Layering, competition and a twist of fate

Deontic modality in dialects of English*

Sali Tagliamonte and Jennifer Smith University of Toronto / University of Glasgow

This paper examines an area of ongoing change in English — deontic modality — and uses an archive of synchronic dialect data from England, Scotland and Northern Ireland to discover new information about its development. History records a cline in this system from *must* to *have to* to *have got to*. By taking a cross-dialectal perspective and utilizing comparative sociolinguistic methods we present a possible reconstruction of the later steps in this process. The results reveal dialectal contrasts in the proportion of older and newer forms, but similar patterns of use. *Must* is obsolescent and there is an unanticipated resurgence of *have to* alongside pan-dialectal grammatical reorganization: (1) *have to* is being used in contexts traditionally encoded by *must* and (2) *have got to* is specializing for indefinite reference. Young women are the leading edge in these developments suggesting that systemic adjustments in grammar combine with sociolinguistic influences to advance linguistic change.

Keywords: language change, comparative sociolinguistics, deontic modality, grammaticalization, dialect

Introduction

One of the fundamental axioms of language change as well as an "essential ingredient of most work in historical linguistics" (Hopper & Traugott 1993: 38)

^{*} The first author gratefully acknowledges the longtime support of the Economic and Social Research Council of the United Kingdom (the ESRC) for research grants between 1996–2003, #R000221842, #R000238287, #R000239097. We extend our gratitude to Anthony Warner and Karen Corrigan for many insightful discussions about English modals. Laurel Brinton, Elly van Gelderen, and the editor provided their expertise in the review and revision process which substantially improved the final product. To them all, our hearty thanks!

is the Uniformitarian Principle — the idea that: "knowledge of processes that operated in the past can be inferred by observing ongoing processes in the present." (Christy 1983: ix). Contemporary dialects offer an important adjunct to this, particularly those spoken in isolated communities. Such communities, because of their peripheral geographic location or isolated social and/or political circumstances, tend to preserve features typical of earlier stages in the history of a language. They are essentially relic areas as far as the progress of linguistic change is concerned (Anttila 1989: 294, Hock 1986: 442) and their use in tracking historical change follows from a long tradition begun in Germany and continued by dialectologists in the 20th century (Kurath 1949, Orton & Halliday 1963). In this paper, we take up this research practise by (1) focusing on an area of continuing change in English grammar — deontic modality, as in (1) — and (2) considering the distribution and patterning of forms used for this function in eight dialects in the British Isles.

(1) No, you have to do something ... you've got to do something. (BCK:3)

The communities in which the dialects are spoken range from highly isolated and conservative to relatively urban and proximate to the mainstream. Our aim is to use these materials to tap the underlying cycle of loss and renewal in the deontic modality system. We speculate that this cross-dialectal perspective will encompass (some of) the "later recorded steps" in the historical change (Whitney 1867, quoted in Christy 1983: 84). In essence, deontic modality in synchronic dialects offers a window into diachrony.

Dialects

The dialects in our archive come from the United Kingdoms of Great Britain and Northern Ireland (Tagliamonte forthcoming), as shown on the map in Figure 1.

Each of these eight communities has evolved in varying conditions of cultural and geographic isolation. Thus, for example the small fishing villages in Northern Ireland, Culleybackey and Portavogie and North East Scotland, Buckie, represent the most remote locales with dialects isolated from mainstream developments. In contrast the small city of York in North Yorkshire, within commuting distance of the metropolitan areas of Manchester and Leeds, hosts a dialect that is taken to mirror a relatively standard variety of northern British English (see Tagliamonte & Smith 2002, Tagliamonte, Smith & Lawrence 2005). These materials give us access to 155 speakers of the oldest



Figure 1. Communities in Great Britain and Northern Ireland

living generation in each community circa 1997–2000 and provide contrasting contexts for loss and renewal in recent and ongoing grammatical change in Britain.1

Layers

The dialects exhibit the full range of old and new forms of deontic modality, including *must*, *have to, have got to* and *got to* as in (2–5).²

- (2) a. If you join the club, you must go to church. (MPT:s)³ b. *Next time I'm in the doctors I must ask to see the physio.* (TIV:d)
- (3) a. They have to keep up with the Jones's now. (MPT:y)
 - b. *It has to be natural to work.* (BCK:q)
- You're told you've got to speak properly. (CLB:a) (4) a.
- 1. For further details on these materials see Tagliamonte (2005) and Smith (2000).
- 2. Unless otherwise indicated our use of these terms include full, enclitic and phonologically reduced variants.
- 3. Each community is represented by a three letter code as follows: CMK=Cumnock, Scotland, TIV=Tiverton, Devon, MPT = Maryport, England, BCK = Buckie, WHL = Wheatley Hill, NI (Northern Ireland) = Portavogie [PVG] and Cullybackey [CLB], YRK = York, HEN = Henfield, Sussex. The single digit codes which appear directly after the community code are speaker codes which identify the individuals in each corpus.

- b. You've got to have a vice of some kind. (CMK:v)
- (5) a. You got to be careful. (YRK:H)
 - The plants got to drag it out. (TIV:h)

Such variability in the grammar can be interpreted as layering where different forms reflect grammatical change (Hopper 1991: 23, 124). On the other hand, variation of this type may simply be the result of functional distinctions where different forms reflect unique functions. In other words, the distribution and patterning of these forms in contemporary dialects may provide us with the ability to assess whether innovation, obsolescence or stability best explains the facts.

Constraints

In order to assess the grammatical function(s) of forms and their status in the community, we examine the effects of linguistic features associated with one form over the other as the linguistic change progresses. We then correlate these contextual factors with the different variants in the data using quantitative techniques such as distributional analysis to assess their patterning. Analysis using multiple regression techniques can assess the direction of effect of these patterns, their significance and relative importance when all factors are considered simultaneously. The comparative method is then used to assess similarities and differences across age groups in the community. For further discussion, see Poplack & Tagliamonte (2001:ch. 5), Tagliamonte (2002).

The information that can be gleaned from cross-variety comparative analysis sheds light on community-based, regional norms, and situates linguistic change at a particular point in its trajectory. We argue that the constraints in variation provide an indication of the contemporary status of the grammatical system (Poplack & Tagliamonte 2001:95-100). Moreover, internal linguistic constraints (or correlations) on variable forms can be used to infer the underlying grammar of variant choice, and in particular, the point of development of the areas of grammar in its trajectory of change (see also Poplack & Tagliamonte 1999, Tagliamonte 2003). Finally, our analysis proffers an example of the utility of variationist sociolinguistic methods in the analysis and interpretation of linguistic patterns and the role it serves in their evaluation (Poplack & Tagliamonte 2001:88–112).

We now turn to consider in greater detail the complex and unsettled system of deontic modality in contemporary English.

Deontic modality: Form and function

Deontic modality encompasses a range of meanings including obligation, permission and necessity (Coates 1983: 32).4 Obligation "reports the existence of external social conditions compelling an agent to complete the predicate action" (Bybee, Perkins & Pagliuca 1994: 186). However, the overall system is really a 'cline' ranging from readings which might translate as "it is imperative or important that ...", to those which mean something more like, "it is necessary or a requirement that ..." (Coates 1983: 32).

Within this system, the various forms are said to encode the strength of the obligation. Must is thought to embody the strongest obligation since its use is "directly applied and irresistible" whereas have to is distinguished by the fact that it is "resistible" under certain circumstances, as in (6) (quoted in Hopper & Traugott 1993: 79, Sweetser 1988: 54).

(6) I have to! must get this paper in, but I guess I'll go to the movies instead.

In other words, differences in meaning in the system relate to gradations in the strength of the obligation. For example Bybee writes (1994:186): "If a weak obligation is not fulfilled, the consequences are not too serious; but the consequences of not fulfilling a strong obligation are much more severe."

Thus discussions of the deontic system typically associate one form or another to a particular reading: must is thought to be strong; while the other forms are weak. This contrast has also been conceived as 'core' vs. 'periphery' (Coates 1983:32). Indeed, most accounts group have to and have got to together, as toned down choices in contrast to must, distinguishing neither subtle meaning differences nor contrastive strength to the choice between them (e.g., Huddleston & Pullum 2002:183).

Unfortunately, a strong-weak distinction, either as contrastive or as a continuum, is virtually impossible to categorize impartially. To do so inevitably leads to circularity from the imposition of the analyst's own subjective interpretations. The main point of relevance here is that there is a cline of meanings and a range of intensities encompassed by deontic modality. Moreover, as we shall see, the variants *have to* and *have got to* represent the vast majority of uses in dialect data. This suggests, at the outset, that there is something more going on than a contrast (functional or otherwise) between must on the one hand, and have to grouped with have got to on the other.

^{4.} Coates (1983: 32) uses the term "root modality" to refer to the basic meaning 'it is necessary for'.

First, however, it is important to contextualize where the extant forms and functions have come from. Thus, we turn to the recent history of deontic modality in the history of English.

Diachronic development

Table 1 briefly summarizes the historical trajectory of the deontic modality system of English (abstracted from e.g. Biber et al. 1999:487, Brinton 1991, Jespersen 1913, Plank 1984, van der Gaaf 1931, Visser 1963-73, Warner 1993).

