Valentine Hacquard

Linguistics Department, University of Maryland, College Park

1401 Marie Mount Hall, College Park, MD 20742

Phone: (301) 405-4935

hacquard@umd.edu

August 14th, 2009

58. Modality

Keywords: possible worlds, modality, intensionality

58. Modality

Abstract: *Modality is the category of meaning used to talk about possibilities and necessities, essentially, states of affairs beyond the actual. This article reviews the approach to modals inherited from modal logic, in terms of quantification over possible worlds, with particular attention to the seminal work of Angelika Kratzer. In addition, it introduces more recent work on the interaction of modals with other elements, in particular with tense and subjects, which challenges classical approaches, and present new directions.*

1. Introduction

Modality is the category of meaning used to talk about possibilities and necessities, essentially, states of affairs beyond the actual. We can talk about what we *must* do, if we are to obey the law (we *must* pick up after our dogs), or what we *may* do to fulfill our desires (we *may* go on sabbatical), what *could* happen if global warming isn't abated (the world as we know it *could* disappear), or what *would* have been if Cleopatra's nose had been shorter (the face of the world *would* have been changed). All of these hypothetical states of affairs may never come to be, yet we are able to talk about them, by using modal words. Modality is

expressed by many categories of lexical items: adverbs like *maybe*, nouns like *possibility*, adjectives like *possible*, or auxiliary verbs like *must*, *may*, *should* or *have to*. This article focuses on modal auxiliaries, since their relatively well-established properties serve as a good basis to present issues and theories of modality. We start by reviewing the approach to modals inherited from modal logic, in terms of quantification over possible worlds, with particular attention to the seminal work of Angelika Kratzer. We then turn to more recent work on the interaction of modals with other elements, in particular with tense and subjects, which challenges classical approaches, and present new directions.

2. Properties of modals

Natural language modals seem to vary along (at least) two dimensions: 'force' (whether they express possibility or necessity), and type of interpretation, or modal 'flavor'. In English, *possibility* modals include *may, might, can,* and *could*. *Necessity* modals include *should, must, would,* and *have to*. Rather than considering possibility or necessity with respect to *all* non-actual states, natural language modals often seem to be relative to a certain body of laws, desires, or information, giving rise to the various 'flavors' of modality. *Epistemic* modality (from Greek *episteme 'knowledge'*) expresses possibilities and necessities given what is known, based on what the available evidence is; *deontic* modality (from Greek *deon 'obligation'*), possibilities and necessities given a body of laws or rules, i.e., permissions and obligations; *abilitive* modality, possibilities given the subject's physical abilities; *teleological* and *bouletic* modality, possibilities and necessities given particular goals and desires (from Greek *telos 'goal'* and *boule 'wish'*). The following examples illustrate:

(1) a. Epistemic

(In view of the available evidence,) John *must/might/may* be the murderer.

b. *Deontic*

(In view of his parents' orders,) John *may* watch TV, but he *must* go to bed at

8pm.

c. *Ability*

(In view of his physical abilities,) John *can* lift 200 lbs.

d. *Teleological*

(In view of his goal to get a PhD,) John *must* write a dissertation.

e. *Bouletic*

(In view of his desire to retire at age 50,) John *should* work hard now.

While certain modal auxiliaries are restricted in the kinds of interpretation they can receive (*might*, for instance, only has epistemic interpretations), many others can express various kinds of flavors: *may* and *must* have epistemic or deontic interpretations, *have to* epistemic, deontic, circumstantial, teleological, or bouletic ones, etc. This is not a peculiarity of English. Instead, this multiplicity of modal meanings is quite pervasive across languages (cf. Fleischman 1982, Traugott 1988, Bybee, Perkins & Pagliuca 1994, Palmer 2001, though see Nauze 2008 for statistical evidence that this multiplicity is not as frequent as originally thought). To cite just a few examples, French *pouvoir* (*can*) and *devoir* (*must*), or Italian *potere* (*can*) and *dovere* (*must*), can all express circumstantial, deontic, teleological, bouletic and epistemic modality. Similarly, Malay *mesti (must)* (Drubig 2001), Cairene Arabic *laazim* (*must*), and Tamil permission and debitive suffixes (Palmer 2001) receive both epistemic and deontic interpretations.

A standard classification separates epistemic modals from all others, subsumed under the label 'root' modals (Hoffmann 1966). As we will see, several semantic and syntactic factors correlate with this distinction: epistemics deal with possibilities that follow from the speaker's knowledge, whereas roots deal with possibilities that follow from the circumstances surrounding the main event and its participants; epistemics are taken to be *speaker*-oriented, roots *subject*-oriented (Bybee, Perkins & Pagliuca 1994); epistemics tend to take widest scope whereas root modals take narrowest scope with respect to each other,

and to various scope bearing elements.

