Modality in Language

Eric Swanson
ericsw@umich.edu

Abstract: This article discusses some of the ways in which natural language can express modal information—information which is, to a first approximation, about what could be or must be the case, as opposed to being about what actually is the case. It motivates, explains, and raises problems for Angelika Kratzer's influential theory of modal auxiliaries, and introduces a new approach to one important debate about the relationships between modality, evidentiality, context change, and imperative force.

Natural language provides many ways to express modal information—information which is, to a first approximation, about what could be or must be the case, as opposed to being about what actually is the case. In some cases it's controversial whether a sentence expresses modal information. But as a starting point, consider the following examples of modality in language:

(1) It's not possible that this table be made of ice.

(2) If this table is made of wood, then it couldn't have been made of ice.

(3) If this table had been made of ice, then it wouldn't have been possible for it to be made of wood.

(4) The table she ordered should be made of ice.

(5) The table she ordered is to be made of ice.

(6) Most of the tables here might be made of ice.

(7)   A table made of ice wouldn't last very long in the summer.

(8)   It's at least somewhat likely that a table made of ice would be hard to carry.

Although these sentences all make modal claims, the claims they make are quite different in character. For example, the respects in which (1) and (6) are about 'what could be the case' are quite different: (1) is most naturally read as being about what is metaphysically possible, and (6) is most naturally read as being about what is epistemically possible. By the same token, sentences of very different syntactic form—like (4) and (5)—can be used to make quite similar modal claims.

In most formal work on the semantics of modal expressions the nature of 'modal information' is assumed to be fairly uniform. For example, it's generally presupposed that whatever 'flavor' of modality is targeted by a given reading of (1)–(8), that reading expresses a *proposition.* Among other things, propositions at least determine truth conditions. On this way of thinking about the meaning of modal expressions, sentences like (1)–(8) rule out any possible worlds with which they are inconsistent. In a sense they thus bear information about worlds that are merely possible as well as information about the actual world; indeed, they bear information about the actual world *by* bearing information about which worlds are not actual.

The assumption that such sentences all express propositions has been criticized from a range of perspectives, and with good reason. For example, philosophers who think that certain kinds of normative claims do not express propositions typically also think that (4), for example, expresses a proposition on its epistemic reading but not on the relevant normative readings. What its meaning *is,* and how it is engendered, will of course depend on the particular non-standard semantics at issue.[1]

Even the assumption that all such sentences are *meaningful* has been called into question. In particular, philosophers have raised worries about so-called 'de re' modality. Quine's influential criticisms in this vein center on the 'third grade of modal involvement'—the "expression of necessity by a sentence operator." Such modal expressions can be attached "not only to statements but also to open sentences, such as '$x > 5$', preparatory to the ultimate attachment of quantifiers" (1953, 159). Quine argues that any logic that allows modal operators to combine with open sentences demands a "reversion to Aristotelian essentialism": any object "must be seen as having

---

1. For details, see especially GIBBARD 1990 and BLACKBURN 1993.

some of its traits necessarily and others contingently" (1961, 155). Because he thinks such a commitment indefensible, Quine insists that any cogent notion of necessity makes it a property of sentences (1953, 176). He takes his arguments to show that there is no other legitimate form of modal information, despite the broad swaths of natural language in which 'quantifying in' to the scope of modal operators seems to yield distinctive meanings.[2]

However one regards Quine's arguments, and his animus toward "essentialism," philosophers and linguists now tend to be reluctant to discount apparently useful fragments of natural language. Modal language in particular was rehabilitated through the galvanizing influence of Kripke's *Naming and Necessity* (1980)[3] and the explanatory power of the Stalnaker/Lewis analyses of counterfactuals (STALNAKER 1968 and LEWIS 1973a), among much other work. With emerging appreciation of the indispensability and intrinsic interest of modal language came the search for compositional semantic theories that could explain (among many other things) the very interactions between modal operators and quantifiers that Quine dismissed. This article is devoted to such theories and to some of the problems they face.

Many distinctions have been drawn between different kinds of modality. It's common, for example, to distinguish between

- deontic possibility (what is compatible with the dictates of morality)

- logical possibility (what is compatible with the laws of logic)

- bouletic possibility (what is compatible with a person's desires)

- nomological possibility (what is compatible with the laws of nature)

- epistemic possibility (often thought of as what is compatible with what is known or believed)

---

2. Quine is not alone here. Compare Kant: "The modality of judgments ... contributes nothing to the content of the judgment ..." (1997, A74/B99–B100). And Frege: "If I call a proposition necessary, I thereby give a hint as to my grounds for judgement. *But ... this does not affect the conceptual content of the judgement ...*" (1879, 55). More recently, VON FINTEL & IATRIDOU 2003 argues for the "descriptive generalization" that quantifiers cannot scope over epistemic modals. But see SWANSON 2008 for several kinds of counterexamples (e.g., 'Every inch of the floor might have paint on it').

