Embedding epistemic modals in English: A corpus-based study*

Valentine Hacquard

University of Maryland

Alexis Wellwood

University of Maryland

Abstract The question of whether epistemic modals contribute to the truth conditions of the sentences they appear in is a matter of active debate in the literature. Fueling this debate is the lack of consensus about the extent to which epistemics can appear in the scope of other operators. This corpus study investigates the distribution of epistemics in naturalistic data. Our results indicate that they do embed, supporting the view that they contribute semantic content. However, their distribution is limited, compared to that of other modals. This limited distribution seems to call for a nuanced account: while epistemics are semantically contentful, they may require special licensing conditions.

Keywords: modality, epistemic, corpus

1 Introduction

Epistemic modals such as may and must below allow speakers to express various degrees of certainty. (1a) expresses a low degree of certainty that John is the murderer and (1b) a higher one. How exactly epistemic modals make this certainty contribution is a matter of active debate: is it part of the asserted content of sentences like (1), or is it a side comment from the speaker?

(1) a. John may be the murderer.
   b. John must be the murderer.

Modal accounts in the Kratzerian tradition (Kratzer 1981, 1991) treat epistemics on a par with other modals, as quantifiers over possible worlds restricted by an accessibility relation. An epistemic accessibility relation picks out worlds compatible with what is known in the world of evaluation, a deontic accessibility relation picks out worlds compatible with certain laws in the world of evaluation. Under this view,

* We are very grateful to Pranav Anand, Tim Hunter, Yakov Kronrod, Darryl McAdams, Ewan Dunbar, Dave Spencer, and Nate Ycas for their help with this project, and to Denis Filimonov for providing us with our parsed corpus. Many thanks to Katrin Schulz and three anonymous S&P reviewers for very helpful comments and suggestions.
Epistemics contribute truth conditional content: they express possibilities (1a) or necessities (1b) given a body of knowledge, usually the speaker’s.

Under speaker’s comment approaches, on the other hand, epistemics are taken not to contribute to the truth conditions of the sentence they appear in. They are not modals per se, but rather illocutionary markers which express a speaker’s comment about, or commitment to, the proposition expressed by the prejacent (cf. Halliday 1970, Bybee et al. 1994, Bybee & Fleischman 1995, Palmer 2001):

(2) “[Epistemic modality] is the speaker’s assessment of probability and predictability. It is external to the content, being a part of the attitude taken up by the speaker: his attitude, in this case, towards his own speech role as ‘declarer.’” (Halliday 1970; emphasis ours)

This intuition has been formalized in various ways. Some argue that epistemics modify or perform a different speech act, e.g., providing a kind of ‘doxastic advice’ (Swanson 2006). Others treat epistemics as evidentials (e.g., Westmoreland 1998, Drubig 2001). Evidentials are often said not to contribute to the truth-conditional content of the sentence they combine with, but rather to indicate the speaker’s grounds, or source of evidence, for expressing that sentence, e.g., sensory, hearsay or inferential evidence (cf. Faller 2002, Aikhenvald 2004). Likewise, accounts of epistemics as evidentials (which assume evidentials lack content) take epistemics to lack content: instead, they simply indicate the source of evidence as a deduction (Drubig 2001).

Central to this debate is whether epistemics can be embedded, as both camps have used embedding to support their approach. Epistemics often take widest scope: within their clause, they tend to scope over tense (Groenendijk & Stokhof 1975, Iatridou 1990, Abusch 1997, Stowell 2004, Hacquard 2006), negation (Coates 1983, Drubig 2001), and strong quantifiers like every or most (von Fintel & Iatridou 2003). They further have been reported not to occur in questions, antecedents of conditionals, or complements of attitude verbs (cf. Leech 1971, Jackendoff 1972, Coates 1983, McDowell 1987, Drubig 2001). For instance, Jackendoff (1972) argues that while may can either be interpreted deontically or epistemically in the declarative in (3a), it only receives a deontic interpretation in the question in (3b). Papafragou (2006) gives the examples in (4)-(5) as illicit cases of epistemics in the antecedent of a conditional and the complement of an attitude verb:

(3) a. John may leave early tonight. **epistemic/deontic**
    b. May John leave early tonight? *epistemic/deontic

(4) ? If Max must/may be lonely, his wife will be worried.

1 For an overview, see Hacquard (2011).
Embedding Epistemics

(5) It is surprising that Superman must be jealous of Lois.

Epistemics’ purported inability to embed was viewed as strong empirical evidence for their lack of participation in the asserted content of the sentences in which they appear. There are, however, counterexamples. To name a few, von Fintel & Gillies (2007) and Homer (2010) argue that epistemics can sometimes scope below tense; Cormack & Smith (2002) and Palmer (2001) that at least some epistemics can scope below negation; Tancredi (2007), Huitink (2009), and Gagnon & Wellwood (2011) that epistemics can scope below some strong quantifiers. Similarly, epistemics may sometimes be acceptable in questions (6), antecedents of conditionals (7), or complements of attitude verbs (8):

(6) Must Alfred have cancer?  
(7) If there might have been a mistake, the editor will have to reread the manuscript.  
(8) Sam thinks that it might be raining.

Crucially, in these embedded environments, the modal is interpreted in the scope of the various operators. As von Fintel & Gillies (2007) point out, (7) claims that the editor must reread the manuscript not just if there is an error, but if it is merely possible that there is. Similarly in (8), Sam believes that rain is a mere possibility. This sort of data suggests that epistemics can be interpreted in the scope of other operators, and are treated as serious challenges to speaker’s comment approaches (cf. Papafragou 2006, von Fintel & Gillies 2007). The existence of such data, however, isn’t the end of the story. First, illocutionary approaches can be made to deal with embedding: Swanson (2006), for instance, provides an illocutionary account where epistemics lack ‘substantive’ truth conditions but can nonetheless appear in the scope of other operators. Second, while the data in (6)-(8) show that epistemics can appear in certain embedded contexts, the data in (3)-(5) suggest that their distribution may be restricted.

There should be a fact of the matter about what the distribution of epistemic modals is like. Could the data motivating either camp be artifacts? How natural are the above examples? What kind of patterns do we actually find in naturalistic data? With this corpus-based study, we provide a clearer picture of the kinds of environments epistemics actually appear in, when compared to root (i.e., non-epistemic) modals, for whom the question of embedding is uncontroversial.