Under all types of interpretations, possibility and necessity modals enter into patterns of entailments and logical equivalences similar to those involving universal and existential quantifiers. *Must* and *may* are *duals* of each other, just as *some* and *every* (cf. Horn 1972):

(2) a. John *must* be home ⇒ John *may* be home
b. John *may* be home ≡ It is *not* the case that it *must* be the case that John is *not* home
c. John *must* be home ≡ It is *not* the case that it *may* be the case that John is *not* home

(3) a. *Every student* is home ⇒ *Some student* is home
b. *Some student* is home ≡ It is *not* the case that *every student* is *not* home
c. *Every student* is home ≡ It is *not* the case that *some student* is *not* home

Standard semantic approaches to modals stemming from philosophical modal logic derive these equivalences by giving them a quantificational analysis. Necessity modals are *universal*, while possibility modals are *existential*, quantifiers over *possible worlds*; this is what underlies the difference in *force*. And just as the set of students in *every student* needs to be restricted to a salient subset (we rarely talk about every single student in the universe), the set of worlds modals quantify over needs to be restricted to a particular subset. This subset is in turn what determines the particular *flavor* that the modal receives: if a modal quantifies over worlds compatible with what is known, the modal is interpreted epistemically, if it quantifies over worlds compatible with certain laws, it is interpreted deontically, etc.

In section 3, we review the quantificational approach to modality inherited from modal logic, and turn to Kratzer's theory in section 4. Section 5 and 6 delve into particulars of the two main classes of modals (epistemics and roots), and illustrate some of the challenges each pose for a Kratzerian approach in particular, and quantificational approaches in general, and

discuss the new directions these challenges have opened up. Section 7 looks at the connection between modals and other categories of meaning.

3. Modal logic and the quantificational approach to modals

Though philosophers have been concerned with modality since Aristotle's modal syllogisms, an explicit model theory, in the modern sense, for a modal logic was only made possible in the 1960s with the advent of *possible worlds*, developed in the works of Carnap (1957), von Wright (1951), Prior (1957), Kanger (1957), Hintikka (1961), and Kripke (1963) (for a history of the development of possible worlds, see Copeland 2002). The notion of a possible world can be traced back at least to Leibniz, according to whom the 'universe' (the *actual world*) was one (in fact, the best one) among an infinite number of possible worlds living in God's mind. Possible worlds can be viewed as possible 'ways things could have been' (Lewis 1973). There are many, many ways things could have been: think about the world as it is, but where the Eiffel Tower was destroyed after the World Expo, or one where the Eiffel Tower was never in Paris, but in London, or one where it is one millimeter taller, one where it is two millimeters taller, etc. You will see that we can conceive of a potentially infinite number of different worlds. Note that any change, however small, from one world to the next may require differing chains of events leading to this change, and may further have unavoidable repercussions, so that it may not be possible to find two worlds differing *only* in where the Eiffel Tower is located (think about all the Eiffel Tower postcards grandparents around the world would be receiving from London). Yet, there are still countless ways the world could be, and each of these ways represents a different possible world. While the ontological status of possible worlds is a topic of serious debate in the philosophical literature, linguists usually do not worry about such metaphysical issues; they assume that we have the capacity to represent alternative states of affairs, and that it is this capacity that we are referring to when we talk about possible worlds.

There are many concerns of modal logic which we cannot go over here (for an introduction to modal logic, see Hughes & Cresswell 1996; for a detailed overview of modal logic's contribution to the semantics of modality, see GAMUT 1991, Kaufmann, Condoravdi & Harizanov 2007, Portner 2009; and for an overview of the model theoretic and possible worlds semantics assumed in this section, see article 33 (Zimmermann) *Model-theoretic semantics*. We will dive right in by assuming a propositional logic, composed of atomic sentences (*p*, *q*, *r*…) and sentential connectives (∧, ∨, →, ¬), with the addition of the possibility (◊ 'diamond') and necessity (□ 'box') operators, which combine with formulae to form new formulae (◊*p*, □*p*…). The introduction of possible worlds was crucial in allowing the extension of the model-theoretic apparatus to modal logic, by having the valuation of a sentence not be *absolute* (either true or false), as in standard propositional logic, but *relative* to a possible world: a sentence is true or false *in a world w*, depending on the facts in *w*. It may be true in one world, and false in another. The truth of modalized formulae (e.g., □p) is likewise relative to a possible world, but in such a way that their valuation depends on the truth of *p* itself in *other* possible worlds – modals have a displacing effect. In a possible worlds framework, ◊ and □ can be viewed as an existential and a universal quantifier over possible worlds respectively. ◊*p* is true if *p* is true in *some* world, and □*p* is true if *p* is true in *all* worlds. This quantificational treatment explains the logical equivalences in (2) (◊*p* ⇔ ¬□¬*p* and □*p* ⇔ ¬◊¬*p*). However, it doesn't yet capture the contingency of modal statements: just as a sentence *p* can be true or false in a world *w*, we want ◊*p* and □*p* to be relative to a world. Moreover, this kind of pure (unrestricted) modality, called *alethic* modality (from Greek *aletheia 'truth'*), is just one of many types of modalities, such as *deontic*, *epistemic*, or *temporal* modalities, which we want to model. Both the contingency and the relativization of modals to a particular type of modality are achieved by having the set of worlds the modal quantifies over be restricted to a particular subset, relative to a world of evaluation. This is done via an *accessibility relation*.

Accessibility relations are binary relations over a set of worlds W, which pick out for each world *w* of W, a set of accessible worlds *w'*. Various kinds of accessibility relations can be defined: an epistemic relation picks out for each world *w* a set of worlds *w'* in which all of the facts known in *w* are true, a deontic relation picks out for each world *w* a set of worlds *w'* in which all of the rules of *w* are obeyed, etc.