Table 1. Summary	of the development	t of DEONTIC MODALITY
------------------	--------------------	-----------------------

Old English	Middle English	Early Modern English	19th century	20th century
mot: permission or possibility	mot → must: develops fuller range of senses, including DEONTIC and	must	must	must: "old, established"
	EPISTEMIC		have to	have to
	permission reading lost			have got to got to considered
	have to 1st attestations	have to Fully established	have got to got to 1st attestations	informal colloquial
			considered colloquial	gotta appears
				considered a vulgar pronunciation

Must

Must has been around since Old English when its form was mot.⁵ It had originally been a preterit-present verb; however, it came to express permission and possibility (Traugott 1999, Warner 1993:160-161). By the Middle English

^{5.} In more northern areas, including Scotland, the form mon (and maun from around 1450) was used instead of must (OED s.v. maun).

period, the permission uses became restricted and a wider range of meanings developed, including deontic readings, as in (7).

(7) Oon of vs two moste bowen doutelees. (c1395 Chaucer Wife of Bath's Tale 440) "One of us two must submit, doubtless"

By the end of the Middle English period must replaced mot⁶ and the permission reading of must was lost (e.g. Denison 1993:303). From that point forward must functions as the English modal of obligation, compulsion or necessity which can be interpreted as "imposed from without, either by circumstances, regulations, legal prescriptions, etc., or by the will of a person" (Visser 1963-73:1805). From the last decades of Middle English into Early Modern English a number of idiomatic uses develop, including sayings such as I must say, I must confess, etc. and the expression of satire or indignation (Visser 1963–73: 1809-1810), as in (8).

(8) She must always be doing something extraordinary. (George Eliot, Daniela Deronda I, I, Ch., &, O, 1874-1876)

Epistemic readings, as in (9) which dates from the late 14th century, are said to have arisen from these deontic readings (Traugott 1999:2-4, Warner 1993:180).⁷ The developmental trajectory of epistemic modality out of an original deontic modality is a common path in grammatical change, both in English and crosslinguistically (Bybee, Perkins & Pagliuca 1994, Krug 2000, Traugott 1997).

(9) Al most we deye; therto so lete vs lowt. (c1450 C. D'Orelans Poems (OED s.v. must)

"We must all die, so let us submit to it"

Have to

There has been ample discussion of the dating of the "obligation" meaning of have to in the literature (e.g. Fischer 1994, van der Gaaf 1931, Visser 1963-73: 1476–78) with conflicting views resulting from semantic ambiguity form (e.g.

^{6.} Some scholars argue that the present day English modal semantic system was established by Shakespeare's time (Ehram 1966:97). Others argue that there has been "a continued semantic focussing of modality" in which "the modals continue to lose past-referring uses of their preterit forms and there is continued reduction in the relevance of subject-oriented uses" (Warner 1993:181). The latter refers to uses involving the speaker's attitude or evaluation.

^{7.} In contemporary northern varieties *must* is used primarily as an epistemic marker (Miller 1993; Trousdale 2003).

Brinton 1991:12). While some scholars believe this function was present in Old English (Plank 1984:320), others argue against this (Mitchell 1985:950-953). Most place the development of deontic meaning in Middle English, as in (10-11) (e.g. van der Gaaf 1931), although its frequency is said to have been rare (Visser 1963-73:1478).

- (10) *Gramaire ferst hath forto* teach to speke upon congruite. (Gower, CA, 7.1530, c.1393) (cited in Denison 1993:317) "Grammar first has to teach to speak with correctness"
- (11) Every man ... hath greatly to marvel that any man would say the loss of two so noble duchies ... is but trespass. (Paston Letters, No 191 I. 260, c1465) (cited in Crowell 1955)

Putting this timeframe in context with the development of obligative *must*, which is also in place by Middle English, we can assume that have to and must competed for this function from this point onwards, as clearly indicated by one of Brinton's (1991) example in (12).

(12) I moot go thider as I have to go. "I must go thither as I have to go" (Chaucer, Canterbury Tales Pardd.C. 749, 1386–1400)

Have got to

The construction have got to, as in (12a) entered the English language much later — not until the 19th century (Biber et al. 1999:487, Visser 1963–73:1479), a full millennium later (Krug 2000:53). The variant with got to by itself, as in (12b), appears around the same time. Once again, this suggests a time period of intensive layering of forms, now with a third layer.

- (12) a. The first thing **I've got to** do is to grow to my right size again. (Lewis Carroll, Alice in Wonderland iv. 54, 1865) (from Visser 1963-73:1479)
 - b. All you got to do is to wait. (Mark Twain, Tom Sawyer, ch. 1, 1876)

Later, orthographic variants of *got to*, i.e. *gotta*, appear, as in (13) (e.g. Denison 1998: 173).

- (13) a. They're wide open, an' all you gotta do is hit 'em. (Jack London. Valley of the Moon, 1911)
 - b. He ... went forward as if to take her arm. "You gotta come along", I heard. Buchan, Three Hostages xviii. 2663, 1924) (OED gotta)

What underlying linguistic processes were happening during the course of this history? We now turn to a review of what distinguishes these constructions grammatically.

Structural change

Perhaps one of the most interesting components of the evolution of deontic modality is that alongside the renewal of forms, the more recent layers, i.e. have to and have got to, involve grammaticalization(e.g. Brinton 1981, Mitchell 1985:401, van der Gaaf 1931:180-182, Visser 1963-73:1477-1478). In the next section we provide a brief overview of the structural implications in order to provide a backdrop for interpreting and understanding ongoing linguistic change.

Have to

The form have to developed its modal obligative meaning out of an original lexical meaning of possession. This longitudinal process is a complex and debated change in the history of English (e.g. Brinton 1991, Fischer 1992, 1994, Krug 2000, Lightfoot 1979, van der Gaaf 1931, van Kemenade 1999, Warner 1993) but is said to involve the semantic and syntactic reanalysis of the construction have+object+to+V to have to V+object (e.g. van der Gaaf 1931, Visser 1963-73). According to van der Gaaf (1931:841) have to transforms into a modal auxiliary "when it expressed nothing but duty, obligation, compulsion, necessity, gradually had the place assigned to it occupied by auxiliaries, namely before the infinitive, while at the same time the object began to be placed after the infinitive".

However, as Brinton (1991) points out "the morphosyntactic status of have to is considerably less clear than that of must". For some researchers, it is an auxiliary (van der Gaaf 1931:184, Visser 1963-73:1478), while others claim that it is at best a 'quasi- or semi-auxiliary' (Bolinger 1980, Coates 1983:54, Quirk et al. 1985:137). This is because its defining characteristics fall between auxiliary and full verb. It is like a full verb in that it has non-finite forms of a full verb, it can occur with other modals (may have to ... etc.) and it has 3rd person present forms. These characteristics allow it to "provide forms missing in the paradigm of the semantically similar but morphologically defective must" (Brinton 1991:5). However, it can undergo phonological reduction including to contraction: This is only found in a small number of forms, all of which have modal or aspectual meaning. According to Krug, the first mention of the contraction of *have to* is from the mid 20th century American English (Francis 1958:258, cited in Krug 2000:57), which, he argues provides evidence for the bondedness of have to and thus that it functions as a "unit lexeme" (Heine 1993:42). The phonological reduction also allows have to to fit into the modal auxiliary slot in its contracted form.

Despite the conflicting evidence on the status of *have to* (Brinton 1991:6) "most would agree that have to is a least partially grammaticalized in Modern English".

Have got to

The introduction of have to go for the same semantic function brought yet another form (and complexity) into the mix. In comparison with have to, have got to appears to have entered the language directly without the complex semantic and structural reanalysis with have to. Krug (2000:73) argues that "speakers used enclitic have followed by got to from the start synonymously with modal have to". Have got to exhibits more of the formal properties of modal auxiliaries when compared to have to: It has operator properties, it cannot occur with other modal auxiliaries and it has no non-tensed forms. However, have got to is biclausal, making it structurally different to the monoclausal must and contracted have to/hafta.

Got to/gotta

The newest variant, got to, is said to most "resemble a single modal auxiliary" (see also Krug 2000, Quirk et al. 1985:142). As with must and other modals, it lacks inflections, non-finite forms, imperative constructions and passive, negation is avoided and it also undergoes to-contraction (Bolinger 1980:107, Curme 1931:360-361, see Table 3.13, Krug 2000:107, Visser 1963-73:2205-2206). Indeed, Krug (2000:109) argues that "more auxiliary criteria apply to gotta than to its historical predecessor have got to".

Summary

Thus, while *must* has long been a central modal of English (e.g. Biber et al. 1999:483-484), the newcomers, have to and have got to are not quite there yet: They are typically considered semi-modals, quasi modals or periphrastic modals. In other words, neither of them currently exhibit all the defining characteristics of a 'modal'. In fact, the incoming variants for deontic modality can be viewed as representing a range of auxiliary-like qualities. As we shall see, this may be a factor in the current evolution of the deontic modality system.

This scenario of longterm evolution of forms for the same function, their contrasting morphosyntactic classifications alongside a documented cycle of loss and renewal presents an interesting case study. First, because the major variants (must \rightarrow have to \rightarrow have got to \rightarrow got to) entered the language at different points in time, their distribution across dialects may shed light on the stages of development of the deontic modality system. At the same time the forms used to express deontic modality have varying degrees of auxiliary-hood ranging from *must* which is a full-fledged modal to the other contenders which each varying degrees of this status. Tracking the synchronic status of this system — form and function — across dialects should add to the existing knowledge base on this system. This, in turn, should inform us of the nature of this area of English grammar as well as its status in the larger history of its development (Bybee, Perkins & Pagliuca 1994).

However, there is still another facet of the forms used for deontic modality that we must consider.

Social evaluation

The newer layer of forms used for deontic modality, i.e. have got to and got to, have long been infused with social meaning. Krug (2000) argues that have got to originated in nonstandard speech. Both Visser and the OED label it as colloquial, along with got to, while gotta has the added characterization as "vulgar". Present-day English grammars still consider have got to and got to "informal", equating them primarily with spoken data (Biber et al. 1999:487-489, Palmer 1979). In contrast, must is associated with formal registers and written exposition (Biber et al. 1999).