We examine the distribution of various English modals (might, can, must and semi-modal have to) in questions, antecedents of conditionals, complements of attitude predicates. Comparing possibility modals might and can allow us to draw

2 Thanks to an anonymous S&P reviewer for pointing this out.
generalizations about modal flavor over large samples. We assume that *might* only receives epistemic interpretations and *can* only root ones (Kratzer 1991). Thus if *might* doesn’t appear in all the environments that *can* does, we have evidence that any gap in the distribution of *might* isn’t due to a general ban against embedding modals but may instead be tied to epistemic modality. Examining the distribution of the necessity modal *must*, which can receive both epistemic and root interpretations, allows us to consider whether there are differences in distribution of epistemic meanings due to modal force. Finally, investigating the distribution of semi-modal *have to* in complements of attitudes will allow us to probe for epistemic modal meanings not only in finite, but in infinitival clauses as well.

Of course, the assumption that *can* only expresses root possibility oversimplifies somewhat, as it can receive epistemic interpretations in the scope of negation (e.g., *John can’t be home*; cf. Cormack & Smith 2002, Palmer 2001). Hence, some of our ‘root’ estimates for *can* may be slightly inflated. Assuming that *might* only expresses epistemic possibility also oversimplifies, as some argue that it can also receive ‘metaphysical’ (or ‘ontic’) interpretations (e.g., Condoravdi 2001, Schultz 2008). We mostly ignore this potential interpretation (except where noted in the text) for simplicity, and leave a systematic investigation of the distribution of metaphysical vs. epistemic *might* for future research.

2 Corpus data

We consider *might*, *can*, and *must* in antecedents of conditionals, questions, and complements of attitude predicates, as well as finite and infinitival *have to* within the last category. To examine distributions for these modals across these embedding contexts, we chose the New York Times section of the English Gigaword Corpus. After custom scripts tokenized, segmented, and excluded irrelevant material, and the data was parsed using Huang & Harper’s (2009) parser, the resultant data set contained 15,691,859 sentences. Out of these, 149,219 contained *might*, 88,859 *must*, and 475,590 *can*.

Anticipating a more complete presentation below, what we find is that epistemic modal meanings occurred in the environments we looked at, though not always to the same extent and in the same way as root modal meanings. In particular, epistemics are rarer in antecedents of conditionals and matrix, but not embedded, questions. They appear in the complements of some attitude verbs, but not others: in particular, they seem restricted from complements of attitudes expressing desires or commands.

The relative frequency of *might* vs. *can* in our three embedded environments is compared in Table 1. Relative to their own distributions, *might* is significantly

---

3 Where we examined sentences individually, we indicate potential cases of metaphysical *might*.

Table 1 Distribution of *might* and *can* in various environments. ** *p* < 0.001, Fisher’s exact test. Compares the distribution of *might* and *can* in each environment against their wider distribution in the corpus.

<table>
<thead>
<tr>
<th>Environment</th>
<th><em>might</em></th>
<th><em>can</em></th>
<th><em>p</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total corpus</td>
<td>149219</td>
<td>475590</td>
<td>100.00%</td>
</tr>
<tr>
<td>Antecedents of conditionals</td>
<td>30</td>
<td>9292</td>
<td>1.95%  **</td>
</tr>
<tr>
<td>Matrix questions</td>
<td>523</td>
<td>17971</td>
<td>3.78%  **</td>
</tr>
<tr>
<td>Complements of attitudes</td>
<td>17252</td>
<td>36864</td>
<td>7.75%  **</td>
</tr>
<tr>
<td>Total embedded</td>
<td>17805</td>
<td>64127</td>
<td>13.48%</td>
</tr>
</tbody>
</table>

less frequent in antecedents of conditionals and matrix questions than *can*, but more frequent in complements of attitude verbs.  

For *must*, we first established the relative frequency of its epistemic vs. root interpretations in matrix declaratives to serve as a baseline for comparison. Taking a random sample of 400 tokens in this environment, we found that it received epistemic interpretations 17% of the time.  As shown in Table 2, epistemic *must* is significantly less frequent than root *must* in antecedents of conditionals and questions when compared to this baseline. In the complements of attitude verbs, however, their distribution does not differ significantly from the baseline.

We examine these environments in detail in the rest of §2, turning to antecedents of conditionals in §2.1, questions in §2.2, and complements of attitude predicates in §2.3. In §3 we discuss potential factors underlying epistemics’ limited distribution, and conclude in §4.

2.1 Antecedents of conditionals

While *might* appears in antecedents of conditionals, it is exceedingly rare, as shown in Table 1: we found .02% of all *might* tokens here, in contrast to 1.95% of all *can* tokens. Relative to their wider distributions in the corpus, *can* is significantly more likely to appear in an *if*-clause than *might*. Out of the 30 instances of *might*, 7 seem to involve a conventionalized *might* of politeness (illustrated in (9)). One arguably

---

5 The *p* values reported are the result of comparing the respective distributions of each modal in a given environment relative to its distribution in the remainder of the corpus (*might* versus *can* in Table 1) or relative to the distribution in matrix declaratives (epistemic versus root *must* in Table 2).

6 This is consistent with Biber et al. (1999) and De Haan (2011) who show that root interpretations are more frequent than epistemic ones in written corpora. Looking at spoken corpora would be interesting, as the frequencies of epistemic and root *must* (and *have to*) are reversed. Our motivation for using our written corpus was its large size (thanks to an anonymous S&P reviewer for pointing out these references).

7 Please see Appendix A for details on the methodology we used to determine modal flavor.
receives a metaphysical/counterfactual interpretation (10). The rest seem to receive a genuine epistemic interpretation (as illustrated in (11) and (12)):

(9) If I might say, on behalf of John McCain, I believe he’s the veterans’ candidate.

(10) If any sector of society outside the military might have formed a political opposition, the Iraqi middle class would have been the only hope, a diplomat said.

(11) If one out of every thousand cases might be less than pure, maybe that’s the price you have to pay,” said Robert Carey, vice president of resettlement for the International Rescue Committee, a relief organization.

(12) Yet if his credibility might have been in jeopardy before, it most certainly is now.

We found a total of 213 must in if-clauses. We inspected each individually to determine its interpretation, and found that in all but one case (shown in (13)$^8$) the modal could only receive a root interpretation. Must is thus significantly less likely to receive an epistemic interpretation in an if-clause than in a matrix clause (compared to the baseline).

