Jo Close (j.close@ucl.ac.uk) & Bas Aarts (b.aarts@ucl.ac.uk), UCL
ISLE, Freiburg, 8th October 2008

In this paper we investigate recent change in the use of the semi-modals *HAVE TO* and *HAVE GOT TO* and the core modal *MUST* in a corpus of spoken English, with the aim of answering the following questions:

Is there support for the idea that core modals are decreasing?

If *MUST* is found to be decreasing, can the decrease be related to (an increase in) the use of ɪfij a î ge i-modal forms?

Is there support for the theory that modals are becoming monosemous?

What is driving the change in the use of the three forms?

We will compare the results from the spoken data with results from written corpora (LOB and F-LOB) presented in Leech et al. (forthcoming).

A further aim of this work is to contribute to the understanding of recent change, an aim we share with Mair, Hundt, Leech and Smith, the authors of the forthcoming CUP book *Change in contemporary English*.

The *Diachronic Corpus of Present-Day Spoken English* (DCPSE) contains around 400,000 words from the British component of the *International Corpus of English* (ICE-GB) collected in the early 1990s and around 400,000 words from the *London-Lund Corpus* (LLC) collected between the late 1960s and early 1980s (http://www.ucl.ac.uk/english-usage/projects/dcpse/). Unlike the *FLOB* and *FROWN* corpora, compiled by Christian Mair at Freiburg, DCPSE contains only spoken English, because spoken English is where changes are likely to occur first.

ɪS e daha [Å] afe hhe gihe f figi f a gh a -prestige i j ahi g i a gi ageî (Mair, forthcoming 2008).

All the sentences in DCPSE have been grammatically analysed and have been given a detailed parse tree.

PU	CL		SU	NP		NPHD	N
main	cop pres					com sing	

Business

VB	VP		MVB	V
cop pres			cop pres	

is

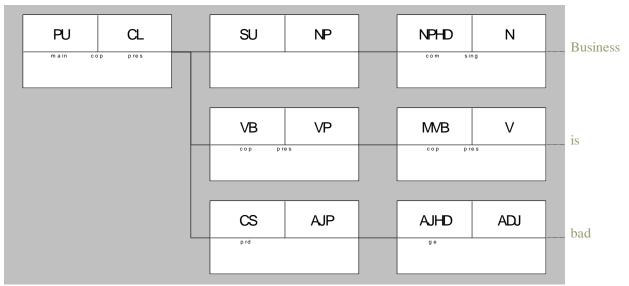| CS	AJP		AJHD	ADJ
prd			ge	

bad

Figure 1. The grammatical analysis of the sentence *Business is bad* in DCPSE. PU=
Parsing Unit, SU = Subject, VB = Verbal, MVB = Main verb, CS = Subject
Complement, AJHD = Adjective Phrase Head, ADJ = Adjective.

Using the International Corpus of English Corpus Utility Program (ICECUP), it is possible to search DCPSE in a number of ways from simple text searches to more detailed grammatical queries using *Fuzzy Tree Fragments* (FTFs), which will retrieve matching examples from the corpus (Aarts, Nelson and Wallis 1998; Wallis and Nelson 2000; Nelson, Wallis and Aarts 2002). The FTF in figure 2 searches the corpus for the string *got to* with the position for *HAVE* left unspecified. We use the label *fuzzy* in Fuzzy Tree Fragments to indicate that users can be as precise or as vague as to what they wish to search for.
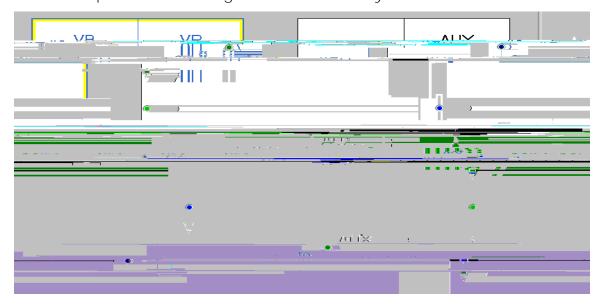

Figure 2. FTF for *(HAVE) GOT TO.*

Retrieval of data from corpus using text searches and FTFs.

Manual semantic coding (use of sound files to disambiguate).

Statistical tests on data (chi-square, log likelihood).

Data excluded from the study in order to study only variable contexts:

Negated forms (not semantically equivalent: *you mustn't go* "you are not allowed to go" vs. *you don't have to go* "you are allowed to not go").

Interrogatives (only one example with *MUST*).

Past tense forms of *HAVE TO* (no past tense of *MUST*).

Non-finite forms *HAVE TO*, including future forms (no non-finite form of *MUST* or *HAVE GOT TO*).

Unfinished/interrupted utterances.

Manual semantic coding is necessary to test claims which refer to semantic meaning of the modals/semi-modals, such as the following:

(3)

increase of just under 4% (from 47.78% in LLC to 51.74% in ICE-GB). There is little evidence that *MUST* is becoming monosemous.

Source corpus	Epistemic		Root		Performative		Ambiguous		Total
	N	%	N	%	N	%	N	%	N
LLC	43.96	47.78	36.63	39.81					

SHOULD	LLC frequency		ICE-GB frequency		Change in frequency %
	raw	per 100,000 words	raw	per 100,000 words	
EPISTEMIC	34	7.33	34	8.07	+10.1
ROOT	226	48.70	230	54.58	+12.07
OTHER	125	29.94	41	9.73	-67.5
TOTAL	385	85.97	305	72.38	-15.81

Table 5. Frequencies of SHOULD bm ge a hic hm e i DCPSE (Ï hhefÐi c i deg formulaic expressions, cases where *should* is being used for *would* and ambiguous cases).

Leech (2004) suggests a scale of intensity for modal meaning:

(1)   You must get a haircut

(2)   You need to get a haircut

(3)   You ought to get a haircut

W11Ø 38213 Tm[(W11Ø 38 38213 Te)-3(ofB ʌMCI4 38 B.21ʌ.94 13.44 reW*nBTh(i)4(t)

Aarts, Bas, Gerry Nelson and Sean Wallis (1998) Using Fuzzy Tree Fragments to explore English grammar. *English Today* 14, 52- 56.

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

Krug, Manfred (1998) *Gotta* ì The tenth central modal in English? Social, stylistic and regional variation in the British National Corpus as evidence of ongoing grammaticalization. In Lindquist, Hans, Staffan