Contemporary analysts also report that there are distinct regional associations for the forms: have got to is said to be the most common form in spoken British English (Coates 1983), while in American English got to and gotta are typically thought to be "standard" (Denison 1998:173). Have got to and got to have been found to be increasing dramatically in frequency in contemporary British and American English, so much so in fact that Krug (2000:63) regards

the change towards have got as a riveting success story.8 Because the forms have their own sociolinguistic interpretations as well as a distinct regional diffusion, their distribution may shed light on social influences in language change.

Now with the diachronic and synchronic picture in perspective and the structural and sociolinguistic aspects in mind, we turn to the dialect data.

Methodology

The corpora were exhaustively searched for every instance of must, have to, have got to, got to which encoded deontic meaning, i.e. "it is imperative/necessary for ..." (Coates 1983). In ensuring a strict form/function correspondence for the analyses, cognate forms with divergent functions were excluded, as we detail below.

Epistemic modality,⁹ as in (14), which encodes inferred certainty, was explicitly excluded from this analysis of deontic modality. This is a later development which has largely been grammaticalized to must in contemporary English (Tagliamonte 2004).

- (14) a. I says there **must** be a better life than this somewhere. (CMK:l)
 - And he must care for the sheep too because he always names them in the bible. (CLB:e)

Epistemic readings may sometimes be encoded with have to and have got to, as in (15); however in these dialects they are rare. 10

- (15) a. There has to be something more than this. Is there somewhere that we're missing? (YRK:d)
 - I says 'Thou's got to be Dearham lass'. (MPT:a)

- 9. Biber et al. (1999:485) use the terms 'extrinsic' for epistemic modality and 'intrinsic' for deontic modality.
- 10. There were a total of 106 contexts of epistemic modality with must, have to, have got to or got to; 96% were lexicalized with must.

^{8.} Most of Krug's data come from written sources: The British National Corpus, the Brown Corpus which consists of American English texts from 1961, the LOB corpus, which is the Lancaster/Oslo-Bergen Corpus, a similarly constituted corpus of British English. He also examines the 'Frown' and 'FLOB' corpora, short forms for the Freiburg versions of Brown and LOB, which match their predecessors but are texts from 1991-1992. He also considers the relatively smaller spoken section of the British National Corpus (for further discussion see Krug 2000:31-36).

The preference for *must* with inferred necessity readings is a distributional pattern reported by Coates (1983:48) and Palmer (1979:53) for standard contemporary (British) English, and for York English (Tagliamonte 2004).

Contexts involving past or future tense, as in (16) were excluded. Consistent with descriptions of the English modal system these were categorically lexicalized with have to (Palmer 1979:114, Tagliamonte 2004).11

- (16) a. So, I says, I'll have to look for a job. (MPT:a)
 - You'll have to go to Shoreham. (DVN:b)

Lexicalized expressions such as in (17) were also excluded as they had heightened use of *must* (71%, N=31).¹²

- (17) a. I haven't watched a lot I must admit. (CMK:o)
 - b. *Oh, monotonous job, I must say.* (MPT:n)
 - I had agorophobia, **I must tell you**. I couldn't go on holidays. (YRK:c)

We also removed infrequent tokens of negation and questions for the simple reason that they were rare [N=19, N=12]. 13

Thus, we circumscribe our analysis of deontic modality to its most productive area, which as it happens is highly circumscribed in the grammar — affirmative declarative contexts in which the forms of must, have to and have got to may be used to express obligation/necessity. In total, the data provided 602 contexts. In the next section we turn to a quantitative analysis.

Distributional analysis

Table 2 provides the overall distribution of forms. The most frequent form used for deontic modality in these dialects is have to, representing a full 44.7% of

^{11.} See Palmer (1979:114) "have got to has no non-finite forms ... have to must be used".

^{12.} When these constructions occur in initial position they are potentially ambiguous as to whether they are formulaic or retain their full lexical content. However, they were in the minority and could be disambiguated on other cues in the data. See Krug (2000:97-102) for a discussion of how such expressions may have contributed to the grammaticalization of have to, in that have to generalized from verbs of saying to all verbs. In these data 4/14 of the say formulae were have to, 4 were 've/'s got to and 5 were must.

^{13.} The negated forms were divided between mustn't (N=6) and don't have to (N=7), haven't to (N=3), and haven't got to (N=3). The questions were have to (N=7) and have got to (N=5).

101AL N = 002									
must		have to	have to		have/'ve/'s got to		got to/gotta		
%	N	%	N	%	N	%	N		
10	60	44.7	269	35.9	216	9.5	57		

Table 2. Overall distribution of variants of DEONTIC MODALITY

the data. Thus, despite the perceived notion that *have got to* is the most common form for expressing obligation/necessity (Bybee, Perkins & Pagliuca 1994, Coates 1983, Krug 2000), this overall view shows that *have to* is predominant in these dialects. Note the marginal status of *must*: It is only used 10% of the time. *Got to* is also very infrequent at only 9.5%. Perhaps the most striking result however, is how robust the variation between *have to* and *have got to* really is — 44.7% *have to*; 35.9% *have got to*.

However, this view of the data mixes eight different dialects and three countries. If dialect data can reveal the tracks of ongoing change, we might find meaningful differentiation amongst them. Figure 2 separates the data so that the frequency of each surface morphological construction is shown as a proportion of the total number of deontic modality contexts in each community. This reveals how each dialect patterns individually with respect to the variants. ¹⁴

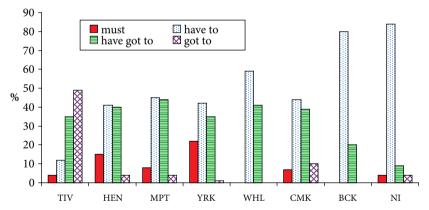


Figure 2. Distribution of variants of forms for DEONTIC MODALITY across communities.

^{14.} The numbers and proportions can be found in Table 5 in the Appendix. The data from Cullybackey and Portavogie have been combined. They are relatively smaller data sets and were found to pattern similarly in this area of grammar as well as in others (see also Tagliamonte forthcoming; Tagliamonte & Smith 2002). This is not surprising since the populations in both locales have a common ancestry. Both are descended from contact varieties where the superstrates were Scots vernaculars.

First, the rarity of must (the shaded columns) is visible in all the communities with slightly elevated proportions in Henfield, where it is used 15% of the time [14/95] and York, at 22% [22/100], where it is still a relatively minor variant. Second, note the distribution of forms in Tiverton which are quite unique. Here, the variant got to is the most frequent variant used 49% of the time [33/68]. Third, two varieties distinguish themselves with an overwhelming preference for the variant have to — Buckie [20/25, 80%] and the northern Irish communities [47/56, 84%]. We can isolate the robust variation between have to (the dotted columns) and have got to (the striped columns) to Henfield, Maryport, York Wheatley Hill and Cumnock.

The challenge now becomes how to interpret these synchronic similarities and differences in the frequency of forms from an informed diachronic and synchronic perspective.

Must

The virtual absence of must in Northern Ireland and Scotland confirm Macafee's (1992) claim that 'Scottish speakers avoid must in the sense of obligation, reserving it to express logical necessity'. In fact, its absence even in the most relic community, Buckie, leads us to suggest that deontic must may never have been present in these northern dialects in the first place.¹⁵

More generally, however there is a documented decline in the use of deontic must. Indeed, the semi-modals (have to and have got to) are becoming more frequent and better established in English more generally (Biber et al. 1999:489, Collins & Blair 1989, Leech 2003, Smith 2003). A number of other findings are supportive of the interpretation of ongoing obsolescence for *must*. Recall that formulaic expressions, as in (18) were excluded earlier on due to the fact that these were disproportionately rendered with must (71%) (behaving antithetically to the other contexts). This is typical of obsolescing features — they often get left behind in islands in the grammar as discourse 'chunks'. In this case, these islands are comparable across the board, suggesting that deontic must has retreated to highly circumscribed contexts across dialects (see also Trousdale 2003). Further, the only communities to use deontic *must* to any degree are Henfield and York. These are the only two communities with sizable numbers of middle class speakers (as represented by education and professional level occupations) (Tagliamonte 2003). This is consistent with the idea that must has developed an association with formality. Whatever the reason,

^{15.} The earlier cognate *mon/maun*, found in these northern areas, not found in these data.

we conclude that deontic *must* is both linguistically and socially restricted in these spoken dialect materials, a good indication of its weakening status in the (synchronic) grammar.

Have to/have got to

Recall that the prevailing interpretation of descriptions of deontic modality in the literature would lead us to expect substantial use of have got to for deontic modality in British English (e.g. Biber et al. 1999:489). However, these materials reveal unexpectedly high rates of have to. Not surprisingly, the heightened frequencies of *have to* are in the Northern Irish communities and Buckie. This can be explained by the comparable socio-historical situations of these two locales. Buckie and the Northern Irish communities have been consistently shown to be the most conservative in our archive in terms of linguistic behaviour (e.g. Tagliamonte & Smith 2002, Tagliamonte, Smith & Lawrence 2005). Given that have to is older than have got to by some centuries, the high rate of have to in these communities may be a reflection of ongoing retention of an older layer in the development of forms. Indeed, these findings concur with Corrigan's (2000:37) Northern Ireland research which shows clearly that have to is the most frequent marker of modality overall. 16

Got to

Contemporary research suggests that the got to variant is among the forms currently on the rise in present day varieties of British English (Krug 1998, 2000). Thus, it is not surprising that elderly speakers in conservative varieties do not use a form that is considered to be newer in the system. However, the question then becomes: Why is it strong amongst the elderly Tiverton speakers?

