May or *Might*? Strength, Duality and Social Meaning*

Hadil Karawani[1] and Brandon Waldon[2]

[1] Leibniz-ZAS, Berlin, Germany karawani@leibniz-zas.de
[2] Stanford University, Stanford, California, U.S.A. bwaldon@stanford.edu

This paper addresses one longstanding claim about epistemic *must*, namely that its semantic dual is the modal auxiliary *might*. While this is a "usual assumption" in the natural language semantics literature (Lassiter 2016: 14), it is not universally accepted: Crespo, Karawani, and Veltman (2017) (henceforth CKV) propose a theory of epistemic modality whereby the dual of *must* is *may*, the non-subjunctive counterpart of *might*. According to their view, an assymetric entailment relation holds between *may* and *might*, whereby *might* is weaker than *may* with respect to speakers' expectations about the likelihood of the prejacent. Consistent with this view, the results of one experiment suggest that English speakers consider *may p* to be stronger than *might p* on two metrics: perceived speaker certainty of *p* and inferred likelihood of *p*. In two subsequent experiments, we were unable to find conclusive evidence that speakers distinguish between epistemic *may* and *might* in discourse contexts, including in epistemic contradiction contexts of the form *must p, but might/may not p*. We interpret this as evidence that the "usual assumption" at a minimum needs to be revised to accommodate *may*. We conclude with a discussion of the diachronic, cross-linguistic, and social meaning facets of epistemic modality that are uniquely accounted for by the framework of CKV.

1 Background

Much of the recent literature on epistemic modality in English has centered around the strength of *must p*, which on some accounts is claimed to be true iff *p* is true in all epistemically possible worlds (c.f. von Fintel & Gillies 2010, Willer 2013); and the strength of *might p*, which on some accounts is claimed to be true so long as there is one epistemically accessible world in which *p* is true (c.f. Rudin 2016). These two claims are theoretically attractive: if the first is correct, then *must* is the natural language expression corresponding to the epistemic necessity operator of modal logic; if the second is correct, then *might* is the natural language counterpart of epistemic possibility. If we accept these claims in tandem, then the infelicity of the assertions in (1) can easily be explained in terms of semantic ill-formedness. Namely, (1) is infelicitous for precisely the same reasons that (2) is a contradiction in modal logic:

(1) #It must be raining, but it might not be raining.

(2) $\Box$ A $\wedge$ $\Diamond$ $\neg A$

Namely, *must* and *might*, along with $\Box$ and $\Diamond$, are semantic duals: that is, *might p* is semantically equivalent to $\neg$*must*$\neg p$, just as $\Diamond$A is semantically equivalent to $\neg\,\Box\,\neg$A. However,

*We thank Heather Burnett, Cleo Condoravdi, Judith Degen, Patrick Elliot, Anastasia Giannakidou, Nicole Gotzner, Stefanie Jannedy, Manfred Krifka, Beth Levin, Guillermo del Pinal, Josep Quer, Uli Sauerland, Stephanie Solt, Jack Tomlinson, Hubert Truckenbrot, Frank Veltman, and Hedde Zeijlstra, as well as audiences at the semantics reading group at ZAS and at the Words With Friends reading group at Stanford for generous and helpful feedback. All remaining mistakes are, of course, our own. We also thank Alexandre Cremers, Floris Roelofsen and the anonymous reviewers of Amsterdam Colloquium.