(4) $R_{epis}(w,w')$ = {*w'*| *w'* is a world in which all of the facts known in *w* hold}

$R_{deontic}(w,w')$ = {*w'*| *w'* is a world in which all of laws of *w* are obeyed}

Formally, the accessibility relation is taken to be a parameter of a *model* (sometimes called a *Kripke model*). A model M consists of a pair <F,V>, where F is a *frame*, consisting of a pair <W,R>, with W, a set of worlds, and R, an accessibility relation. V is a valuation function which assigns truth values (1 and 0) to every atomic sentence at each world in W. R determines for each world *w* of W a set of accessible worlds *w'*, in which the proposition *p* is evaluated. Modals quantify over the worlds determined by the accessibility relation: ◊*p*/□*p* are true if *p* is true in some/all of the worlds picked out by the accessibility relation.

(5) a. $V_{M,w}(\Diamond p) = 1$ iff in some world w' in W, such that R(w,w'), $V_{M,w'}(p) = 1$

b. $V_{M,w}(\Box p) = 1$ iff in every world w' in W, such that R(w,w'), $V_{M,w'}(p) = 1$

What patterns of inference are valid for various types of modal reasoning can be explained and captured in terms of the properties that the accessibility relation for the corresponding modal logic should have. For example, accessibility relations can be serial, reflexive, transitive, etc. Different modalities are differentiated via the different properties that their accessibility relations have:

(6) a. R is *serial* iff for every w in W there is a world w' in W such that R(w,w')

b. R is *reflexive* iff for every w in W, R(w,w)

c. R is *transitive* iff for every w,w',w'' in W, if R(w,w') and R(w',w'') then R(w,w'')

Seriality, for instance, corresponds to *consistency*: it implies that the set of worlds picked out by the accessibility relation is not empty. This is an important property to prevent modals from quantifying vacuously. *Reflexivity* corresponds to *realism*: with a reflexive accessibility relation, $\Box p$ implies p. Epistemic relations are reflexive, but deontic ones aren't. Reflexivity further differentiates epistemic (knowledge-based) from doxastic (belief-based) accessibility relations. With an epistemic relation the world of evaluation is accessible from itself, but not with a doxastic relation: the world of evaluation may not be compatible with what is *believed* to be true. We will see in section 4.2 that there is some controversy surrounding the 'realistic' status of natural language epistemic modals: if their accessibility relation is reflexive, a sentence such as *it must be raining* should entail *it is raining*; yet intuitively, the former is somehow 'weaker' than the latter.

Modal logic is concerned with patterns of inferences in various modalities, independently of each other, and certainly independently of any idiosyncrasies of the natural language words that correspond, perhaps imperfectly, to these notions. Yet, some of the insights there have been crucial to our understanding of natural language modals, and the formal apparatus of quantification over possible worlds and accessibility relations is central to current semantic accounts of modality.

4. Kratzer's unifying account

Moving back to the realm of natural language modals, recall from section 2 that, sometimes, the same modal auxiliaries can receive several interpretations. Take the ambiguous sentence *John may watch TV*: it either expresses a deontic possibility (*John is allowed to watch TV*), or an epistemic one (*for all we know, it is possible that John is a TV watcher*). Is this due to a lexical ambiguity of the modal? Do English speakers store two different *may*s in their lexicon (either as homonyms or polysemes)? Such an ambiguity is tacitly assumed in semantic analyses that focus on particular subtypes of modality (Groenendijk & Stokhof

1975 for epistemic *may*, Kamp 1975 for deontic *may*), and, perhaps, postulating ambiguity for certain modal words may not be too big of an issue, given that there are modal words which are never ambiguous (e.g., *might*). However, this multiplicity of modal meanings is common enough cross-linguistically, and in languages from different families, so as to make a lexical ambiguity account unlikely: it is highly improbable that the same lexical accident should be found in language after language. Rather, it seems that we should give a single meaning for those modals that show an ambiguity, and derive the variety of flavors via some contextual factors (providing lexical restrictions for unambiguous modals like *might*). This is exactly what Kratzer proposes, in a series of influential papers.

In section 3 it was shown that different kinds of modalities can be explained as different accessibility relations. But how are these accessibility relations associated with a particular modal word? One possibility is that they are hard-wired in the denotation of modals. Epistemic *may* would differ from deontic *may* by combining with an epistemic vs. a deontic accessibility relation *in the lexicon*. This is what is usually assumed, for instance, for attitude verbs, which are traditionally analyzed as universal quantifiers over a set of worlds determined by an accessibility relation hard-wired in the semantics of each verb (e.g., *believe* takes a doxastic accessibility relation; *want* a bouletic one; cf. Hintikka 1962). Could modals work the same way? One crucial difference between modals and attitude verbs is that, while a modal like *may* can be associated with at least two different accessibility relations, the accessibility relation of an attitude like *believe* seems fixed: *believe* never gets a bouletic interpretation. Thus, if we were to hard-wire the accessibility relation in a modal's lexical entry, we would end up with as many lexical entries as there are possible interpretations, which seems undesirable, especially since this ambiguity is found in language after language.

Kratzer was the first to point out the improbability of an ambiguity account for natural language modals, and to give them a linguistically realistic semantics. We turn to her

account now, focusing on two main ideas: (i) that context partly determines the meaning of modals; (ii) that modals are 'doubly-relative', which avoids shortcomings of previous accounts.

