in CanE) with *have to* associated with the young and females. Using the same methodological approach in Tyneside English, however, Fehringer and Corrigan (2015) do not find the same social factors at play. In contrast, they find that in their most recent data *have to* is no longer associated with females and (quasi-)modal use does not have a significant relationship to speaker sex or level of education (2015: 374).

In terms of linguistic (or internal) factors, Tagliamonte and colleagues' research has found distribution to depend on subject type (grammatical person and generic vs. definite reference, e.g. *you* meaning 'one' vs. second person sg./pl.) used to quantify the pragmatic force of the obligation (Tagliamonte 2004, 2012; Tagliamonte and Smith 2006; Tagliamonte and D'Arcy 2007). Furthermore, there was also a relationship to the source of the obligation

(subjective or objective). Subjective use relates to speakers creating the obligation/necessity whereas objective uses are external to them and function as a report or assertion (Collins 2009a: 28) (e.g. compare, in ordinary circumstances, a friend saying *you have to share your cake with you have to pay a 10% tax*: one is not obligated to pay the tax because the friend said so). In Tagliamonte and colleagues' analyses, these measures are combined in the multivariate analysis to approximate modal use strength (e.g. Generic vs. Definite+Objective vs. Definite+Subjective [Tagliamonte and D'Arcy 2007] or categories such as Definite+1<sup>st</sup> & 3<sup>rd</sup> person+Objective [Tagliamonte and Smith 2006] ) with these providing significant differences in modal use) (see also Denis and D'Arcy 2019). Fehringer and Corrigan (2015), on the other hand, did not find this with such combinations of linguistic factors non-significant in the multivariate analysis. For instance, there was little difference in the split of subjective and objective tokens for *have to* and (*have*) *got to* suggesting that this is not a factor affecting their use in Tyneside English, although it was statistically significant for *need to*.

Although most of the previous variationist studies on (quasi-)modals of obligation and necessity have utilised the previously described approach to investigate the linguistic constraints of subject and obligation type, there is also another possible way to examine these. In their investigations of CanE, Hoffman and Walker (2014) and Walker and Hoffman (2016) have analysed the modal forms *have to*, *have got to* and *got to* alongside their equivalent possessive forms, demonstrating shared constraints as part of a larger (lexical) variable. Put differently, *I have to go now* (root modality) and *I have a hat* (possessive) operate under the same linguistic and social constraints as the variable of HAVE. This sort of treatment does not utilise the linguistic constraints discussed above, instead relying on the shared grammatical properties of the subject as a generic pronoun, a personal pronoun or a noun phrase. This simpler classifying of forms by subject type thus offers an alternative way to ascertain the importance of linguistic factors without reducing the number of tokens in groups by excessively subdividing the data and introducing additional ambiguities which result in exclusions.

### 3 This study

The studies discussed in Section 2 raise a number of questions with which a sociolinguistic study of AusE can engage, as new source of data and approach for examining (quasi-)modal use in AusE, and in providing another variety to consider alongside other sociolinguistic studies and their contrasting findings. In order to do this, we address the following research questions:

- RQ1. How does the distribution of *must*, *have to*, *have got to*, *got to* and *need to* in these data compare to previous studies of AusE?
- RQ2. How does the distribution of the root variants compare to similar studies of other varieties of English?
- RQ3. Do the social factors of speaker age and sex affect the distribution of the root (quasi-)modal forms? Do these distributions support an interpretation of change in progress?
- RQ4. Is the variation in the root forms conditioned by linguistic factors related to subject and/or obligation types?

The data and methods of analysis used in the study to explore these research questions are described in the following sub-sections.

#### 3.1 Data

The study is based on interview data from two locations within eastern Australia: Blayney, a country town in the state of New South Wales (about 250 kilometres inland and west of the state capital of Sydney), and Melbourne, the capital city of Victoria (the southern-most mainland state). The second of these sites contributes solely to one of the age groups discussed below (Young Adult, 17 of 18 participants in this group). The two datasets were collected independently but were both designed to elicit casual spoken data from the Australian-born, local speakers. The interviews were in part based on well-established sociolinguistic interview schedules and the aim was to make the participants feel comfortable and encourage them to speak “freely”. Despite this, we would not describe the

resulting data as completely informal throughout, with some obvious differences between participants and interview sections.

The interviews were with 87 participants aged 14 to 80 with 56% females.<sup>8</sup> In total, the combined corpus consists of approximately 60 hours of talk and around 439,400 transcribed words from participants. Age was divided into four groups, commensurate to the life stages of adolescence, young adulthood, middle-age and older adulthood. Table 2 provides a summary of the participant information by age group and sex.

**Table 2.** Distribution of speaker age and sex in the dataset

Age group	(years)	Female	Male	Total
Older	(57–80)	14	7	21
Middle	(40–55)	14	9	23
Young	(20–30)	10	8	18
Adolescent	(14–18)	11	14	25
Total		49	38	87

### 3.2 Analysis

Analysis began with identifying all instances of the (quasi-)modals included in the study: *must*, *have to*, *have got to*, *got to* and *need to*. Given that the research presented here is the first sociolinguistic analysis of these (quasi-)modals in AusE, we did not want to risk obscuring potential distinctions so opted to separate forms which are sometimes combined as one variant (e.g. *(have) got to*). For similar reasons, and also because of its increasing use (Nokkonen 2010; Fehring and Corrigan 2015), *need to* was included even though it has been excluded in some earlier analyses. Note that the full forms are used to refer to a range of realisations such as *has to*, *'ve got to* and *gotta*.

<sup>8</sup> We believe gender, the social role, is the important factor in language variation and not biological sex but use *sex* and *male* and *female* following conventions in variationist studies.

The analysis of the data is two-fold. First, we take a corpus-based comparison approach in which we include all examples of (quasi-)modals for obligation and necessity. For this, we coded all the relevant (quasi-)modals for type of modality: “root” (non-epistemic) or “epistemic”. From this part of the analysis, we removed any instances with unclear audio ( $N = 8$ ), those which were part of uncompleted intonation units and utterances which were restated ( $N = 27$ ), as exemplified in (8) and (9).

(8) I **have to** th- I had to think when I saw that (G1\_F04\_O)

(9) No. You **have to**, well, I’m in there usually (G3\_F03\_A)

In the second part of the analysis in which we used the variationist approach, our coding decisions largely followed those of Tagliamonte and colleagues (i.e. Tagliamonte 2004, 2012; Tagliamonte and Smith 2006; Tagliamonte and D’Arcy 2007) to ensure comparability of forms/uses and to enable comparisons between varieties (also the case for Fehringer and Corrigan [2015], in real-time data). This meant excluding a large number of occurrences to ensure that each of the forms was potentially possible in that context. As part of this, tokens were coded for a range of syntactic criteria. All uses included in the analysis were:

1. in present tense,
2. not preceded by another modal,
3. not modified by emphatic *do*,
4. not interrogative and
5. without negation in the clause.