this is not the only analysis that provides us a story of the infelicity of (1) via semantic dualism: a variety of accounts on the market achieve this by strengthening the semantics of *might* and correspondingly weakening the semantics of *must*: for example, Kratzer's (1991) denotation of *must* is universal quantification over only the maximally-normal subset of epistemically accessible worlds, while *might* is existential quantification over this same set; Lassiter (2016) gives *must* and *might* a probabilistic semantics, where the probability threshold for *must* is defined as the inverse of that of *might*. The intuition, nevertheless, is still that the semantic duality relationship of *must* and *might* explains the contradictory feel of the sentences in (1). Willer (2013), on the other hand, argues for a strong denotation of *must* according to which *must p* entails *p*, but he also argues that *might* is strong: *might p* expresses that *p* is a live possibility, with 'live possibilities' defined as "possibilities that are compatible with the agent's evidence and that the agent takes seriously" (2013: 5). In his dynamic semantics of *must* and *might*, they are formally duals, and (1) is predicted to be infelicitous: if a context update of *must p* is incompatible with a subsequent update that *p* is epistemically possible (given that *p* is entailed), then *must p* is certainly incompatible with the suggestion that *p* is a possibility that we should seriously consider in future deliberations. Conversely, Veltman (1986) argues for a weak *must* (in which *must p* does not entail *p*) but also, in Veltman (1996), for a weak *might* (in which *might* is simply existential quantification over epistemically possible worlds). Note that Veltman does not treat *must* and *might* as duals. Moreover, (1) on Veltman's account is not predicted to express a semantic contradiction; the infelicity must be explained through other extra-semantic means. What is the dual of *must*, though, if not *might*? On Willer's (2013) story, another conceivable dual is *possible*, assuming that *possible* means something along the lines of it being only a bare epistemic possibility that the prejacent is true. This solution seems less tenable for weak theories of *must*: if *must* explicitly leaves open the epistemic possibility of the negation of the prejacent, then it would be surprising to find that *possible* is its dual (again assuming this rough sketch of the semantics of *possible*). In fact, in arguing for weak *must*, Lassiter (2016) provides numerous examples from the web where speakers make statements to the effect of *must p, but possible not p*. Given a weak theory of *must*, the dual (if it exists) needs to be stronger than bare possibility and thus intuitively will be a lexical item whose semantics are stronger than those of *possible*.

2 Introducing *may*

CKV propose a dynamic analysis of epistemic modality that builds on the assumption that *might* is weaker than *may*: in their theory, *might* expresses existential quantification over epistemically possible worlds, while *may* expresses existential quantification only over the *most likely* of those epistemically possible worlds – those that are consistent with one's expectations.[1] Hence, two kinds of possibility are introduced, with a distinction between knowledge and expectation:
May: *likely possibility, something to reckon with, consistent with one's expectations.*
Might: *unlikely possibility, consistent with everything one knows but not with expectations.*

In CKV's framework, *may*, and not *might*, is the dual of *must*, as *must* according to their analysis expresses universal quantification over the set over which *may* expresses existential quantification. One empirical prediction of this view is that while (1) is not necessarily contradictory, (3) most certainly is, for the familiar reason that the sentence contains both a lexical

[1] The relative strength of *may* and *might* is derived in the following way: *might* is equal to *may* plus (fake) past tense resulting in an unlikelihood inference. Past tense is defined in terms of non-actual veridicality (Karawani 2014, in the spirit of Giannakidou's 1997 non-veridicality notion and Iatridou's 2000 exclusion feature). The unlikelihood inference is argued to be presuppositional.

item (*must*) and the negation of that item's dual (*may not*):

(3) #It must be raining, but it may not be raining.

As we show below, empirical support for this analysis of *may* relative to *might* is clear in one experiment, where speakers were asked directly to assess the strength of *may* and *might*. A complete story of epistemic modality in English - including what the dual of *must* is - will need to reconcile this empirical finding. First, however, we recap the distribution of *may* and *might* in English. As is well known, both *may* and *might* have epistemic readings:

(4) Jack may/might be in the office.

As (4) demonstrates (and following Condoravdi 2002), both *might* and *may* can quantify over the present epistemic (information) state. *May* and *might* are closely related in the sense that the latter is the subjunctive/(fake) past counterpart of the former, yet the inflectional marking of *may* and *might* only obscures the similarity of these lexical items' respective distributions, at least in the epistemic domain. Although *might* seems to be inflected for past tense, it is acceptable in present contexts such as (4); moreover, present-tense *may* is acceptable in past contexts such as (5): contrary to expectations, *may* seems to be licensed in real past tense situations and appears to be able to quantify over past information states. Even in sequence-of-tense constructions (e.g. *said that...*) where one expects *may* to not be licensed, native-speaker intuitions regarding examples such as (6a) and (6b) are unclear: according to some speakers, *may* is slightly dispreferred, but no one we consulted considered it to be completely ungrammatical. Similar things can be said for the case of subjunctive conditionals (7).

(5) I was so sick that I thought that I might/may not make it to school yesterday.

(6) a. He said he might go. / b. (?) He said he may go.

(7) If Mr. Smith were to win the election, he (?)may/might appoint a new sheriff.