4.1. The role of context

Kratzer (1977) shows the improbability of an ambiguity account for natural language modals as follows. Not only do modals come in various flavors, but each flavor itself seems to come in many subflavors. Take our deontic statement *John may watch TV*. It could be understood as a permission with respect to various rules: his father's, his mother's, a dorm or a prison. An ambiguity account becomes hopeless: not only would we need as many *may*s as there are possible flavors, but each of these *may*s would itself be ambiguous between various *may*s.

Now, a modal flavor can be specified unambiguously with an '*in view of*' phrase, as in the following examples:

(7) a. *In view of his vast knowledge of celebrity gossip*, John *may* watch TV.

b. *In view of his father's orders*, John *may* watch TV.

This phrase doesn't seem redundant. This means, Kratzer argues, that the *may* it combines with cannot be specified for a particular interpretation, and must instead be a kind of 'neutral' *may*, which needs to be added to our growing list of homonyms.

What if, instead, we took neutral *may* to be the only *may*? The '*in view of* ' phrase would itself provide the modal flavor. This would solve the hopeless homonymy problem. But what about cases in which such a phrase is missing? Kratzer proposes that its content is then supplied by the *context of utterance*, via what she calls a *conversational background*. Sentences are always uttered against a conversational background, which, Kratzer argues, can fill in information for modals that isn't explicit. Formally, a conversational background is represented as a function from worlds to sets of propositions. These propositions correspond to bodies of information, facts, rules, etc., responsible for determining the modal

flavor. Take the sentence *Mary must be the culprit*. Imagine that this sentence is uttered in a context where we are discussing a recent crime. In the course of our conversation, we discuss facts related to the crime, such as the fact *that the crime was committed yesterday, that Mary has no alibi, that she has a good motive, that no one else has a motive,* etc. All these propositions together form the set of facts known in our world. This set of facts is contingent. Things could have been different: Mary could have had an alibi, Paul, a motive, etc. Thus, what is known in this world may be different from what is known in some other world. What a conversational background does, then, is assign a (different) set of propositions to each world of the domain. An *epistemic* conversational background is a function f_{epis}, which assigns to each world *w* in W the following set of propositions:

(8) f_{epis}(w) = {p| p is a proposition that expresses a piece of established knowledge in w – for a group of people, a community…}

There is a tight connection between a conversational background and an accessibility relation. Recall section 2's epistemic relation:

(9) $R_{epis}(w,w')$ = {*w'*| *w'* is a world in which all of the facts known in *w* hold}

A proposition *p* corresponds to a set of worlds, namely, the set of worlds in which *p* is true. A set of propositions *A* corresponds to a set of sets of worlds, and its *intersection* to a set of worlds, namely, the worlds in which all of the propositions of *A* are true. Thus, from a conversational background *f* (which assigns to each world a set of propositions), we can derive the corresponding accessibility relation R_f by intersecting, for each world *w*, the set of propositions that *f* assigns to that world:

(10) $R_{epis}(w,w')$ = $\cap f$(w) = {*w'*| *w'* is a world in which all of the propositions *p* (such that *p* expresses a piece of established knowledge in *w*) hold}

Thus Kratzer's system is formally equivalent to previous quantificational accounts. The novelty is that the determination of the set of accessible worlds is not hard-wired in the lexical entry of the modal. Rather, it arises from a contextually-provided conversational

background *f*, formally represented as a parameter of the interpretation function, as in the following lexical entries, adapted from Kratzer (1991):

(11) For any world *w,* conversational background *f*:

a. $[[\text{must}]]^{w,f} = \lambda q_{<st>} . \forall w' \in \cap f(w): q(w') = 1$ (*in set talk* $\cap f(w) \subseteq q$)

b. $[[\text{can}]]^{w,f} = \lambda q_{<st>} . \exists w' \in \cap f(w): q(w') = 1$ (*in set talk* $\cap f(w) \cap q \neq \varnothing$)

Modal statements of the form '*must p*' or '*can p*' are true relative to a conversational background *f* if and only if *p* is true in *all* or *some* of the worlds in which the propositions of the conversational background are true. Note that because the conversational background is treated as a parameter, iterated modals, as in *John might have to leave*, should be evaluated against the same conversational background, and thus receive the same interpretation. Kratzer (1978) proposes a dynamic way to allow modals in the same sentence to be relative to different conversational backgrounds. Alternatively, one could represent the conversational background in the object language as a covert argument of the modal (we will see a case of this in section 5.5).

To sum up, Kratzer's introduction of conversational backgrounds preserves the main insights of traditional quantificational accounts of modals, and explains why modals can receive various kinds of interpretations without having to postulate massive lexical ambiguity.

4.2. The double relativity of modals

Consider the following sentence, uttered in a context where John murdered someone:

(12) John must go to jail.

(12) should say that in all of the worlds in which the law is obeyed, John goes to jail. But surely, in all of the worlds in which the law is obeyed, there is no murder! So here is our conundrum: how can we talk about worlds where the law is obeyed, when the law has been broken? The problem with the semantics outlined above for sentences like (12) is that it treats the cold fact that John committed murder and the content of the law on a par. We need

to separate facts from (moral) ideals. If we cannot get around the fact that murder was committed, we can still talk about moral obligations, *given* the resulting morally imperfect state of affairs. What (12) should say, then, is that the best way to obey the law in the imperfect world in which John committed murder is to have John go to jail.