The first three of these relate to not all forms being possible, particularly due to syntactic restrictions in *must* in current use and not all varieties of English allowing the same number of alternatives (Mair 2014). For instance, AusE does not allow double modals nor *do* with *have got to* or *got to*. Interrogative and negated forms are excluded by Tagliamonte and colleagues due to their infrequency, rather than factors relating to the possibility of variation. However, Fehringer and Corrigan exclude them due to not all variants taking negation in the same way (2015: 368). To ensure comparability with most of the research

overviewed in Section 2.2 and to avoid including elements which complicate future comparisons in a wide variety of Englishes, negated tokens were removed (N = 61; 87% with *have to*). There were few instances of interrogative sentences, all using *have to* with *do*-support (N = 7). Following previous studies, we also excluded formulaic expressions such as *I must say*, as they were suggested to potentially preserve use of *must* in their fixedness (N = 10) and there was evidence to support this in our data.<sup>9</sup> This isolated all the relevant tokens in affirmative, declarative utterances in the present tense.

To explore linguistic constraints on the variation, the remaining tokens were coded for definite or generic reference, grammatical person of subject, and objective or subjective obligation. These were later combined following established models discussed in Section 2.2; however, these divisions did not work well with our data in the multivariate analysis and as the results of a preliminary analysis supported Hoffman and Walker's (2014) approach of shared grammatical constraints with possessive forms (Penry Williams and Korhonen 2018), we used the categorisation of generic pronoun, personal pronoun and noun phrase in the variationist analysis. The other distinctions are still explored in Section 4.2.2.

The coding was completed by the authors and in each of the categories uncertain cases were checked by a second author/coder. Any instances of ambiguity that we could not resolve were removed from the relevant analysis.

In coding tokens, we also noted if they were part of constructed dialogue, that is apparent direct reported speech and similar phenomena such as reported internal dialogue or imagined speech (see Tannen 1986). Such talk may differ in important ways from other speech, including in trying to represent speech that belongs elsewhere or to others (Penry Williams 2019a, 2019b). Smith (2003: 253) has briefly noted something similar in his analysis

---

<sup>9</sup> The category of formulaic expressions involved more interpretation and potential ambiguity than the other exclusions. There was one example excluded as the range of forms could not occur in this situation (*we just have to see. see how it goes*). A further nine were instances using *say* or *admit*. Although there was variation in both verbs between *must* and *have to*, there was support for excluding these as not the same as other data. given that (1) they all occurred with *I* and (2) there was no intraspeaker variation (although small counts). For analysis of these forms in more detail, see Tagliamonte and D'Arcy (2007: 73–74) and, as *performative modals*, Close and Aarts (2010).

of (quasi-)modals, finding that “direct speech quotation” increased use of *have to* (i.e. representations of speech within writing). It can also complicate coding of other elements, for example, for subjective obligation it may be that the actual speaker is not the author of the utterance, rather just the animator of another’s words (see Goffman 1981 on these distinctions). A separate analysis of some of these data found *have to* least used in constructed dialogue as a percentage of overall tokens, with all forms used a similar number of times. This means that these metapragmatic uses may inflate apparent use of less common forms, here *got to* and *must* (Penry Williams 2018). While there may not be a large number of tokens, even in large studies, it could have an effect on the quantitative patterning of less frequent forms in a similar way that promotes the exclusion of formulaic usage. While we kept the tokens to preserve comparability with previous studies, we suggest that, in future studies, tokens within constructed dialogue should be analysed and potentially removed.

The multivariate analysis included the linguistic factors in one model and the social factors in a second. All forms were included in the analysis with each being compared against all others, that is, for example, the analysis for *have to* compared it against *must*, *have got to*, *got to* and *need to*. These analyses utilised Rbrul (Johnson 2017) and when used, chi-square analyses employed the Vassar Stats site (Lowry 2011–2017) with significance set at  $p \leq 0.05$ .

## 4 Results and discussion

This section reports the results of the study, divided according to the two types of analyses used.

### 4.1 Corpus analysis of distribution

This first analysis captures all root and epistemic uses from all 87 participants (N = 760, root forms N = 716). The overall distribution of the forms is presented in Table 3.

**Table 3.** Overall distribution of all root and epistemic forms (N = 760)

	<i>must</i>		<i>have to</i>		<i>have got to</i>		<i>got to</i>		<i>need to</i>		Total
	N	%	N	%	N	%	N	%	N	%	N
<b>Root</b>	36	5.0	448	62.6	65	9.1	55	7.7	112	15.6	<b>716</b>
<b>Epistemic</b>	27	61.4	8	18.2	9	20.5	0	-	0	-	<b>44</b>

Collins' (2009a) data from ICE-AUS include a broader range of forms of talk and we have separated *(have) got to* into *have got to* and *got to*, so the alignment of the data is imperfect; however, the frequency of highest to lowest *have to*, *(have) got to*, *need to*, *must* in root meanings is consistent across analyses. Comparing Table 3 with Table 1, *have to* and *must* figures are very similar (within 2% of one another). The difference in distribution shows less use of *(have) got to* in our data (16.8% when collapsed, cf. 24.1% in ICE-AUS) and higher use of *need to* (15.6%, cf. 9.2% in ICE-AUS). In these data, combined *(have) got to* then occurs at a similar rate to *need to* rather than occurring more than 2.6 times more often as it does in Collins (2009a). The difference between the two datasets was confirmed significant by a chi-square analysis (chi-square = 27.38, d.f. = 3,  $p < 0.0001$ , effect size Cramer's  $V = 1.2$ ) with percentage deviations showing *(have) got to* at -20% and *need to* +33% in relation to expected frequencies.<sup>10</sup> This potentially presents some evidence of shift in the AusE system in real time (early 1990s vs. mid to late 2000s) to be examined further in Section 4.2 via apparent time and patterns of language change associated with speaker sex.