3. For an introduction to issues surrounding modality and proper names, see CAPLAN 2006.

and still others (see PALMER 2001, 4–18). It's important to draw distinctions like these because there are often dramatic differences between (say) what is compatible with the laws of nature and what is compatible with what is known. Notice that differences in the kinds of modality that a claim is about are not always marked by differences in the words used to make the modal claims. For example, 'must,' 'have to,' and 'could' all have epistemic, deontic, and bouletic uses. As a result it's often important (and sometimes difficult) to get clear about what kind of modal claim a particular sentence expresses. This is not to say that the distinctions are always clear-cut: sometimes a given modal expression can be plausibly interpreted as having any of several modal flavors, or as involving a 'mix' of two or more modalities.

The striking ease with which a single modal expression can target different modalities is sometimes taken to suggest that context alone determines which modality is targeted—that there is no lexical difference between epistemic and deontic 'must,' for example, but only a difference in some parameter or parameters supplied by the context. An analogy may be helpful here. We should not think that 'every bottle' is lexically ambiguous just because it can mean *every bottle in the house,* or *every bottle in the fridge,* or *every bottle in the city,* or . . . . Similarly for modals: clearly 'must' can target different kinds of modality, but this is not reason enough to think that 'must' has a distinct lexical entry for each of these kinds of modality. Rather—the thought goes—just as quantifiers require contextual supplementation to fix a domain of quantification, modals require contextual supplementation to fix on a particular kind (or kinds) of modality. In principle one could posit lexical ambiguities instead of taking this strategy. But this would at least prima facie violate the injunction not to multiply senses beyond necessity—what Grice calls "Modified Occam's razor." Moreover, we want to explain the often surprising features shared by expressions that target different flavors of modality. For example, knowing only that the same modality is targeted by

(9) It could be that $\neg\phi$.

and

(10) It must be that ϕ.

we can tell that the sentences contradict each other. What modality it is that they both

4

target is immaterial. Positing lexical ambiguities for different modal flavors makes it harder to explain facts like to this one.

A simple and at least initially attractive way to implement the idea that modals require contextual supplementation is to exploit the analogy to quantifiers mentioned above, treating modals as quantifiers over contextually determined domains of possible worlds (see, e.g., Lyons 1977, 787–791). We might say that 'might,' 'could,' 'can,' and other possibility modals express existential quantification over a domain of possible worlds, and that 'must,' 'should,' and so on express universal quantification. Qua quantifier, a modal can bind the variable for the world relative to which expressions within its scope are evaluated. On a deontic reading of 'It should be that ϕ,' for example, the sentence expresses the proposition that in *all* worlds compatible with what ought to be the case, it's true that ϕ. A deontic reading of 'It could be that ϕ' expresses the proposition that in *some* world compatible with what ought to be the case, it's true that ϕ. The fact that (for a given modality) (9) and (10) contradict each other is then a special case of the fact that (for a given domain of quantification) $\exists x \neg Px$ and $\forall x Px$ contradict each other. The apparent duality of certain modals is explained by the duality of $\exists$ and $\forall$.

Angelika Kratzer—the linguist whose semantics for modals is the starting point and benchmark for most formal work on them—wants to hold on to the hypotheses that modals of different flavors have a common semantics and that the targeting of a particular modality depends on contextual supplementation. But she does not think that modals should be simply treated as quantifiers. Indeed, she gives three of the most influential reasons for rejecting such approaches (see her 1977, 1981, and especially 1991).

1. Suppose that a pair of equipotent judges issue contradictory rulings. Then the domain of worlds over which jurisprudential modals quantify will be empty, because there are no worlds in which the contents of both rulings come out true. This has absurd results: sentences headed by necessity modals will be vacuously true, and sentences headed by possibility modals will be vacuously false.

2. The standard analyses of modals and conditionals have trouble handling sentences like

    (11)   If a murder occurs, the jurors must convene.

In such sentences the necessity modal scopes over either the consequent alone or over the entire conditional. (For illustrative purposes, Kratzer takes the material conditional analysis as her foil (1991, 643); the reader is invited to work out how other analyses of conditionals fare with such sentences.)

Suppose the necessity modal scopes over the consequent alone. Then if no murder occurs, the conditional is true. Worse still, if a murder does occur, then the conditional comes out true iff the jurors must convene simpliciter. But the law needn't make that kind of demand; it can demand simply that they must convene in the event that a murder occurs.

Suppose the necessity modal scopes over the entire conditional. Suppose further that the law demands both that no murders occur and that if a murder occurs, then the jurors must convene. Then with 'must' interpreted in the standard way, and as targeting the law's demands,

> (12)   If a murder occurs, then it must be the case that ϕ.

will be true for *any* ϕ. This is because the wide scope 'must' makes the conditional it embeds be evaluated only at worlds in which no murder occurs, and material conditionals with antecedents that are false at the world of evaluation are true simpliciter.