To capture this, Kratzer (1981, 1991) proposes that modals be relative not to just one but two conversational backgrounds. The first is what she calls a *modal base*. It is made up of a set of facts, which is always consistent. In our example (12), the modal base notably contains the fact that John committed a murder. The second conversational background, dubbed the *ordering source*, consists of a set of ideals, moral or other (which may or may not be consistent), which imposes an ordering on the worlds of the modal base. Modals end up quantifying over the *best* worlds of the modal base, given the ideal set by the ordering source.

Both types of conversational backgrounds are functions from worlds to sets of propositions. For the modal base *f*, these propositions are relevant facts (e.g., *that John murdered Bill*). For the ordering source *g*, these propositions are ideals (e.g., *that murderers go to jail*). From the set of propositions *g(w),* Kratzer proposes an ordering $\leq_{g(w)}$, which ranks worlds according to how close they come to satisfying the ideal given by *g*:

(13) The ordering $\leq_{g(w)}$:

For all u,z $\in$W, for any g(w) $\subseteq \wp(\wp(W))$:

$u \leq_{g(w)} z$ iff $\{p: p \in g(w) \text{ and } z \in p\} \subseteq \{p: p \in g(w) \text{ and } u \in p\}$

The ordering states that for any pair of worlds *u, z, u* is closer to the ideal set by *g(w)* if the set of propositions true in *z* is a subset of the set of propositions true in *u*. Imagine two worlds *u* and *z* in which John committed a murder, and where John goes to jail in *u*, but not in *z*. Take a deontic ordering source containing two propositions: *that murder is a crime* and *that murderers go to jail*. Both worlds violate the first law, but *u* is closer to the ideal set by

the ordering source than *z*, since in *u*, the murderer John goes to jail, but not in *z*: the set of propositions of the ordering source true in *u* is a superset of the set of propositions true in *z*.

Our doubly-relative necessity modal looks as follows:

(14) For any world *w*, modal base *f* and ordering source *g*,

[[**must p**]]^{w,f,g} is true iff:

For all u $\in \cap$f(w), there is a v $\in \cap$f(w) such that v $\leq_{g(w)}$ u and

For all z $\in \cap$f(w): if z $\leq_{g(w)}$ v, then z $\in$ p

A necessity modal requires that for all worlds *u* of the modal base, there is a world *v* that comes closer to the ideal imposed by the ordering source, and in all worlds *z* closer than *v* to the ideal, the proposition *p* expressed by its complement is true: *p* is true in all of the most ideal worlds of the modal base. We can simplify this definition by making the so-called 'limit assumption', i.e., by assuming that there always are accessible worlds that come closest to the ideal, call these worlds $Best_{g(w)}(\cap f(w))$ (for arguments in favor of the limit assumption, see Stalnaker 1984; the 'Best' operator is from Portner 2009). We obtain the following lexical entries:

(15) For any world *w,* and conversational backgrounds *f*, *g*:

[[**must**]]$^{w,f,g} = \lambda q_{<st>} . \forall w' \in \text{Best}_{g(w)} (\cap f(w))$: q(*w'*) = 1.

[[**can**]]$^{w,f,g} = \lambda q_{<st>} . \exists w' \in \text{Best}_{g(w)} (\cap f(w))$: q(*w'*) = 1.

where $Best_{g(w)}(X)$ *selects the most ideal worlds from X, given the ordering given by g(w)*

This doubly-relative system allows Kratzer to solve a problem with previous accounts, namely the problem of 'inconsistencies'. Standard (singly-relative) quantificational accounts break down when the set of propositions that the modal is relative to is inconsistent, that is, when two of the propositions cannot both be true in a world. When a conversational background is inconsistent (i.e., when the corresponding accessibility relation is not serial),

its intersection is empty, and the modal quantifies over an empty set. In this case, a sentence comes out as trivially true if the modal has universal force, and trivially false, when it has existential force. To see why, take the singly-relative lexical entries in (11). In set talk, a *necessity* modal requires that the worlds provided by the conversational background be a subset of the set of worlds that make up the propositional complement ($\cap f(w) \subseteq q$). Given that the empty set is a subset of any set, any necessity statement comes out as trivially true. A *possibility* modal requires the non emptiness of the intersection of the set of worlds provided by the conversational background and the set of worlds that make up the propositional complement ($\cap f(w) \cap q \neq \varnothing$). Since the intersection of the empty set with another set is always empty, any possibility statement comes out as trivially false.

This type of problem typically arises with deontic modality, in cases where laws conflict with one another, or bouletic modality, in cases of conflicting desires. Consider a toy example from Kratzer (1977, 1991), where the law consists of judgments handed down by various judges. One uncontested judgment states that murder is a crime. Two other judgments (from different judges), however, conflict: one states that goat owners are liable for the damage caused by their goats, while the other states that they aren't. The law thus consists of three propositions: *that murder is a crime*, *that goat owners are liable, that goat owners are not liable.* This set is inconsistent. Thus, the necessity statement in (16a) is wrongly predicted to be true, and the possibility statements in (16b) and (16c) wrongly predicted to be false:

(16) a. Murder must not be a crime.

b. Goat owners may be liable for damage caused by their goats.

c. Goat owners may not be liable for damage caused by their goats.