(13) “If there must be a gray area in making serious and difficult decisions,” they wrote, “how would it ever be deciphered who would be in the right and who would be in the wrong?”

Thus might is very rare and epistemic must virtually absent from antecedents of conditionals in our corpus, while root can and must are relatively more frequent.

$^8$ A paraphrase that brings out the epistemic interpretation might be If there is likely to be a gray area when making serious and difficult decisions.... We, however, consider this a marginal case.
2.2 Questions

While we find instances of epistemics in matrix questions, they are quite rare. Can is significantly more likely to appear in a matrix question than might (3.78% vs. .35%). We examined each might question individually. Most seem to receive genuine epistemic interpretations (examples (14)-(15) illustrate). One (16) possibly receives a ‘metaphysical’ interpretation:

(14) With the owners and the players on opposite sides philosophically and economically, what might they talk about at the next bargaining session?
(15) Might he be blackballed by all institutions of higher learning?
(16) What might the Grizzlies have been like if their leading scorer and rebounder, 6-foot-10 center Brent Smith, had not missed his third straight game because of a sprained ankle?

Epistemic interpretations of must are attested in this environment (34 instances), but they are significantly less common than root ones. We found only four instances of epistemic must in a yes/no question, two in a tag question:

(17) And mustn’t it tell something about durable intentions?
(18) Having represented so many of these men, Shargel must like them, must he not?

The few instances of wh-questions show an interesting pattern: 23 out of the 30 consist of questions where the speaker wonders about someone else’s thoughts or feelings. An anonymous reviewer suggests that these examples could reflect rhetorical questions:

(19) How must it be for a teen-age welfare mom to hear that she and her baby have caused most of the ills of society?
(20) Conversely, what must they think of him, after seeing the way Coach Cal responded to the Camby situation?

Note that such sentences occurred with might as well, but less frequently (7 instances out of 294 matrix wh-questions):

(21) So what might it be like to upset the real thing, Dream Team III, in tonight’s semifinal at the Georgia Dome?

9 Epistemic must may be more permissive than Italian epistemic dovere (must), which, according to Rocci (2007) cannot appear in questions, unless it is a tag or echo question.
10 An anonymous reviewer notes that the awkwardness of these examples could suggest they are production errors.
We conclude that epistemics are more rare in matrix questions than their root counterparts, although they are definitely attested. A possible reason for their rarity in questions may be pragmatic. If, as is often assumed, epistemics are anchored to the speaker’s knowledge, it may be strange for the speaker to ask about her own epistemic state (Papafragou 2006, Dorr & Hawthorne 2010). To see whether epistemics are freely compatible with questions once pragmatic considerations are factored out, we turned to embedded questions.

If the low number of epistemic must in matrix questions is due to pragmatic factors rather than an incompatibility with questions per se, we should find more epistemics in embedded questions. However, various types of predicates formally embed questions, without necessarily reporting an inquisitive act: verbs of knowledge (know) or decision (decide) take embedded questions as complements, but are not question reports. To see whether epistemics can appear not only in a question form, but in question reports, we looked specifically at modals in complements of verbs that describe question reports (Karttunen’s (1977) “inquisitives” class: e.g., wonder, ask). Table 3 summarizes the results.

<table>
<thead>
<tr>
<th>Environment</th>
<th>might</th>
<th>can</th>
<th>p</th>
<th>Epi must</th>
<th>Root must</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embedded questions</td>
<td>1716</td>
<td>0.15%</td>
<td>**</td>
<td>86</td>
<td>21.77%</td>
<td>309</td>
</tr>
<tr>
<td>Under inquisitive verbs</td>
<td>1367</td>
<td>0.92%</td>
<td>**</td>
<td>19</td>
<td>35.85%</td>
<td>34</td>
</tr>
</tbody>
</table>

Table 3 Distribution of modals in embedded questions. ** p < 0.001, - p > 0.05. Fisher’s exact test. For might and can, compares each environment to their respective wider distributions in the corpus. Flavors of must were compared to the distribution found in matrix contexts (17% epistemic).

In embedded questions, might is in fact significantly more frequent than can (1.15% vs. .80%). Examples of epistemic might are given in (22a-d).

(22) a. “We really have no plans to do that,” says John Barker, a spokesman for American Greetings, when asked if the company might leave Nasdaq.
b. Regina’s mother, Elizabeth Hershberger, wondered what her daughter might expect from a marriage.
c. Another option is a management buyout, but Havenstein said it is too soon for him to discuss whether he might want to do that.
d. He added, however, that he hadn’t had a chance to study the number and didn’t know what the components of that increase might be.

Examples with epistemic must are also attested here (86 cases), but the difference in distribution with root must does not differ significantly from the baseline.
Embedding Epistemics

(23) a. A sixth-grader at the time, he looked at the luxury cars parked there, Mercedes and BMWs, and thought how out of place he and his father must look.

b. I can’t help but wondering what the people in Rwanda or Bosnia-Herzegovina must be thinking about this.

Epistemics are thus not incompatible with a question form. Turning to modals in complements of ‘inquisitives’, we find that might was in fact significantly more frequent than can: .92% of might appear in a complement of an inquisitive vs. only .04% for can. We found 19 cases of epistemic must in complements of inquisitives, but no significant difference with root must.

Thus, while epistemic might and must are much rarer than their root counterparts in antecedents of conditionals and matrix questions, these differences level out in embedded questions.

2.3 Attitude contexts

Finally, we examined the distribution of the modals might, must, can, and the semi-modal have to in declarative complements of attitude predicates. To get a sense of their distribution, we sorted the various embedding predicates into the following semantic classes based on classifications in Villalta (2000) and Anand & Hacquard (m.s.) (Appendix B contains the complete list of predicates for each class):

I. Predicates of ‘acceptance’: those said to be correct if their complement proposition turns out to be true (Stalnaker 1984): predicates of argumentation (argue, explain), communication (say), doxastics (think) and semifactives (learn, realize).

II. Predicates of possibility (possible) and certainty (certain, likely).\textsuperscript{11}

III. Emotive predicates,\textsuperscript{12} including emotive factives (anger, bother) and emotive doxastics (hope, fear).

\textsuperscript{11} Sometimes modals embedded under predicates of possibility and certainty “agree” with the embedding possibility/necessity operators, and are hence not interpreted (so-called ‘modal concord’; Geurts & Huitink 2006, Zeilstra 2008). Here we assume that every instance of a modal is interpreted and is hence included in our counts (see footnote 13).