3. Just as there are many determiners with intermediate quantificational strength—quantifiers that are in some sense "between" 'all' and 'some'—there are many modal expressions 'between' 'necessarily' and 'possibly'. Kratzer gives examples like 'it is easily possible,' 'it is barely possible,' 'it is more desirable that ϕ than that ψ,' and so on. Proportional quantificational determiners like 'most' are provably undefinable in terms of $\forall$ and $\exists$; "graded modals" are likewise undefinable in terms of $\Box$ and $\Diamond$.

We can overcome these problems, Kratzer suggests, by hypothesizing that modal expressions are considerably more complex than quantifiers. On her view, they require supplementation by *two* contextually determined factors: a **modal base** and a **partial ordering on worlds.**[4] Kratzer's partial orderings play a role like that played by selec-

---

4. In her dissertation Anette Frank develops an alternative to the use of ordering sources—"multiply

tion functions in Stalnaker's theory of counterfactuals, and by comparative similarity relations in LEWIS 1973a and 1973b. In particular, they let Kratzer's semantics talk not just about "what is true in the actual world, in all possible worlds, or in at least one, unspecified world" but "also about what is true in *particular* non-actual possible situations" (STALNAKER 1968, 46). I will treat the modal base as a function from worlds (in particular, from the world of evaluation) to sets of worlds, but it could just as well be formalized using an accessibility relation R (such that for functional modal base $f(\cdot)$, $Rww' \leftrightarrow w' \in f(w))$.[5] For example, an epistemic modal base might be a function that takes a world w as argument and yields the set of all worlds compatible with what is distributed knowledge at w, and a jurisprudential modal base might be a function that takes a world w as argument and yields the set of all worlds that are compatible with the laws of w. And the ordering source orders the sets of worlds yielded by the modal base: some worlds in that set better approximate some 'ideal' represented by a set of propositions. For any worlds w_1 and w_2, relative to an ideal set of propositions P, w_1 comes at least as close to the ideal given by P as w_2 does iff there are no propositions in P that are true in w_2 but false in w_1. Formally:

$$w_1 \leq_P w_2 =_{df} \forall p(w_2 \in p \in P \rightarrow w_1 \in p)$$

$\leq_P$ needn't be a total order because it's possible for there to be a pair of propositions in P, p and p', and a pair of worlds w and w', such that p is true in w and false in w', and p' is true in w' and false in w.

Kratzer then gives "doubly relative" semantics for all the modal expressions she considers. Her intuitive, very rough treatment of 'is necessary' is "that a proposition is a necessity if and only if it is true in all accessible worlds which come closest to the ideal established by the ordering source" (644). This paraphrase suggests a very simple semantics:

FIRST PASS: p is necessary iff $\forall u \forall v (u \leq v \rightarrow u \in p)$

That is, p is necessary iff all worlds at least as good as any worlds are p worlds.

---

relative modality" (1997, 39–46).

5. In most of her 1991 paper, Kratzer treats the modal base as a function from worlds into sets of propositions (i.e., into sets of sets of worlds), and there are conceivable applications for the extra information that such functions could carry.

But Kratzer does not want to assume that there must be worlds that come closest to the ideal, or that are at least as good as any world. This is because she thinks that an ordering source needn't rule that there is some world, in any set of worlds closed under comparability, such that none are better than it.[6] For example, an ordering source could value the number of leaping lizards in a world in a strictly increasing way (holding other facts fixed). One way to get this kind of ordering is to include in P for every natural number n the proposition that there are at least n leaping lizards. Then given a modal base that includes, for every n, a world with n leaping lizards (and which, again, holds other facts fixed), for each world there will be a better world. Complicating the semantics as follows handles such cases:

SECOND PASS: p is necessary iff $\exists v \forall z (z \leq v \rightarrow z \in p)$

Intuitively, this says that p is necessary iff as we consider better and better worlds, eventually we hit a world such that all worlds at least as good as it are p worlds. This semantics gets the nice result that, given the 'leaping lizards' base and ordering source, (13) is true relative to any world of evaluation.

(13)   There ought to be one more leaping lizard, holding other facts fixed, than there is.

The problem with this semantics is that the ordering on worlds may be merely partial. Suppose that there is a world w which is worse than all the worlds in two sets closed under the 'better than' relation. Suppose further that p is true in all the worlds in one of the two sets, and false in the other. This semantics would make 'p is necessary' true relative to such an ordering when, intuitively, it should come out false, because of the $\neg p$ worlds that are better than w and also better than any worlds comparable to them.

To avoid these problems, Kratzer gives her official semantics as follows: with respect to modal base $f(\cdot)$ and ordering source $g(\cdot)$, and with the domain of quantification restricted to $f(w)$,

FINAL PASS: p is necessary in w iff $\forall u \exists v \big(v \leq_{g(w)} u \wedge \forall z (z \leq_{g(w)} v \rightarrow z \in p) \big)$

---

6. An assumption that is roughly analogous to the one Kratzer rejects is called the 'Limit Assumption' in the literature on conditionals. LEWIS 1973a and STALNAKER 1984 are standard entreés into the debate over its plausibility.