Let's see how the doubly-relative system avoids this problem. The law consisting of the three judgments, make up our deontic ordering source. Let's assume the modal base is empty (the ordering orders all worlds in W). We find four types of worlds: type 1 worlds, where

murder is a crime and goat owners are liable; type 2 worlds, where murder is a crime and goat owners are *not* liable; type 3 worlds, where murder is *not* a crime and goat owners are liable; type 4 worlds, where murder is *not* a crime and goat owners are *not* liable. The worst worlds are those in which murder is not a crime (type 1 and 2 worlds are respectively more ideal than type 3 and 4, since one of the propositions of the ordering source hold in type 1 and 2 but not in type 3 or 4). Worlds of type 1 and 2 cannot be ordered with respect to each other, since the set of propositions true in each cannot stand in a subset relation, and both types make up the 'best' worlds, i.e., those that the modals quantify over. It is true that goat owners are liable in some of these worlds (type 1), not liable in some others (type 2), and that murder is a crime in all of them.

A further advantage of the doubly-relative system is that it provides an explanation for the problematic intuitions we get about epistemic modals, mentioned in section 2. Recall that we should expect sentences with epistemic necessity modals (17b) to entail their unmodalized counterparts (17a), given the reflexivity of their accessibility relation (i.e., the world of evaluation—here the actual world—should be one of the accessible worlds). Yet, intuitively, (17b) doesn't seem to entail (17a):

(17) a. It is raining.

b. It must be raining.

In Kratzer's doubly-relative system, it doesn't need to. Indeed, the modal could take an ('stereotypical') ordering source, which would force the modal to quantify only over the *most normal worlds* of the (epistemic) modal base. Thus, while the world of evaluation would be one of the worlds selected by the modal base, given that this modal base is realistic (i.e., it corresponds to a reflexive accessibility relation), it could well be atypical, and hence not be among the most *normal* of these worlds.

Before we turn to particularities of conversational backgrounds, let's mention a final benefit of the ordering source, discussed at length in Kratzer (1981, 1991), which is that it gives us a

means of deriving graded notions of modality (e.g., *slight possibility*), by invoking more or less far-fetched possibilities. Graded modality is however a complex topic that may require technologies beyond the doubly-relative system. The interested reader should consult Yalcin (2007) and Portner (2009).

4.3. Modal bases and ordering sources

According to Kratzer, there are two kinds of modal bases. The *epistemic* modal base picks out worlds in which *what is known* in the base world holds. The *circumstantial* modal base picks out worlds in which *certain circumstances* of the base world hold. The difference may seem subtle, but it, in fact, involves reasoning from qualitatively different kinds of premises, and leads to truth conditional differences.

Circumstantial modality looks at the material conditions which cause or allow an event to happen; epistemic modality looks at the knowledge state of the speaker to see if an event is compatible with various sources of information available. The following example illustrates this contrast with *might* and *can*, each of which have idiosyncratic constraints that force *might* and disallow *can*, to select an epistemic modal base:

(18) a. Hydrangeas might be growing here.

b. Hydrangeas can grow here. Kratzer (1981)

The sentence in (a) is evaluated against an epistemic modal base: to the best of my knowledge, it is possible that hydrangeas are growing here. The sentence in (b) is evaluated against a circumstantial modal base, which includes circumstances such as the quality of the soil, the climate, etc. (a) and (b) differ truth conditionally: if I know for a fact that there are no hydrangeas in this part of the world, (a) is false; however, if the circumstances are still conducive to hydrangeas' growth, the sentence in (b) is true.

Epistemic modal bases combine with ordering sources related to information: what the normal course of events is like (stereotypical ordering source), reports, beliefs, rumors, etc.

Circumstantial modal bases combine with various kinds of ideals, yielding the various root interpretations: deontic (laws), bouletic (wishes), or teleological (aims). Note that the ordering source may also be empty, as with the circumstantial modal in (18b).

To sum up, in a Kratzerian system, a modal is a quantifier over possible worlds, restricted by a **modal base** (*circumstantial* or *epistemic*), which returns a set of accessible worlds, which can then be ordered by an **ordering source**, to yield the most ideal worlds of the modal base. Both modal bases and ordering sources are contextually determined (when not overt). This allows for a single lexical entry for *must* and for *can*, and their counterparts in various languages, which differ only in **force** of quantification (universal vs. existential).

Kratzer (1981, 1991) treats modality as an autonomous system, mostly putting aside the way it interacts with other elements such as tense or negation. While this tack was very useful in isolating general properties of modality and providing a unified theory, we will see how the interaction of modality with various elements requires that this account now be expanded.

5. The *epistemic vs. root* distinction

We have seen that, cross-linguistically, the same modal words can express both epistemic and root modality, a fact which Kratzer's context-sensitive account successfully captures. We will now see the flipside of the coin: just as systematically, epistemics and roots *differ* from each other in ways, which cast doubt on a unifying account. In the typology literature, epistemics and roots are sometimes taken to differ in that the former express 'propositional modality' (i.e., the speaker's judgment about a *proposition*), and the latter, 'event modality' (i.e., the speaker's attitude towards a potential *event*) (Jespersen 1924, Palmer 2001). Even more frequently, epistemics are said to be *speaker*-oriented, and roots *subject*-oriented (cf. Bybee, Perkins & Pagliuca 1994). This difference is sometimes formally captured by having *root* (but not *epistemic*) modals enter into a thematic relation with the subject (Perlmutter 1971, Ross 1969, Jackendoff 1972). In Kratzer's system, the difference between roots and