The real-time reduction in the use of *(have) got to* is of particular interest given that it has been suggested to be an increasingly dominant form in the shift away from *must* use. The larger change to *need to* shows that it is, in part, filling the space created by declining use of *(have) got to*, given that other forms are similar in percentage use. An examination of Fehringer and Corrigan's (2015) similarly dated early 1990s and late 2000s data also shows a decrease in *(have) got to* and increase in *need to* suggesting that this could have been a change

<sup>10</sup> This is calculated for each cell by taking away the expected count from the observed and dividing it by the expected count then multiplying it by 100 to show it as a percentage, thus -20% means that the observed frequency is 20% lower than expected by the chi-square calculation (Lowry 2011–2017). The effect sizes are reported alongside chi-square to give a fuller account of the data in response to recent criticisms of presenting significance alone.

in this period on an inter-varietal scale. The prominence of *need to* might continue to increase if democratisation is a driving force behind change in the (quasi-)modals of obligation and necessity, as *need to* is claimed to be understood as more in the interests of the interlocutor(s) and therefore less authoritarian (compare *you must pay tax* and *you need to pay tax*), partly via its strong associations with subjective obligation (Smith 2003).

In terms of the much smaller number of epistemic forms, *must* accounts for 61.4% of occurrences rather than its 94.2% domination in Collins (2009a). There were no instances of the low frequency *need to* in the data with the space created by lower *must* use being filled by instances of *have got to* (20.5%) or *have to* (18.2%). Although the counts are very low, so the display is speculative, there is a potential pattern created by the examination of age-based usage as shown in Table 4.

**Table 4.** Overall distribution of epistemic forms

	%	<i>have got to</i>	<i>have to</i>	<i>must</i>	N
Older		58.3		41.7	12
Middle		25.0		75.0	8
Young		11.0	27.8	61.5	18
Adolescent			16.7	83.3	6

The distribution in Table 4 shows a movement towards *must* as the minority form in the Older speakers (unshaded cell). The oldest two groups use only *have got to* and *must*. The Young participants add *have to* into the system and continue the apparent shift away from *have got to* and finally the Adolescent tokens are only *have to* or *must*. The caveat is that varying perceptions around the formality of the situation could be part of the data differences. The analysis supports the notion of *must* specialising to epistemic meanings. Trousdale (2003: 281) argues that this is occurring via simplification, with the removal of polysemy in modals and thus “semantic focussing” of *must* as epistemic (in Tyneside English, 20 speakers). If this is a larger pattern, a detangling of epistemic and root meanings, it could further be a force in promoting increasing use of root *need to* which only occurs once in its epistemic meaning in Collins’ (2009a) 1990s data and not at all in these data from the 2000s.

Comparing the 61.4% use of epistemic *must* to other sociolinguistic studies, our results are much closer to the *Toronto English Archive's* 55% (Tagliamonte and D'Arcy 2007) than results from BrE regional varieties. In DECTE, *must* accounts for 98.4% of epistemic uses (Fehringer and Corrigan 2015) and 96% in the (later named) *Roots Archive* (Tagliamonte and Smith 2006). While there may be specialisation in relation to *must*, this does not appear to equate to dominance in AusE in the way seen in the BrE data. Indeed, Coates (1983: 57) describes epistemic *have to* as “an Americanism” and notes it as associated with young people. It appears well-established in AusE from these data at over 18% and, speculating based on age-based patterns, it could be set to be the second epistemic variant in AusE overtaking *have got to* (both occur at 18% in the *Toronto English Archive*). Although at this stage, even over twenty years after Coates' comment, it is still associated with young speakers in AusE.

#### 4.2 Variationist analysis

The second part of our analysis deals with the 494 tokens of root modals left after we had made the linguistic exclusions needed for variationist analysis. The restrictions left tokens from 81 of the 87 participants. Comparisons between Table 3 and Table 5 below demonstrate how the exclusions listed in Section 3.2 shaped the overall distribution. In particular, the flexibility of *have to* is apparent in the huge reduction of these tokens. Note though that the order of frequency is maintained and *have to* still dominates.

**Table 5.** Overall distribution of root variable forms (N = 494)

<i>must</i>		<i>have to</i>		<i>have got to</i>		<i>got to</i>		<i>need to</i>	
N	%	N	%	N	%	N	%	N	%
12	2.4	274	55.5	64	13.0	55	11.0	89	18.0

Comparing these AusE results to those from CanE (Tagliamonte and D'Arcy (2007), *have to* is used at a considerably lower percentage (55.5%, cf. 72% in CanE) and *need to* 10% more with the other modals not presenting large differences (N = 1 314, 152 speakers). In another study of CanE, Walker and Hoffman (2016) report very similar overall percentages to those

found here, differing by less than 5% for all except for *have got to* which is given as 10% lower than found here (N = 1 202, 69 speakers<sup>11</sup>). In summary, many differences are not robust across the CanE studies but it can be said, looking at the two together, that AusE displays similar proportions of *got to* usage but a higher use of *have got to* than found by Walker and Hoffman (3%), and Tagliamonte and D'Arcy (6%).

When comparing the findings with the percentage of use for the studies of BrE from DECTE (Fehringer and Corrigan 2015) and the *Roots Archive* (Tagliamonte and Smith 2006; Tagliamonte 2012), allowing for the areal diversity within the latter, there is a lower use of *have got to* in these AusE data (except for compared to in Northern Ireland) (N = 602, 155 speakers). Use of *have to* at 55.5% is higher than found by Fehringer and Corrigan (2015) (31%, 17%, 43% over the three time periods) (N = 604, 160 speakers) but this varies hugely in the *Roots Archive* by variety (45% overall [Tagliamonte and Smith 2006]). This picture is complicated by the fact that previous studies do not have five forms in focus as is the case in our study (either collapsing (*have*) *got to* or not including *need to*).

While these differences are not suitable for detailed analyses due to the differences in the interview populations and the included variants, comparing them as percentages is of value. For instance, it highlights the fact that competition between *have got to* and *have to* appears pertinent across varieties. It is noteworthy, that AusE, which is the youngest of the varieties discussed here, seems to sit between the BrE norms and those of CanE in its use of these (quasi-)modals, suggesting it could be finding its own path in the layering of these variants. For *have got to*, it is much closer to CanE's overall percentage but is not a minor variant as it is there, with use still on par with *need to* and *got to* and much higher than *must*. For *have to*, the use is either much lower or similar to CanE figures, depending on the study, and higher than in the BrE studies. Understandings of potential directions of change within AusE require analyses that considers the social factors implicated models of language change.

---

<sup>11</sup> The participants in Walker and Hoffman's (2016) study were from a variety of ethnic backgrounds, with ethnic groups a focus of the study.

#### 4.2.1 Social factors

This section deals with two central social factors, speaker age and sex, both fundamental to models of language change. Modelling use in apparent time, for instance, may indicate increasing use of newer forms/norms. An analysis of sex can be insightful since females usually lead prestigious incoming changes while young males often engage more with innovations with covert prestige (Labov 1963, 2001 *inter alia*).