\textsuperscript{12} Emotive factives describe an emotive state w.r.t. a state of affairs, and presuppose that their subject knows that the proposition expressed by their complement is true. Emotive doxastics express a preference, but they also involve a doxastic component: the complement proposition has to be a doxastic possibility for the subject. If John hopes that Mary is home, John has to believe that it is possible that she is home. This doxastic component differentiates emotive doxastics from desideratives (Truckenbrodt 2006, Scheffler 2008, Falau 2010, Anand & Hacquard m.s.), which is why we separate the two classes.
IV. **Desideratives** (*want, wish*) and **directives**, including **orders** (*command, order*), **permissions** (*allow, permit*), and **prohibitions** (*forbid, ban*).

V. **Futures**, including **commissives** (*promise*) and **conjectures** (*expect, guess*); predicates of **perception** (*hear, see*); and **fiction** predicates (*dream, imagine*).

V. **Aspectuals** (*begin*); **causatives** (*cause*); **existence** predicates (*happen*).

As we will see in §2.3.1 and §2.3.2, the vast majority of occurrences of *might*, *can*, and *must* are found in complements of attitudes of acceptance, and there are very few occurrences in complements of desideratives and directives. This is consistent with Anand & Hacquard’s (2009) claim that epistemics cannot occur in complements of desideratives and directives. However, since predicates that express desires and commands usually take infinitival complements in English, this result may well be due to the fact that modal auxiliaries can only occur in finite complements. Thus, we consider and compare the distribution of the semi-modal *have to*, which can appear in both finite and infinitival complements, in §2.3.3.

### 2.3.1 Distribution of *might* in attitude contexts

Table 4 shows the number of *might* and *can* in complements of attitude verbs. Looking at percentages, one can see that *might* and *can* have similar distributions over the various attitude contexts. Most *might* and *can* occur in complements of attitudes of acceptance (significantly more so for *might*). Proportionally, there are significantly more *might* in complements of emotives and more *can* in complements of possibility/certainty predicates. The relative distribution of *might* and *can* doesn’t differ significantly for other attitudes. We discuss occurrences of *might* in these various attitude contexts below.

<table>
<thead>
<tr>
<th>Classification</th>
<th><em>might</em></th>
<th><em>can</em></th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Acceptance</td>
<td>14635</td>
<td>30641</td>
<td></td>
</tr>
<tr>
<td>II. Possibility, certainty</td>
<td>199</td>
<td>1348</td>
<td></td>
</tr>
<tr>
<td>III. Emotive</td>
<td>1791</td>
<td>2943</td>
<td></td>
</tr>
<tr>
<td>IV. Desideratives, directives</td>
<td>15</td>
<td>77</td>
<td>-</td>
</tr>
<tr>
<td>V. Futures, perception, fiction</td>
<td>504</td>
<td>1771</td>
<td></td>
</tr>
<tr>
<td>VI. Other</td>
<td>22</td>
<td>92</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>17166</td>
<td>36872</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4** Distribution of *might* and *can* across attitude contexts. **p** < 0.001, - **p** > 0.05, Fisher’s exact test. Compares the distribution of *might* and *can* in each class to their respective distribution in matrix declaratives (*might* 75526, *can* 267130).
Embedding Epistemics

Most instances of *might* occur under *acceptance* verbs (85%). (24) and (25) illustrate instances under *doxastic think* and *argumentation suggest*:

(24) Lainey said he had owned it for only a few weeks, and the police said they *thought* it might have tiny marijuana seeds in it somewhere.

(25) Lange said testimony appeared to *suggest* Simpson might have used a 15-inch knife that witnesses said Simpson bought in May.

The second largest category for *might* are the *emotives* (10%), and in particular emotive doxastics. The examples below show *might* under the *emotive doxastic fear*, and the *emotive factive be surprised*:

(26) Cardenas and other critics also say they *fear* National Action might acquiesce to a questionable Zedillo victory.

(27) ‘Having been to Cuba and knowing how repressive it is, I was a little *surprised* that he might not want to stay (in America)’.

Finally, although much less frequent, instances of *might* are found in complements of predicates of *perception, certainty, possibility*, and *fiction*, some of which are illustrated below:

(28) He’s *heard* the Mexico City Tigers might like to sign him.

(29) Many TCU players remain *convinced* that tonight’s game might be the most critical of the year.

(30) Residents of the building said it was difficult to *imagine* one of their neighbors might have thrown away a baby.

The large majority of *might* in *futures* were under complements of *conjectures* (31), with only one instance in complements of *commissives* (32):

(31) Phillip Adrian, the marketing manager of Driscoll Strawberries Associates in Watsonville, a major grower and shipper, *expected* there might be higher prices only in the next couple of weeks due to a slight shortfall.

(32) In addition, the U.S. had *vowed* for more than a year it might only offer conditional access to the U.S. market if other countries didn’t reciprocate, said William Hawley, a Washington vice president of Citicorp.

*Might* is exceedingly rare in complements of *desideratives* and *directives*. We examined each instance individually, to see whether they receive genuine epistemic

---

13 An anonymous reviewer points out that (26) might involve a case of ‘modal concord’. For discussion of modal concord in complements of emotive doxastics, see Anand & Hacquard (m.s.).
interpretations. Out of the 7 cases of desideratives, 4 were misparses. The remaining 3 occur with wish and arguably receive a metaphysical rather than epistemic interpretation:

(33) “It is the thrill of the moment,” he said, “Most of them probably look back and wish they might not have done it.”

(34) I wish we might have entered this new century with the ability to assert, without question, that the trend toward eliminating confrontation is irreversible.

(35) Vows to “return to the people’s business” have echoed from all corners of the Capitol for weeks, beginning with Clinton’s State of the Union address on Jan. 19, in which he wished later generations might look back and say, “We put aside our divisions and found a new hour of healing and hopefulness, that we joined together to serve and strengthen the land we love.”

Out of the 8 might sentences in directives, 5 were misparses. The remaining 3 are given below:

(36) Heaven forbid they might have to wait a few seconds to continue their progress.

(37) History dictates the Cowboys might not make a move to sign both.

(38) After learning that a federal judge had ruled California might be liable for up to $500 million in damages over its issuance of IOUs during a budget crisis in 1992, Wilson lashed out at Congress for having approved the Depression-era Fair Labor Practices Act.