This says that p is necessary iff for every world there's some world that's at least as good as it (call it 'v') such that every world at least as good as v is a p world. As before, there's a point at which we do not see any $\neg p$ worlds as we go to better and better worlds, but the first such point needn't be the same for every world in the modal base. So this analysis doesn't presuppose that there are worlds that are at least as good as any worlds, and doesn't presuppose that the ordering source induces a total ordering.

In a moment I'll explain how Kratzer applies her analysis to contradictory rulings, conditionals, and graded modality. But first I want to note that it's important to Kratzer that *all* flavors of modality demand both a modal base and an ordering source. If graded modality, say, were her only motivation for introducing ordering sources, then the fact that it's difficult if not impossible to construct examples of graded metaphysical, logical, and nomological modality might be taken to suggest that ordering sources aren't needed for all modals. This result would be worrying since, although Kratzer's analysis is more complicated than purely quantificational approaches, she too imputes to modals a common lexical core. That said, Kratzer allows that apparent syntactic differences between epistemic and root modals "...may correlate with a difference in argument structure" (650). (The putative syntactic differences are complex and controversial; for recent work on the topic, see HACQUARD 2006 and the citations therein.) But without significant progress on the argument structure of modals, it won't be clear to what extent Kratzer's semantics would need to be changed to compensate for such a difference.

How do ordering sources help solve the problems Kratzer raised for the simple analysis? The problem engendered by equipotent but conflicting judges was that their inconsistent rulings would make the domain of quantification empty. In her sketch of a solution to the problem Kratzer assumes that their inconsistent rulings *do not* make the domain of quantification empty. (Indeed, for simplicity's sake she assumes that the modal base is maximally unopinionated: "for each world, the set of accessible worlds is the set of all possible worlds" (647).) Then as long as there are accessible worlds in which no murders occur, and no propositions in P contradict the proposition that there are no murders, then all the *best* worlds will be worlds in which no murders occur. By Kratzer's semantics for 'is necessary,' it will be necessary that no murders occur, given a jurisprudential reading of 'is necessary.'

It's worth noting here that Kratzer has trouble with demands that, intuitively,

should make both 'Must ϕ' and 'Must $\neg\phi$' true relative to the same index and context. Given her semantics, and the truth of 'Possibly ψ' for some ψ, the supposition that both must ϕ and must $\neg\phi$ leads to a contradiction: we can derive (1) that in every set of worlds in the modal base closed under comparability there is a world such that every world better than it is a ϕ world *and* (2) that in every set of worlds in the modal base closed under comparability there is a world such that every world better than it is a $\neg\phi$ world. For some modal flavors—for example, for nomological, metaphysical, and logical modalities—it's not plausible that 'Must ϕ' and 'Must $\neg\phi$' can together be true. But deontic, jurisprudential, bouletic, and some other modalities *can,* at least prima facie, demand both that ϕ and that $\neg\phi$. Prima facie I can have equally strong but incompatible moral obligations that make it the case both that I must lie and that I must not lie. At the same time, if we give a wholly unified treatment of the different 'flavors' of modality that allows such incompatibilities in one, it will also allow incompatibilities in the others.

One also might worry that it's not legitimate to assume that, in this case, there will be accessible worlds—let alone accessible worlds in which no murders occur. But I think Kratzer should be read more sympathetically than this: we should let our judgments about the truth value of modal claims tell us about the modal bases and ordering sources that are supposed to be relevant to them, rather than presupposing that some simple theory tells us all there is to know about their features. After all, Kratzer's aim is to give a relatively unified semantics for the diverse range of modal expressions, not to systematize all the (perhaps messy) facts about how context affects their semantic values. Looking at things this way, what we learn from the fact that, intuitively, 'Must ϕ' and 'Must $\neg\phi$' *can* together be true may be that when judges make incompatible rulings (to take one example), sometimes neither ruling will have any net effect on the contextually determined ordering source and modal base. But if that's ultimately Kratzer's favored treatment, then it's not clear that this example provides a good reason to reject the simple, quantifier-only analysis of 'is necessary.' The advocate of that analysis needs more than Kratzer does—rather than having *some* accessible no-murder worlds, he needs for *all* accessible worlds to be no-murder worlds—but if neither judge's ruling constricts the domain of worlds quantified over, that's easy to achieve. Kratzer does not take a stand on exactly how the modal base and ordering source are determined, and so does not rule out such an approach.

In fact, however, Kratzer herself needs some constraints on how context affects the modal base and ordering source. In the toy example we just considered, for example, the ordering source needs to rank worlds in which no murders occur as best. Plausibly, it comes to have that property through rulings of judges. But if that's right then it seems that contradictory rulings of judges might put incompatible demands on the *ordering* of worlds. (For examples of this sort of thing consider a case in which some laws of country one cannot be followed without violating some laws of country two, and vice versa. We could imagine judges from the two countries issuing rulings of the form 'It's worse to violate the laws of country x than to violate the laws of country y.') I already explained why I'm not convinced that cases involving contradictory absolute rulings give us good reason to posit ordering sources. But if Kratzer is right about how the dialectic goes there, it would seem that parallel arguments about contradictory rulings involving *relative* orderings might give us good reason to posit yet another dimension of context dependence, or to make ordering sources more complicated.