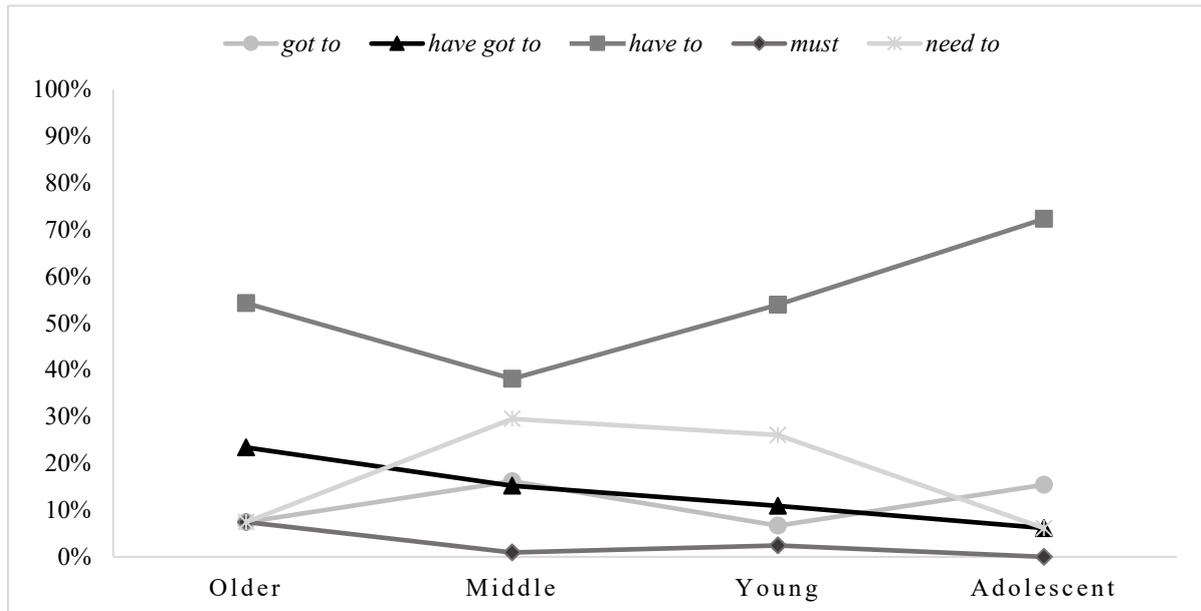
Starting with speaker age, Table 6 displays the use of each form by age group.

**Table 6.** Distribution of root variable forms by age group (N = 494)

	%	<i>must</i>	<i>have to</i>	<i>have got to</i>	<i>got to</i>	<i>need to</i>	N
Older		7.4	54.3	23.4	7.4	7.4	<b>94</b>
Middle		1.0	38.1	15.2	16.2	29.5	<b>105</b>
Young		2.4	53.9	10.9	6.7	26.1	<b>165</b>
Adolescent		0	72.3	6.2	15.4	6.2	<b>130</b>

Concentrating on age groups, there do seem to be differences in the systems used. Still, *have to* is the most frequently used (quasi-)modal in all age groups and *must* least or equally infrequent as *got to* and *need to* in the case of the Older group. In contrasting age groups, there is support for analysing *have got to* and *got to* separately: if these were tallied together the shift in usage would be obscured.

Figure 2 presents an apparent time perspective by charting these figures against the x-axis, representing change over time, and making trends clearer through the visualisation.

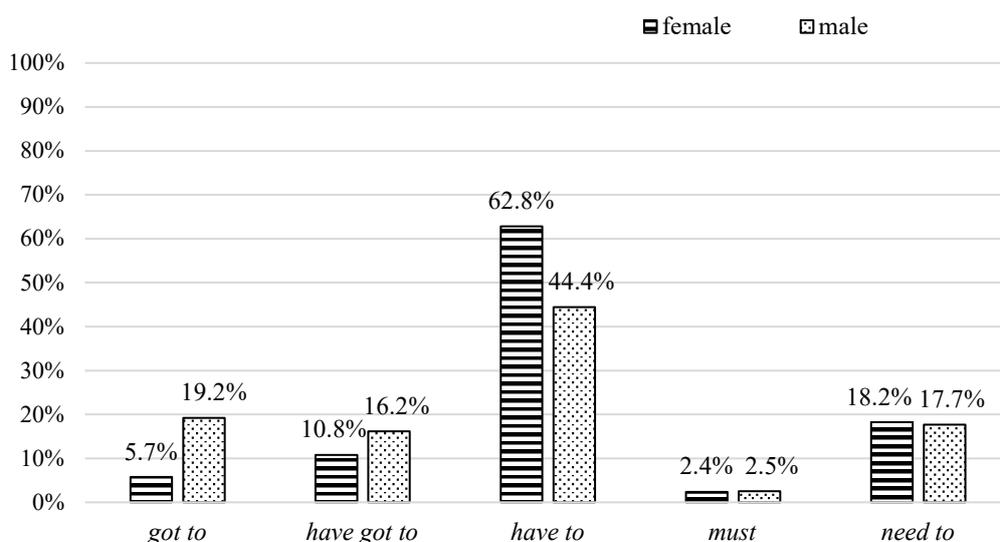


**Figure 2.** Distribution of root variable forms over apparent time

From Figure 2, it is clear that root *must* is dipping from already low percentages to being obsolescent amongst the youngest group. This parallels the larger change in the system of English which has been well-documented. If this trend continues, root modal uses of *must* will be very rare outside of fixed phrases. In contrast, *have to* is on the increase, after recovering from a dip in use from the oldest to middle-aged group. The progression from Middle, Young to Adolescent shows steady increases amounting to 34%. As discussed above, for adolescents *have to* truly dominates their system of modals of obligation and necessity. Despite claims of *have got to* as a great success story (Krug 2000), here it shows steady decline (also found in possessive forms [Penry Williams and Korhonen 2018]). It seems to be moving towards becoming a minor form in the variation as it is in CanE. The decline from the older to adolescent group equates to a 17% decrease. The pattern for *got to* is less clear with an overall slight rise-fall-rise pattern. Although age is still significant for this form (see Section 4.2.3), it appears not to be the full story in its distribution as further discussed below. *Need to* use is highest in the middle two age groups but then decreases from being the second most common to a minority form amongst the adolescents. In fact, looking at Figure 2, the line for *need to* is close to a mirror image of that of *have to*, with a suggestion of increasing use of *need to* that was not continued by younger participants, with ground lost amongst the Adolescents. Given that the analysis in Section 4.1 suggested overall increased use of *need to* in real-time it is possible that this distribution relates to age-

grading in use. Nokkonen (2010: 67) finds clear age-based patterning for *need to* in the *British National Corpus* with use low for 15–24-year-olds, highest amongst 25–34-year-olds and decreasing after this with increasing age and lowest for people over 60 years of age. She explains this through appealing to the differences in communication needs and the types of responsibilities and authority people hold at different ages. *Need to* encodes a strong obligation more indirectly (Smith 2003: 263–264) and indirectness in English has associations with politeness and notions of “softening” impositions. Nokkonen (2010) links this age-based patterning to levels of authority and responsibility, stating that “[p]erhaps the middle-aged have more authority, and, therefore, do not need to use equally strong expressions” (2010: 67). A further consideration of this might be strengthened by considering broader age-graded patterns around engaging with societal norms around politeness, extending the model of U-shaped (inverted here) engagement in the “standard language” market (Downes 1998) to the pragmatic norms bound to “speaking well”. Alternatively, it could be part of a larger pattern of change not yet visible with recent increases in *need to*, as it does seem to dip and recover in Denis and D’Arcy’s (2019) real-time data.