However, the respective distributions of *may* and *might* diverge more substantially in non-epistemic contexts. *May*, unlike *might*, is licensed in indicative deontic contexts, (8). *Might*'s distribution as a deontic modal seems to be restricted to question contexts, and semantic intuitions in relation to *may* are crisper here than for the epistemic contexts discussed above. Namely, White (1975) remarks that *may* and *might* are not synonymous in contexts such as (9), presented below: *might he take it?* seems to express "a more tentative request" than does *may he take it?* (1975: 49).

(8) You may/#might go now!

(9) May/might he take it?

Returning to epistemic uses, it appears as though *may* and *might* are not clearly distinguished, at least from the standpoint of grammatical distribution. However, there is a persistent claim in the descriptive literature on English that *may* and *might* are distinguishable on semantic grounds. For example, Leech (2004) claims that "The effect of the hypothetical auxiliary [*might*], with its implication 'contrary to expectation', is to make the expression of possibility more tentative and guarded [relative to *may*]." Similarly, Nuyts (2001: 209) notes that "Obviously *might* used to be the past of *may*, but the past tense has been entirely reinterpreted as an 'epistemic past', i.e. a weakener of the epistemic qualification." This intuition is taken into serious consideration by CKV: in their theory of epistemic modality, *may* is formally stronger than *might*. This feature of their analysis also leads them to revise the 'normal' duality assumption of *might* and *must*.

3 Experiment 1

In our first experiment, we were interested in determining whether speakers' intuitions about the strength of *may* indeed varied relative to strength intuitions for *might*. To create our stimuli, we first gathered data from the Corpus of Contemporary American English (Davies 2008-) where *may* and *might* were employed in epistemic contexts. We identified such 10 sentences, of which 5 contained a *might* phrase while the other 5 contained a *may* phrase. We then modified these sentences to improve discourse coherence. We systematically manipulated these sentences such that they contained either *may*, *might*, *could*, or *must*. The result was 40 sentences: 10 sentences with four possible modal configurations. An example paradigm is presented below:

Confounding factors {*may/might/could/must*} *have skewed the results of the doctors' study.*[2]

We then recruited Amazon Mechanical Turk workers (n = 61) to provide responses to the following two questions about these sentences:

i. *"According to the speaker of the above sentence, how likely is it that* [paraphrase of the prejacent]*?"*

ii. *"How certain is the speaker that* [paraphrase of the prejacent]*?"*[3]

In the vein of a recent experimental investigation into the semantics of epistemic *must* by Scontras et al. (2016), our participants assessed likelihood and certainty by answering the above questions on a slider scale. For the first question, the left end of the scale read "0% likely (impossible)" and the right end read "100% likely (guaranteed to be the case)". For the second question, the left and right ends were "0% certain (certain it is not the case)" and "100% certain (absolutely certain)", respectively. We collected 15 responses for our 40 sentences (600 responses total). We allowed participants to answer for as many sentences as they liked: the mean number of responses per participant was 9.84. Participants were paid $0.05 per response.

3.1 Results and Discussion

	Mean certainty rating	Mean likelihood rating
may	54.71	60.21
might	46.67	54.83
could	46.32	54.97
must	75.49	79.04

The left end of each slider scale was coded "1" in our analysis, while the right end was coded "100". The above table shows the mean response values for our two questions, broken down by modal condition. We performed pairwise linear mixed effects regression analyses using the lme4 package (Bates et al. 2014) in R to predict certainty ratings from a fixed effect of modal as well as by-participant and by-item random intercepts. Certainty ratings were higher in the *may* condition than in the *could* condition ($\beta = 9.07$, $SE = 2.23$, $t = 4.06$, $p < 0.001$). However, there was no difference in certainty ratings between the *might* and *could* conditions ($\beta = 1.29$, $SE = 2.13$, $t = 0.607$, $p < 0.55$). A similar pattern of results held for the likelihood ratings, which were higher in the *may* condition than in the *could* condition ($\beta = 5.91$, $SE = 1.87$, $t = 3.16$, $p < 0.005$); however, there was no difference in likelihood ratings between the *might* and *could* conditions ($\beta = 0.63$, $SE = 1.74$, $t = 0.365$, $p < 0.72$). In this experiment, the difference between *might* and *may* was significant in pairwise comparison tests for likelihood

[2]For some stimuli, we extraposed *must* in order to avoid deontic interpretations, for example:
In order to stay competitive, the company {*may/might/could*} *need to outsource its production.*
It must be the case that in order to stay competitive, the company needs to outsource its production.

