Form and function co-variation: Obligation modals in Australian English

Abstract

Shifts in the frequencies of English modals of obligation have been linked to shifts in modal function and changing interpersonal authority. Interpretation of over 2,000 tokens in spontaneous speech data recorded in Sydney, Australia in the 1970s and 2010s establishes a replicable classification of obligation meanings, based on source of obligation, with a three-way distinction between *Hierarchical* authority/norms, *General* circumstances/conditions, *Personal* choice/opinion. Competing expressions for these obligation types, besides *have to, have got to* and older *must*, include *should* and, recently, *need to*. Two sets of regression analyses provide evidence of co-variation of form and function: first, the linguistic and social conditioning of forms, with meaning as one of the predictors; and second, the conditioning of function, with modal form as a predictor. *Need to* rises in real-time and so does talk of personal obligation. However, the change in modal function is concomitant with, but independent of, shifting modal forms.

1   Introduction: Co-variation of form and function

The set of competing modal forms for the expression of obligation in English has been the subject of considerable study, with variability between older forms *must* and *should*, long standing *have to* and *have got to*, and newer *need to* being reported across the English-speaking world. Shifts in the relative frequencies of modal forms have been tied to a shift in modal function, with a move away from expression of obligation based on social order and general principles to functions that are more personal and "interactive", as when offering advice or making an emotional appeal (Myhill 1995: 160). Specifically, a decline in *must* and an increase in *need to* has been observed and linked to notions such as anti-authoritarianism, subjectivity and colloquialization (e.g., Collins 2005: 260; Fehringer and Corrigan 2015: 377; Leech 2013: 111; Penry Williams and Korhonen 2020: 280; Tagliamonte and Smith 2006: 372).

But how can we disentangle change in function from change in form, or talk of different kinds of obligation from use of different modal forms? Change proceeds through the rise of new expressions in conjunction with shifts in the conditioning of existing expressions with which the new forms come to compete (Labov 1982: 20) (cf., Poplack and Malvar 2007; Szmrecsanyi 2016; Torres Cacoullos 2012). To examine the co-variation of form and function (Labov 2008: 2), we probe rates and conditioning not only of competing forms expressing similar meanings but also of functions—the distributions (proportions) of the data according to particular meaning nuances expressed.

Tracking change in function requires a meaning classification that is independent of other linguistic factors that may impact modal choice. Though there has been much discussion of

modal meanings, such a replicable classification for actual usage has been lacking. Here, based on close examination of over 2,000 tokens in spontaneous speech data, we are able to establish that source of obligation provides a reliable basis for distinguishing modal meanings and we put forward a three-way classification—*Hierarchical* authority / norms, *General* circumstances / conditions, and *Personal* choice / opinion. The modals *must, have to, have got to, should* and *need to*, illustrated in examples (1), (2) and (3), exist as competing variants, and this meaning classification allows us to disentangle modal form and function, and thus independently track change in both over time.

(1)
Beatrice: *we were told too that if --*
              we must be on time for meals.

[Bcnt_AEF_158, 01:07:09-01:07:12][1]

(2) (re children using nicknames to refer to teachers)
Belinda: *I think puts,*
              the children on a more ... equal footing with them,
              especially as they're getting older and,
              have to go through school till .. they're eighteen,
              they've got to be treated like young adults,
Angela: *hm.*
Belinda: *...(1.0) where .. as,*
              we were definitely treated like children,

[SSDS_AAF_164, 32:37-32:51]

(3) (re unsolicited advice offered to people suffering depression)
Anna: *it's not just,*
              oh you need to cheer up,
              or you need to get it over it,
              or you just need some sunshine,
              or you should drink more water.
              <@ like the @> --
              none of that is going to help the neurochemical imbalance,

[SydS_AYF_033, 49:26-49:34]

We first describe the speech data for the study, which comprises recordings made in Sydney, Australia, at two time points: 1970s-1980s and 2010s-2020s. We then present the meaning classification of Hierarchical, General or Personal source of obligation and specify the variable context within which modals of obligation compete. As a first overview, we compare the relative frequencies in the two time periods of both the modal forms and functions. This allows us to narrow down the relevant comparisons for the statistical analyses, as not all modals compete equally in the same semantic-structural space. Two sets of regression analyses provide evidence of the co-variation between form and function: first, of the linguistic and social conditioning of

---

[1] Examples are produced verbatim from the transcripts. In brackets following each example are the corpus name, speaker social code (Ethnicity, Age group, Gender), speaker number, and beginning-ending time stamps of the excerpt; all names given are pseudonyms.

modal forms over time, for which meaning is one of the set of the predictors, and then, of the conditioning of function over time, for which modal form is a predictor. We confirm the change in form reported in other studies, most recently with the advent of *need to*. But going beyond previous work, we establish change in function, with an increase in the relative frequency with which speakers express obligation stemming from personal choice, independently of the modal form used.

2   Data for the study of parallel formal and functional change

The data come from three sub-corpora of Australian English, capturing the speech of 259 Australians, recorded in two time periods separated by some forty years, and compiled for the *Sydney Speaks* project (Travis 2014-2022) (Table 1). What we will refer to as the 1970s data comprises recordings made in the 1970s and 1980s with three age groups, Elderly Australians (NSW Bicentennial Oral History Project, 1987), Adults, and Teenagers (Sydney Social Dialect Survey, Horvath 1985), who were born in the 1900s, 1930s and 1960s respectively. The 2010s data comprises recordings made 2016-2022 with two age groups, Adults born in the 1960s— thus, the same generation as the 1970s Teenagers, recorded 40 years on—and Young Adults, born in the 1990s. Based on a composite score calculated from occupation, level of education and high school type, participants are placed into three social class groups, labelled Middle, Lower Middle, and Working Class.

Participants are all speakers of Australian English who were either born in Australia or arrived before the age of five. They reflect the changing ethnic makeup of Australia—Anglo-Celtic (all corpora), Greek and Italian (1970s Teens and 2010s Adults and Young Adults) and Chinese Australians (2010s Young Adults) (cf., Grama, Travis and Gonzalez 2020: 349-352). We are able to consider the ethnic groups together since, overall, they share conditioning of both modal form and function, though we will make note of some differences that we observe, which relate to social class.

Table 1   Distribution of participants

Time period	1970s							2010s				
Corpus	Bicentennial Oral History Project BCNT		Sydney Social Dialect Survey SSDS					Sydney Speaks 2010s SydS				
Age group	Elderly		Adults		Teens			Adults		Young Adults		
Ave. birth decade	1900s		1930s		1960s			1960s		1990s		
Ave. age	85 years		44 years		15 years			53 years		24 years		
Gender	Women	Men	Women	Men	Women	Men	**TOTAL**	Women	Men	Women	Men	**TOTAL**
Social class												
Middle	2	4	4	2	12	8	32	13	10	16	5	44
Lower Mid	10	5	4	4	9	14	46	14	11	15	17	57
Working	5	5	4	2	14	15	45	11	7	4	13	35
TOTAL							123					136

For each participant, approximately 30 minutes of speech (or some 5,000 words) were transcribed following the transcription conventions outlined in Du Bois (1993) (see Appendix), giving a total of some 1.3 million words of spontaneous speech for analysis.

Though rarely commented on, genre differences may arise in cross-corpora comparisons of the kind that are regularly drawn on for the study of change in real-time. Here, the primary

differences between the data from the two time periods are that, in the 1970s, the participants were generally not known to the interviewers and the interviews were more structured, while in the 2010s, interviewers recorded people in their same social network and the recordings are more conversational. There is a corresponding difference in data distributions (cf., Travis and Lindstrom 2016). Relevant to modal choice are data distribution differences with respect to tense and subject. A greater occurrence of narratives in the 1970s data results in more Past tense, while the denser dialogic interactions in the 2010s gives rise to a higher proportion of Present tense. And the topic of childhood games in the earlier data contributes to a higher proportion of generic subjects (75% 1970s vs. 39% 2010s), as such descriptions tend to occur with non-specific subjects, for example in (4). The distribution is, however, similar across the two time periods in the proportion of first person singular, an indication of their comparability with respect to speaker involvement (Torres Cacoullos and Travis 2020: 260). We deal with data distribution differences by incorporating tense and subject in assessing the linguistic conditioning of variation (cf., Torres Cacoullos and Travis 2018: 6-9, 148-149, and references therein on contextual distributions with respect to rates and conditioning of variant forms). Linguistic conditioning includes both the variable context (circumscribed to the Present tense) and the probabilistic constraints (testing subject as a predictor).

(4)
Mike: *two persons stand opposite each other,*
       and <u>you gotta throw </u>the knife and .. stretch their legs out,
       you just --
       ... wherever the knife lands,
       <u>you gotta move</u> your foot to,
       .. pick up the knife,
       and you throw it back to the other person.

[SSDS_ATM_126, 23:08-23:16]

3   Obligation meanings on the ground

The modal meaning of obligation, sometimes called necessity, expresses the existence of conditions, social or physical, that compel the participant to carry out the action indicated by the predicate (cf., Bybee, Perkins and Pagliuca 1994: 177). In reviewing studies of English modals of obligation we are confronted with a plethora of meaning classifications.