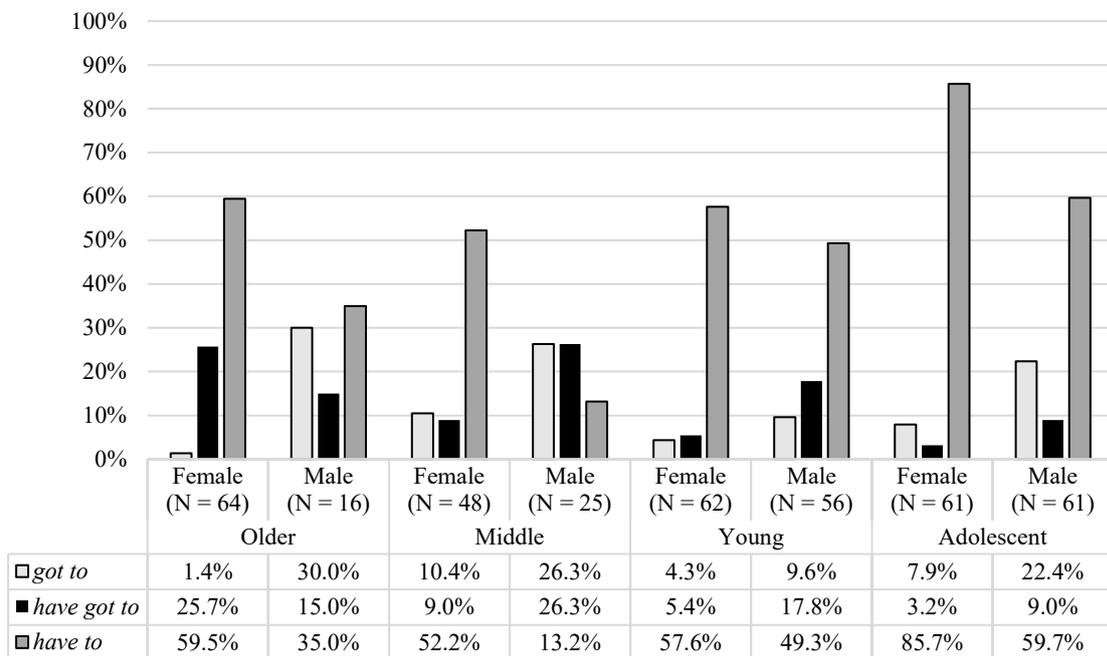
In addition to age, speaker sex displayed clear differences in the data. The higher number of female participants led to 296 tokens from them as compared to 198 from the males. Figure 3 shows the distributions by form as percentages by speaker sex.



**Figure 3.** Distribution of root variable forms by speaker sex

Figure 3 illustrates that *must* and *need to* are used with very similar frequencies by both sexes, while *got to*, *have got to* and *have to* present sex-based difference: females use considerably more *have to* while males use more *got to* and *have got to*. These results support *have to* as increasing overall and being a more prestigious form.

To look for further evidence that might support the understanding of *have got to* as previously increasing, in line with much of the literature but then *have to* becoming the preferred variant, Figure 4 breaks down the distributions for the key variants by speaker age group and sex.



**Figure 4.** Distribution of root *got to*, *have got to* and *have to* by speaker age group and sex (*N* shows total tokens)

Although some cell counts are low, Figure 4 shows that, in the Older group, females use more *have got to* than males but less after this. Male use of *have got to* increases in the Middle group but then reduces following this. An apparent time reading of this pattern suggests that it was a female-led change but then females moved away from this, with males lagging behind a generation and continuing the change, a common phenomenon found in variationist work (Labov 2001). In all age groups, males use more *got to* than females, supporting previous discussion of this as more common in male speech. The age-based break down of sex-based use also shows that females lead males in the use of *have to* across

age groups while it still is a pattern of increased use by all speakers, supporting the interpretation of it as an incoming change. Before measuring the actual impact on these social factors on the distribution, the linguistic factors in this analysis need to be accounted for.

#### 4.2.2 Linguistic factors

Although there are a range of strengths possible with the same form such as *must* (Collins 1991), previous research has operationalised the pragmatic force or strength of the particular instances via the properties of obligation type, subject grammatical number and definiteness/genericness of the subject. As discussed in Section 2.2, these have been found to be meaningful distinctions in some studies but not others.

Categorising tokens as subjective or objective obligation is not always straightforward and Figure 5 excludes ambiguous tokens. Forms are presented as a percentage of obligation type to remove the impact of the much larger number of objective uses (modelled on Tagliamonte and Smith [2006], Figure 5b).