[3]Note that we take likelihood to be a property of the sentence while certainty of p is a subjective assessment of speaker's certainty.

ratings ($\beta = -5.34$, $SE = 1.62$, $t = -3.30$, $p < 0.001$) and certainty ratings ($\beta = -8.07$, $SE = 2.08$, $t = -3.89$, $p < 0.005$). Additionally, as shown above, we found evidence that *may* was stronger than another epistemically-construed modal (*could*) where *might* was not, on both the metric of likelihood and the metric of certainty.[4]

We interpret this as evidence of a distinction between *may* and *might*, in the sense that *may* appears to encode higher likelihood of the prejacent as well as a higher degree of speaker certainty of the prejacent. This result is consistent with claims from the descriptive literature on English - discussed above - that epistemic *might* is the 'tentative' or 'qualified' counterpart of epistemic *may*. Furthermore, this result is predicted by the analysis offered by CKV.

4 Experiment 2

The results of Experiment 1 suggested that speakers of English are sensitive to a difference in strength between *may* and *might*. With Experiment 2, we began our investigation of whether and how this distinction is relevant in discourse. For Experiment 2, we took as our point of departure the *Dismissive Agreement* paradigm of Rudin (2015), in which a speaker B may assert *might p* even when *p* is only of marginal possibility given B's epistemic state, as in the example below (Rudin 2015: (1)):

(10) **A:** Paul might come to the party.
B: Yeah, he might, but it's extremely unlikely.

Consistent with CKV's analysis, Rudin (2015) invokes (10) to build the case for a semantics of *might* whereby *might* semantically encodes nonzero likelihood of the prejacent but may give rise to the implicature that "the prejacent is likely enough to be relevant" (Rudin 2015: 596). For our purposes, the key intuition of examples such as (10) is that B's response is a natural way to communicate that for all B knows, the likelihood that Paul comes to the party is extremely low but nonzero. One additional prediction of CKV, however, is that because *may* encodes the 'likely possibility' of *p*, B's response in (11) should be disfluent relative to her response in (10):

(11) **A:** Paul may come to the party.
B: (?)Yeah, he may, but it's extremely unlikely.

To test this, we produced three *Dismissive Agreement* discourse contexts, systematically manipulated such that the featured modal was either *may* or *might*. An example paradigm is shown here: *Context: Bill and Simon are discussing the future of their city's public zoo.*

Bill: *The local zoo {might/may} be shut down by city government this year.*

Simon: *Yeah, it {might/may} be, but it's extremely unlikely.*

The discourses always included an interlocutor "Simon" responding to an interlocutor "Bill". In an online acceptability judgment experiment, we asked workers on MTurk (n = 240, US IP Addresses, Approval Rating > 80%, compensation = $0.15) to "rate the extent to which you think, in the given scenario, that Simon's reply is a natural response to Bill's statement". Participants rated on a scale of 1 to 7, with 1 being "completely unnatural", and 7 being "completely natural". Each participant provided a rating for just one of our six items (three discourse contexts × two modal conditions = six items), and we used the UniqueTurker script (https://uniqueturker.myleott.com/) to prevent multiple participation. We had 240 ratings.

[4] We also coded each of our ten sentence contexts according to whether the sentence originally contained a *may* or a *might* in the COCA. Including this variable as a random or fixed effect did not improve the models for likelihood ratings ($p < 0.67$) or for certainty ratings ($p < 0.77$).