The next cluster of problems Kratzer raised for the simple analysis had to do with modalized indicative conditionals of the form

(11)   If a murder occurs, the jurors must convene.

Kratzer takes the surprising but not unprecedented view that English 'if…then' is not a two place connective but a way to restrict a modal base and make a modal claim with that restriction in place.[7] In other words, the antecedent of a conditional is a device by which a speaker can restrict the modal base of a (possibly covert) modal in the consequent. So at some underlying level of representation, the form of (11) is the same as the form of

(14)   The jurors must convene.

The crucial difference between them is that the modal base of the 'must' in (11) is explicitly restricted to worlds in which a murder occurs. The modal base of (11) has no such restriction. Formally:

$$[\![\text{if } \alpha, \text{ must } \beta]\!]^{f,g} = [\![\text{must } \beta]\!]^{f',g}, \text{ where for all } w \in W, f'(w) = f(w) \cap \{[\![\alpha]\!]^{f,g}\}$$

---

7. Lewis 1975 is indisputably an important predecessor of this view. It's arguable that some 'conditional assertion' accounts also have affinities with it.

On Kratzer's theory, then, (11) means that for every set of worlds in the modal base that is closed under comparability and in which a murder occurs, as we look at better and better worlds within that set, eventually we come to a world such that there are no better worlds than it in which the jury does not convene. It's as though the process of 'perfecting' a world in which a murder occurs would, from some point onward, never involve the jury's failing to convene. Nevertheless, such an ordering can allow that some worlds in which a murder occurs and the jury does not convene are better than some worlds in a which a murder occurs and the jury does convene.

But Kratzer's treatment is not the only game in town: if the antecedent of an indicative conditional effects a context shift, then in principle the modal could take narrow scope without requiring an ordering source. Thony Gillies defends such an analysis of epistemic modals in indicative conditionals, holding that in an indicative conditional, the consequent is evaluated not relative to the global context, but relative to "the subordinate or derived context got by adding the [proposition expressed by the antecedent] to it" (GILLIES 2007, 24; see also (e.g.) STALNAKER 1988, 156). This approach works only if the relevant modal base *just is* the context (as it might be for epistemic modals) or by its nature cannot include worlds outside the context, since otherwise adding the information that p to the context isn't the same as adding the information that p to the modal base. So Gillies' approach can't be directly extended to non-epistemic modals. Kratzer's theory, on the other hand, elegantly unifies not only diverse modals, but also modals and conditionals. As we will see, however, differences between epistemic modality and other 'flavors' of modality give us some reason to think that, even if they have a common core of meaning, their complete treatments will have significant differences.

The final application for ordering sources that Kratzer discusses—and, probably, the most obvious application for them—is in giving semantics for "graded modals" like 'probably,' 'there is a good possibility that,' 'there is a slight possibility that,' 'is more likely than,' and so on. Ordering sources let us compare possibilities, for example: for p to be at least as good a possibility as q is for there to be, for every q world, at least as good a p world; for p to be a better possibility than q is for p to be at least as good as q and for q not to be at least as good as p. Kratzer puts this treatment of comparative possibility to work in her semantics for 'probably': on her account 'probably p' means that p is a better possibility than $\neg p$. The ordering source can induce

quite a bit of structure: for example, it can represent that p is better than q is better than r is better than s by including exactly p, $p \vee q$, $p \vee q \vee r$, and $p \vee q \vee r \vee s$. For many modal flavors this kind of treatment seems adequate. It's illegal to torture one innocent and illegal to torture two; torturing two is worse, in the eyes of the law, than torturing one; but it's not *twice as illegal*.[8] But Kratzer's ordering sources do not carry enough information to handle quantitative characterizations of subjective uncertainty like

(15) Rain was about 60% likely.

(16) It was twice as likely to rain as it was to snow.

(17) If it rained, then they were twice as likely to get in an accident.

This suggests that, for all their similarities, there are profound differences between epistemic modals and at least some non-epistemic modals—differences which Kratzer's theory cannot accommodate. (See Swanson 2006 and 2008 and Yalcin 2007a and 2007b for more on quantitative expressions of subjective uncertainty.)

Another problem for Kratzer's treatment of epistemic modals, discussed at length in Yalcin 2007a and 2007b, is that it makes the wrong predictions about how they will embed under 'suppose that'. Although

(18) #Suppose that it's raining and it might not be raining.

is infelicitous,

(19) Suppose that it's raining and for all we know it isn't raining.[9]

is perfectly felicitous. Another way to see this contrast is to imagine a speaker who has just learned about Bond's penchant for reverse psychology uttering the conditionals

---

8. Perhaps it is twice as bad, or twice as wrong—judgments are mixed.

9. On Kratzer's semantics 'might' does not mean quite the same thing as 'for all we know,' but the relevant judgments persist even if we replace 'for all we know' with a more accurate paraphrase of Kratzer's semantic value for epistemic 'might'. (For one thing, if 'might' meant 'for all we know' then, given that 'might' and 'must' are duals, Kratzer's account would predict that 'Must p' entails 'p', an entailment which many have found counterintuitive. To avoid this Kratzer allows that the ordering source for epistemic modals may not consist of facts alone: see 1991, 645. For arguments that 'Must p' does entail 'p', see von Fintel & Gillies 2008.)