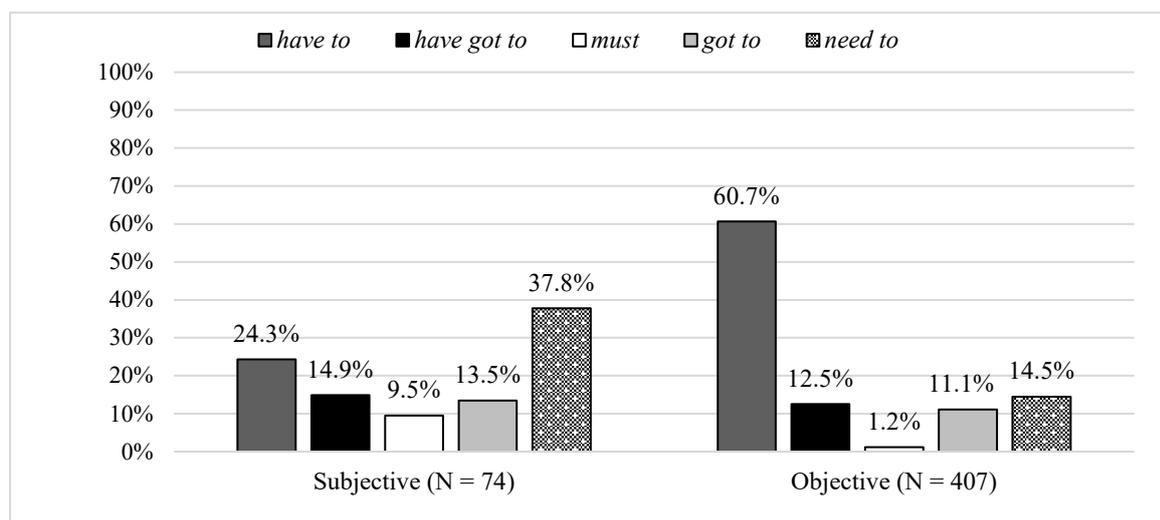
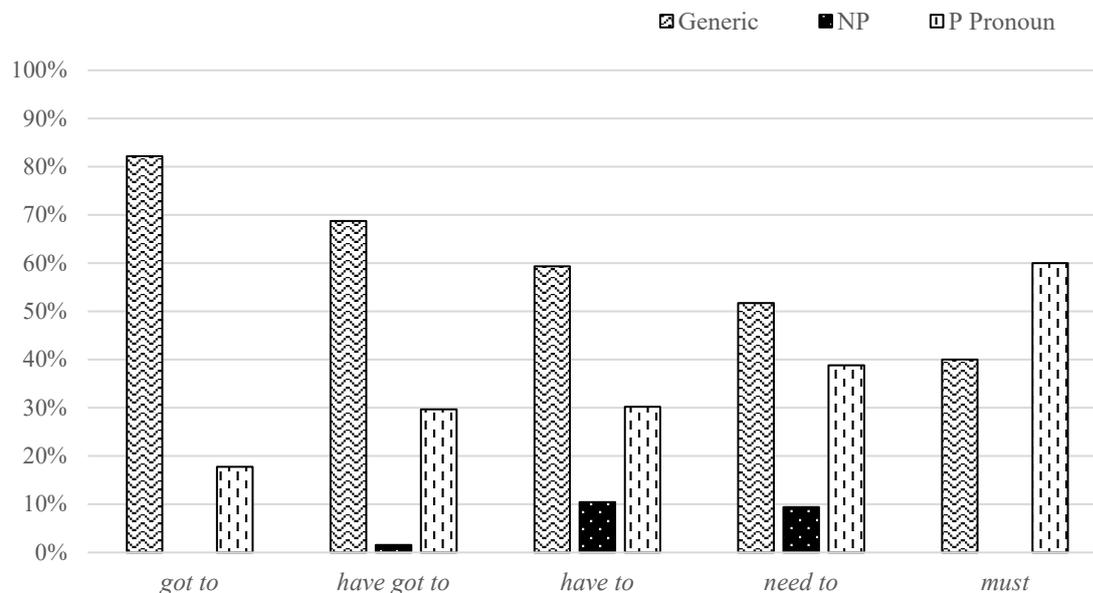


Figure 5. Forms as a percentage of obligation type

As found in Tagliamonte and Smith (2006), these data suggest a relationship between *have to* and objective obligation and *must* and subjective obligation. These relationships were not supported in the analysis by Fehringer and Corrigan (2015), but their study did find a (significant) relationship between *need to* and subjective obligation which is corroborated in

these results. In fact, a chi-square analysis which excluded the 12 tokens of *must* found this distribution was significant and an examination of the percentage deviations showed *need to* in subjective use at +125% and *have to* -52% (chi-square = 36.2, d.f. = 3,  $p < 0.0001$ , effect size Cramer's  $V = 2.8$ ). This provides a further explanation for the increasing use of *need to* in real-time (Section 4.1) alongside the change in progress of use of *have to* (Section 4.2.1), with each associated with a particular type of obligation. This analysis also shows that it is the quasi-modal *need to* that is most directly taking the place of modal *must*.

Figure 6 shows the data by subject type, incorporating subjective in comparison to objective meanings given that generic contexts do not allow for subjective readings (Tagliamonte and D'Arcy 2007).



**Figure 6.** Forms by subject type

In removing the obfuscation of differences in token numbers (which mean *have to* dominates most divisions), this analysis shows a clear relationship between forms and generic pronouns and inverse relationships with personal pronouns, supporting intersecting linguistic constraints on likelihood of use. This analysis, adds to that displayed in Figure 5, showing a distinctive relationship for *got to* with grammatical forms in strong association with generic pronouns, followed by *have got to*. This accords with specialising to these types of subjects also found in previous studies (e.g. Tagliamonte and Smith 2006, re. *have got to*).

This potentiality reserves a place for *have got to* in AusE despite overall decrease in use. It further exemplifies the integration of these variants into distinct roles in the system (the variable) regardless of broader discourse equivalence.

#### 4.2.3 Multivariate analysis

In order to further explore the relationships between (1) the social and (2) linguistic factors and to check if relationships were significant and their relative strength, we carried out multivariate analyses of the data, the results of which are presented in Table 7. Note that the analysis for *must* needs to be viewed with caution due to the low token numbers and is included only for the sake of completeness.<sup>12</sup>

---

<sup>12</sup> For readers unfamiliar with factor weights (FWs) and this sort of display of results, the figure measures the relationship from 0 to 1. The closer to 0 the more disfavoured and the closer to 1, the stronger the relationship.

**Table 7.** Results from the multivariate analyses (significant results bolded)

	<i>have to</i>		<i>have got to</i>		<i>got to</i>		<i>must</i>		<i>need to</i>	
	FW	%	FW	%	FW	%	FW	%	FW	%
<b>Age group</b>										
Older	<b>.45</b>	54	<b>.71</b>	23	<b>.46</b>	7	<b>.99</b>	7	<b>.32</b>	7
Middle	<b>.32</b>	38	<b>.56</b>	15	<b>.63</b>	16	<b>.96</b>	1	<b>.72</b>	30
Young	<b>.50</b>	54	<b>.44</b>	11	<b>.35</b>	7	<b>.98</b>	2	<b>.68</b>	26
Adolescent	<b>.72</b>	72	<b>.29</b>	6	<b>.56</b>	15	<b>.01</b>	0	<b>.28</b>	6
<i>Range</i>	40		42		28		99		44	
<b>Sex</b>										
Female	<b>.62</b>	63	<b>.40</b>	11	<b>.33</b>	6	.43	2	.51	18
Male	<b>.38</b>	44	<b>.60</b>	16	<b>.67</b>	19	.57	3	.49	18
<i>Range</i>	24		20		34		14		2	
<b>Subject type</b>										
Generic pronoun	<b>.42</b>	55	<b>.66</b>	15	<b>.99</b>	13	.99	1	.42	15
Personal pronoun	<b>.42</b>	55	<b>.62</b>	13	<b>.99</b>	5	.99	4	.53	22
Noun phrase	<b>.65</b>	76	<b>.24</b>	3	<b>.01</b>	0	.01	0	.53	22
<i>Range</i>	23		42		99		99		11	