4.1 Results and Discussion

Condition	Average naturalness rating
may	5.72
might	5.62

As the above table shows, naturalness ratings for *may* and *might* were virtually indistinguishable, contrary to what one might expect given the strength asymmetry of *may* and *might* argued for by CKV and suggested by the results of Experiment 1. That is, in contexts where the responding interlocutor Simon wished to assert that the prejacent is an extremely unlikely possibility, *may* appeared to be no worse than *might*.[5]

5 Experiment 3

In Experiment 3, we explored the potential implications of a strength asymmetry between *may* and *might* for semantic duality hypotheses in the epistemic modal domain. First, note that if *may* is indeed stronger than *might*, and if *may* is the dual of *must*, then we are committed to the idea that *must* is weak - because its dual is "stronger" than the existential epistemic possibility modal. Moreover, if we accept CKV's claim that *may* is the dual of a weak *must* and *might* is indeed the bare existential epistemic modal, then we make the following empirical prediction: whereas *must p, but may not p* should always be contradictory given their duality relationship, there may be contexts in which *must p, but might not p* is not contradictory. As a first step towards investigating this, we considered Lassiter (2016), who reports a series of naturally-occurring examples of epistemic *must* in non-maximal certainty contexts, in which the *must* claim is preemptively (12) or subsequently (13) hedged:

(12) I don't know for sure, sweetie, but she must have been very depressed. A person doesn't do something like that lightly. (Lassiter 2016:7, ex. 16)

(13) This spot might be good for fishing, I've always thought, though I haven't seen a soul out there trying. The land must be private, I'm almost certain. (ex. 28)

Lassiter (2016) uses these data to argue, following Karttunen (1972), Veltman (1985) and Kratzer (1991), that *must* is weak. We were particularly interested in these data because they were contexts where a claim of *must p* allowed for the epistemic possibility of *not p*. Thus, they seemed to us to be particularly good contexts to investigate whether *must p* could be felicitously hedged by *might p*, to similar effect as the *I don't know for sure* and *almost certain* hedges of (12) and (13). Because we predict that *must p, but may not p* should always be a contradiction, however, we also predicted that it should be considerably worse than *must p, but might not p* in these same contexts.

For our experiment, we elicited acceptability judgments (MTurk workers, US IP Addresses, Approval Rating > 80%, compensation = \$0.15) of sentences – presented in a 3-4 sentence discourse frame – based on Lassiter's (2016) original examples and modified to achieve our own experimental aims. We tested the acceptability of hedging an assertion of *must p* with *may*

[5] It is possible that a distinction between *may* and *might* in this experiment was obfuscated by the fact that Simon's response always echoed the modal in Bill's statement. Taking this a step further, perhaps the echoic usage of the modal signified something weaker than full agreement with Bill and hence cannot be construed as Simon's full commitment to the modal claim. On this view, *Yeah, Paul may come to the party, but it's extremely unlikely* is potentially not interpeted as a contradiction, as Simon's commitment to the first conjunct is weaker than his commitment to the second.

not p, *might not p*, or *not p*.[6] Judgments were provided on a 1 to 7 acceptability scale - with 1 being "completely acceptable" and 7 being "completely unacceptable". The critical sentence was always bolded, and participants were asked to read the entire discourse frame but to rate just the bolded sentence. An example paradigm is shown below:

I think I've found my dream car at a used car dealership down the road: a beautiful 1964 white Ford Mustang. The body, paint, and suede interior look pristine. I checked the speedometer, and it shows 38,000 miles. ***The mileage must actually be 138,000, but {it might not be / it may not be / it isn't}.*** *At any rate, the car drove beautifully during the test drive!.*

Additionally, for every discourse context, we included two modifications where there was no second conjunct and where the *must* was replaced by either a *may* or *might*. These conditions were included to investigate whether any difference in acceptability between *must p, but may not p* and *must p, but might not p* could be attributed to a default preference for one of the two modals in these discourse contexts.

I think I've found my dream car at a used car dealership down the road: a beautiful 1964 white Ford Mustang. The body, paint, and suede interior look pristine. I checked the speedometer, and it shows 38,000 miles. ***The mileage {may/might} actually be 138,000.*** *At any rate, the car drove beautifully during the test drive!.*

We had three discourse contexts with five manipulations each, for a total of 15 items. Each participant provided a rating of just one item, and we used the UniqueTurker script to prevent multiple participation. We had 40 ratings for each of our 15 items, for a total of 600 responses.