(20) #If Bond might be in London, then he's not in London.

(21)   If for all we know Bond is in London, then he's not in London.

Evaluating (20) involves supposing both that Bond might be in London and that Bond isn't in London—which, as Yalcin's original sentences suggest, is difficult if not impossible.

Many English epistemic modals carry an evidential feature that indicates that the speaker has derived the prejacent—the clause that is the modal's complement—via an inference. For example, if I believe that it is raining on the basis of watching it rain from the window, my reporting this belief to you (over the phone, say) using (22) sounds much better than my using (23).

(22)   It is raining.

(23) #It must be raining.

(23) strongly suggests that I have inferred that it is raining, and that I do not have direct perceptual evidence to that effect. So 'must' carries an "inferential" evidential feature (Karttunen 1972).

English also distinguishes between inferences that the speaker takes to establish (or likely establish) the truth of the modal's prejacent, and inferences that the speaker needn't take to establish the truth of the prejacent. To see this consider the contrast between

(24)   They left an hour ago, and there isn't any traffic. So they should be here by now. But they're not.

(25) # They left an hour ago, and there isn't any traffic. So they must be here by now. But they're not.[10]

Examples like these show that while 'should' and 'must' do both carry an evidential

---

10. 'Should' is often given as an example of an epistemic modal, especially in the typological and historical literature, but scant attention has been paid to its evidential feature (though see Behre 1955, 17, Huddleston & Pullum 2002, 186, and Copley 2006). The *OED* offers helpful illustrations. Byron, *Cain* (1821): "I have heard it said, That seraphs *love most*—cherubim *know most*—And this should be a cherub—since he loves not."; Kingsley, *Westward Ho!* (1855): "That should be Barbados...unless my reckoning is far out."

feature signalling inference, they can convey quite different attitudes on the part of the speaker toward the inferred conclusion.[11] 'Should ϕ' is compatible with the denial of 'ϕ'—indeed, that is perhaps its paradigmatic use. It often indicates simply that a salient body of information that doesn't entail that ϕ would typically be taken to be a good reason for one to infer that ϕ. In addition to its inferential signal, 'Must ϕ' indicates that the speaker has at least very high credence that ϕ. In short, 'Must ϕ' signals an inference that the speaker thinks is (at least) likely sound; 'Should ϕ' doesn't.

This contrast between epistemic 'should' and 'must' parallels a contrast between deontic 'should' and 'must':

(26)   He should go, but he won't.

(27) # He must go, but he won't.[12]

Notice that, again, 'Should ϕ' is compatible with the denial of 'ϕ', and 'Must ϕ' is not. NINAN 2005 argues that it's quite difficult for Kratzer to explain the contrast, and suggests that "An utterance of 'S must ϕ' asserts that S must ϕ and adds the proposition that S ϕ's to someone's To-Do List" (in the sense of PORTNER 2004). This is in contrast to an utterance of 'S should ϕ,' which generally won't change anyone's To-Do List (NINAN 2005, 20). It is pragmatically infelicitous for a speaker to put something on someone's To-Do List only to assert that it won't get done. So (27) is infelicitous while (26) may be perfectly fine.[13]

Unfortunately, Ninan's account does not help explain the contrast between *epistemic* 'should' and 'must,' and it would be nice to have a single explanation of the two contrasts if possible. Here is a rough sketch of such an explanation. It seems that 'must' has an "imperative-like force" that distinguishes it from some other modals

---

11. This difference is blurred somewhat in the standard typologies of evidentials (e.g., WILLETT 1988 and AIKHENVALD 2004, chapter 2), which abstract away from the fact that some evidentials signal that the speaker does not herself vouch for the claim. Aikhenvald briefly discusses commitment attenuation in Quechua non-direct evidentials (162) and typical reported evidentials (176–177).

12. This example is from LYONS 1977, 846. 'Ought to' patterns with 'should,' as Lyons notes, and 'have to' and 'can' (under wide scope negation) pattern with 'must':

  (i) #You can't break that promise, but you're going to.

13. For other discussions of the contrast between deontic 'should' and 'must', see PALMER 1986, 100 and WERNER 2003, 125–126.