The significant results, in bold, further support analyses presented in previous sections. They demonstrate and measure the robustness of the relationship between speaker sex and forms and includes age and sex in the one analysis. The results further corroborate the association between males and *got to* and *have got to* and females and *have to*, also found by Tagliamonte and Smith (2006) and Tagliamonte and D’Arcy (2007) in CanE but contrasting with those of Fehringer and Corrigan (2015) in Tyneside. In terms of age, the factor weights (FW) are over .70 for the relationships identified previously for the Adolescent group and *have to* and the Older group and *have got to*, lending further support to the analysis of this “reversal” in status. Furthermore, the FWs ascend or descend in the predicted way. There is some disruption to this pattern in the Middle group and their strong relationship to use of

*need to* is also shown here (.72). Future analyses will be needed to see if this age difference is a fluctuation in a process of change or is related to the life stage.

The analysis of the linguistic factors only includes the one factor which entails the relationships to obligation type and genericness within it. It further shows the specification of some forms for subject type, for instance generic for *got to*, supporting the analysis in Section 4.2.2. The analysis shows the significant elements the inversion pattern seen in Figure 6, with FWs for *have to* in comparison to *have got to* showing them holding distinct roles when analysed at the micro-level.

## 5 Conclusions

This paper has investigated the use of (quasi-)modals of obligation and necessity, *must*, *have to*, *have got to*, *got to* and *need to*, in AusE from two perspectives: corpus-based and variationist. The first of these showed that layering of the forms is present in AusE, but with *have to* dominating while the only true modal *must* occurs very infrequently. In comparison to earlier frequencies in corpus studies in AusE, usage of *need to* is higher and (*have*) *got to* lower, suggesting reorganisation that could relate to larger processes which, in turn, promote the use of indirect laying down of obligations. In line with previous studies, the findings broadly place AusE between BrE and the North American variety of CanE in terms of its uptake of the newer quasi-modal forms. It is important to remember though, that AusE is not simply moving from a BrE system to a North American one and that relationships between varieties are much more complex than this. Still, comparisons allow us to understand inter-varietal differences in more detail and uncover shared and divergent norms.

The variationist analysis showed AusE displaying some similar tendencies in the use of the (quasi-)modals under investigation as have been reported in other varieties. The analysis of social constraints finds an increase in the use of *have to* led by females, whereas *have got to* and *got to* are more often used by male speakers. Apparent-time modelling highlights the increasing presence of *have to*; after recovering from a dip in frequency in the middle-age

group, it shows a steady rise to become very robust in both young adult and adolescent use. Of the other forms, *have got to* presents a steady decline across the age groups, while *got to* presents a minor rise-fall-rise pattern, surpassing the use of *have got to* in middle-aged and adolescent usage. *Need to* use increases then decreases and shows a strong relationship to middle-aged participants. This study therefore shows a clear change in progress in AusE favouring *have to* over *have got to*, contrasting with predictions of *have got to* as the new dominant form for expressing root modality (Krug 2000) and some projections for the direction of change in BrE (Tagliamonte 2012).

Constraining linguistic factors investigated included obligation type (subjective/objective), and the type of subject (including type of reference). The findings place AusE in line with Tagliamonte and Smith's (2006) study of BrE regional varieties in that *have to* more often occurs in cases of objective obligation whereas *must* prefers subjective obligation. *Need to* also favours subjective readings, and is thus more directly replacing the true modal form.<sup>13</sup> In terms of subject via grammatical person, our decision to follow Hoffman and Walker (2014), and Walker and Hoffman (2016) proved to be fruitful in revealing a significant relationship between generic pronouns and *got to* and *have got to*, and the inverse relationship for *have to* and noun phrases. This demonstrates the interconnection of linguistic constraints regarding definite/generic reference and obligation type.

We have followed and directly compared our results to earlier studies in the hope of presenting a clearer big picture for the changes in (quasi-)modal use in relation to obligation/necessity. We hope that other authors will build on this, finding the potential different trajectories and possible endpoints of this large-scale change across Englishes, which may differently value social influences such as colloquialisation and democratisation. While many varieties of English, including those beyond large, old and "native-speaker" varieties have been the focus of corpus studies on this variation, it is difficult to assess how AusE exactly fits into a picture of broader change without more sociolinguistic studies to understand directions of change and social and linguistic conditioning. These explorations

---

<sup>13</sup> We thank an anonymous reviewer for highlighting the importance of this and encouraging us to consider it further.

are further revealing for their insights into processes of grammaticalization and specialisation. We recognise a number of limitations on this research, including that sample size made some analyses rather speculative. Future research should also consider the use of individuals, not able to be modelled here due to the high ratio of participants to tokens and, as we identified, check for the impact of constructed dialogue in the data which may, like formulaic expressions, give an inaccurate account of use. Despite its limitations, this paper has provided new insights into an area of grammar that has been vigorously studied in multiple varieties of English but had not received variationist attention in AusE, where sociolinguistic studies of morphosyntactic features are still rare. It further identifies points of interest and comparison for future explorations into the changing use of (quasi-)modals of obligation and necessity in Englishes, and assessing how much diversity exists.

## Acknowledgements

The authors would like to acknowledge and thank La Trobe University's Discipline Research Program Grant Scheme for its financial support of this project (CPW) and the Research Unit for Variation, Contacts and Change at the University of Helsinki for financial support during material collection and data analysis (MK). We also thank James Walker for reading and commenting on a draft of this paper.