5.1 Results and Discussion

must p, but $\neg p$	*must p, but might* $\neg p$	*must p, but may* $\neg p$	*might p*	*may p*
4.01	4.12	4.36	5.57	5.69

The above table provides the average ratings for the five manipulations across our three discourse contexts.[7] Hedges of *must* with either *might not* or *may not* are more or less equally degraded in our data: we performed pairwise ordinal logistic regression analyses using the MASS package (Venables and Whitney 2002) in R to predict acceptability ratings from a fixed effect of discourse manipulation, and we found that these two aforementioned conditions did not differ significantly ($\beta = 0.23$, $SE = 0.23$, $t = 1.01$, $p < 0.32$). It appears to the extent that *must p, but may not p* is felt to be contradictory, so is *must p, but might not p*. Moreover, the *may p* and *might p* conditions did not differ significantly ($\beta = -0.26$, $SE = 0.24$, $t = -1.12$, $p < 0.27$). Thus, in Experiment 3, we could find no evidence that speakers were distinguishing between *might* and *may* in these discourse contexts: accounts arguing in favour of *might* existing in a unique duality relationship with *must*, must reconcile the data here and offer an account of why *must* and *may* do not exist in a duality relationship.[8]

[6] The *not p* condition – a guaranteed contradiction – was included as a baseline comparison against which to assess the acceptability of the other two continuations.

[7] Two data points were excluded due to the failure of the participant to answer the question.

[8] It is also notable – and perhaps unexpected – that there was no significant difference in acceptability between the *must p, but not p* condition and the *must p, but might not p* condition ($\beta = 0.15$, $SE = 0.228$, $t = 0.663$, $p < 0.51$) also between *must p, but not p* and *must p, but may not p* ($\beta = 0.39$, $SE = 0.228$, $t = 1.69$, $p < 0.1$). This is surprising given that a continuation of *not p* certainly feels 'more' contradictory than a continuation of *might p* in these contexts. Thus, it could be that there is **some** distinction between the *might not p* and *may not p* continuations of *must p* but that our paradigm was not sensitive to finer distinctions of acceptability. In any case, there was no obvious dropoff in acceptability from *may not p* to *might not p*.

6 Discussion - Strength, Duality, and Social Meaning

Though we have found some evidence of a strength asymmetry between *may* and *might*, consistent with the analysis of CKV, we did not find any direct empirical support for the relevance of this asymmetry in contexts of language use: Experiment 2 suggests that *may* is not significantly worse than *might* to express extremely low but nonzero likelihood of the prejacent; Experiment 3 suggests that *must p, but might not p* is not any less disfluent than *must p, but may not p* - at least not in the discourse contexts we explored. How do we reconcile these results with claims from the descriptive literature on English that *might* is the weak epistemic possibility counterpart of *may*? One potential explanation is that *might* and *may* have lost a distinction in epistemic strength as *may* disappears in spoken English. In a corpus study conducted by Bowie et al (2013), it is reported that there has been a significant drop in the use of *must* (-54%) and *may* (-39%) but a rise in the use of *might* (5%). On this story, we might reasonably have had a debate as to what the dual of *must* is several decades ago, when both *might* and *may* were salient (and hence semantically distinguished) members of the spoken lexicon.

There is, however, more to be said here from the diachronic perspective: the shift in the usage of modals in informal spoken registers is a phenomenon that has been observed in languages other than English, including by Gonzales et al. (2017), who report that modals encoding low speaker certainty are generally supplanting markers of high certainty in informal spoken registers of Catalan. A concurrent shift favoring weak epistemic modals is taking place in German, where *könnten* is increasinlgy preferred to *dürften* in everyday discourse.[9] Generally, the trend appears to be a shift away from markers of high certainty. Given this general trend, we straightforwardly account for the rise of *might* and concurrent fall of *may* in English on a semantic account of modals - such as CKV's - whereby *might* is weaker than *may*.[10]

Another story (not incompatible with the diachronic facts mentioned above) makes reference to sociolinguistic theory and to the distributional differences of *might* and *may* across registers of English. Biber et al (1999) report a dichotomy between the use of *may* and *might*: in formal (e.g. academic) registers the use of *may* is significantly higher, as opposed to a higher use of *might* in informal registers. If *may* and *might* are tied to a social register, then we might expect them to encode different **social meaning**: whereas sentence meaning (as analyzed by semanticists) conveys information about the world (facts or thoughts), social meaning (as analyzed by sociolinguists) conveys information about the identity of the speaker, which in turn may enrich sentence meaning. In the following, we will sketch how pragmatic enrichment via social meaning can help us make sense of the empirical picture of *may* and *might*. To begin, consider the following contexts a doctor may find herself in. In both (14) and (15), she makes a statement regarding the sickness afflicting her patient and raises the possibility that it is cancer:

(14) *Doctor, to colleagues:* "This may/might be cancer."