Within the broad meaning of obligation, one distinction is according to a gradation of strength, ranging between strong and weak obligation, paraphrasable as "it is imperative" (a demand) and "it is important" (a suggestion), respectively (Coates 1983: 31-32, 58). Strong versus weak obligation is commonly associated in English with *must* versus *should*. On this basis, *should* has not been counted in studies of variation between obligation modals (e.g., Tagliamonte and D'Arcy 2007: 52). However, as seen above in example (3), *should* exists in competition with *need to*, and thus excluding it precludes identifying the conditioning of *need to* and tracking the introduction of this newest form.

A second distinction that has been applied to obligation meaning is speaker involvement along a subjective-objective cline (Lyons 1977: 792-793) (cf., Palmer 1986: 16). Subjective instances are those in which "the speaker has authority" as opposed to objective ones in which "the authority comes from no particular source" (Coates 1983: 55), as when relaying rules and

regulations (Huddleston and Pullum 2002: 183). *Have got to* and *must* are considered to span the subjective-objective cline whereas *have to* is said to completely occupy objective necessity, "it is necessary for" (Coates 1983: 53). A corresponding proposed contrast is that between participant-oriented modality, which applies to a specific participant, and event-oriented modality, where the obligation stems from "general rules of conduct" (Hengeveld and MacKenzie 2008: 176). Neither the subjective-objective nor the strong-weak distinction allow the assessment of modal meaning and subject as independent factors, as the former has been bound to subject specificity, and the latter to subject person (Tagliamonte and Smith 2006: 365).

Cross-linguistically obligation is often expressed through auxiliaries grammaticalized to greater or lesser degrees from lexical verbs meaning 'owe', 'need', 'be fitting', or from general auxiliaries, as with English *have to*, in particular constructions (Bybee, Perkins and Pagliuca 1994: 181-184). Thus, another proposed distinction considers the source of the compelling conditions as internal vs. external to the participant, in a three-way classification between "need" vs. "necessity", with as a subtype of the latter, "deontic necessity", which may arise from practical circumstances or be imposed by some authority, whether a specific authority figure or social norms and decorum (van der Auwera and Plungian 1998: 80-81). In this view, source constructions from 'need' yield participant-internal modal meaning as with English *need to*, those from 'have', general participant-external necessity as with *have to*, and ones from 'owing' and 'duty', deontic necessity as with *should* (van der Auwera and Plungian 1998: 95).

The array of classification schemes, largely supported by cherry-picked or constructed examples, is indicative of the difficulty of implementing them in a replicable and exhaustive way. We introduce here a replicable meaning classification that is independent of both degree of strength and subjectivity understood as speaker authority, and is based instead on source of obligation, with a three-way distinction between *Hierarchical* authority / norms, *General* circumstances / conditions, and *Personal* choice / opinion.

Hierarchical obligation is due to an unequal power relationship, typically between employer and employee, teacher and pupil, or parent and child, where the hierarchical relationship provides the basis with which the superior demands something of the inferior, as in (1) above and (5) below. Included in this category are societal norms that may be due to precepts directed at a particular sector of the population that is minoritized or a minority, as with decorum for women and culture-specific mores, as in (6). In contrast, obligation related to General circumstances / conditions commonly expresses regulations or conventions that do not target a defined group, but standardly apply to everyone, as in processes followed at work seen in (7), at school (as in *have to go through school till .. they're eighteen* in (2) above), or in play (as with the rules of a game in (4) above). Situations that are rationally attributed to wider circumstances or that come about as logical consequences also belong to this second type, as in (8), where the person has to work *to support her family*. Third, obligation of the Personal choice / opinion type stems from personal preferences, such as a personal desire or wish, as in (9), personal limitations, such as with different learning styles in (10), or opinion or advice, as seen above in (3).

(5)
Simon: *maybe because their parents,*
           ... maybe it's how they're brought up to --
           .. too much pressure?
Brenda: *hm.*
Simon: *say,*
           <u>you have to go</u> to Greek school,
           ... <u>you have to do</u> this,
           maybe it's because of that,
           the pressure put onto them?

<div align="right">[SSDS_GTM_029, 29:35-29:46]</div>

(6)
Stuart: *like the Asian culture's like,*
           results results @results.
Jill: *... Hm.*
Stuart: *You know,*
           <u>you gotta get</u> the best and,
           you know,
           <u>you gotta get</u> the --
           .. um,
           certificates.
           <u>You gotta get</u> those awards.

<div align="right">[SydS_CYM_030, 01:01:59- 01:02:08]</div>

(7)
Aida: *.. And then <u>we have to keep track</u>,*
           .. of what's s- sold,
           and what's not sold.

<div align="right">[SydS_IOF_042, 49:08-49:12]</div>

(8)
Sophia: *because there's this .. moment in the film,*
           when this Bangladeshi lady who's a factory worker,
           ... and like,
           ... she's getting paid barely anything,
           but <u>she has to do</u> it,
           because she has to support her family.

<div align="right">[SydS_AYF_035, 50:25-50:32]</div>

(9)
Caterina: *I'm like,*
              I need to be here,
              .. It's my brother's wedding,
              next year,
              in May,
              I was like,
              I'm not missing that.

<div align="right">[SydS_IYF_111, 38:22-38:26]</div>

(10) (a teacher describing games she uses in class to help her students learn)
Federica: *some of them have to do,*
              .. have [to move] around,
Tania: *[Mhm].*
Federica: *.. in order to learn.*

<div align="right">[SydS_IOF_053, 44:30-44:33]</div>

Coding for obligation type is a matter of contextual interpretation of each instance, and relies on sociocultural familiarity with the community. There are sometimes congruent co-occurring cues, such as, for General obligation, *if* clauses, as in (11), or *because* clauses as in (8), and for Personal obligation, hedges or speaker stance frames such as *I think*, seen in (12). Mostly, however, we rely on stretches of discourse surrounding the clause of interest, taking account of the broader interaction, including contributions from the interlocutor. In (13), for example, we can establish that Sarah's comment that her brother "needs to" apply for more jobs is her personal opinion on the basis of her clarifying in previous and following talk that it is not a matter of general job market circumstances (as Simon suggests) but of her brother's disinterest.

(11)
Douglas: *Cause there's certain requirements.*
              If you join a car club,
              you gotta --
              ...(1.0) you know,
              do specific things and,

<div align="right">[SydS_AOM_097, 42:46-42:50]</div>

(12)
Naomi: *It's not going to happen soon according to our current Prime Minister.*
              .. unfortunately.
              ...(1.5) But I think,
              .. I think we need that to shift.
              I think we need to become a republic.
              ... I think there's a lot of ... unfinished business,
              .. in our .. national ... narrative.

<div align="right">[SydS_AOF_071, 01:16:50-01:17:05]</div>

(13)
Sarah: *and he's not real- that that motivated to find a a job.*
Simon: *yeah [yeah yeah].*
Sarah: *   [he's been apply]ing,*
           but not very --
Simon: *it's difficult.*
           the job market's crazy.
Sarah: *.. It is,*
           but he's not trying.
           he [needs to apply for] more jobs.
Simon: *   [yeah].*
           yeah,
Sarah: *[he just can't be bo]thered,*
Simon: *[yeah].*
Sarah: *he's having too much fun.*

[SydS_CYF_048, 48:55-49:07]

This three-way classification between Hierarchical, General and Personal obligation responds to the recognition of different sources of obligation, corresponding respectively to "some person or institution to whose authority [someone] submits; […] some more or less explicitly formulated body of moral or legal principles; […] no more than some inner compulsion" (Lyons 1977: 824). This classification also fits with a deontic necessity, participant-external, and participant-internal obligation typology (van der Auwera and Plungian 1998: 80-81) and the assignment of subfunctions such as "obligation motivated by social norms", "habitual obligation" and "obligation motivated by emotion" (Myhill 1995: 163). Above all, the three meaning types are sufficiently broad for replicable coding of large numbers of tokens in actual language use.

4   Counting modals

To identify the tokens of modals of obligation, we prioritize functional over structural criteria, recognizing that "the forms which enter into contextual, stylistic, or social complementarity of distribution do not generally originate as related syntactic structures" (Sankoff 1988a: 156). Thus, we set aside form-based divisions, such as between "core" or "central" and "semi-", "quasi-"or "emerging" modals (e.g., Krug 2000; Leech 2013: 95), and instead search the transcripts for modal forms with the broad function of expressing obligation. Extracting all occurrences of these forms yielded a total of 4,299 tokens, made up of the five forms we have noted so far, as well as three minor variants: *ought to* (N=2), *had better* (N=26)*,* and *be supposed to* (N=103).[2] We set these aside in order to be able to accurately identify usage patterns, and study here the five main modals: *must, have to, (have) got to, should, need to* (N=4,168).[3] We

---

2   *Be supposed to* has an obligation sense with dynamic verbs (58/103), e.g., *you're supposed to do sixteen subjects*, vs. a passive meaning retained with stative verbs, e.g., *people in Sydney are supposed to be tough or something* (Bybee, Perkins and Pagliuca 1994: 186).
3   *Need to* includes two instances of *need* (without *to*), both from the Bicentennial corpus: *you just did what needed doing*, *it need only be for three years*.

identify obligation uses (N=3,781) for which all five are competing variants (N=1,697), as explained below.