(Ninan 2005, 9). As Leech puts it, " 'Have to' indicates the general idea of obligation, but 'must' more precisely identifies the speaker as the person who gives the orders. … the speaker is the agent of coercion" (1970, 228). And it seems that 'should' is more descriptive than imperative:

> … 'you should' does not force the speaker's will upon the hearer in the same way as 'you shall' does. If I say "you should work harder" … I am suggesting, tactfully and considerately, the presence of necessity, duty, or obligation, but I am concerned not to give you the impression that I am the one to impose it upon you. If the thing is to be done it is not because I wish it but because it is required by a general standard of propriety or obligation…. (Behre 1955, 16)

Building on these intuitions, my suggestion is that 'must' is (at least usually) used to *change* 'what must be the case' by making changes to the modal base and ordering source—or to whatever aspects of the context in fact play their putative role—and that 'should' is (at least usually) used simply to *describe* standing features of the base and ordering source. It's plausible that it would be pragmatically infelicitous for a speaker to *change* the modal base and ordering source, only to deny that the new ideal will be realized. But one can well imagine such a denial being felicitous after one has merely *described* the base and ordering source.

  So we can now explain the contrast between (26) and (27), and we can also explain the contrast between (24) and (25):

(24)   They left an hour ago, and there isn't any traffic. So they should be here by now. But they're not.

(25) #They left an hour ago, and there isn't any traffic. So they must be here by now. But they're not.

With (24) a speaker notes that on the basis of some body of information one would naturally expect that they would be here by now, but does not add the proposition that they are here to the modal base or to the conversational common ground. With "They must be here by now" the speaker *does* aim to add to the common ground the proposition that they are here (cf. Swanson 2006, 45, 73, and Swanson 2008, 10–13). That is why 'Must ϕ' is incompatible with a denial of 'ϕ'. I am here assuming, with Kratzer, that participants in a conversation are often ignorant of features of the modal base and ordering source, despite the fact that their features are at least partially determined by

the context. I am also assuming that the 'direction of fit' between modal bases and ordering sources, on the one hand, and communication, on the other, is not one-way. For prosaic examples of this, consider the fact that issuing rulings is a way for judges to change what the law demands, and the fact that making promises is a way for people to change what morality demands.[14]

The substantial differences between epistemic and so-called 'root' (non-epistemic) modals make it unclear precisely what one is aiming for in giving a 'relatively unified' semantics of different modal expressions. At the same time, there are enough similarities between the flavors of modality—the contrasts between 'should' and 'must' that we have been considering are an illustrative example—that it is fruitful to look for interesting phenomena involving one flavor of modality where we've already found such phenomena involving another. A plausible explanation of these similarities is that at least some of them are due to shared history.[15] But it's consistent with this that some ways of expressing modal information have come to have quite different features, and that they now demand quite different semantics.

---

14. The fact that 'must'/'have to' signal that the speaker is committed (or nearly committed) to the prejacent, and 'should'/'ought to' do not signal this, suggests to me that their evidential features differ. (It's plausible to attribute this difference to differences in their evidential features because many evidentials have the quite distinctive function of signalling that the speaker does not commit to the prejacent—Aikhenvald 2004, 162, 176–177.) If this is right, then at least some aspect of their evidential features must be 'hardwired'—not derivable on the basis of conversational principles plus an austere semantics. For dissent, see von Fintel & Gillies 2008, who argue that because in many languages epistemic modals "are markers of indirect inference" (19), their evidential feature "should not be a stipulated, arbitrary part of … lexical meaning" (20).

15. For more on the evolution of modality in language, see Anderson 1986, Traugott 1989, Bybee et al. 1994, Bybee 1995, and van der Auwera & Plungian 1998.

References

Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford University Press, Oxford.

Anderson, Lloyd B. 1986. "Evidentials, Paths of Change, and Mental Maps: Typologically Regular Asymmetries." In *Evidentiality: The Linguistic Coding of Epistemology*, Wallace Chafe & Johanna Nichols, editors. Ablex Publishing Corporation, Norwood, NJ.

van der Auwera, Johan & Vladimir A. Plungian. 1998. "Modality's Semantic Map." *Linguistic Typology*, vol. 2: 79–123.

Behre, Frank. 1955. *Meditative-Polemic Should in Modern English That-Clauses*. Almqvist & Wiksell, Stockholm.

Blackburn, Simon. 1993. *Essays in Quasi-Realism*. Oxford University Press, Oxford.

Bybee, Joan. 1995. "The Semantic Development of Past Tense Modals in English." In *Modality in Grammar and Discourse*, Joan Bybee & Suzanne Fleischman, editors, 503–517. John Benjamins, Amsterdam.

Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. University of Chicago Press, Chicago.

Caplan, Ben. 2006. "On Sense and Direct Reference." *Philosophy Compass*, vol. 1 (2): 171–185.

Copley, Bridget. 2006. "What Should *Should* Mean?" Ms., CNRS/Université Paris 8, URL http://copley.free.fr/copley.should.pdf.

von Fintel, Kai & Anthony S. Gillies. 2008. "*Must*…Stay …Strong!" Ms., Massachusetts Institute of Technology and University of Michigan, Ann Arbor, URL http://mit.edu/fintel/fintel-gillies-2008-mss.pdf.

von Fintel, Kai & Sabine Iatridou. 2003. "Epistemic Containment." *Linguistic Inquiry*, vol. 34 (2): 173–198.