(15) *Doctor, to a patient or patient's family:* "This may/might be cancer."

When asked to compare *may* and *might* sentences in (14) and (15), English native speakers that we consulted were able to report a qualitative difference between the two epistemic possibility modals more readily than in what we saw in Experiments 2 and 3. Namely, in the case of the doctor speaking to her colleagues, native speakers felt that *might* conveyed that the doctor is "not so sure that it is cancer." Intuitions differed in the case of talking to the patient or

[9] We thank Hubert Truckenbrot (p.c.) for pointing this out.

[10] For some opinionated thoughts on the cultural forces driving this diachronic change, consider these selected quotes from a recent New York Times article (April 30, 2016) entitled *Stop Saying: I feel like*: "[T]here is a tendency to commit less [...] a reflex to hedge every statement ."

patient's family: the use of *may* in this context is taken to convey a more aloof, sterile, attitude from the doctor, whereas the use of *might* is taken to involve politeness, empathy, breaking the news softly, etc – there is no perception of uncertainty relative to *may*.

The example illustrates that speakers' intuitions regarding the difference between *may* and *might* depend on social context and the intended social meaning of interlocutors – two factors which we did not systematically investigate in our experiments. Note that the contrast is further reflected when the modals are modified: in the doctor to patient('s family) example, a use of "very well" is particularly marked, for the reason that strengthening the possibility clashes with the doctor's choice for *might*.

(16) *Doctor, to patient('s family)*: "This may/#might very well be cancer."

Speakers learn to access implicit attitudes towards speakers of different linguistic varieties and are competent in shifting between styles to achieve being perceived in one way or another (Eckert 2000; Beltrama 2016; Burnett 2017). A speaker using *may* could risk being perceived as arrogant or pretentious, whereas by using *might* s/he has better chances for coming across as more friendly, sincere, and supportive – even at the risk of being perceived as inarticulate, doubtful, or hesitant. Conversely, a speaker may choose to risk being perceived as authoritative, by using *may*, so as to not be perceived as inarticulate, incompetent, or casual for using *might*. The two possibilities can be seen as pragmatically enriched in the following way:
May: *"speaker is committed to a possibility"*
Might: *"speaker is not committed to a possibility"*

The CKV analysis can explain the contrast detected in (14) and (15) in terms of a core semantic strength asymmetry. Being stronger, *may* introduces a possibility to reckon with, while *might* introduces a possibility that can be taken to be unlikely. Note that this unlikelihood inference is conditioned on social context: one can also use *might* when one means the stronger sense but stops short of greater commitment out of social considerations.[11] In other words, one can use *might* even when one does not wish to communicate that something is unlikely, *per se*. This is a particularly available option when you yourself expect p, but your interlocutor does not expect or is averse to the possibility of p (as in (15)): the decision to use *might* reflects face-saving politeness considerations given the incongruence of epistemic states between interlocutors. In the doctor-to-colleagues *it may be cancer* example in (14), the doctor introduces cancer as a likely possibility, one that needs to be reckoned with, while in the *it might be cancer* example, cancer is inferred to be a (unexpected, unlikley) possibility, or the speaker is perceived as hedging her own certainty or authoritativeness. In this professional interaction between colleagues, there is no reason for the doctor to be polite or sensitive in the same way she would be to a patient's family; thus, *might be cancer* reliably communicates "not so sure it's cancer" with colleagues but not with patients.

Clearly, the respective expressive capabilities of *might* and *may* are at least in part a function of their respective abilities to convey social meaning in a given context. Indeed, the influence of social meaning may well have confounded our attempts to disentangle *may* from *might* in the experiments we report here. Future research must take this context sensitivity seriously. That epistemic modals in English are sensitive to register and social meaning also has implications for the question with which we began the paper: what is the dual of *must*? At a minimum, an answer to this question will need to disentangle the semantics of epistemic modals from their distributional patterns across registers and contexts of use, as well as address the fact that *might* appears to be replacing *may* in spoken English. Indeed, the final answer may prove

[11] See Krifka (2015, et seq) for a formal account of commitment.

to be variable depending on the social dynamics of a given discourse context. We leave this exploration of the interaction between social meaning and modality to future work.