(36) involves a formulaic use of the command forbid. (37) and (38) seem to be genuinely good instances, though the attitude verbs they involve sit on the more ‘acceptance’ end of the command class: we classified dictate as a command, yet it does not seem to be interpreted as an order in (37), given the subject’s inanimacy. Indeed, the same example with an animate subject is infelicitous: ??John/bureaucrats dictated that the cowboys might not make a move (contrast with (44) in the next section). Rule is interesting, as comparing it to a more stereotypical command verb like order it could be argued to have a doxastic meaning component: a ruling has to be made on the basis of facts and evidence, while an order can be based on whim. Pranav Anand (p.c.) suggests that when judges rule, they have a kind of metaphysical authority, specifying what is the case with respect to a particular system: when it

Please see Appendix A for details about the reliability of the parser and the nature of the misparses we found examining cases individually.

Alternatively, these may be cases of what Portner (1997) called subjunctive mood marking.
Embedding Epistemics

was ruled that Lance Armstrong was not guilty of doping, the ruling set a fact of the matter, not an order to bring about a state of affairs.

*Can* is also rare in complements of desideratives and directives, but not entirely absent. The sentences below illustrate:

(39) You *wish* your whole career can be like that; it can make everything a lot easier.
(40) The ordinance *requires* that contractors can not discriminate.

To sum up, we find that epistemic *might* is found embedded in attitude contexts. Its distribution looks similar to that of root modal *can*. Both *can* and *might* are frequent in complements of *acceptance* and *emotives* (even more so for *might*); they appear less frequently in complements of *perception, conjecture, certainty* and *possibility* predicates; *can* is rare and *might* practically absent in complements of *desideratives* and *directives*, as predicted by Anand & Hacquard (2009). Individual inspection indicates that the few instances of epistemic *might* in these contexts may be marginal.

2.3.2 Distribution of epistemic *must* in attitude contexts

To estimate the proportion of epistemic interpretations of *must* in complements of attitude predicates, we examined all instances in complements of attitude adjectives (130 tokens), together with a random sample of 400 instances in complements of attitude verbs. The occurrences of *must* and their distribution over the various attitude contexts in this sample are given in Table 5. The distributions of epistemic and root *must* do not differ significantly in these contexts from matrix contexts.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Total must</th>
<th>Epi <em>must</em></th>
<th>Root <em>must</em></th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Acceptance</td>
<td>424</td>
<td>80.00%</td>
<td>61 11.51%</td>
<td>363 68.49%</td>
</tr>
<tr>
<td>II. Possibility, certainty</td>
<td>79</td>
<td>14.91%</td>
<td>20 3.77%</td>
<td>59 11.13%</td>
</tr>
<tr>
<td>III. Emotive</td>
<td>10</td>
<td>1.89%</td>
<td>0 0.00%</td>
<td>10 1.89%</td>
</tr>
<tr>
<td>IV. Desideratives, directives</td>
<td>4</td>
<td>0.75%</td>
<td>0 0.00%</td>
<td>4 0.75%</td>
</tr>
<tr>
<td>V. Futures, perception, fiction</td>
<td>12</td>
<td>2.26%</td>
<td>2 0.38%</td>
<td>10 1.89%</td>
</tr>
<tr>
<td>VI. Other</td>
<td>1</td>
<td>0.19%</td>
<td>0 0.00%</td>
<td>1 0.19%</td>
</tr>
<tr>
<td>Total</td>
<td>530</td>
<td>100.00%</td>
<td>83 15.66%</td>
<td>447 84.34%</td>
</tr>
</tbody>
</table>

Table 5 Distribution of epistemic and root *must* across attitude contexts (random sample). - p > 0.05, Fisher’s exact test. Compares the distribution of epistemic and root *must* in each class to its distribution by flavor in matrix declaratives (17% epistemic).

The examples below illustrate epistemic *must* in complements of *semifactive* realize, *certainty* be convinced, and *conjecture* guess:
(41) They never said why, but after a while we **realized** something must be wrong.

(42) Just when I’m **convinced** that Windows 95 must be the buggiest, slowest, most difficult to use software invented since Windows 3.1, I hear from computer users who maintain they have had no problems with it.

(43) I **guess** the idea must have stuck.

Our sample contained no instances of epistemic **must** in complements of desideratives and directives. To determine the robustness of this generalization, we examined all instances of **must** in complements of desideratives and directives in the entire corpus. We found that **none** received an epistemic interpretation. Example (44) illustrate an instance of deontic **must** in the complement of command **dictate**:

(44) It’s enraging to have bureaucrats order you not to build on your land, tell you whom you must hire, **dictate** how you must advertise.

We further examined all emotives (the second most common class for epistemic **might**). We found only two instances of epistemic **must** in the complement of emotive doxastic worry, illustrated below:

(45) When some parents see this, they **worry** something must be wrong.

(46) Niles said many semiconductor stocks are still trading at low price-to-earnings ratios because Wall Street is nervous that chip orders have been so strong for four years that investors **worry** they must eventually fall.

To sum up, epistemic **must** is found in complements of attitude predicates, most overwhelmingly in attitudes of acceptance, and possibility/certainty. Almost no epistemic occurrence appears in complements of desideratives and directives. Finally, we find an asymmetry for complements of emotives: whereas **might** is relatively frequent in complements of such verbs, epistemic **must** is virtually absent.

**2.3.3 Have to in finite and nonfinite attitude contexts**

We see that epistemics can be embedded in complements of attitudes, but that these occurrences are overwhelmingly from the semantic class of predicates of acceptance and possibility/certainty, and virtually no modal (regardless of flavor) appears in the complements of desideratives or directives. Since many of the attitudes that express desires and commands only take infinitival complements, a syntactic environment that bars modal auxiliaries (e.g., *John wants to can/might/must go*), we consider the distribution of **have to**. As for **must**, we first determined a baseline of epistemic/root
interpretations by examining a random sample of 400 matrix declaratives. This yielded a baseline of 10% of epistemic have to.

We first examine finite complements, to check their consistency with the results of the preceding section. Given the high number of results for this environment (13,960 tokens), Table 6 reports occurrences by modality and proportional distribution across attitude contexts based on a random sample of 400 sentences. As with must, the distribution of epistemic and root interpretations did not differ significantly in these attitude contexts from the baseline. And again for both flavors the majority is found in complements of attitudes of acceptance.