## References

- Baker, Paul. 2011. "Times May Change, but We Will Always Have Money: Diachronic Variation in Recent British English". *Journal of English Linguistics* 39: 65–88.
- Close, Joanne, and Bas Aarts. 2010. "Current Change in the Modal System of English: A Case Study of *Must*, *Have to* and *Have Got To*". In Ursula Lenker, Judith Hurber and Robert Mailhammer, eds. *English Historical Linguistics 2008: Volume 1 the History of English Verbal and Nominal Constructions*. Amsterdam: John Benjamins, 165–182.
- Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries*. London: Croom-Helm/Routledge.
- Collins, Peter C. 1991. "The Modals of Obligation and Necessity in Australian English". In Karin Aijmer and Bengt Altenberg, eds. *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman, 145–165.
- Collins, Peter C. 2005. "The Modals and Quasi-Modals of Obligation and Necessity in Australian English and Other Englishes". *English World-Wide* 26: 249–273.
- Collins, Peter C. 2007. "Modal Wars: Some Ascendant Semi-Modals in Australian English". In Jeff Siegel, John Lynch and Diana Eades, eds. *Language Description, History and Development: Linguistic Indulgence in Memory of Terry Crowley*. Amsterdam: John Benjamins 197–208.
- Collins, Peter C. 2009a. *Modals and Quasi-Modals in English*. Amsterdam: Rodopi.
- Collins, Peter C. 2009b. "Modals and Quasi-Modals". In Peter C. Collins, Pam Peters and Adam Smith, eds. *Comparative Studies in Australian and New Zealand English*. Amsterdam: John Benjamins, 73–88.
- Collins, Peter C. 2014. "Quasi-Modals and Modals in Australian English Fiction 1800-1999, with Comparisons across British and American English". *Journal of English Linguistics* 42: 7–30.
- Denis, Derek, and Alexandra D'Arcy. 2019. "Deriving Homogeneity in a Settler Colonial Variety of English". *American Speech* 94: 223–258.
- Depraetere, Ilse, and Susan Reed. 2006. "Mood and Modality in English". In Bas Aarts and April McMahon, eds. *Handbook of English Linguistics*. Oxford: Blackwell, 269–290.

- Downes, William. 1998. *Language and Society* (2nd ed.). Cambridge: Cambridge University Press.
- Fehringer, Carol, and Karen P. Corrigan. 2015. "'You've Got to Sort of Eh Hoy the Geordie Out': Modals of Obligation and Necessity in Fifty Years of Tyneside English". *English Language and Linguistics* 19: 355–381.
- Goffman, Erving. 1981. "Footing". In *Forms of Talk*. Philadelphia: University of Pennsylvania Press, 124–159.
- Hansen, Beke. 2018. *Corpus Linguistics and Sociolinguistics: A Study of Variation and Change in the Modal Systems of World Englishes*. Leiden: Brill.
- Hoffman, Michol F., and James A. Walker. 2014. "'Toronto Has Everything', 'Toronto's Got It All': Ethnolinguistic Dimensions of *Have* in Toronto English". Paper presented at New Ways of Analyzing Variation 43, University of Illinois at Urbana–Champaign and University of Illinois at Chicago.
- Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Jankowski, Bridget. 2004. "A Transatlantic Perspective of Variation and Change in English Deontic Modality". *Toronto Working Papers in Linguistics* 23: 85–113.
- Johnson, Daniel Ezra. 2017. Rbrul [Computer Software] Version 3.1.1, <<http://www.danielezrajohnson.com/rbrul.html>>.
- Krug, Manfred G. 2000. *Emerging English Modals: A Corpus-Based Study of Grammaticalization*. Berlin: Mouton de Gruyter.
- Labov, William. 1963. "The Social Motivation of a Sound Change". *Word* 19: 273–309.
- Labov, William. 2001. *Principles of Linguistic Change: Social Factors*. Malden: Blackwell.
- Leech, Geoffrey N. 2013. "Where Have All the Modals Gone? An Essay on the Declining Frequency of Core Modal Auxiliaries in Recent Standard English". In Juana I Marín-Arrese, Marta Carretero, Jorge Arús Hita and Johan van der Auwera, eds. *English Modality: Core, Periphery and Evidentiality*. Berlin: Walter de Gruyter, 95–115.
- Leech, Geoffrey N., Marianne Hundt, Christian Mair, and Nicholas Smith. 2009. "The So-Called Semi-Modals". In *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press, 91–117.
- Lowry, Richard. 2011–2017. Vassar Stats [Computer Software]. Vassar College, Poughkeepsie, <<http://vassarstats.net/>>.
- Mair, Christian. 2014. "Do We Got a Difference? Divergent Developments of Semi-Auxiliary (Have) Got (to) in British and American English". In Marianne Hundt, ed. *Late Modern English Syntax*. Cambridge: Cambridge University Press, 56–76.
- Nokkonen, Soili. 2010. "'How Many Taxis There Needs to Be?' The Sociolinguistic Variation of Need to in Spoken British English". *Corpora* 5: 45–74.
- Palmer, Frank. 1990. *Modality and the English Modals* (2nd ed). London: Longman.
- Penry Williams, Cara. 2018. "The Analytic Importance of Metapragmatics in Variationist Methodology: Voicings within Australian English". Paper presented at Sociolinguistics Symposium 22, University of Auckland.

- Penry Williams, Cara. 2019a. "Appeals to Semiotic Registers in Ethno-Metapragmatic Accounts of Variation". *Journal of Linguistic Anthropology* 29: 294–313.
- Penry Williams, Cara. 2019b. *Folklinguistics and Social Meaning in Australian English*. London: Routledge.
- Penry Williams, Cara, and Minna Korhonen. 2018. "Variation and Change in Possession and Root Modality in Australian English: The Declining Fortunes of *Have Got*". Paper presented at Sociolinguistics Symposium 22, University of Auckland.
- Smith, Nicholas. 2003. "Changes in the Modals and Semi-Modals of Strong Obligation and Epistemic Necessity in Recent British English". In Roberta Facchinetti, Frank Palmer and Manfred Krug, eds. *Modality in Contemporary English*. Berlin: Mouton de Gruyter, 241–266.
- Tagliamonte, Sali A. 2004. "Have to, Gotta, Must?: Grammaticalisation, Variation and Specialization in English Deontic Modality". In Hans Lindquist and Christian Mair, eds. *Corpus Approaches to Grammaticalization in English*. Philadelphia: John Benjamins, 33–55.
- Tagliamonte, Sali A. 2012. *Roots of English: Exploring the History of Dialects*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali A., and Alexandra D'Arcy. 2007. "The Modals of Obligation/Necessity in Canadian Perspective". *English World-Wide* 28: 47–87.
- Tagliamonte, Sali A., and Jennifer Smith. 2006. "Layering, Competition and a Twist of Fate: Deontic Modality in Dialects of English". *Diachronica* 23: 341–380.
- Tannen, Deborah. 1986. "Introducing Constructed Dialogue in Greek and American Conversational and Literary Narrative". In Florian Coulmas, ed. *Direct and Indirect Speech*. Berlin: Mouton de Gruyter, 311–332.
- Trousdale, Graeme. 2003. "Simplification and Redistribution: An Account of Modal Verb Usage in Tyneside English". *English World-Wide* 24: 271–284.
- Walker, James A., and Michol F. Hoffman. 2016. "Ethnolinguistic Variation in Toronto English: Possession and Deontic Modality". Paper presented at the Annual meeting of American Linguistics Society, Marriott Marquis Washington, DC.