4.1 Modality functions

Besides that of expressing obligation, three other functions were identified. Epistemic modality indicates the speaker's degree of commitment to the truth of the proposition, as with *must* in (14), where the speaker judges the proposition to be probable (cf., Bybee, Perkins and Pagliuca 1994: 179-180; Coates 1983: 41-46, 64-65; van der Auwera and Plungian 1998: 81). A performative use occurs in several formulaic expressions with a variety of modals, such as *I must admit*, *I have to say, you've got to remember*, *I should imagine* (Coates 1983: 36; Myhill 1995: 171; Tagliamonte and D'Arcy 2007: 73-74). And "quasi-subjunctive" uses in some subordinate clauses are also found, specific to *should*, as in *funny you should say that* (Coates 1983: 67-69) (cf., Bybee, Perkins and Pagliuca 1994: 215-216).

> (14)
> Manuel: *people will say something in Italian and I'm like,*
>   *I know that word,*
>   *and I know that word.*
>   *.. so* <u>*they must be talking*</u> *about that.*
>
> [SydS_IYM_129, 32:52-32:56]

Figure 1 illustrates the proportion of obligation, epistemic, performative, and subjunctive uses, in the 1970s, on the left, and the 2010s, on the right. Epistemic uses are largely restricted to *must*, accounting for 68% of all instances of this modal in the 1970s and increasing to 86% in the 2010s (cf., Fehringer and Corrigan 2015: 360-361; Tagliamonte and D'Arcy 2007: 67-69). Performative uses are infrequent overall (and absent with *need to*), the highest proportion occurring with *must* (around 10% of *must* instances). Subjunctive uses of *should* are also minimal (10% for 1970s and 5% for 2010s). In sum, *must* stands out from the other forms in its low and shrinking obligation uses (from 20% to 6%), being largely limited to epistemic uses, and otherwise surviving in conventionalized performatives as entrenched interactional niches (cf., Bybee and Torres Cacoullos 2015: 144). Overall, these distributions confirm the primacy of obligation uses, which account for a full 90% of all instances (3,781/4,136).[4]

---

[4] 32 instances of unidentifiable modality were generally due to truncated or unclear speech (on *should* instances "indeterminate between Root and Epistemic", see Coates (1983: 77-78)). Of the minor variants, only *be supposed to* is used with epistemic modality (11% of the time).

Figure 1   Distribution of modalities for each form: 1970s (N=2,151) vs. 2010s (N=1,985):
            Obligation vs. Epistemic, Performative, Subjunctive)
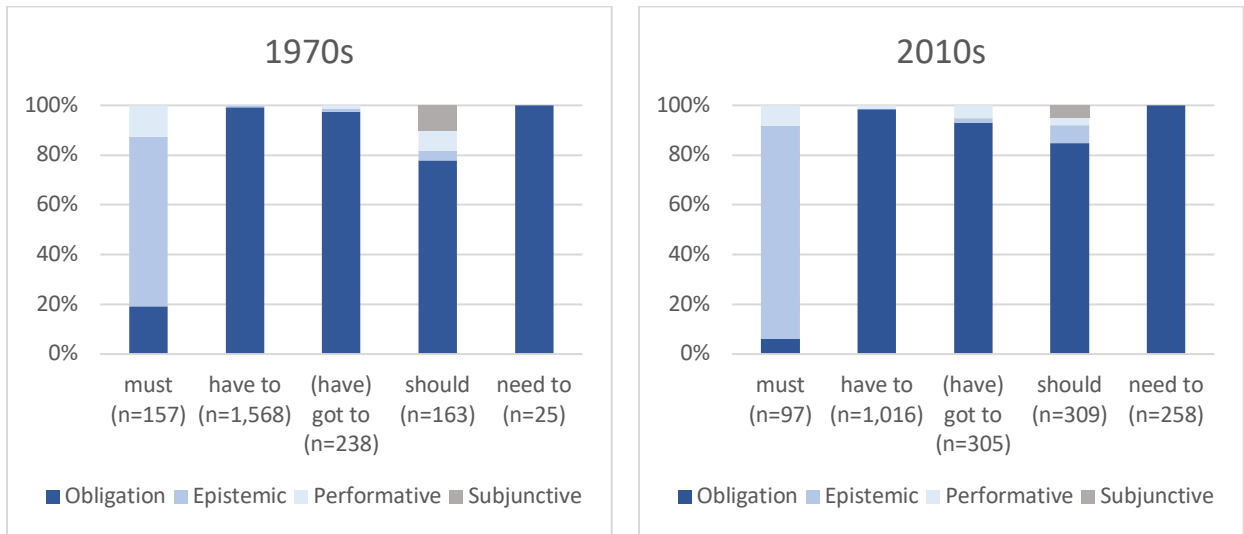
4.2   The variable context for obligation modality

The domain of obligation modality represents the initial broad definition of the linguistic variable, but speakers make choices between forms only within the variable context—where absolute differences between the forms that may exist in other contexts, whether in meaning or distribution, do not come into play (cf., Labov 1972: 72; Labov 2004: 7; Poplack and Torres Cacoullos 2015: 268-270). For obligation modals, the variable context is circumscribed by tense, polarity, and sentence type, because only *have to* occurs robustly in Past tense, negative polarity and interrogative contexts (e.g., Myhill 1995: 165-166; Tagliamonte and D'Arcy 2007: 60-63). Of non-Present instances (almost all simple Past), *have to* makes up 93% (1,756/1,898), with *should* (as *should have*) and *need to* together making up the remainder. While *have to, need to, should* and *must* all occur with negation (N=148/1,883 of all Present tense instances), the scope of the negation with *must* and *should* can lie only over the proposition, not the modality (Palmer 2003: 11), as illustrated in (15), which means that doing favors is a bad thing, not that you are not required to do favors (as would be conveyed by *you don't have to / don't need to*). And lastly, the interrogatives (N=38/1,735 Present tense positive polarity instances) are primarily made up of *have to* and *should*.

(15)
Deborah: *you shouldn't do any favours for friends or family,*
                                                                          [SydS_IOF_029, 32:40-32:42]

    Thus, though speakers may have choices between different variants in defined sub-contexts (e.g. *need to* and *have to* with negative polarity), the choice between all five modals is limited to simple Present, positive polarity declaratives. Barely one half of the instances of these modals expressing obligation occurs within this variable context (1,697/3,781). Were we to rely on overall frequency counts, we would obtain a skewed distribution of the modals, for example an

10

inflated frequency of *have to*, which accounts for over two thirds of all obligation modal instances (2,557/3,781), but just 42% of those in the variable context (707/1,697).

These 1,697 tokens in the variable context were coded for meaning, using the three-way classification described above. Following coding by both authors, cases that appeared not to be straightforward were discussed and resolved by further consideration of the broad discourse context and drawing on sociocultural familiarity with the speech community. Only a small number that occurred as part of a repair or were truncated remained unclassified for meaning (N=18), leaving 1,679 for analysis. Of these, 10% were used to refer to *Hierarchical* norms, 57% to *General* circumstances, and 33% to *Personal* choice.[5]

5 Modal frequencies and meaning correspondences

To obtain a general picture of the use of obligation modals over time, we first consider relative frequencies of forms in the 1970s and 2010s data, and then form-meaning correspondences, comparing relative frequencies of obligation meanings for each modal form.

Figure 2 shows the rate of each of the modal forms within the variable context in the two time periods, the 1970s in the darker shade and the 2010s in the lighter shade. We see here the drop in *must* referred to above (see Figure 1), from a minimal variant in the 1970s to virtually absent in the 2010s. *Have to* and *(have) got to* are the main variants in both time periods; *have to* shows greater reduction in use, from nearly one half down to one third of obligation modal instances, while *(have) got to* drops only slightly, in both periods accounting for one quarter to one third of the total. *Should* and *need to* are the two forms on the rise; for *should* this means that it retains its place within the set, while *need to*, barely present in the 1970s, accounts for almost one fifth of all instances in the 2010s.

Figure 2   Rate of occurrence of different modals of obligation in the variable context, N=1,679 (1970s N=680; 2010s N=999)
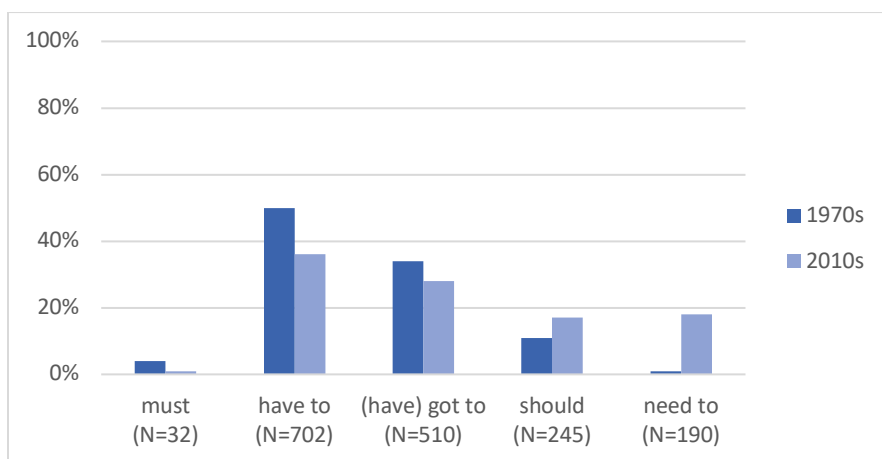
5   In total, more than 2,000 tokens were coded for meaning (all those in the variable context, and a sample of 456 tokens outside the variable context, primarily non-Present 1sg *have to* and *need to*). Outside the variable context the proportion of General obligation is higher, due to the predominance in the Past of *have to*, which is most associated with General obligation: 10% Hierarchical, 82% General, 8% Personal.