epistemics is a matter of modal base: epistemic interpretations arise from an epistemic modal base, root interpretations from a circumstantial one. Kratzer (1991) already suggests that the difference in modal bases could be correlated to Perlmutter and Ross's epistemic vs. root distinction in terms of argument structure. It is not entirely clear, however, how this correlation can be formalized without losing some of the unifying power of her system: how can we encode that epistemic and deontic *must*s differ not just in modal bases, but in argument structure as well, without postulating two different lexical entries? What's worse, epistemics further differ from roots in their temporal relativity: epistemic modals are evaluated at the speech time, root modals at the time provided by the main tense of the sentence. Take the ambiguous sentence *John had to be home*. When the modal receives a root (deontic) interpretation, it expresses a *past* obligation given John's circumstances *then* to be home *then*. When it receives an epistemic interpretation, it expresses a *present* necessity, given what is known *now*, that at some past time John was home.

We will explore the hypothesis that what underlies these systematic differences in time and individual relativity is a difference in height of interpretation: epistemics scope at the 'S-level', roots scope at the 'VP-level', following *Cinque's hierarchy*. Based on a careful cross-linguistic survey of the positioning of adverbs and various functional elements like tense and modals, Cinque (1999) proposes that functional heads are universally organized along a rigid universal hierarchy, in which epistemic modals appear higher than root modals, as shown below:

(19) Cinque's hierarchy (*irrelevant projections omitted*)

$$Mod_{epis} > Tense > Aspect > Mod_{volitional} > Mod_{deontic\ necessity} > Mod_{ability/deontic\ possibility}$$

We now turn to evidence that supports such a hierarchical split.

5.1. Interaction between modals

Recall that in English, modals *may* and *have to* are ambiguous between epistemic and

deontic readings. Interestingly, however, when they are stacked together, the ambiguity disappears:

(20) John may have to watch TV.

This sentence can only mean that it is possible, given what is known, that John has an obligation to watch TV, not that it is allowable that it be epistemically necessary that John watches TV. This restriction in ordering follows easily if modals have dedicated slots, with the highest reserved for epistemics and the lowest for roots.

Alternatively, the unattested ordering could perhaps be ruled out on conceptual grounds: no matter how tyrannical the issuer of a command, he may not be able to demand that a state of affairs be epistemically necessary. Yet, consider the following example from von Fintel & Iatridou (2004), which argues against conceptual impossibility. Imagine a scenario in which "an insurance company will only pay for an expensive test if there is a possibility that the patient may have Alzheimher's". Such a state of affairs can be reported as follows, with a deontic modal taking scope over an epistemic adjective:

(21) For the test costs to be reimbursed, it has to (DEONTIC) be possible (EPISTEMIC) that the patient has Alzheimer's.

Interestingly, this embedding possibility does not seem available with modal auxiliaries, at least in English. (20) doesn't seem able to receive an interpretation where the first modal is read deontically and the second epistemically. German may be different. Kratzer (1976) argues that the following example can receive an interpretation where the embedded modal auxiliary *können* receives an epistemic interpretation while the modal *müssen* is interpreted deontically (though see Nauze 2008, for claims that the embedded modal is not interpreted epistemically).

(22) Und auch in Zukunft muss diese Schnecke […] Saugfüsse haben können.
And also in future must this snail suction.feet have might
And even in the future, this snail must possibly have suction feet.

[translation from Nauze (2008)]

Why should there be a difference between German and English? Why should there be a difference in English between adjectives like *possible* and modals like *may*? What these examples seem to show is that the ordering restriction in sentences like (20) cannot be solely based on conceptual grounds: it seems possible to embed a modal adjective with an epistemic interpretation (if we grant that '*possible*' indeed refers to an epistemic, and not just a circumstantial possibility). Adjectives and modal auxiliaries may have different properties that would allow the former to embed, but not the latter, so that Cinque's hierarchy may not be so much about types of modality, but rather types of modal auxiliaries. It is also possible, more generally, that counterexamples like (22) result from biclausal structures, in which the epistemic modal is part of an embedded clause, though further empirical support and cross-linguistic inquiry is needed (von Fintel & Iatridou 2004).

5.2. Interaction with the subject

As mentioned earlier, a traditional distinction between epistemics and roots is that the former are *speaker-oriented* and the latter *subject-oriented*. We now review some evidence for this distinction, based on the interaction of modals with the subject.

One type of evidence is the way quantificational subjects scope with respect to the modal. Brennan (1993) shows that while epistemics are able to take scope over a quantificational subject, roots, such as ability modals, cannot. Consider the following pair of examples:

(23) a. Every radio may get Chicago stations and no radio may get Chicago stations.

b. #Every radio can get Chicago stations and no radio can get Chicago stations.

With epistemic *may* in (23a), no contradiction arises, suggesting that 'every radio' is interpreted below the modal: *it may be that every radio gets Chicago stations and (it may*

also be that) no radio gets Chicago stations. The contradictoriness of (23b) with root *can*, however, suggests that 'every radio' has to be interpreted above the modal. Note that while certain speakers may find the conjunction in (23a) anomalous (presumably for pragmatic reasons), all agree that the two conjuncts are compatible in a way that the conjuncts in (23b) are not.

While (23) shows that a quantifier like *every* can scope below an epistemic modal, von Fintel & Iatridou (2003) argue that in fact, it must. This is their Epistemic Containment Principle (ECP), according to which a quantifier cannot bind its trace across an epistemic modal. The ECP is illustrated in the infelicitous example below:

(24) #Every student may be the oldest student.