FRANK, ANETTE. 1997. *Context Dependence in Modal Constructions.* Ph.D. thesis, University of Stuttgart.

FREGE, GOTTLOB. 1879. "*Begriffsschrift:* Selections." In *The Frege Reader*, MICHAEL BEANEY, editor. Blackwell Publishers, Ltd., Oxford. Translated by Michael Beaney.

GIBBARD, ALLAN. 1990. *Wise Choices, Apt Feelings.* Harvard University Press, Cambridge.

GILLIES, ANTHONY S. 2007. "Iffiness." Ms., Department of Philosophy, University of Michigan Ann Arbor., URL `semanticsarchive.net/Archive/TI1OGVlY/iffiness.pdf`.

HACQUARD, VALENTINE. 2006. *Aspects of Modality*. Ph.D. thesis, Massachusetts Institute of Technology.

HUDDLESTON, RODNEY & GEOFFREY K. PULLUM, editors. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge.

KANT, IMMANUEL. 1997. *Critique of Pure Reason.* Cambridge University Press. Translated and edited by Paul Guyer and Allen W. Wood.

KARTTUNEN, LAURI. 1972. "*Possible* and *Must.*" *Syntax and Semantics*, vol. 1: 1–20.

KRATZER, ANGELIKA. 1977. "What *Must* and *Can* Must and Can Mean." *Linguistics and Philosophy*, vol. 1: 337–355.

—. 1981. "The Notional Category of Modality." In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, HANS JURGEN EIKMEYER & HANNES RIESER, editors, 38–74. W. de Gruyter, Berlin.

—. 1991. "Modality." In *Semantics: An International Handbook of Contemporary Research*, ARNIM VON STECHOW & DIETER WUNDERLICH, editors, 639–650. W. de Gruyter, Berlin.

KRIPKE, SAUL. 1980. *Naming and Necessity.* Harvard University Press, Cambridge.

LEECH, GEOFFREY N. 1970. *Towards a Semantic Description of English.* Indiana University Press, Bloomington.

Lewis, David K. 1973a. *Counterfactuals*. Basil Blackwell Ltd., Malden, MA.

—. 1973b. "Counterfactuals and Comparative Possibility." *Journal of Philosophical Logic*, vol. 2.

—. 1975. "Adverbs of Quantification." In *Formal Semantics of Natural Language*, Edward L. Keenan, editor, 3–15. Cambridge University Press, Cambridge.

Lyons, John. 1977. *Semantics*, vol. 2. Cambridge University Press, Cambridge.

Ninan, Dilip. 2005. "Two Puzzles About Deontic Necessity." In *New Work on Modality*, Jon Gajewski, Valentine Hacquard, Bernard Nickel & Seth Yalcin, editors. MIT Working Papers in Linguistics, vol. 51. URL `http://web.mit.edu/dninan/www/papers.html`.

Palmer, F. R. 1986. *Mood and Modality*. Cambridge University Press, Cambridge.

—. 2001. *Mood and Modality*. Cambridge University Press, Cambridge, second edn.

Portner, Paul. 2004. "The Semantics of Imperatives within a Theory of Clause Types." In *Proceedings of Semantics and Linguistic Theory (SALT) 14*, Kazuha Watanabe & Robert B. Young, editors. CLC Publications, Ithaca.

Quine, W. V. O. 1953. "Three Grades of Modal Involvement." In *The Ways of Paradox*, 158–176. Random House, New York.

—. 1961. "Reference and Modality." In *From a Logical Point of View*. Harvard University Press, Cambridge, second edn.

Stalnaker, Robert C. 1968. "A Theory of Conditionals." In *Ifs: Conditionals, Belief, Decision, Chance, and Time*, William L. Harper, Robert Stalnaker & Glenn Pearce, editors, 41–55. D. Reidel Publishing Company, Dordrecht.

—. 1984. *Inquiry*. MIT Press, Cambridge.

—. 1988. "Belief Attribution and Context." In *Context and Content*, 150–166. Oxford University Press, Oxford.

Swanson, Eric. 2006. *Interactions with Context*. Ph.D. thesis, Massachusetts Institute of Technology.

—. 2008. "How Not to Theorize about the Language of Subjective Uncertainty." Forthcoming in *Epistemic Modality,* Andy Egan and Brian Weatherson, eds.

TRAUGOTT, ELIZABETH CROSS. 1989. "On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change." *Language*, vol. 65: 31–55.

WERNER, THOMAS A. 2003. *Deducing the Future and Distinguishing the Past: Temporal Interpretation in Modal Sentences in English*. Ph.D. thesis, Rutgers, The State University of New Jersey.

WILLETT, THOMAS L. 1988. "A Cross-Linguistic Survey of the Grammaticalization of Evidentiality." *Studies in Language*, vol. 12: 51–97.

YALCIN, SETH. 2007a. "Epistemic Modals." *Mind*, vol. 116: 983–1026.

—. 2007b. "Nonfactualism about Epistemic Modality." Ms., Massachusetts Institute of Technology.