These trends echo reports from other varieties in the demise of *must*, though there may be community differences in the relative frequency of the other modals (cf. Collins 2005: 253). For the 'have' variants, for example, *(have) got to* holds up in Tyneside (England) but apparently has been nearly ousted by *have to* in Toronto (Canada) (Fehringer and Corrigan 2015: 369; Tagliamonte and D'Arcy 2007: 72). And *should* is reported to have increased again after a decline at the end of the 19th century in American English (Myhill 1995: 162).

The rapid rise of *need to* also matches other varieties. *Need to* was characterized as "relatively rare" in early 1990s British English conversation (Biber, Conrad and Reppen 1998: 209), and its frequency with respect to competing modals was reported at 8% in Toronto speech recordings of the early 2000s, a rate that is likely to be even lower were *should* included as a competing variant (Tagliamonte and D'Arcy 2007: 73-74). Real-time comparisons of the normalized frequency of *need to* suggest that it increased some two-fold in British and American written texts between 1961 and 1991 (Smith 2003: 248-249). Similarly, its rate in interviews roughly doubled between the early 1990s and late 2000s, from 6% to 16% in Tyneside (Fehringer and Corrigan 2015: 365) and 9% to 16% in Australian English (Penry Williams and Korhonen 2020: 272, 279).

We now turn to form-meaning correspondences over time. Figure 3 depicts the proportions of the three obligation meaning types (in different shades), for each of the five modals, in pairs of columns, with the 1970s on the left and the 2010s on the right (with *must* in the 2010s and *need to* in the 1970s lighter, due to sparse token counts). It is evident that there is no singular form-meaning correspondence, as each form hosts all meaning types. However, it is also the case that the modals have different foci and that these are stable over time: *must* is most associated with Hierarchical norms; *have to* and *(have) got to* with General circumstances; and *should* and *need to* with Personal choice. These foci may be explicable by cross-linguistic patterns of the source constructions (cf., Bybee, Perkins and Pagliuca 1994: 177-179; van der Auwera and Plungian 1998: 81).

Figure 3 Distribution of obligation meanings (Hierarchical norms, General circumstances, Personal choice) according to modal form in 1970s (N=680) vs. 2010s (N=999)
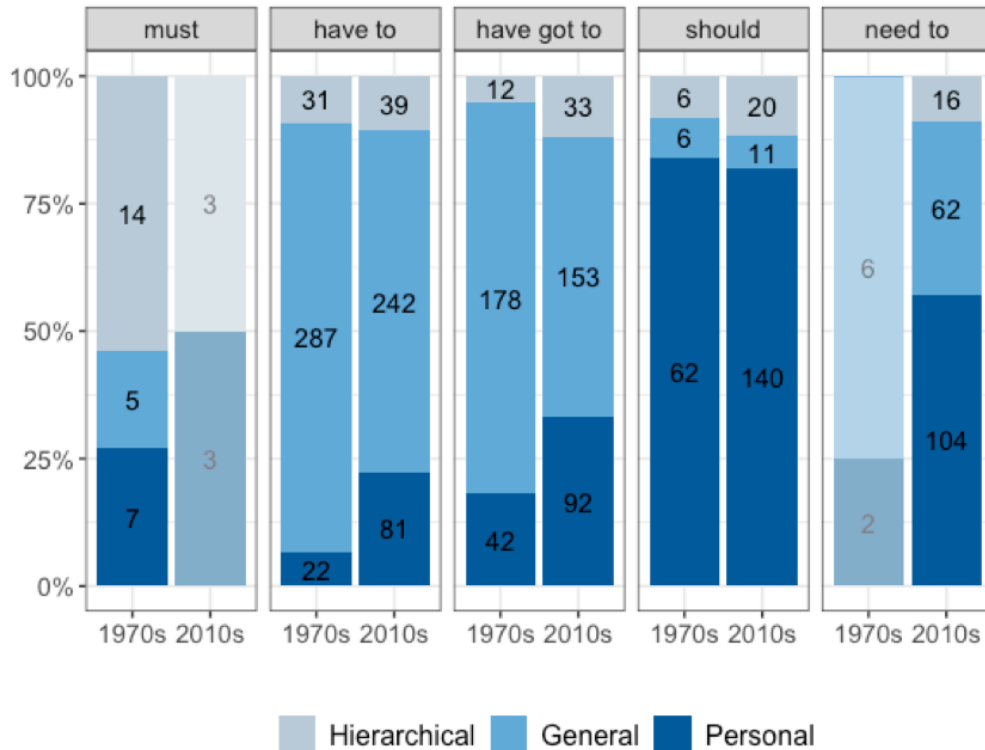
Though the meaning foci are stable, there has been an overall increase in the expression of Personal obligation from the 1970s to the 2010s (20% *vs.* 42%, *p*<0.0001, Fisher's exact test). To determine whether such an increase is attributable to the rise in *need to* and *should* or whether there is a broader, genuine change in the kind of obligation people talk about, we now turn to consider the conditioning, both of modal form and of obligation meaning.

6   Conditioning of modal forms

For the conditioning of modal form, we conduct regression analyses, including modal meaning as a predictor. Another linguistic predictor is Subject person and specificity, features that affect obligation meanings (Coates 1983: 36; Nokkonen 2006: 59-60; Tagliamonte and D'Arcy 2007: 66). We consider 1st and 2nd person together, as speech act participants, and contrast these with 3rd person, and with non-specific referents, which refer to any member of a class of entities (Torres Cacoullos and Aaron 2003: 306-308, and references therein). Social variables are Time period, Social class and Gender and, for the 2010s analyses, Age.

Mixed-effects logistic regression analyses were conducted using the lme4 package in R (Bates et al. 2019; R Core Team 2019), with random intercepts for Speaker and (main) Verb (pooling into a single level verb types that occur only once, or hapax legomena—approximately one half of all verb types, cf. Travis and Torres Cacoullos 2021: 5). We tested for interactions between the predictors, and removed both interactions and main effects that did not improve model fit, reporting here on the final models.

Rather than comparing all modals in a single model, we conduct independent analyses in defined sub-contexts where the variability is most robust. We thus examine the conditioning of the long-term majority variant, *have to*, first in relation to its main competitor, *(have) got to*, in both time periods, and second in relation to the newest form, *need to*, for the 2010s data only. A third analysis tests further the place of *need to* in the 2010s data, comparing it with *should*, as the two main expressions of Personal obligation.

Numbers of tokens and of speakers vary for the different analyses due to the particular variable context and the portion of data under study (whether both time periods, or only the 2010s). Participants produce an average of 6.5 tokens, ranging from none (19 speakers) to a maximum of 31 (one speaker) (Standard Deviation = 5.4). This means that, in each particular variable context, a different number of speakers produce no tokens, and therefore do not feature in that analysis. Although this is insufficient data for comparison of individuals, it does provide a large enough sample for identifying patterns across social groups, as we shall see.

6.1   *(have) got to* vs. *have to*

To understand the variation between *(have) got to* and *have to,* we consider together the two forms of *have got to*, with the auxiliary (*have got to,* N=213, virtually always contracted as *'ve got to* or *'s got to*, 206/213), and without (which we represent as *gotta*, N=297). In these data, *gotta* does not occur with 3sg subjects. In the variable context specific to *have got to* and *gotta* (that is, setting aside 3sg subjects), the two forms share similar obligation meaning distributions, justifying treating them as a single variant (Hierarchical obligation accounts for 8% of *have got to* and 9% of *gotta*; General, 60% and 65%, and Personal 31% and 25%). *Have got to* and *gotta* do, however differ socially, with *gotta* favored by Men, regardless of Social class, and by Working Class Women.[6]

Now comparing *(have) got to* vs. *have to*, we find both social and linguistic conditioning, as summarized in Table 2. Socially, *(have) got to* is favored by Men and by Working Class speakers. Though a significant interaction with Time period suggested that the Social class effect applied to the 2010s only, closer examination revealed these same class differences also operate in the 1970s for the Anglo Australians, but not the Italian and Greek Australians, consistent with different social conditioning for these groups with vocalic variables (Horvath 1985: 93). This social conditioning of *(have) got to* is not solely due to *gotta*, as comparisons between *have got to* and *have to* indicate a similar favoring by Men, though little social stratification.[7] *(Have) got to* is said to have originated as a "colloquial" or "nonstandard" variant in 19th century English literature but to have quickly reached the upper class (Krug 2000: 61-63). An association of *(have) got to* with (less educated) men has been observed in Toronto (Canada), Wheatley Hill (Northern Ireland) and York (England) (Tagliamonte and D'Arcy 2007: 78-79; Tagliamonte and Smith 2006: 369), though not most recently in Tyneside (England) (Fehringer and Corrigan 2015: 374).