To check the emerging generalization that epistemic meanings do not appear in desideratives, directives, or emotives, we examined all instances of have to in complements of these attitudes in the full corpus. We found no case of epistemic have to in this environment. Root have to, on the other hand, is attested in complements of desideratives (26 instances), directives (28 instances), and emotives (101 instances).

Table 7 shows the distribution of have to by modality and across attitude contexts in infinitival complements: none receive an epistemic interpretation. This contrasts significantly with finite complements where 27 out of 400 have to receive an epistemic interpretation (p < 0.001; Fisher’s exact test).

Looking at the distribution of have to across the various attitude contexts, we see that for infinitival complements, the majority of cases occur in complements of desideratives. Examples are shown below:

(47) Many people do not report a domestic worker’s wages because the worker does not want to have to pay taxes on his income, said Stuart Kessler, a senior tax partner at Goldstein Golub Kessler & Company in New York.

(48) “I don’t like to have to get on my knees and beg someone to sign me in,” Ms. Lapine said.
### Table 7: Distribution of *have to* in infinitival complements of attitude verbs.

<table>
<thead>
<tr>
<th>Classification</th>
<th>-fin <em>have to</em></th>
<th>Epi <em>have to</em></th>
<th>Root <em>have to</em></th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Acceptance</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>II. Possibility, certainty</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>III. Emotive</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>IV. Desideratives, directives</td>
<td>311</td>
<td>74.40%</td>
<td>0</td>
<td>311</td>
</tr>
<tr>
<td>V. Futures, perception, fiction</td>
<td>25</td>
<td>5.98%</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>VI. Other</td>
<td>50</td>
<td>11.96%</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>Total (all instances)</td>
<td>418</td>
<td>100.00%</td>
<td>0</td>
<td>418</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.001, Fisher’s exact test. Compares the distribution of epistemic and root *have to* in each class to the distribution of this semi-modal in matrix declaratives.

Given the results with finite complements, we know that *have to* can receive epistemic interpretations. However, they completely disappear in infinitival complements. The crucial difference between finite and infinitival complements is in the semantic class of the embedding verb: the majority of verbs taking infinitival complements come from the desire/command class. This suggests an incompatibility between epistemic modality and desideratives/directives.

### 3 On the limited distribution of epistemics

Our results show that epistemics can appear in embedded environments, supporting theories for which epistemics contribute semantic content. Yet their distribution is more constrained than that of roots. What could be responsible for this? In this section, we sketch possible explanations for their limited distribution in questions and antecedents of conditionals (§2.2) and complements of attitudes (§2.3).

#### 3.1 Questions and the antecedent of conditionals

Epistemics are often taken to express possibilities given what the *speaker* knows. If this is true, the oddness of epistemics in questions and antecedents of conditionals could be pragmatic in nature. As Papafragou (2006) and Dorr & Hawthorne (2010) point out, under normal circumstances it is strange for a speaker to ask about her own knowledge state (*is it possible given what I know that p?*). Similarly for antecedents of conditionals: uttering *if p then q* generally triggers the inference that *p* is not known. Again, under normal circumstances, it is strange for the speaker to be uncertain about her own knowledge state (*if it is possible given what I know that p...*). Thus, epistemics may only be felicitous in questions and antecedents of conditionals only when such introspective meaning is licensed. In newspaper articles
Alternatively, epistemics should be acceptable in questions and antecedents of conditionals if the epistemic claim can be interpreted relative to someone aside from the speaker’s knowledge state, or that of a larger group that includes the speaker. Papafragou (2006) argues that such epistemics anchored to the collective knowledge of the speaker’s community constitute what Lyons (1977) called objective epistemics. Because such epistemics are not anchored to the speaker’s sole knowledge (a subjective epistemic use), the speaker can felicitously use them in questions or suppositions (is it/if it is possible given what the community knows that p). Do we find any evidence for such a view?

A few instances of might in antecedents of conditionals do hint at such an objective stance, where might seems anchored to an implicit generic perspective, triggered by the presence of seem under might. In (49), the modality doesn’t seem merely anchored to the speaker, but to whoever might look at the Texas economy:

(49) If the Texas economy might have once seemed independent, now the fortunes of Tom Hicks, John Muse and almost everyone else here rely heavily on outside investors.

Absent further context, however, it is difficult to state with certainty whether the epistemics we found in questions and antecedents of conditionals have to be interpreted objectively, or otherwise anchored to someone other than the speaker.

Turning to embedded questions, we see that epistemics are just as frequent as their root counterparts (and even more so for possibility modals). Do we find evidence for a subjective/objective distinction there? In question reports, the lexical semantics of the embedding verb indicates whether the question is a solipsistic act (as with wonder), or an information-seeking question (as with ask) that typically requires an addressee. We might expect solipsistic questions to favor a subjective interpretation, if subjectivity involves the speaker/asker’s sole knowledge. On the other hand, addressee-oriented questions might favor an objective interpretation, with the modal anchored either to the addressee’s or some collective knowledge state. Under such a view, if x wonders whether might p, x is asking herself whether p is possible given what she knows; if x asks y whether might p, x is asking whether p is possible, given y’s sole or x and y’s pooled knowledge.

To see if we could detect a subjective/objective contrast, we looked at the distribution of modals in complements of question reports. We focused on Karttunen’s inquisitives, as only these truly correspond to question reports, further classifying verbs as to whether they indicate a solipsistic questioning act (e.g., wonder) or a question that requires an addressee (e.g., ask). An interesting contrast emerges: for
the root modal *can*, three quarters of inquisitives are solipsistic. For *might*, however, the strong bias for solipsistic inquisitives disappears: half of the inquisitives are addressee-oriented. The fact that *might* occurs proportionally more frequently than *can* under such verbs could support the hypothesis that epistemics in questions favor (though do not require) an objective stance, which these verbs might facilitate. The following illustrate instances of *might* in the complement of addressee-oriented *ask* and solipsistic *wonder*:

(50) a. “We really have no plans to do that,” says John Barker, a spokesman for American Greetings, when asked if the company *might* leave Nasdaq.
   b. Regina’s mother, Elizabeth Hershberger, wondered what her daughter *might* expect from a marriage.