Research on the southwest more generally has demonstrated considerable retention of conservative features (Jones & Tagliamonte 2004, Klemola 1996, van den Eynden 1993). Indeed, Tiverton retains many features from earlier

^{16.} The Northern Ireland data (Cullybacky and Portavogie) which comprise over 300,000 words contained not a single token of the modal be+to construction Corrigan (2000:29) reports for her South Armagh English data base, which contained 52,000 words from manuscripts compiled between 1942 and 1974. Of course this area is an entirely different sociocultural context. Evidence for the be to modal is found only in historical English sources, not in Scots which is the main difference between the communities in Corrigan's study compared to Cullybackey and Portavogie which have Scots roots. These differences may account for the discrepancies in forms used for modality.

stages in the history of English (Godfrey & Tagliamonte 1999). Thus, if got to is an innovation currently on the rise, its use here seems contradictory.

On the other hand, a general north/south divide in Great Britain is to be expected (Trudgill 1990). Along these lines, it is interesting to note that the geographic patterns revealed in these data mirror the split between southern and northern British dialects with respect to the use of the got to variant found in Krug's (2000:112, map 3.1) large scale study — where southwest dialects, Essex and Wales favour got to considerably more than the north. Interestingly, Henfield, positioned within the Home Counties, is the single geographic area in the south which has scant use of got to according to Krug's analysis. Thus, the use of got to in Tiverton may be a reflection of regional differences in Britain. However, this may not be a reflection of innovation vs. retention, as is so often the case with regional asymmetries. An alternative scenario is that the favoured form in the evolving layers of deontic modality (i.e. have got to or got to) is selected at the same point in the progression of development rather than consecutively as previously thought, as the frequent use of got to amongst these elderly speakers in Devon suggests. Supportive of this hypothesis is the historical record: Both *have got to* and *got to* are attested around the same time period, as demonstrated in (12), suggesting that they are equally old. Thus, while some research suggests that got to is newer than have got to and currently increasing (Krug 1998, 2000) we suggest that in Tiverton at least, it was selected as the favoured variant of the have got to layer from its inception.

Summary

In sum, the results from the distributional analysis reveal a wealth of information that gains a heightened focus from the cross-variety perspective. This information may be interpreted with a view to the history and development of the deontic modality system. The broad historical context documents a progressing cline from must to have to to have got to. These dialects lay out a reflection of that history in geographic relief.

We suggest that variation provides an indication of the state of development of the grammatical system. Where have to is more frequent, it can be taken as evidence that the variety is more conservative. Indeed, the two most remote communities exhibit the most have to — Northern Ireland and Buckie. In contrast, where have got to is more frequent, it can be taken as evidence that the new layer has made inroads into the system. Interestingly, variation between the two variants is found across dialects spanning Scotland (Cumnock), Northeast England (Maryport, York, Wheatley Hill) and Southeast England

(Henfield). In other words, we have caught the system in flux. Thus, for all intents and purposes, these communities represent the current state of the eldest generation of speakers of British English circa 1997-2002 for deontic modality — robust variation between have to and have got to with neither variant showing a clear lead.

The next question is: What is the underlying mechanism that may be guiding this development? Are the communities operating on their own accord, reflecting independent developments, parametric variation or dialectal divergence? Can an overarching 'drift' within the system be identified? Are supralocal patterns visible and if so, at what level of breadth?

Constraints analysis

Despite the semantic, pragmatic and stylistic facets of *must* that are emphasized in most treatments of deontic modality, the variability observed in the data hardly involves this form. Instead, there is vigorous competition between have to and have got to. Given that the current system has evolved from an earlier one in which *must* is assumed to have been the majority form, at least two hypotheses can be put forward to explain the current competitors. (1) they may have inherited the semantic and pragmatic functions of the earlier system; or (2) they may be differentiating the system along a new functional divide. Either way we may be able to view a contrast between have to and have got to with respect to whatever ongoing reorganization (restructuring) is underway.

In an attempt to answer these questions, we turn to consideration of a series of hypotheses about the choice between one variant or the other derived from the relevant literature on this system.

Type of subject

Recall that researchers describe deontic modality in terms of different degrees of 'strength' of the obligation/necessity reading. Because of the subjectivity of such assessments, we seek a grammatical correlate that will enable the analyst to operationalize this notion objectively. Fortunately, the encoding of grammatical person in English provides a reasonable facsimile. Coates (1983:37), for example, notes that: "It is generally true that examples with second person subjects are stronger than those with first person subjects, while examples with first person subjects are usually stronger than those with 3rd person subjects."

This means that we may be able to expose the relative strength of forms by categorizing each context in our data according to this hierarchy, i.e. 1st person, 2nd person and 3rd person subjects. However, in attempting to do so, we are confronted with a number of issues.

First, there are actually two types of 2nd person subject types. One is definite, as in (18) and one indefinite, generic, as in (19).

2nd person (definite):

- Brian, you've got to come up on the bridge. (MPT:n)
 - You have to go through my old scullery. (CLB:a) b.
 - Come on Davey, thou's to hear this. (CLB:f)

2nd person (indefinite):

- (19) a. You've go to take the blood when the fever's on. (MPT:@)
 - You *have to* believe in ghosts to ever see it. (YRK:%)
 - When you've got a man suddenly plunged into your life you've got to feed him, haven't you! (HEN:d)

In essence, of 2nd person subjects only those that are definite, as in (19), have the possibility of encoding a strong reading of "it is imperative that" Thus, these two types of 2nd person subjects must be kept separate in order to adequately test for whether 2nd person embodies strong obligation.

The same problem arises for 3rd person subjects, which may also be indefinite, as in (20).

- They have to keep up with the Jones' now. (MPT:y) (20)
 - Life has to go on. (CLB:e)

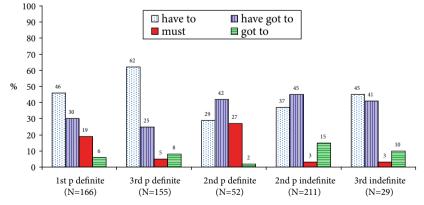


Figure 3. Distribution of forms for DEONTIC MODALITY according to type of subject overall.

Figure 3 displays the proportion of different forms for the varying types of subjects, distinguished by definiteness of the subject.

The figure shows that there are some contexts where must occurs quite frequently (i.e. 2nd person definite and 1st person definite). Indeed, were we to consider definite grammatical subjects only, Coates' observation would indeed be confirmed: 2nd person definite contexts have the highest proportion of must, then 1st person definite contexts, then 3rd person definite contexts, i.e. 27% > 19% > 5%. However, in none of these three key contexts is *must* the preferred choice. Indeed, it is a minor variant in all three. Thus, while must is clearly a hold out in certain contexts, once again its status as a receding form is apparent.

As far as the other variants are concerned, the following over-arching pattern emerges. The form *have to* is the most frequently used form in 1st person definite (46%) and 3rd person definite (62%) contexts. In contrast, 2nd person indefinite contexts show a preference for have got to (45%).

Thus, Figure 3 lays bare two issues: (1) rarity of the form must, despite frequent occurrence of contexts such as 1st and 3rd person definite which may be a prime location for it given their subjective classification and (2) a dramatic contrast in frequency of the different contexts for deontic modality in the data (i.e. 2nd person definite hardly occurs whereas 2nd person indefinite is healthy; 3rd person definite is strong but 3rd indefinite is rare). Therefore, we now focus in on the categories representing the largest proportions of the data — 1st person definite, 3rd person definite and 2nd person indefinite. Furthermore, we abstract away from the rare instances of must and focus in on the forms carrying out the vast majority of the system's functional load in the

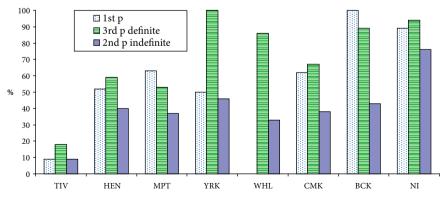


Figure 4. Distribution of have to for DEONTIC MODALITY according to type of subject across communities.

grammar — have to and have got to/got to. Figure 4 presents the findings separated by community.

The striking result visible in Figure 4 is how consistently 2nd person indefinite subjects (the darkly shaded columns) have a lower frequency of have to compared to the other contexts. Indeed, with this view of the data it now becomes clear that 2nd person indefinite is the more innovative context — as indicated by greater use of *have got to*. Moreover, this is consistent across all of the communities. This suggests that 2nd person indefinite contexts may have been the entry, or expansion-point, for have got to into the system. In contrast, 1st and 3rd person definite subjects preserve the older form, *have to* as in (21) and (22).

1st person:

- (21) a. I have to live with it. (MPT:%)
 - *I have to take pills every day.* (HEN:d)

3rd person:

- (22) a. *She's working till late 'cos she has to bide and count up all the money.*
 - When they're writing essays and things it's to be in English.(CMK:h)

In sum, by dividing the data according to this fundamental distinction, a conspicuous division in the use of have to and have got to has been revealed. Have to has become established in contexts where the out-going form must would have been predicted to occur (i.e. 1st and 3rd person definite contexts). In contrast, the newer layer of the system — have got to — is the frontrunner with 2nd person indefinite.