As to the linguistic conditioning, there is a significant effect for Subject, such that *(have) got to* is disfavored in the third person (due to the absence of *gotta* with 3sg subjects), and favored

---

[6]   Rates according to gender and social class for *gotta* vs. *have got to* (setting aside 3sg subjects, which are categorically *have got to*) - Women (N=145): Middle Class (MC) 17%, Lower Middle Class (LMC) 31%, Working Class (WC) 63%; Men (N=310): MC 61%, LMC 72%, WC 76%.

[7]   For *have got to* (excluding *gotta*) vs. *have to* - Women (N=407): MC 18%, LMC 13%, WC 21%; Men (N=346): MC 29%, LMC 23%, WC 26%.

with Non-specific referents (primarily, *you* 76%, 591/774), which make up approximately two-thirds of the subjects in this analysis. In particular, there is a preponderance of non-specific *you* with *gotta*, seen in (16) (and in (4), (6) and (11) above). Non-specific *you* accounts for 70% (207/297) of all tokens of *gotta*, compared with 32% (69/213) of *have got to*, and 40% (667/1,679) for the data overall. Thus, non-specific *you gotta* is a candidate chunk, a complex unit of memory representation, which contributes to distribution patterns according to subject person and specificity (cf., Bybee 2010: 33-56).

(16)
Fiona:　　　*if you don't .. bring your sports uniform or your sandshoes,*
　　　　　　... or you don't wear the right shirt you know,
　　　　　　.. <u>you gotta run</u> round the oval till you drop?

[SSDS_ATF_159, 34:51-34:57]

Of particular interest to us here is the effect for Meaning, which exists in interaction with Time period, as depicted in Figure 4. It has been stated that *(have) got to* and *have to* differ according to aspect, whereby both occur in a future context, as in the constructed example, "I have to / I've got to get up at 7 a.m. tomorrow", but only *have to* in a habitual context, as in "I have to get up at 7 a.m. every day" (Coates 1983: 54) (cf., Myhill 1995: 172). This would predict a favouring of *have to* with General obligation, the meaning type most compatible with habitual aspect, since conditions that apply to people in general are habitual occurrences, customarily repeated and holding for extensive periods of time (cf., Comrie 1976: 27-28). There is no such favouring effect, however, and instead, the meaning difference between them is that *(have) got to* is favored for Personal obligation in the 1970s (as in the opinion *they've got to be treated like young adults* in (2)), consistent with earlier attribution of "obligation motivated by emotion" to this form (Myhill 1995: 175, 178). This favoring of *(have) got to* over *have to* for Personal obligation no longer applies in the 2010s, ascribable to the rise of *need to,* and to an overarching shift in obligation meanings, to which we will return.

Table 2   Output of logistic mixed effect model predicting *(have) got to* (vs. *have to*)

	Estimate	Std. Error	Z	p	N	% (have) got to
(Intercept)	-1.886	0.371	-5.08	**< 0.001**		
Time Period (ref: 1970s)					572	41%
2010s	0.376	0.331	1.136	= 0.256	640	43%
Meaning (ref: General)					860	38%
Hierarchical	-0.438	0.499	-0.877	= 0.380	115	39%
Personal	1.959	0.454	4.314	**< 0.001**	237	57%
Social Class (ref: Middle)					374	35%
Lower Middle	0.206	0.353	0.583	= 0.560	445	41%
Working	0.681	0.387	1.761	**= 0.078**	393	50%
Gender (ref: Women)					554	28%
Men	1.431	0.303	4.722	**< 0.001**	658	54%
Subject (ref: non-specific)					774	43%
1st and 2nd person	-0.229	0.230	-0.996	= 0.3194	315	46%
3rd person	-1.474	0.330	-4.475	**< 0.001**	123	26%
Time Period x Meaning						
1970s General					465	38%
1970s Hierarchical					43	28%
1970s Personal					64	66%
2010s General					395	39%
2010s Hierarchical	1.131	0.628	1.801	= 0.072	72	46%
2010s Personal	-1.113	0.535	-2.082	**< 0.05**	173	53%

Overall rate of *(have) got to* = 42% (510/1,212)
225 speakers, variance = 2.72 (SD = 1.65); 159 verb types, variance = 0.09 (SD = 0.29)
Log likelihood: -648.6; AIC: 1323.1; BIC: 1389.4

Figure 4   Predicted rate of *(have) got to* (vs. *have to*) for Meaning by Time Period
               (from model in Table 2)
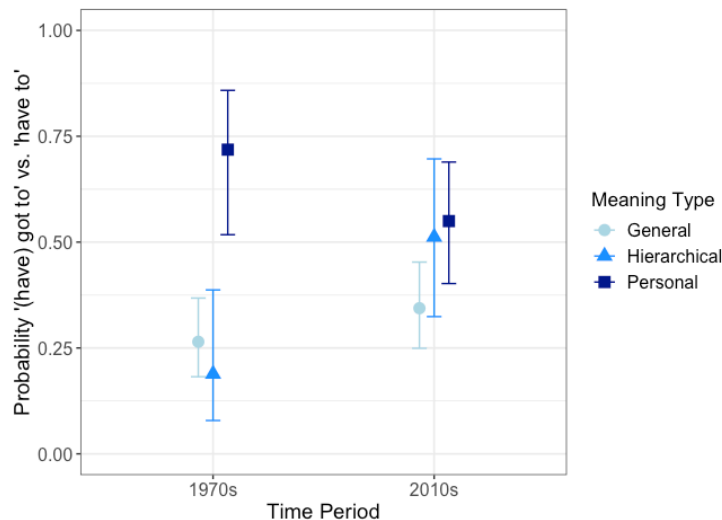
6.2   *Need to* vs. *have to*: 2010s

Let us now turn to the incoming form. *Need to* is virtually absent in the 1970s (8 tokens in the variable context, Figure 2), and therefore we compare *need to* with *have to* just for the 2010s data. Though *need to* remains restricted by tense and sentence type (with rates of *need to* (vs. *have to*) of 8% (51/620) in non-Present tense contexts and 6% (1/16) in interrogatives), there is variability under negation (28%, 21/74). However, while the relevant variable context for *need to* vs. *have to* extends to negative polarity instances, polarity is not significant in the model (Table 3), suggesting some generalization of *need to*, as negative polarity contexts are usually conservative (cf., Torres Cacoullos 2012: 106).

     Unlike *(have) got to*, *need to* shows no social conditioning in variation with *have to*, as neither Gender nor Social class has a significant effect (cf., Fehringer and Corrigan 2015: 374, on a congruent finding in Tyneside). It is notable that neither is Age significant, with only a small increase in the rate of *need to* vs. *have to* from the 2010s Adults to the Young Adults (Adults: 28%, 59/209; Young Adults: 35%, 144/409). Considering *need to* as a proportion of all modals, the evident real time difference (1970s: 1%, 2010s: 18%, Figure 2) and smaller apparent time difference (Adults: 13%, 57/435; Young Adults; 22%, 125/564) indicates that the 2010s Adults are less like their same birth year 1970s counterparts and more like their time-period counterparts, the Young Adults. This suggests that *need to* may be a case of "communal change", with uniform adoption in the community (cf., Labov 1994: 84), or a "Zeitgeist" effect characterizing change during a specific period of time (Fruehwald 2017: 22).

     Subject is configured here by collapsing 3rd person and non-specific subjects since there was no difference between these (unlike for *have to* vs. *(have) got to*). "Generic" subjects with *need to* have been previously considered to have "not very subjective" meaning as opposed to an "internal compulsion" meaning with 1sg subjects (Nokkonen 2006: 59-61), or more generally have been taken to "preclude subjective readings" (Tagliamonte and D'Arcy 2007: 65-66). Here there is no categorical overlap between non-specific subjects and obligation meaning (cf., Sankoff 1988b: 986 on dependence among predictors), though non-specific subjects are, as we

may expect, heavily used with General obligation (71%, 668/937). For example, (17) illustrates a non-specific subject with Personal obligation, as an opinion is expressed, and (18) (and (6) above) illustrates a non-specific subject under Hierarchical obligation. Importantly, there is a significant interaction between Meaning and Subject, visualized in Figure 5, where we see that the newer modal *need to* is most favored in Personal obligation contexts—especially for 1st and 2nd person, expressing preference (as in (9) above) and giving advice (in (3) above).