Bates, D., Mächler, M., Bolker, B., & Walker, S. 2014. Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1-48.
Beltrama, A. 2016. Bridging the gap: intensifiers between semantic and social meaning. PhD dissertation. University of Chicago.
Biber, D., S. Johansson, G. Leech, S. Conrad, & E. Finegan. 1999. The Longman grammar of spoken and written English. London: Longman.
Bowie, J, Wallis, S, Aarts, S. 2013. Contemporary change in modal usage in spoken British English: mapping the impact of "genre". In: M. Carretero, J. Arús Hita, J. Van der Auwera, JI. Marín-Arrese (eds.) English modality: core, periphery and evidentiality. 57-94. De Gruyter: Berlin and New York.
Burnett, H. 2017. Sociolinguistic Interaction and Identity Construction: The View from Game-Theoretic Pragmatics. Journal of Sociolinguistics, 21, 2.
Crespo, I., H. Karawani, and F. Veltman. 2017. Expressing expectations. In D. Ball and B. Rabern (eds.), The Science of Meaning: Essays on the Metatheory of Natural Language Semantics. Oxford Univerisity Press.
Davies, Mark. 2008. The Corpus of Contemporary American English (COCA): 520 million words, 1990-present.
Degen, J. 2015. Investigating the distribution of some (but not all) implicatures using corpora and web-based methods. Semantics and Pragmatics 8(11). 1-55.
Degen, J., Scontras, G., Trotzke, A., & Wittenberg, E. 2016. Definitely, maybe: Approaching speaker commitment experimentally. MS, Stanford University.
Eckert, P. 2000. Language variation as social practice: The linguistic construction of identity in Belten High. Wiley-Blackwell.
von Fintel, K. & A. Gillies. 2010. Must...stay...strong! Natural Language Semantics 18(4), 351–383.
Giannakidou, A. 1997. The Landscape of Polarity Items. Ph.D. thesis. U Groningen.
Gonzales, M., Roseano, P., Borras-Comes, J., Prieto, P., (2017). Epistemic and evidential marking in discourse: Effects of register and debatability. Lingua 186-187. 68-87.
Iatridou S. 2000. The grammatical ingredients of counterfactuality Linguistic Inquiry. 31: 231-270.
Karawani, H. 2014. The Real, the Fake, and the Fake Fake in Counterfactual Conditionals, Crosslinguistically. Doctoral dissertation, University of Amsterdam. LOT Dissertation Series 357, 2014.
Karttunen, L. 1972. Possible and must. In John Kimball (ed.), Syntax and semantics, vol. 1, Seminar Press.
Kratzer, A. 1991. Modality. In Arnim von Stechow and Dieter Wunderlich (eds.), Semantics: An international handbook of contemporary research. 639–650. de Gruyter.
Lassiter, D. 2016. Must, knowledge, & (in)directness. Natural Language Semantics 24(2). 117-163.
Leech, G. N. 2004. Meaning and the English verb. Pearson Education.
Krifka, M. 2015. Bias in Commitment Space Semantics: Declarative questions, negated questions, and question tags. Semantics and Linguistic Theory (SALT) 25, 328-345.
Lambert, W.E., R.C. Hodgson, R.C. Gardner, and S. Fillenbaum. 1960. Evaluational reactions to spoken language. Journal of Abnormal and Social Psychology 60.1: 44-51.
Nuyts, J. 2001. Epistemic modality, language, and conceptualization: A cognitive-pragmatic perspective. John Benjamins Publishing.
R Core Team 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Rudin, D. 2015. Deriving a variable-strength might. Sinn und Bedeutung 20. 587–603.
Veltman, F. 1985. Logics for Conditionals. Ph.D. thesis, University of Amsterdam.
Veltman, F. 1986. Data semantics and the pragmatics of indicative conditionals. In E. Traugott et al. (eds.), On conditionals. 147–167. Cambridge University Press.
Venables, W. N. & Ripley, B. D. 2002 Modern Applied Statistics with S. Fourth Edition.
White, A. R. 1975. Modal thinking. Cornell University Press, Ithaca, N.Y.
Willer, M. 2013. Dynamics of epistemic modality. Philosophical Review 122(1). 45-92.