Subjective vs. objective

Another way of viewing the nuances within the English deontic modality system is pragmatic. Huddleston & Pullum (2002:183) invoke the distinction between subjective vs. objective obligation. Subjective obligation is considered prototypical deontic modality (Coates 1983:32). This is when the speaker imposes authority on his- or herself or on others, as in (23).

- (23) a. Excuse me Jennifer, I've got to watch the football. (CMK:L) [speaker imposes authority on themselves]
 - b. You must look after Amy. (MPT:h) [speaker imposes authority on another]

Objective contexts, on the other hand, arise when the authority comes from some other source, external to the speaker, as in (24).

- (24) a. *I've to sell my one now.* (BCK:n)
 - Aye, when I go and visit him I have to go outside. (CMK:A)

This type of obligation is most obvious in reports of rules and regulations, as in (25).

- When they sink a pit shaft, you've got to leave about a quarter of a (25) a. mile of solid coal round about it. (CMK:d)
 - b. You've got to take the blood when the fever is on. (MPT:@)

Figure 5a depicts the distribution of have to and have got to according to the contrast between subjective and objective obligation in the data across communities.17

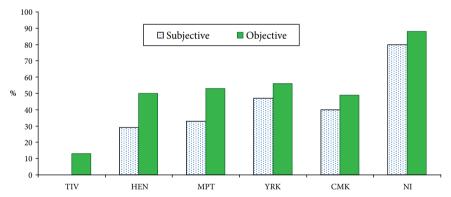


Figure 5a. Distribution of have to for DEONTIC MODALITY according to type of obligation across communities.

We observe that *have to* is more frequent for objective readings than for subjective readings. However, due to the relatively small number of subjective contexts in the data overall (N=117), the findings in Figure 5a may be more revealing if we combine the dialects and examine the distribution of all forms in each context, as in Figure 5b.

When the data are collapsed together and all four forms (must, have to, have got to and got to) are considered, we can now see that subjective contexts are highly layered. While must is a dominant form as predicted the

^{17.} We have removed from consideration two communities (Buckie and Wheatley Hill) which had fewer than 5 tokens of subjective obligation.

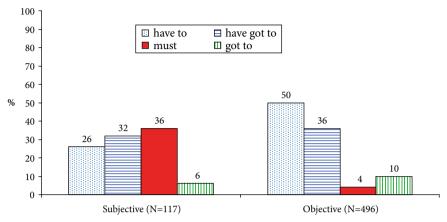


Figure 5b. Distribution of forms for DEONTIC MODALITY according to type of obligation overall.

conspicuous result is that it has stiff competition from both have to (26%) and have got to (32%). Indeed the extent of layering in an area of the system that is widely agreed to be the core of deontic meaning is striking. In contrast, have to stands out as the preferred form for objective contexts, representing 50% of all forms.

In sum, by dividing the data according to subjective vs. objective obligation, we notice yet another contrast in the system. Objective contexts show a clear preference for have to. Use of must for subjective modality may be present, particularly in some varieties, but even there have got to and have to are closing in. Where must once held the role of marking strong subjective obligation, other forms are encroaching.

By operationalizing these grammar internal constraints and testing them in the data we have uncovered at least three potential contextual constraints underlying the selection of forms within the deontic modality system — grammatical person, definiteness and type of obligation.

However two-dimensional views of the data will obscure interaction amongst factors. Let us now attempt to disentangle potential cross-cutting influences. Figure 6 shows a cross-tabulation of grammatical person/definiteness by type of obligation for all the data combined. The subjective readings are shown in the solid shaded columns and the objective readings are shown in the lined columns. This presents quite a complex perspective. Note that the first two columns are have to and the second two columns are have got to (first in subjective, second in objective contexts).

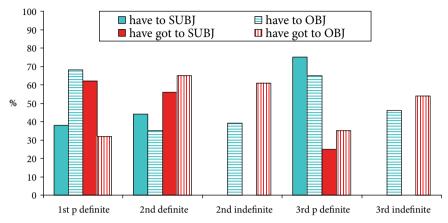


Figure 6. Distribution of have to and have got to according to type of obligation and type of subject.

When subjective contexts are viewed independently of the objective contexts as well as separated by definiteness of reference, further subtleties in distribution of forms are revealed. As we have observed earlier, have to (the first two columns under each category), is the front-runner for definite 3rd person contexts regardless of type of obligation. However, notice that for 1st person contexts no such trend is found. Instead, there is an inexplicable mix of forms that goes in the opposite direction to the pattern in the other grammatical persons, e.g. have to represents the most frequent form with objective readings; have got to with subjective readings. This contrasts with the other contexts in which both subjective and objective readings occur — namely 2nd definite and 3rd definite. Here, have to is more frequent in subjective readings and have got to is more frequent in objective readings. In other words, the intersection of the two contextual factors shows some clear trends, but there are some areas of mixture which prevent us from being able to decide whether grammatical person, definiteness or type of obligation is the definitive constraint guiding developments in the deontic modality system.

Disentangling interaction

Like many features that involve multidimensional influences, proportional distributions alone make it difficult to say absolutely which of the factors (grammatical subject; type of reference; subjective vs. objective meaning) are the explanatory ones. It is more likely to be the case that they are all contributing factors, particularly in a system that is in the midst of ongoing reorganization. A more nuanced view of the situation can be gleaned from a multivariate analysis of these factors in which each hypothesis about the origin and trajectory of development of the variants forms is treated simultaneously (Poplack & Tagliamonte 2001:91). In this way, the more subtle mechanism(s) that are underlying or incipient may emerge from the apparently erratic choice of form. The evidence from how competing factors manifest, i.e. their significance, relative strength and patterning, will provide us with a microscopic view of the variable grammar in each community (see Poplack & Tagliamonte 2001: 93).

However, because these particular factors are so intertwined, it is difficult to construct a straightforward model for statistical testing. We opted for a series of analyses in which we tested one configuration and then another, a typical approach when dealing with interacting cross cutting factors (Tagliamonte 2006a). In other words, we first tested the pragmatic factor (subjective vs. objective), then the grammatical factors, first grammatical person, then definiteness, etc (see Poplack, Tagliamonte & Eze 2000). In the end, an interaction factor group with a three-way split/division provided the best model: Definite subjective 1st and 3rd person contexts, definite objective 1st and 3rd person contexts, and generic objective 2nd person contexts. We first test this on the total data set (i.e., with all the communities combined). This enables us to test for the effect of community, i.e. how each dialect performs, vis-à-vis the variant forms, as in Table 3.18

As we might have expected given the extreme dialect differences in form frequencies, the most significant effect operating on the choice of forms is community.¹⁹ The strength of this factor is high as indicated by the high range values for have to, 78, and have got to, 48. Notice however, that community is not significant for *must*. Further, the values for the factor weights for each form are consistent with the distributional results — high factor weights for have to in Buckie and Northern Ireland where have to reigns supreme and higher factor weights for have got to in Henfield, Maryport, York, Wheatley Hill and Cumnock where have got to has made substantial inroads. However, it is the statistical significance of the carefully configured internal factor (grammatical person/definiteness), in the face of such widely divergent factor weights for the forms across communities that is most relevant. Despite these differences the

^{18.} We have not considered got to here as it is only found with any frequency in one community (Tiverton).

^{19.} Proportions and numbers may vary slightly from the distributional analyses due to the exclusion of knockout factors in some cases (e.g. data from Buckie and Wheatley Hill were removed from the analysis of must because must did not occur). Similarly, we did not treat the rare occurrence of got to because they are virtually restricted to Tiverton.

Table 3. Variable rule analysis of the contribution of factors to the probability of must, have to, have got to for deontic modality

	must		have	have to		got to	
Corrected Mean	.05		.51		.33		_
Overall proportion	10%		44.79	44.7%		6	
	%	FW	%	FW	%	FW	Ns/cell
Community							
TIV	4	.40	12	.14	35	.42	65
HEN	15	.61	48	.42	40	.61	81
MPT	8	.44	49	.45	44	.62	108
YRK	22	.65	54	.49	35	.59	78
WHL	0	_	59	.63	41	.53	17
CMK	7	.47	48	.48	39	.54	125
BCK	0	-	80	.83	20	.30	25
NI	4	.34	87	.92	9	.11	54
Range				<i>78</i>		48	
Grammatical person							
Definite (1st & 3rd) [objective]	4	.42	66	.65	27	.41	213
Definite(1st & 3rd) [subjective]	38	.91	38	.37	34	.49	53
Generic (2nd) [objective]	3	.36	39	.37	45	.60	205
Range		55		28		19	
Sex							
Female	15	.63	64	.66	26	.37	277
Male	5	.35	37	.34	46	.65	276
Range		28		32		28	
TOTAL N							553

Factors selected as significant in bold

multiple regression selects the contextual constraint as statistically significant predictor of form across the board. Moreover, the hierarchy of factor weights comprising this factor group draws out a critical pattern: Definite objective contexts favour have to, with a factor weight of .65; 2nd person indefinite contexts favour have got to, with a factor weight of .60. Finally, the rare instances of must are the primary choice for the much rarer definite 1st and 3rd person subjective contexts where it is favoured at .91.20

^{20.} Separate analysis distinguishing 1st from 3rd person revealed that there was no statistically significant difference between them.