(17)
Tessa: *it is something that we will do,*
            ...(1.5) when we're a bit older,
            we kind of are of the --
            of the opinion that --
            ...(1.5) <u>you need to use</u> your travel legs while you've got them?
                                                            [SydS_AOF_006, 51:29-51:39]

(18)
Carmel: *and so there was like a --*
            .. a --
            .. a feeling of,
            you know,
            <u>you need to be punished</u>,
            because,
Stella: *hm.*
Carmel: *you didn't go to church on Sunday.*
                                                            [SydS_GOF_142, 32:27-32:35]

Table 3   Output of logistic mixed effect model predicting *need to* (vs. *have to*), 2010s

|	Estimate	Std. Error	Z	p	N	*% need to*
(Intercept)	-1.761	0.285	-6.184	**< 0.001**		
Meaning (ref: General)					362	22%
Hierarchical	-0.715	0.757	-0.945	= 0.345	60	28%
Personal	1.254	0.345	3.632	**< 0.001**	196	55%
Subject (ref: 3rd and non-specific)					360	28%
1st and 2nd person	-0.278	0.355	-0.783	= 0.434	258	40%
Subject x Meaning						
3 & nonspec x General					255	23%
3 & nonspec x Hierarchical					18	17%
3 & nonspec x Personal					87	44%
1 & 2 x General					107	19%
1 & 2 x Hierarchical	1.999	0.932	2.146	**< 0.05**	42	33%
1 & 2 x Personal	1.431	0.536	2.668	**< 0.001**	109	64%

Overall rate of *need to* = 33% (203/618)
122 speakers, variance = 1.13 (SD = 1.06); 109 verb types, variance = 0.95 (SD = 0.97)
Log likelihood: -339.0; AIC: 693.9; BIC: 729.3
(no significant effect for Gender, Social class, Age, Negation)

Figure 5   Predicted rate of *need to* (vs. *have to*) for Meaning by Subject (from model in Table 3)
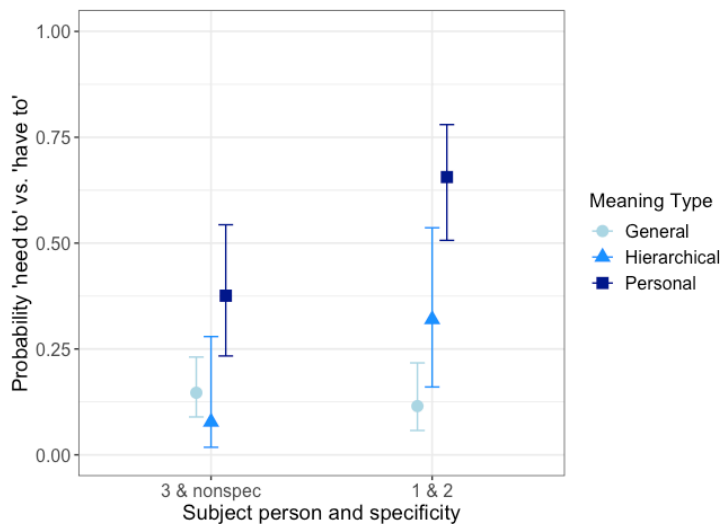
6.3   *Need to* vs. *should*: 2010s, Personal obligation

Long-standing *should* is reported to have shifted in meaning between the 19th and 20th century in American English, from "general principles applying to generic subjects", as in *a girl should be a lady* ((21) below), to "future good idea" (what we code as Personal), as in *you should drink more water* ((3) above) (Myhill 1995: 175, 178). In the data examined here, *should* is the major

expression of Personal obligation in the 1970s and continues in the 2010s as the modal with the greatest proportion of Personal meaning (82%, 140/171 of *should* tokens; Figure 3). In this, it is similar to *need* to, which we have seen differs from *have to* mainly by being favored in Personal obligation contexts. We can therefore discover what conditions the use of *need to* when expressing Personal obligation by comparing it with *should* (which has been previously placed outside the variable context due to its categorization as a modal of "weak" obligation). We thus analyze these two modals in the 2010s in Personal obligation contexts only (Table 4).

Need to in variation with *should* in Personal contexts displays no significant effect for Age (with only a small difference, between the Adults and the Young Adults, 37% vs. 45%), as was also the case for *need to* vs *have to*. Here, though, there is an effect for Gender in interaction with Social Class. This is depicted in Figure 6, which shows a lack of social stratification by women, but a favoring of *need to* by Middle over Lower Middle and Working class men. *Need to* is thus most favored by women and by Middle Class men. There is also a compatible ethnic difference here, in that Chinese Australians favor *need to* vs. *should*, which is consistent with a general adoption by this group of variants favored by the Middle Class and women; they similarly favor *have to* over *(have) got to* (e.g., Grama, Travis and Gonzalez 2021: 307-309; Travis, Grama and Purser To Appear).

Linguistically, there is an effect for subject corresponding with that seen in variation with *have to,* confirming that *need to* is favored with 1st and 2nd person subjects (as in (19)). The bulk of 1st and 2nd person subjects are 1sg (74%, 105/142) (which is the case in the data overall; 78%, 412/528). As 1sg subjects mark linguistic subjectivity, co-occurrence of *need to* with 1sg would be a measure of the expression of speaker involvement, in accordance with a source meaning of internal need (Aaron and Torres Cacoullos 2005: 616). Non-specific subject referents (as in (17) above) are not significantly different from 1st and 2nd person, favoring *need to*, and therefore disfavoring *should*. This is consistent with the earlier meaning shift away from general principles (associated with generic subjects, as in *a girl should be a lady*) that has been conjectured for *should*.

(19)
Sophia: *I've never really been like,*
              I need to find someone.
              I need to get married.
Heather: *... That's [good].*
Sophia: *[Like that's never been] a priority in my life?*

                                                    [SydS_AYF_035, 47:47-47:53]

Should is favored by 3rd person subjects. 3sg subjects with *should* are mostly inanimate, and often co-occur with a 1sg speaker stance frame (Thompson 2002), such as *I think* (20). The regular co-occurrence of 3rd person *should* with such discourse formulas suggests the availability of a construction of the form [*I think / feel like* + 3rd person *should*] and a meaning of personal opinion (cf. Myhill 1995: 178 on co-occurrence of *should* with hedging devices).

(20)
Arthur: *I think the shops should close on Boxing Day.*

                                                    [SydS_AOM_120, 54:37-54:39]

Table 4   Output of logistic mixed effect model predicting *need to* (vs. *should*), 2010s,
         Personal meaning only

|	Estimate	Std. Error	Z	p	N	*% need to*
(Intercept)	0.319	0.694	0.46	=0.645		
Subject (ref: 1&2)					142	47%
Nonspecific	-0.428	0.496	-0.863	= 0.388	59	44%
3	-2.001	0.655	-3.054	**< 0.001**	43	26%
Gender (ref: Women)					139	49%
Men	1.053	1.024	1.029	= 0.304	105	34%
Social Class (ref: Middle)					79	44%
Lower Middle	0.468	0.908	0.515	= 0.607	111	42%
Working	0.423	1.082	0.391	= 0.695	54	41%
Gender x Social Class						
Women Middle					48	40%
Women Lower Middle					60	50%
Women Working					31	61%
Men Middle					31	52%
Men Lower Middle	-2.561	1.425	-1.797	**= 0.072**	51	33%
Men Working	-4.334	1.808	-2.397	**< 0.05**	23	13%

Overall rate of *need to* = 43% (104/244); 87 speakers, variance = 3.77 (SD = 1.942)[8]
Log likelihood: -138.6; AIC: 295.3; BIC: 326.8
(no significant effect for Age)

---

[8]   It was not possible to run Verb as a random effect, as out of a total of 62 verb types, 36 occurred only
      once in this subset, and only 18 exhibited variability. Verb aspect and class have been said to affect
      need to readings (cf., Nokkonen 2006: 59-60); we found a (non-significant) tendency toward favoring
      need to over *should* with dynamic verbs, and stative *be* was particularly associated with *should.*

Figure 6   Predicted rate of *need to* (vs. *should*) 2010s Personal meaning
            for Social Class by Gender (from model in Table 4)
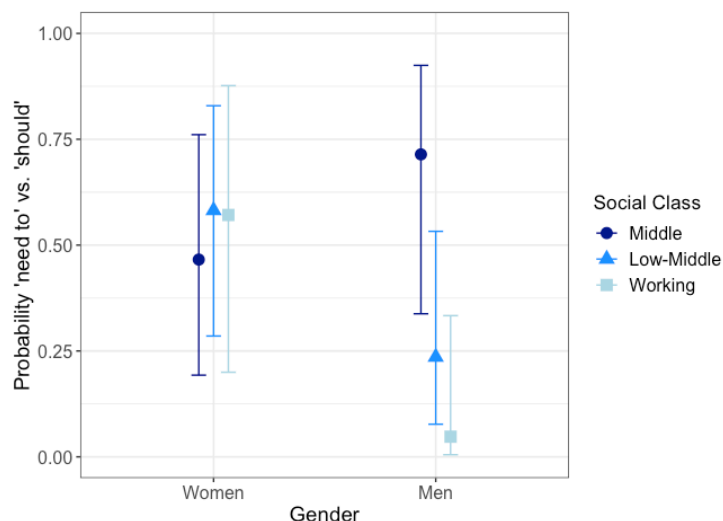
7   Conditioning of modal functions

In the overview of modal frequencies and meaning correspondences we saw that obligation modals have distinct meaning foci, and that the modal most associated with Hierarchical obligation, *must*, has virtually disappeared, while those that most express Personal obligation—*need to* and *should*—have grown in relative frequency. This does not tell us, however, whether there has been a genuine shift in meaning across the obligation modal system, or whether an increase in Personal obligation is simply due to the increase in *need to* and *should*. To answer this, we conduct an analysis with obligation Meaning as the dependent variable (predicting Personal, in contrast to General and Hierarchical obligation), and incorporate Modal form as a predictor.[9] Also included are Subject and the extralinguistic variables of Social class, Gender and Time period. Results are given in Table 5.