Note, however, that epistemic *must* doesn’t show the same bias. The few instances of epistemic *must* in embedded questions (i.e., (23)) all appear in complements of solipsistic inquisitives. Out of these, 10 report someone wondering about someone else’s thoughts or feelings, echoing the behavior of epistemic *must* in matrix questions.

(51) The Cowboys, therefore, have to be wondering what in the name of Rod Hill *must* their new guy be thinking.

But if one can really ask whether *it is possible given what WE/YOU know that p*, as the addressee-oriented inquisitive data suggests for *might*, why doesn’t it happen in matrix questions? That is, why don’t we find more questions with (objective) epistemic *might*? This could be an artifact of our corpus: a newspaper article is not a dialogue, there is no addressee who can actually answer a question. Questions here, in general, may not be truly information-seeking but rather mere rhetorical devices. If the epistemics occurring in matrix questions in such a corpus can only be subjective or used rhetorically, we do not expect them to be very frequent.

Thus epistemics can appear in questions and antecedents of conditionals, but their distribution may be limited by pragmatic considerations. However, the fact that they seem to appear in embedded questions to the same extent (or more so) than roots supports the view that the limited distribution in matrix questions is due to pragmatics, rather than a general incompatibility between questions and epistemicity.

### 3.2 Attitude contexts

Our results show that epistemics occur in the complements of some attitudes but not others: epistemics are found in abundance in complements of attitudes of acceptance, but not under desideratives or directives. Indeed, there seems to be an incompatibility
between epistemics and such attitudes, which could be pragmatic (i.e., it is strange to desire or order that something be epistemically possible or necessary) or semantic, as Anand and Hacquard (2009, m.s.; A&H) argue.

A&H propose that there is a fundamental semantic difference between classes of verbs that pattern with such phenomena as mood selection in Romance: representational vs. non representational attitudes (Bolinger 1968). Attitudes of acceptance are representational: they “convey a mental picture”, or describe the content of a propositionally consistent attitudinal state. Desideratives and directives are non representational: they do not describe the content of a propositionally consistent attitudinal state. Instead, they express a preference for a state of affairs, captured, for instance, by Villalta’s (2000, 2008) comparative semantics. A&H argue that epistemic modals are anaphoric to the information state associated with the embedding attitude verb (cf. Hacquard 2006, Hacquard 2010, and Yalcin 2007). Representational attitudes are associated with such an information state and hence can license epistemics, but non representational attitudes are not so associated and hence do not license epistemics.

As we saw, we found virtually no cases of might or epistemic must and no instance of epistemic have to in the complements of desideratives or directives (not even in infinitival complements). Beyond attitudes of acceptance, epistemics are acceptable with fiction, perception, certainty, possibility and conjecture predicates, which all arguably involve a representational semantics. Epistemic might was relatively common in complements of emotive doxastics like hope, but there were virtually no instances of epistemic must in such complements. This result is consistent with A&H’s findings (in Romance languages) that possibility epistemics, but not necessity epistemics, are acceptable in the complements of hope and fear. A&H explain this contrast by arguing, following Truckenbrodt (2006), Scheffler (2008) and Falaus (2010), that such attitudes have a representational meaning component in addition to their preference component: if John hopes that p, p has to at least be a doxastic possibility for John. This doxastic component differentiates emotives from desideratives, and licenses epistemic possibility modals. Epistemic necessity modals are ruled out by an incompatibility between the certainty of a necessity claim, and the uncertainty tied to considering the several alternatives induced by preferences.

The limited distribution of epistemics in attitude contexts thus could arise from an incompatibility between their meaning and the meaning of certain attitudes. If A&H are right, this incompatibility may be semantic in nature and arise from categorical semantic differences between various classes of attitude verbs.
4 Conclusion

This corpus study has shown that epistemics can be found in various embedded contexts. This supports theories of epistemic modality according to which they contribute semantic content. Yet, their distribution seems to be restricted when compared to other modals: they are rare in questions and antecedents of conditionals, and absent in complements of certain attitude verbs. We have suggested that the limited distribution of epistemics could be due to a combination of semantic and pragmatic factors, though we leave a detailed account for future research.

Our data raises further questions. First, how freely available is an objective interpretation of a modal? What constrains its distribution? If it was completely free, we might have expected to find more epistemics in antecedents of conditionals and questions. Second, is there a unified explanation for the limited distribution of epistemics across these various contexts, or is it a combination of factors, as we have suggested? Finally, our results show a modal force asymmetry: while might was relatively frequent in complements of emotives, must was virtually absent. What underlies this asymmetry? Is it pragmatic or semantic? While we cannot settle such questions here, we hope to have shown that embedded epistemics do appear in naturalistic data, but that their distribution exhibits surprising gaps that warrant semantic and/or pragmatic explanations.

A Annotation and parser reliability

Interannotator agreement. For determining modal flavor of must and have to, each co-author examined each sentence individually, and determined whether the modal was interpreted as ‘epistemic’ or ‘root’ based on the sentence context and using paraphrases (it is probable/likely/obvious that p for epistemic, it is required/the laws or circumstances require/X is obliged to...’ for roots). Cases where either interpretation was possible were classified as epistemic, as our research question was the extent to which epistemic interpretations were possible in various contexts. Interannotator agreement was very high (κ= 0.84). ‘Either’ was possible for 15/530 of must in complements of attitudes, 1/213 in if-clauses, 2/277 in matrix questions, 1/400 in embedded questions, 20/400 in declaratives and 10/400 have to in finite complements. We discussed every instance of disagreement, resolving these mismatches and using the numbers post resolution for our analyses.

The parser. For details on the general accuracy of the parser, see Huang & Harper 2009. Because the corpus was automatically parsed, some sentences were misparsed and had to be excluded. For questions and antecedents of conditionals, exclusions consisted of misparses identified by individual inspection of all might and must cases. Because the number of can sentences was so large, we estimated the number of
Embedding Epistemics

misparses by visual inspection of a random sample of 300 sentences for each of these environments. This led to an exclusion rate of .13% and .04% for questions and if-clauses respectively. For sentential complements of attitudes, we excluded sentences on the basis of whether the main predicate could take a sentential complement. Visual inspection of some of the excluded data showed that they were the result of the parser misparsing relative clauses as complement clauses. This led to an exclusion rate of 1.05% of might, 4.19% of can, and .73% of must in these environments. For embedded questions, we excluded sentences on the basis of whether the main predicate could take a question complement. Our exclusion rate in this environment is very high (45.2%), due to the parser picking up free relatives.