Summary

Each dialect can be distinguished based on its unique distribution of frequencies of forms for deontic modality (Figure 2). However, there is a significant underlying, pan-dialectal, grammar-internal constraint which guides their selection (Figures 3–6, Table 3) — something that could not be unambiguously decided from the distributional analysis. In the next section we explore this further

External correlates

The historical trajectory predicts ongoing encroachment of have got to within the English deontic modality system. Yet, even in the earliest stages of this development, variation between the oldest form *must* and the second layer, *have* to, was apparent (e.g. Brinton 1991) and must have endured for a fairly long stretch of time. The dialect data under consideration here confirm ongoing layering, yet there is a new quality to the competition — variation between have to and have got to. Of course, this may be, at least in part, an artefact of our data, which comprises elderly speakers only and spoken vernacular language rather than formal written texts. The stability of have to may simply be the result of the fact that we have accessed an earlier stage of vigorous competition, which in more mainstream varieties has resolved itself towards increasing have got to. Recall that this observation is made repeatedly in the literature, but is there any way to substantiate this hypothesis?

In fact, we have a consummate opportunity to test it. In three of the dialects — Buckie, Wheatley Hill and York, we have access to three generations of speakers, not just the oldest speakers. In the analyses that follow we add these middle and younger generations to the data. This permits an assessment of change in apparent time. For the purposes of inter-variety parity we have grouped the speakers into a broad three-way categorization: Under 40, 41-69 and 70 and above.21

If the varying rates of *have to* and *have got to* reflect ongoing developments, then we would expect to see a decrease in the use of have to in apparent time and an increase in the use of have got to. Further, we should be able to assess if this development is being driven by the same underlying process.

^{21.} The youngest speakers are mostly between 20–40; however, three are 16–18.

Table 4 tests these hypotheses by subjecting the data to the same statistical modelling as earlier in Table 3, however, this time we test the model in each dialect separately including speaker age as an independent factor in the analysis and this time exclude the rare instances of *must*.

Table 4. Three independent variable rule analyses of the contribution of external and internal factors to the probability of have to in Buckie, Wheatley Hill and York.

	Buckie		Wheatley I	Hill	York	
	.60		.32		.49	
	152		165		369	
	F/W	%	F/W	%	F/W	%
FACTORS:						
Age						
> 70	.74	80	[.72]	59	.52	54
41-70	.39	47	[.46]	39	.40	40
< 40	.47	58	[.49]	29	.59	55
Range	35				20	
Sex						
Female	[.54]	63	.72	58	.55	54
Male	[.45]	54	.34	26	.41	40
Range			38		12	
Type of reference						
Definite	.61	70	.81	65	.66	65
Objective						
Generic	.37	46	.32	17	.41	40
Objective						
Definite	.40	52	.22	16	.39	40
Subjective						
Range	36		59		27	

Factors selected as significant in bold.

The remarkable finding is that the constraints on the competition between *have* to and have got to are virtually the same across communities, age, sex and type of subject. Let us consider each of these effects in turn.

Speaker age

Consider first the effect of speaker age (generation). Given the well-known historical trajectory with regard to the development of the forms for deontic modality, we would expect to find a statistical correlation moving from higher to lower probability of have to across generations. However, this is not the case. Have to is the most highly favoured form amongst the oldest generations in Buckie and Wheatley Hill, and even exhibits a favouring factor weight in York (.52). The middle aged group all show the expected decrease in use from the older speakers. In contrast, from middle to young speakers, the factor weights actually increase across all communities. This finding is not what we would have predicted from the trajectories of change that appeared to be underway from the diachronic perspective.

Speaker sex

Social factors have not figured prominently in the study of deontic modality. However, the association of *have got to* with informal language and the putative stigma associated with it predicts that females might shy away from it. Indeed, in each of the three communities, females eschew the use of have got to, while males favour it. Moreover, this effect is statistically significant in Wheatley Hill and York.

Type of subject

The type of subject constraint is consistent across all three communities — definite subjects (primarily 1st person and 3rd person pronouns) with objective readings clearly favour have to across the board. In contrast, indefinite subjects (mostly generic you), favours have got to. The core deontic context of strong obligation — definite subjects with subjective readings — are the lowest for have to in Wheatley Hill.

Thus despite widely varying rates of have to and have got to across highly differentiated communities, the constraints operating on the use of forms within the deontic system, whether social or linguistic, are remarkably parallel. We conclude that whatever is ongoing in the deontic modality system involves subtle readjustments on a grammatical level and these are shared by all dialects.

However, there are also good indications that sociolinguistic influences are involved. The fact that age and sex are hallmarks of substantive generational readjustments, lead us to delve deeper into these patterns of distribution in the data since it is well-known that the two are critical indicators of linguistic change (Labov 2001). For example, males tend to use stigmatized forms and females prestige forms (Labov 1970, 1972, and subsequent work). Further, in the context of linguistic change females tend to lead in the use of standard innovations while males tend to lag behind and use more non-standard or stylistically-marked forms (Labov 1990, 2001). Figure 7 provides a graphic depiction of the frequency of have to across the three generations of speakers in Buckie, Wheatley Hill and York.

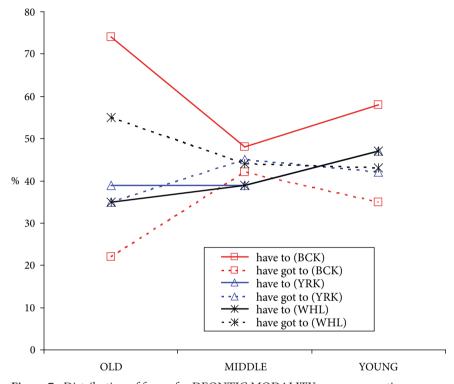


Figure 7. Distribution of forms for DEONTIC MODALITY across generations

The figure provides a corroborating picture to the variable rule analysis results that highlights the fact that the recent history of the forms used for deontic modality is more complicated than expected. There is a dramatic decrease in use of have to from the old to middle aged speakers in Buckie, suggesting that have got to was a newcomer to this dialect for the middle-aged generation. But there is no advance of *have got to* during this same time period in either York or Wheatley Hill. Instead, there appears to be a levelling in these dialects in the middle-aged generation. However, it is the increase in use of have to between the middle and younger age groups across all dialects that is unexpected. Essentially, this suggests that there has been a reversal in the trajectory of change. Although *have got to* had made substantial in-roads early in the 20th century it is not increasing anymore. Have to is on the rise and further, this is happening in all the dialects in the same way at the same time.²²

We now divide the data by age and sex in the middle aged and younger speakers where this reversal is occurring as in Figure 8.

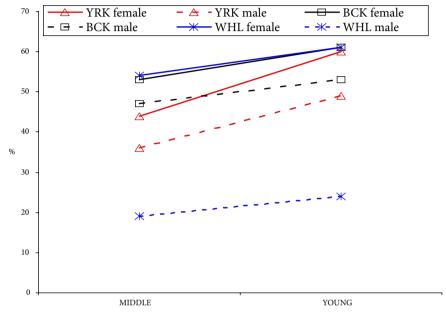


Figure 8. Distribution of have to for DEONTIC MODALITY by age and sex in middle-aged and younger speakers.

Another trend is revealed. It is the female speakers in the younger generation in every community who are leading in the change toward have to, while the males lag behind.

These new findings reveal that: (1) the direction of change in the deontic modality system is not what we expected (have got to was thought to be increasing, not have to), and (2) all three communities exhibit a rise in frequency of have to from middle aged to younger speakers, and finally (3) young women are on the leading edge of change in each case.

^{22.} Interestingly, this contrasts with questionnaire findings for Australian English which shows a rise in have got to in the youngest speakers (10-24 years old) but stability between have to and have got to in the older generations (Collins & Blair 1989:265).

Discussion

What do these findings tell us about the deontic modality system in dialects of English in Britain?

First, we suggest that the layering of must, have to and have got to provide important insights into the status of the grammar at a particular point in time as well as an indication of how changes have been taking place in the language. The varying effects of social context and geographic location across these eight different communities can be viewed in the different frequencies of forms (Figure 2, Table 5). Moreover, we suggest that the contrast between the proportion of older and newer forms provides a particularly interesting mirror of the pathways of change, in this case the putative transition from must, to have to, to have got to with ongoing grammaticalization processes, such as specialization and phonological coalescence, taking place in tandem.

In the case of *must* the dialect data reveal that (1) it is not now and may never have been firmly established in some dialects; and (2) whether must was present in all the dialects or not, it is clearly obsolescent now. Why is deontic *must* in decline? Some researchers argue that it may be the loss of the particular stylistic register associated with must (e.g. Biber et al. 1999) — in Conradie's (1987:179) terms "an anti-authoritarian development" which also jibes with Mair's idea of increasing "colloquialization" (Mair 1997, see also Mair & Hundt 1999), particularly of American English (see also Leech 2003:236-237), as well as Myhill's (1995) notion of 'democratisation'. Others suggest that it involves ongoing reduction in auxiliary uses that express the speaker's attitude or evaluation (Warner 1993). Whatever the reason it seems clear that such developments must be rooted in changes external to the system, likely in distinctions of style, genre or register, which is beyond the scope of the present study though this clearly bears further investigation (see also Facchinetti, Krug & Palmer 2003). For now, it is clear that these spoken dialect data exhibit deontic must rarely and where it is used it seems to have retreated to the core — contexts of subjective obligation with definite subjects. Supportive to the idea of ongoing retreat is the fact that it is used most often in specialized formulaic utterances and in communities that include middle class speakers (York, Henfield (Tables 3, 4, Figures 3, 8). Thus, where Standard English prescribes deontic must, spoken synchronic dialects show us that its use is highly restricted in vernacular language.