Personal obligation is favored with the modals *should* and *need to*; with 1st and 2nd person subjects; and in the 2010s. Though the model returns a significant interaction between Meaning and Time Period, this is due specifically to the behavior of *should*. As depicted in the visualization of this interaction in Figure 7, the greater expression of Personal obligation in the 2010s than in the 1970s applies to all modal forms, including the main modal forms *have to* and *(have) got to*. It does not apply to *should* because this form is already used to express Personal obligation at a very high rate in the 1970s. We should not be misled by the lack of a main effect for *need to*, which is due to the lack of token numbers in the 1970s (N=8, Figure 3), and represented in the large error bars in Figure 7 (as is also the case for *must* in the 2010s).

---

[9]   Frequency shifts in General and Hierarchical obligation depend on subject type. Overall, General obligation declines over time (Personal: 20% 1970s vs. 42% 2010s, $p<0.0001$; General: 71% vs. 47%, $p<0.0001$; Hierarchical 9% vs. 11%, $p=0.254$; 1970s N=680; 2010s N=999), while for specific-referent subjects only, the greatest decline is with Hierarchical obligation (Personal: 34% vs. 50%, $p<0.001$; General 45% vs. 36%, $p=0.072$; Hierarchical 21% vs. 13%, $p<0.05$; 1970s N=146; 2010s N=596).

This across-the-board rise is evidence that the increased talk of Personal obligation is independent of the decline in *must* and upsurge of *need to*. Here, then, is a case of "language change [...] affected by changes in the relative frequencies with which certain functions are expressed" (Myhill 1995: 206), or a genuine change in discourse.

Table 5   Output of logistic mixed effect model predicting Personal vs. General and Hierarchical meaning

	Estimate	Std. Error	z value	*p*	N	% Personal
(Intercept)	-3.485	0.329	-10.581	**< 0.001**		
Time period (ref: 1970s)					680	20%
2010s	1.393	0.351	3.967	**< 0.001**	999	42%
Modal (ref: *have to*)					702	15%
have got to	1.251	0.359	3.48	**< 0.001**	510	26%
must	1.954	0.655	2.982	**< 0.01**	32	31%
need to	1.464	1.069	1.369	=0.171	190	56%
should	5.057	0.496	10.196	**< 0.001**	245	82%
Subject (ref: 3 & non-specific)					1151	26%
1st & 2nd person	0.985	0.167	5.91	**< 0.001**	528	49%
Time Period x Modal						
2010s x *have to*					(see Figure 3 for Ns)	
2010s x *have got to*	-0.575	0.43	-1.337	= 0.181		
2010s x *must*	-0.837	1.219	-0.687	= 0.492		
2010s x *need to*	0.3164	1.096	0.289	= 0.772		
2010s x *should*	-1.849	0.555	-3.329	**< 0.001**		

Overall rate of Personal Obligation vs. General & Hierarchical = 33% (555/1,679)
240 speakers, variance = 1.29 (SD = 1.13); 173 verb types, variance = 0.292 (SD = 0.54)
Log likelihood: -754.5; AIC: 1535.1; BIC: 1605.6
(no significant effect for Gender, Social class)

Figure 7   Predicted rate of Personal obligation (vs. General and Hierarchical) for Time Period by
            Modal form (from model in Table 5)
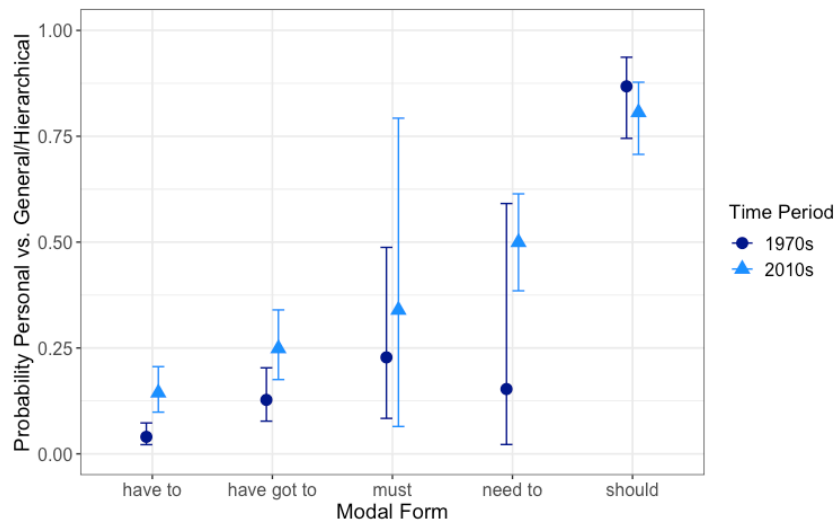
   As to social differences in modal meaning, there may be a tendency toward greater
expression of Personal obligation by women relative to men in the 2010s only (2010s women
46%, 244/536; men 38%, 176/463, *p*<0.05 by Fisher's exact test; 1970s: women 20%, 60/294;
men 19%, 75/386, *p*=0.771). Remember that we saw a Gender difference in interaction with
Social class in the regression analysis for *need to* vs. *should*, where *need to* is favored by women
(but for men, Middle class only; Table 4 and Figure 6). Though not achieving significance in the
regression analysis here, such a tendency would indicate that the change toward more talk of
Personal than other types of obligation was being led by women. This would be expected from
reports of other grammatical changes related to interpersonal relations. For example, such a
gender asymmetry has been invoked for changes in Dutch causatives between the 18th and 20th
centuries, which have been linked to a decrease in the social role of interpersonal authority
(Verhagen 2000: 276). Gender differences have also appeared with the rise of the English
progressive in the 19th c., which has been correlated with the psychosociological factor of
interlocutor intimacy (Arnaud 1998: 142).
   A leading role for women in both the formal change (using *need to*) and functional change
(expressing Personal choice / opinion) in English modals of obligation would also be consistent
with events of the second half of the 20th century such as the civil rights and women's
movements. Asymmetries in the role of authority according to gender was a topic raised by some
speakers, as in (21), from an 88-year-old woman born in 1900, talking about her father's "austere
outlook" toward women. Data from disciplines other than linguistics would be needed to
demonstrate that social changes or movements have impacted interpersonal relations. On that
foundation, linguists could then probe systematic correlations between such language-external
change and linguistic factors or features of the interaction between speakers (cf., Myhill 1995:
201).

(21)

Gladys: *I had a h- father who was ... very austere in his outlook.*
              ...(1.5) a --
              .. <u>a girl should be a lady</u>,
              she shouldn't be permitted to ... soil her fingers --
              ...(1.0) for pay,

[Bcnt_AEF_079, 01:02:15-01:02:27]

8   Conclusion

Variability among English modals of obligation is long-standing, with renewal of forms over the centuries via grammaticalization. The newest form, *need to,* appears to have been incorporated into the system of modals of obligation very rapidly, within one generation in Australia, as a communal change adopted uniformly (cf., Labov 1994: 84). Thus, while it is possible that the rise of *need to* is a language-internal change, it could also be a change from above, originating outside of the community, as with widely discussed changes in English quotatives and intensifiers (Lee 2020; Plehwe and Travis 2020; Tagliamonte, D'Arcy and Rodríguez Louro 2016). Determining the mechanism and source of the change, whether internal or diffusion from one community into another, or both (cf., Arnaud 1998: 144) calls for real and apparent-time studies and identification of the leaders of change within the community (cf., Labov 2018). Comparisons of the social and linguistic conditioning of obligation modals across communities can advance this work by attending to motivated variable contexts and to obligation meanings expressed. This would involve taking into consideration the distinct sub-contexts in which different modals compete; including *should* as a competing variant; and classifying obligation meaning in a way that is replicable and independent of subject specificity-person. Now is the ideal time to conduct such studies, while this change is still recent or unfolding.

    English modals of obligation provide a good demonstration of form-function covariation: shifts in function—the relative frequency of obligation meanings expressed—and shifts in form—the relative frequency of the modals expressing them. The changing of the guard of modals has been widely observed; what we have additionally shown is change in the nature of the obligation expressed, or the types of obligation speakers spend more time talking about. Whether this is connected to changes in interpersonal relations remains to be investigated, but the variation patterns discovered here demonstrate increased talk of obligation stemming from personal preferences, not solely attributable to the rise of *need to*. Change in modal function is *concomitant with, but independent of*, change in modal form.

Appendix: Transcription Conventions (Du Bois et al. 1993)

Carriage return	new Intonation Unit
.	final intonation contour
,	continuing intonation contour
--	truncated intonation contour
-	truncated word
..	short pause (≈ 0.2 secs)
...	medium pause (≈ 0.7 secs)
...(N.0)	timed pause (1 sec or longer)
[ ]	overlapped speech
@word	word uttered while laughing
<@ @>	speech uttered while laughing
Capital initial letter	higher initial pitch level

References

Aaron, Jessi Elana and Rena Torres Cacoullos. 2005. Quantitative measures of subjectification: A variationist study of Spanish *salir(se)*. *Cognitive Linguistics* 16(4): 607-633. https://doi.org/10.1515/cogl.2005.16.4.607

Arnaud, René. 1998. The development of the progressive in 19th century English: A quantitative survey. *Language Variation and Change* 10: 123-152. https://doi.org/10.1017/S0954394500001265

Bates, Douglas, Martin Mächler, Ben Bolker, Steve Walker, Rune Haubo Bojesen Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, Gabor Grothendieck, Peter Green and John Fox. 2019. lme4: Linear mixed-effects models using 'Eigen' and S4. R package, version 1.1-121.