B Classification of attitude verbs and adjectives

Note that * marks lexemes that embed might at least 5 times.
<table>
<thead>
<tr>
<th>I. Acceptance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ARGUMENTATION</strong></td>
<td>accept, accepting, acknowledge*, adamant, admit*, advise, advocate, affirm, agree*, allege, appealing, argue*, assert, assure, attest, caution*, challenge, charge, claim*, coach, concede*, concur, condemn, confirm, contend*, contest, convince*, counsel, counter, criticize, decide*, deliberate, demonstrate, deny*, develop, emphasize, endorse, ensure, establish, explain*, follow, grant, guarantee, highlight, hint*, imply*, insist*, insistent, insure, maintain, mean*, offer, oppose, outline, persuade, petition, pitch, plead, point, portray, posited, presuppose, propose, protest, prove*, provide, rationalize, recommend, reiterate, signify, stipulate, suggest*, suppose*, teach, unanimous, underscore, urge, vote, warrant</td>
</tr>
<tr>
<td><strong>COMMUNICATION</strong></td>
<td>add*, address, advertise, announce*, answer, apologize, articulate, beg, boast, brag, call, certify, chant, cheer, chime, chirp, clarify, cluck, comment, communicate, complain, confess, confide, contribute, convey, curse, declare, decline, detail, disclose, document, elaborate, excuse, expose, express, extrapolate, illustrate, indicate*, inform*, informed, interrupt, intimate, invite, joke*, mention, notice, notify, proclaim, pronounce, publicize, publish, quip, rejoin, release, remark, remind, reply*, report*, respond, restate, return, reveal, say*, scoff, scream, scribble, shout, show*, showcase, sign, signal*, speak, specify, state, stress, tell*, thank, threaten, unsaid, warn*, whine, whisper, write*, yell</td>
</tr>
<tr>
<td><strong>SEMIFACTIVES</strong></td>
<td>conceal, discover, forget, learn*, miss, note*, notice, overlook, realize*, recall, rediscover, remember</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II. Possibility</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>POSSIBILITY</strong></td>
<td>conceivable, far-fetched, inconceivable, possible*</td>
</tr>
<tr>
<td><strong>CERTAINTY</strong></td>
<td>certain*, clear*, confident*, convinced*, correct, doubt, doubtful, dubious, judge, likely*, obvious, positive, proof-positive, sufficient, sure*, true, uncertain*, unclear, unconvinced, unlikely, unreasonable, unsure*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>III. Emotive</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DOXASTIC</strong></td>
<td>anxious*, apprehensive, baffle, careful, comfortable, concern, concerned*, confuse, deserve, emotional, encourage, encouraged, fear*, fearful*, felt, fret*, frightened, frustrated, frustrating, hope*, hopeful*, impress, incredulous, marvel, mystified, nervous*, optimistic*, overcome, paranoid, pessimistic, sanguine, satisfy, satisfying, stand, tempting, troubled, trust, wary*, worried*, worry*</td>
</tr>
</tbody>
</table>
Embedding Epistemics

**FACTIVE** affect, affected, aggravating, anger, angry*, annoyed, antagonistic, applaud, appreciate, ashamed, bad, bear, bother, content, crazy, crushed, dandy, delighted, destructive, dig, disappointed, disappointing, disgusted, disillusioned, dismayed, distressed, dread, embarrassed, embarrassing, embittered, endure, enjoy, enraged, evil, excite, excited, exuberant, fantastic, fortunate, furious, glad, gleeeful, good, grateful, happy, hate, horrorstruck, humiliated, hurt, impressed, indignant, intense, irrational, irritate, lament, like*, livid, love, lucky, mad, magical, mind, oppressive, outraged, please, pleased, proud, regret, relieved, relish, resent, respect, ridiculous, rubbish, sad, sadder, shame, shock, shocked, sick, sorry, staggered, substandard, terrify*, thankful, thrill, tragic, uncomfortable, unconcerned, understandable, uneasy, unexceptionable, unfair, unfortunate, unhappy, unnerved, unsettled, unusual, upbeat, upset*, withstand, wonderful, amaze, amazed, astounded, attractive, awed, care, count, critical, curious, deliver, exploit, help, important, matter, problematic, relevant, right, strange, stun, surprise, surprised*, surprising, unbelievable, valuable, value, welcome

**IV. Desideratives, directives**

**DESIRE** afford, eager, need, pray, prefer, want, wish, determined

**ORDER** command, decree, demand, dictate, direct, force, impose, induce, influence, legislate, mandate, order, provoke, regulate, request, require, required, rule, steer, supersede, tolerate, uphold

**PERMISSION** allow, approve, authorize, pardon, permit

**PROHIBITION** ban, bar, censor, constrain, disallow, forbid, limit, prevent, prohibit, restrict, undermine

**V. Futures, perception, fiction**

**COMMISSIVE** promise, promising, swear, vow

**CONJECTURE** anticipate*, bet*, estimate*, expect*, forecast, foresee, gamble, gauge, guess*, overestimate, portend, predict*, project, underestimate

**PERCEPTION** detect, discern, feel*, glimpse, hear*, observe, overhear, perceive, read, regard, see*, sense*, smell, taste, witness

**FICTION** dream*, envision, fathom, imagine*, pretend, visualize

**VI. Other**

**ASPECTUAL** attempt, become, begin, change, continue, finish, start, try, use

**CAUSATIVE** make, cause

**EXISTENCE** happen, occur, rare, remain, result, tangible, turn out

**References**


Anand, Pranav & Valentine Hacquard. m.s. Epistemics and Attitudes. UCSC & UMD.


De Haan, F. 2011. Disambiguating modals: constructions and must. m.s.

Drubig, Hans Bernhard. 2001. On the syntactic form of epistemic modality. ms, University of Tuebingen.


Swanson, Eric. 2006. Something *might* might mean. Ms, University of Michigan.
Truckenbrodt, Hubert. 2006. On the semantic motivation of syntactic verb movement
In Brendan Jackson & Tanya Matthews (eds.), *Proceedings of SALT X*, 239–256.
Villalta, Elisabeth. 2008. Mood and gradability: an investigation of the subjunctive
Westmoreland, Robert. 1998. *Information and Intonation in Natural Language
Modality*: Indiana University dissertation.
Zeiljstra, H. 2008. Modal concord is syntactic agreement. In M Gibson & T Freidman