The use of have to and have got to on the other hand reveal that deontic modality is the site of vibrant ongoing variation and change (Tables 3, 4, Figures 2-8). While the literature predicts ongoing change towards have got to

and gotta, what we discovered is that dialects in Northern Ireland, Scotland and England are holding on to have to. If our spoken data are any indication, the vernacular trends we have found may herald what will come in written registers. However, unlike Krug, who argues for the rise of gotta as the new modal for English deontic modality we would like to argue for have to as the latest renewal of the deontic modal category. Although have to started down this path way back in Middle English, it gained a newer competitor, have got to, in the 19th century. We cannot speculate on why have got to came in to replace a form that was already undergoing grammaticalization as a deontic modality marker, but we can speculate on the resurgence of have to (Figure 7 and 8). We suggest that both linguistic and social factors intervened to favour its return. In the case of have to and have got to, the former is often considered monoclausal (e.g. Heine 1993, Krug 2000, van Gelderen 2004), while the latter, biclausal; these structural differences might explain why there is a visible resurgence of the simpler have to form — it more readily fits into the modal auxiliary slot. But if purely linguistic factors were involved, then we would predict that *got to*, also monoclausal, and the most auxiliary-like of all three competitors, would win out in the end. Yet, this form has only made inroads in one community. We propose that social factors must also be implicated in the patterns of use across the communities investigated here. The pivotal point for the reversal in the trajectory of change is among the youngest generation (20-40 year olds). Moreover, the fact that female speakers are leading this development suggests that the social evaluation of forms is involved. Sociolinguistic research has shown that women tend to adopt forms that are legitimized in the standard language. Moreover, women tend to advance these changes at about a generation ahead of men (Labov 2001). This fits the pattern of change we have uncovered here. With *must* in decline and *have got to* a more structurally complex item which is also imbued with undesirable social stigma, women in the later part of the 20th century began to favour the already grammaticalizing have to and avoiding the structurally simple, but socially stigmatized form with got. Now they appear to be leading the way in its increasing use as the favoured deontic modal, at least in British English.

Yet how was this change made possible in the grammar? Examination of the underlying internal constraints on the variation among forms reveals that these dialects show a consistent adjustment for the specialization of forms. This is best demonstrated in the large cross-generational York English Corpus, as in Figure 9.

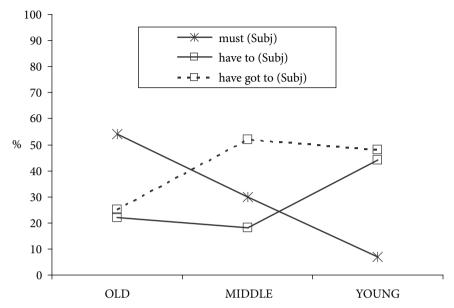


Figure 9. Distribution of forms *for* DEONTIC MODALITY in contexts of subjective obligation by generation in York.

Figure 9 shows ongoing restructuring of the core contexts of deontic modality — subjective obligation meaning — in apparent time. Note that while *must* drops for this function, *have to* is increasing, particularly among the youngest generation. The building evidence suggests that *have to* is steadfastly taking over the auxiliary slot left behind in the wake of the (dramatic) recession of *must*.

But what will happen next? Will British dialects continue their present course toward increasing use of *have to*? Will *have to* become more and more auxiliary-like as time goes on, evidencing increasing concomitants of advanced grammaticalization? Broader study of the system of deontic modality across other major varieties of English will present a unique opportunity to track the development of this sub-system of grammar at a time when contemporary British dialects expose an unusual twist in the course of its history. Also worthy of more in-depth consideration are the tantalizing sociolinguistic qualities of this variable. Will the sociolinguistic correlations persist? And if so, how will they figure in the ensuing steps of this historical process? Indeed, the finding that the middle-aged generation in Britain in the late 20th century is the locus of pivotal change is not unique to developments in the deontic modality system. This result must be situated in the context of other studies in which this particular generation has been identified as a watershed of linguistic change more

generally (e.g. Chambers 2002, 2004, Tagliamonte 2006b). There are undoubtedly developments in progress even now that are laying down new processes of change for us to examine in the future.

References

- Anttila, Raimo. 1989. An Introduction to Historical and Comparative Linguistics. Amsterdam & New York: John Benjamins.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. Longman Grammar of Spoken and Written English. Harlow: Longman.
- Bolinger, Dwight. 1980. Language: The loaded weapon. London: Longman.
- Brinton, Laurel J. 1981. The historical development of aspectual periphrases in English. PhD dissertation, University of California, Berkeley.
- Brinton, Laurel J. 1991. "The origin and development of quasimodal 'have to' in English". Paper presented at The Workshop on Verbal Periphrases, Amsterdam. http://faculty.arts. ubc.ca/lbrinton/.
- Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. The Evolution of Grammar: Tense, aspect, and modality in the languages of the world. Chicago: University of Chicago Press.
- Chambers, J.K. 2002. "Patterns of variation including change". The Handbook of Language Variation and Change ed. by J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes, 349-372. Malden & Oxford: Blackwell.
- Chambers, J.K. 2004. "Canadian Dainty': The rise and decline of Briticisms in Canada". The legacy of Colonial English: Studies in transported dialects ed. by Raymond Hickey, 224-241. Cambridge: Cambridge University Press.
- Christy, Craig. 1983. Uniformitarianism in Lingusitics. Amsterdam & Philadelphia: John Benjamins.
- Coates, Jennifer. 1983. The Semantics of the Modal Auxiliaries. London: Croom Helm.
- Collins, Peter & David Blair. 1989. Australian English: The language of a new society. Queensland: University of Queensland Press.
- Conradie, C. J. 1987. "Semantic change in modal auxiliaries as a result of speech act embedding". Historical Development of Auxiliaries ed. by Martin B. Harris & Paolo Ramat, 171-189. Berlin: Mouton de Gruyter.
- Corrigan, Karen P. 2000. "What bees to be maun be": Aspects of deontic and epistemic modality in a Northern Dialect of Irish English". English World-Wide 21.25-62.
- Crowell, Thomas L. 1955. "Predating 'have to', 'must'?" American Speech 30.68-69.
- Curme, George O. 1931. A grammar of the English Language, III: Syntax. Boston: Heath.
- Denison, David. 1993. English historical syntax. Harlow, Essex: Longman.
- Denison, David. 1998. "Syntax". The Cambridge History of the English Language, 1776-present day ed. by Suzanne Romaine, 92-329. Cambridge: Cambridge University Press.

- Ehram, Madeline Elizabeth. 1966. The Meanings of the Modals in Present-day American English. (= Janua Linguarum, Series practica, 45). The Hague: Mouton.
- Facchinetti, Roberta, Manfred Krug & Frank Palmer, eds. 2003. "Modality in contemporary English". (= Topics in English Linguistics, 44.) Berlin & New York: Mouton de Gruyter.
- Fischer, Olga. 1992. "Syntax". The Cambridge History of the English Language ed. by Norman Blake, 207-408. Cambridge: Cambridge University Press.
- Fischer, Olga. 1994. "The development of quasi-auxiliaries in English and changes in word order". Neophilologus, 78.137-164.
- Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2000. The Syntax of Early English. Cambridge: Cambridge University Press.
- Francis, W. N. 1958. The structure of American English. New York: Ronald Press Co.
- Godfrey, Elisabeth & Sali Tagliamonte. 1999. "Another piece for the verbal -s story: Evidence from Devon in Southwest England". Language Variation & Change 11.87-121.
- Heine, Bernd. 1993. Auxiliaries: Cognitive forces and grammaticalization. Oxford: Oxford University Press.
- Hock, Hans Henrich. 1986. Principles of Historical Linguistics. Amsterdam: Mouton de
- Hopper, Paul J. 1991. "On some principles of grammaticization". Approaches to Grammaticalization Vol. 1: Focus on theoretical and methodological issues ed. by Elizabeth Closs Traugott & Bernd Heine, 17–35. Amsterdam & Philadelphia: John Benjamins.
- Hopper, Paul J. & Elizabeth Closs Traugott. 1993. Grammaticalization. Cambridge textbooks in linguistics. Cambridge: Cambridge University Press.
- Huddleston, Rodney & Geoffrey K. Pullum. 2002. The Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.
- Jespersen, Otto H. 1913. A Modern English Grammar on Historical Principles Part II: Syntax, vol. VII. London: George Allen & Unwin.
- Jones, Megan & Sali A. Tagliamonte. 2004. "From Somerset to Samaná: Pre-verbal did in the voyage of English". Language Variation & Change 16.93-126.
- Klemola, Kaarlo Juhani. 1996. Non-standard Periphrastic do: A study of variation and change. Essex: PhD dissertation.
- Krug, Manfred. 1998. "Gotta The tenth central modal in English? Social, stylistic and regional variation in the British National Corpus as evidence of ongoing grammaticalization". The Major Varieties of English ed. by Hans Lindquist, Staffan Klintborg, Magnus Levin & Maria Estling, 177-191. Växjö: Växjö University.
- Krug, Manfred. 2000. Emerging English Modals: A corpus-based study of grammaticalization. Berlin & New York: Mouton de Gruyter.
- Kurath, Hans. 1949. A Word Geography of the Eastern United States. Ann Arbor: University of Michigan Press.
- Labov, William. 1970. "The study of language in its social context". Studium Generale 23.30-
- Labov, William. 1972. Sociolinguistic Patterns. Philadelphia: University of Pennsylvania Press.