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

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan and Rena Torres Cacoullos. 2015. Grammaticalization and variation of *will* and *shall* in Shakespeare's comedies. In Rena Torres Cacoullos, Nathalie Dion, and André Lapierre (eds), *Linguistic variation: Confronting fact and theory*, 131-146. New York / London: Routledge.

Coates, Jennifer. 1983. *The semantics of the modal auxiliaries*. London: Croom Helm.

Collins, Peter C. 2005. The modals and quasi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide* 26(3): 249-273. https://doi.org/10.1075/eww.26.3.02col

Comrie, Benrard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Du Bois, John W., Stephan Schuetze-Coburn, Susanna Cumming and Danae Paolino. 1993. Outline of discourse transcription. In Jane Edwards, and Martin Lampert (eds), *Talking data: Transcription and coding in discourse*, 45-89. Hillsdale: Lawrence Erlbaum Associates.

Fehringer, Carol and Karen 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(2): 355-381. https://doi.org/10.1017/S1360674315000131

Fruehwald, Josef. 2017. Generations, lifespans, and the zeitgeist. *Language Variation and Change* 29(1): 1-27. https://doi.org/10.1017/S0954394517000060

Grama, James, Catherine E. Travis and Simon Gonzalez. 2020. Ethnolectal and community change ov(er) time: Word-final (er) in Australian English. *Australian Journal of Linguistics* 40(3): 346-368. https://doi.org/10.1080/07268602.2020.1823818

Grama, James, Catherine E. Travis and Simon Gonzalez. 2021. Ethnic variation in real time: Change in Australian English diphthongs. In Hans Van de Velde, Nanna Haug Hilton, and Remco Knooihuizen (eds), *Studies in Language Variation*, 292-314. Amsterdam: John Benjamins. https://www.jbe-platform.com/content/books/9789027259820-silv.25.13gra

Hengeveld, Kees and J. Lachlan MacKenzie. 2008. *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure*. Oxford: Oxford University Press.

Horvath, Barbara. 1985. *Variation in Australian English: The sociolects of Sydney*. Cambridge: Cambridge University Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Krug, Manfred. 2000. *Emerging English modals: A corpus study of grammaticalization*. Berlin and New York: Mouton de Gruyter.

Labov, William. 1972. *Sociolinguistic patterns*. Oxford: Basil Blackwell.

Labov, William. 1982. Building on empirical foundations. In Winfred P. Lehmann, and Yakov Malkiel (eds), *Perspectives on Historical Linguistics*, 17-92. Amsterdam: John Benjamins.

Labov, William. 1994. *Principles of linguistic change: Internal factors*, vol. 1, 3 vols. Oxford: Basil Blackwell.

Labov, William. 2004. Quantitative analysis of linguistic variation. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, and Peter Trudgill (eds), *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society*, vol. 1 (2nd ed.), 6-22. Berlin: Mouton de Gruyter.

Labov, William. 2008. Quantitative reasoning in linguistics. https://www.ling.upenn.edu/~wlabov/Papers/QRL.pdf

Labov, William. 2018. The role of the Avant Garde in linguistic diffusion. *Language Variation and Change* 30(1): 1-21. https://doi.org/10.1017/S0954394518000042

Lee, Esther. 2020. Quotatives over time: A study in ethnic variation. Honours thesis, School of Literature, Languages and Linguistics, Australian National University.

Leech, Geoffrey. 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*, 95-115. Berlin and New York: Mouton de Gruyter.

Lyons, John. 1977. *Semantics*, 2 vols. Cambridge: Cambridge University Press.

Myhill, John. 1995. Change and continuity in the functions of the American English modals. *Linguistics* 33: 157-211. https://doi.org/10.1515/ling.1995.33.2.157

Nokkonen, Soili. 2006. The semantic variation of *need to* in four recent British English corpora. *International Journal of Corpus Linguistics* 11: 29-71. https://doi.org/10.1075/ijcl.11.1.03nok

NSW Bicentennial Oral History Project. 1987. *NSW Bicentennial oral history collection*. Council on the Ageing NSW Branch and NSW Oral History Association of Australia, housed at the National Library of Australia.

Palmer, Frank. 1986. *Mood and modality*. Cambridge: Cambridge University Press.

Palmer, Frank. 2003. Modality in English: Theoretical, descriptive and typological issues. In Roberta Facchinetti, Frank Palmer, and Manfred Krug (eds), *Modality in Contemporary English* 1-18. Berlin, Boston: Mouton de Gruyter. https://doi.org/10.1515/9783110895339.1

Penry Williams, Cara and Minna Korhonen. 2020. A sociolinguistic perspective on the (quasi-)modals of obligation and necessity in Australian English. *English World-Wide* 41(3): 267-294. https://doi.org/10.1075/eww.00051.pen

Plehwe, Renate and Catherine E. Travis. 2020. Variation and change over 60 years of intensification in Australian English. Paper presented at the Australian Linguistic Society Annual Conference, 2020, Griffith University.

Poplack, Shana and Elisabete Malvar. 2007. Elucidating the transition period in linguistic change: The expression of the future in Brazilian Portuguese. *Probus* 19: 121-169. https://doi.org/10.1515/PROBUS.2007.005

Poplack, Shana and Rena Torres Cacoullos. 2015. Linguistic emergence on the ground: A variationist paradigm. In Brian MacWhinney, and William O' Grady (eds), *The Handbook of Language Emergence*, 267-291. Malden, MA: Wiley-Blackwell.

R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, http://www.R-project.org

Sankoff, David. 1988a. Sociolinguistics and syntactic variation. In Frederick Newmeyer (ed), *Linguistics: The Cambridge survey (Vol. 4, Language: The socio-cultural context)*, 140-161. Cambridge: Cambridge University Press.

Sankoff, David. 1988b. Variable rules. In Ulrich Ammon, Norbert Dittmar, and Klaus J. Mattheier (eds), *Sociolinguistics: An international handbook of the science of language and society*, vol. 2, 984-997. Berlin: Walter de Gruyter.

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* 241-266. Berlin, Boston: Mouton de Gruyter. https://doi.org/10.1515/9783110895339.241

Szmrecsanyi, Benedikt. 2016. About text frequencies in historical linguistics: Disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory* 12(1): 153-171. https://doi.org/10.1515/cllt-2015-0068

Tagliamonte, Sali and Alexandra D'Arcy. 2007. The modals of obligation/necessity in Canadian perspective. *English World-Wide* 28(1): 47-87. https://doi.org/10.1075/eww.28.1.04tag

Tagliamonte, Sali, Alexandra D'Arcy and Celeste Rodríguez Louro. 2016. Outliers, impact, and rationalization in linguistic change. *Language* 92(4): 824-849.

Tagliamonte, Sali and Jennifer Smith. 2006. Layering, competition and a twist of fate: Deontic modality in dialects of English. *Diachronica* 23(3): 341-380. https://doi.org/10.1075/dia.23.2.06tag

Thompson, Sandra A. 2002. 'Object Complements' and conversation: Towards a realistic account. *Studies in Language* 26(1): 125-163. https://doi.org/10.1075/sl.26.1.05tho

Torres Cacoullos, Rena. 2012. Grammaticalization through inherent variability: The development of a progressive in Spanish. *Studies in Language* 36(1): 73-122. https://doi.org/10.1075/sl.36.1.03tor

Torres Cacoullos, Rena and Jessi Elana Aaron. 2003. Bare English-origin nouns in Spanish: Rates, constraints and discourse functions. *Language Variation and Change* 15(3): 289-328. https://doi.org/10.1017/S0954394503153021

Torres Cacoullos, Rena and Catherine E. Travis. 2018. *Bilingualism in the community: Code-switching and grammars in contact*. Cambridge: Cambridge University Press.

Torres Cacoullos, Rena and Catherine E. Travis. 2020. Code-switching and bilinguals' grammars. In Evangelina Adamou, and Yaron Matras (eds), *The Routledge Handbook of Language Contact*, 252-274. London/New York: Routledge.

Travis, Catherine E. 2014-2022. *Sydney Speaks*. Australian Research Council Centre of Excellence for the Dynamics of Language, Australian National University. https://dx.doi.org/10.25911/m03c-yz22

Travis, Catherine E., James Grama and Benjamin Purser. To Appear. Stability and change in (ing): Ethnic and grammatical variation over time in Australian English. *English World-Wide*.

Travis, Catherine E. and Amy M. Lindstrom. 2016. Different registers, different grammars? Subject expression in English conversation and narrative. *Language Variation and Change* 28(1): 103-128. https://doi.org/10.1017/S0954394515000174

Travis, Catherine E. and Rena Torres Cacoullos. 2021. Categories and frequency: Cognition verbs in Spanish subject expression. *Languages (Special Issue 'Revisiting Language Variation and Change: Looking at Metalinguistic Categories Through a Usage-Based Lens', eds. Brown & Rivas)* 6(3): 126. https://doi.org/10.3390/languages6030126

van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2(1): 79-124. https://doi.org/10.1515/lity.1998.2.1.79

Verhagen, Arie. 2000. Interpreting usage: Construing the history of Dutch causal verbs. In Michael Barlow, and Suzanne Kemmer (eds), *Usage-based models of language*, 261-286. Stanford University, Palo Alto: CSLI Publications.