- Laboy, William. 1990. "The intersection of sex and social class in the course of linguistic change". Language Variation & Change 2.205-254.
- Labov, William. 2001. Principles of Linguistic Change. Vol. 2: Social factors. Malden & Oxford: Blackwell.
- Leech, Geoffrey. 2003. "Modality on the move: The English modal auxiliaries 1961-1992". Modality in Contemporary English ed. by Roberta Facchinette, Manfred Krug & Frank Palmer, 223-240. Berlin & New York: Mouton de Gruyter.
- Lightfoot, David. 1979. Principles of Diachronic Syntax. Cambridge: Cambridge University Press
- Macafee, Caroline. 1992. Characteristics of non-standard grammar in Scotland, accessed 15 December 2005 http://www.abdn.ac.uk/~enl038/grammar.htm
- Mair, Christian. 1997. "Parallel corpora: A real-time approach to the study of language change in progress". Corpus-based studies in English: Papers from the seventeenth international conference on English language research on computerized corpora (ICAME 17) ed. by Magnus Ljung, 195-209. Amsterdam: Rodopi.
- Mair, Christian & Marianne Hundt. 1999. "'Agile' and 'uptight' Genres: The corpus-based approach to language change in progress". International Journal of Corpus Linguistics 4.221-242.
- Miller, James. 1993. "The grammar of Scottish English". Real English: The grammar of English dialects in the British Isles ed. by James Milroy & Lesley Milroy, 99-138. New York: Longman.
- Mitchell, Bruce. 1985. Old English syntax. Oxford: Clarendon Press.
- Myhill, John. 1995. "Change and continuity in the functions of the American English modals". Linguistics 33.157-211.
- Orton, Harold & Wilfrid J. Halliday. 1963. Survey of English dialects. Leeds: E. J. Arnold.
- Palmer, Frank R. 1979. Modality and the English modals. New York: Longman.
- Plank, Frans. 1984. "The modals story retold". Studies in Language 8.305-364.
- Poplack, Shana & Sali A. Tagliamonte. 1999. "The grammaticalization of going to in (African American) English". Language Variation & Change 11.315-342.
- Poplack, Shana. 2001. African American English in the Diaspora: Tense and aspect. Malden: Blackwell Publishers.
- Poplack, Shana, Sali A. Tagliamonte & Ejike Eze. 2000. "Reconstructing the source of Early African American English plural marking: A comparative study of English and creole". The English History of African American Englished. by Shana Poplack, 73-105. Oxford & Malden: Blackwell.
- Quirk, Randolph, Sidney Greenbaum, Geoffry Leech & Jan Svartvik. 1985. A Comprehensive Grammar of the English Language. New York: Longman.
- Smith, Jennifer. 2000. Synchrony and Diachrony in the Evolution of English: Evidence from Scotland. D. Phil. dissertation, University of York.
- Smith, Nicholas. 2003. "Changes in the modals and semi-modals of strong obligation and epistemic necessity in recent British English". Modality in Contemporary English ed. by Roberta Facchinette, Manfred Krug & Frank Palmer, 241–266. Berlin & New York.

- Sweetser, Eve E. 1988. "Grammaticalization and semantic bleaching". Berkeley Linguistics Society: General Session and Parasesson on Grammaticalization Berkeley Linguistics Society, 14 ed. by Shelley Axmaker, Annie Jaisser & Helen Singmaster.
- Tagliamonte, Sali A. 2002. "Comparative sociolinguistics". Handbook of Language Variation and Change ed. by Jack Chambers, Peter Trudgill & Natalie Schilling-Estes, 729-763. Malden & Oxford: Blackwell Publishers.
- Tagliamonte, Sali A. 2003. "Every place has a different toll": Determinants of grammatical variation in cross-variety perspective". Determinants of Grammatical Variation in English ed. by Günter Rhodenberg & Britta Mondorf, 531-554. Berlin & New York: Mouton de Gruyter.
- Tagliamonte, Sali A. 2004. "Have to, gotta, must: Grammaticalization, variation and specialization in English deontic modality". Corpus research on grammaticalization in English ed. by Hans Lindquist & Christian Mair, 37-55. Amsterdam & Philadelphia: John Benjamins.
- Tagliamonte, Sali A. 2006a. Analysing Sociolinguistic Variation. Cambridge: Cambridge University Press.
- Tagliamonte, Sali A. 2006b. "So cool, right?": Canadian English entering the 21st century". Canadian English in a Global Context. Theme issue of Canadian Journal of Linguistics ed. by Peter Avery, Alexandra D'Arcy, Elaine Gold & Keren Rice.
- Tagliamonte, Sali A. Forthcoming. Roots of English: Exploring the history of dialects. Cambridge: Cambridge University Press.
- Tagliamonte, Sali A. & Jennifer Smith. 2002. "Either it isn't or it's not: NEG/AUX contraction in British dialects". English Word Wide 23.251-281.
- Tagliamonte, Sali A., Jennifer Smith & Helen Lawrence. 2005. "No taming the vernacular! Insights from the relatives in northern Britain". Language Variation & Change 17.75-
- Traugott, Elizabeth Closs. 1997. "Subjectification and the development of epistemic meaning. The case of promise and threaten". Modality in Germanic Languages: Historical and comparative perspectives ed. by Toril Swan & Olaf Westvik, 185-210. Berlin: Mouton de Gruyter.
- Traugott, Elizabeth Closs. 1999. "Why must is not moot". Paper presented at Fourteenth International Conference on Historical Linguistics. Vancouver, British Columbia.
- Trousdale, Graeme. 2003. "Simplification and redistribution: An account of modal verb usage in Tyneside English". English World-Wide 24.271-284.
- Trudgill, Peter J. 1990. The Dialects of England. Oxford: Blackwell Publishers.
- van den Eynden, Nadine. 1993. Syntactic Variation and Unconscious Linguistic Change: A study of adjectival relative clauses in the dialect of Dorset. Frankfurt am Main: Peter Lang.
- van der Gaaf, W. 1931. "Beon' and habban' with inflected infinitive". English Studies 13.180-
- van Gelderen, Elly. 2004. Grammaticalization as Economy. Amsterdam & Philadelphia: John Benjamins.

van Kemenade, Ans. 1999. "Sentential negation and clause structure in Old English". Negation in the History of English ed. by Ingrid Tieken-Boon van Ostade, Gunnel Tottie & Wim van der Wurff, 147-165. Berlin: Mouton de Gruyter.

Visser, Fredericus T. 1963-1973. An Historical Syntax of the English Language. Leiden: Brill. Warner, Anthony. 1993. English Auxiliaries: Structure and history. Cambridge: Cambridge University Press.

Whitney, William Dwight. 1867. Language and the Study of Language. New York: Scribner.

Résumé

Le système de modalité déontique en anglais contemporain fait preuve de changements en cours. Cet article a pour but d'étudier son évolution à la lumière de données dialectales anglaises, écossaises et d'Irlande du nord. L'histoire nous montre la prédominance de 'must' mise en cause par 'have to' et ensuite par 'have got to'. Notre perspective interdialectale et la méthodologie de la sociolinguistique comparative nous permettent de présenter une reconstruction hypothétique des étapes les plus récentes dans cette évolution. Nos résultats démontrent que la proportion de chaque forme employée varie d'un dialecte à l'autre, mais que le contexte où chacune s'emploie reste constant. 'Must' est en voie de disparition et la restructuration du système pan-dialectal est accompagnée d'une résurgence inattendue de 'have to': (1) 'have to' s'emploie dans les contextes traditionnellement réservés à 'must' et (2) 'have got to' est en train de se limiter aux référents indéfinis. Ce sont les jeunes femmes qui se trouvent en tête de cette évolution, ce qui suggère que la réorganisation grammaticale systématique se combine avec des forces sociolinguistiques pour faire avancer les changements linguistiques.

Zusammenfassung

Diese Arbeit untersucht ein der ständigen Veränderung unterliegendes Gebiet der englischen Sprache — deontische Modalität — und benutzt einen synchronischen Dialektkorpus aus England, Schottland und Nordirland zur Gewinnung neuer Informationen über die Entwicklungen der Modalität. Es ist geschichtlich belegt, dass in diesem System must zu have to zu have got to wechselt. Durch eine dialektal übergreifende Perspektive und die Verwendung komparativer soziolinguistischer Methoden präsentieren wir eine mögliche Rekonstruktion der späteren Schritte in diesem Prozess. Die Ergebnisse zeigen dialektale Unterschiede im Verhältnis von älteren und neueren Formen, aber ähnliche Anwendungsmuster. Must ist (zunehmend) veraltet, während have to neben pan-dialektaler, grammatischer Neueinteilung wiederbelebt wird. Diese Neueinteilung ist zweiteilig:

- 1. have to wird in Kontexten traditionell mit must kodiert und
- 2. have got to ist eine Spezialisation für einen unbestimmten Hinweis.

Junge Frauen sind die Vorreiter in dieser Entwicklung, was auf eine Verbindung von grammatischer Systemanpassung mit soziolinguistischer Beeinflussung als treibende Kraft der sprachlichen Veränderung hinweist.

Authors' addresses

Sali Tagliamonte Department of Linguistics

University of Toronto

130 St. George Street, rm 6076

Toronto, Ont. M5S 3H1

Canada

E-mail: sali.tagliamonte@utoronto.ca

Jennifer Smith

Department of English

12 University Gardens

University of Glasgow

Glasgow, G12 8QQ

Scotland, UK

E-mail: jennifer.smith@englang.arts.gla.ac.uk

Appendix

Table 5. Overall distribution of forms by community.

Community	must		have to	have to		've/'s got to		gotta
	%	N	%	N	%	N	%	N
TIV	4	3	12	8	35	24	49	33
HEN	15	14	41	39	40	38	4	4
YRK	22	22	42	42	35	35	1	1
WHL	0	0	59	10	41	7	0	0
MPT	8	9	45	51	44	50	4	4
NI	4	2	84	47	9	5	4	2
CMK	7	10	44	60	39	52	10	13
BCK	0	0	80	20	20	5	0	0