The infelicity of (24) indicates that the only possible interpretation is one where the modal takes scope over the quantifier (*#it is possible that all of the students are the oldest*), while the felicitous surface scope is ruled out by the ECP (*For every student x, it is possible that x is the oldest*) (for discussion and refinements, see Tancredi 2007 and Huitink 2008).

The scopal facts, then, argue that epistemics can (and perhaps must) take scope over quantifier subjects, but that root modals cannot. Why should this be? One common explanation is that epistemics and roots differ in argument structure: roots are control predicates which enter into a thematic relation with the subject, while epistemics are raising predicates, with no particular relation to the subject. Supporting evidence comes from idioms and expletives (Brennan 1993). As idiom chunks lose their idiomatic meaning in control constructions (*e.g., #the shit wants to hit the fan*), the example in (25) suggests that root '*can*' takes an individual and a property as arguments, while epistemic '*might*' takes a proposition:

(25) The shit might/#can hit the fan.

However, while there does seem to be some connection between the subject and a root modal, Bhatt (1998), Hackl (1998) and Wurmbrand (1999) argue that this connection cannot

be due to a control configuration. Focusing on deontic modals, Bhatt and Wurmbrand independently show that obligations do not necessarily fall on the subject. Consider the following examples:

(26) Jonny must brush his teeth. Bhatt (1998)

(27) The plants must be watered.

The obligation is likely to be on the addressee of (26) rather than on Jonny, if he is a small child (for instance, his babysitter), and on the implicit agent, rather than on the plants, in (27), suggesting that the purported thematic relation between the modal and the subject is not syntactic. In Wurmbrand's terminology, deontics do not involve 'syntactic control', but rather 'semantic control'. This is corroborated by the fact that deontics do not necessarily require an agentive subject, and allow expletive subjects:

(28) There have to be fifty chairs in this room. Bhatt (1998)

It should be noted, however, that while these examples show that some deontics are best analyzed as raising predicates, they do not necessarily show that *all* deontics, let alone all *roots* are raising. Brennan (1993), for instance, argues that deontics split into two categories: what she calls *ought to be* deontics and *ought to do* deontics, using Feldman's (1986) terminology. Brennan argues that the former are S-level modals, just like epistemics, and the latter VP-level modals, like other roots. If such a distinction really is grammaticized, *must* in (26) could be ambiguous between an *ought to be* (obligation on the addressee) and *ought to do* (obligation on the subject) interpretation, with two different argument structures underlying this ambiguity. Evidence that *all* modals are raising predicates come from examples involving the most root-like type of modality, namely, ability modals. Hackl (1998) shows that some ability modals allow expletive *it* subjects (29a). Moreover, they do not always force their *subject* to enter into a particular relation with them: *can* in (29b) seems to express a capacity of *the pool*, rather than that of *a lot of people*:

(29) a. It can rain hard here. Hackl (1998)

b. A lot of people can jump in this pool.

The supporting evidence for a control analysis of modals is thus controversial at best. If we need a raising analysis for *some* roots, then by Occham's razor, we should avoid postulating two different kinds of argument structures for the same modals. While obligees and permittees are usually identified with an overt argument of the verb, the above examples show that they do not always need to; the context may be able to provide salient individuals around which the modality is centered. A possible explanation for the lack of idiomatic meaning with roots, as in (25), is that root modality somehow needs to centered around *some* participant of the VP event, but not necessarily its subject (Hacquard 2006). In most cases, the main participant *is* the subject, and hence properties of the subject are highlighted. In other cases, however, the location or properties of other participants of the event are more relevant (*here* or *the pool*). The fact that modal statements involving idiom chunks seem to improve when a location is added corroborates the intuition that the relevant factor is not argument structure, but rather whether the modality can be anchored to one of the VP event's participants:

(30) The shit can really hit the fan in this part of the world. Hacquard (2006)

To sum up, the interaction of modals with subjects shows some differences between roots and epistemics, although the evidence doesn't seem to warrant a control vs. raising split. In principle, this could have been otherwise. There is nothing about modality *per se* that would prevent this. In fact, predicates like *able to* are modal and require a control analysis. The purported thematic relation between a root modal and the subject instead seems to highlight the fact that root modals are centered around the circumstances of the event described by the main predicate, and especially, but not necessarily, those of its agent, confirming Palmer's intuition that root modality is 'event' modality.

Recall that we started this section with the traditional speaker vs. subject orientation of epistemic and root modals. We have seen that root modals do not always center around the

subject, but rather around some participant of the main event. I would like to close this section by showing that epistemics are not always tied to the speaker either. When an epistemic appears in the complement of an attitude verb, the epistemic state that the modal seems to report is that of the attitude holder, not of the speaker:

(31) a. Every boy$_1$ thinks he$_1$ must$_1$ be stupid. Stephenson (2007)

b. Every contestant$_1$ thinks he$_1$ might$_1$ be the winner. Speas (2004)

We can thus refine the traditional subject vs. speaker orientation split as follows: *roots* are anchored to a *participant of the main event*, *epistemics* to the local 'attitude' holder: the *speaker* when the modal is in a matrix (though cf. section 6.2), the *attitude holder* when the modal is in the complement of an attitude verb.

5.3. Interaction with negation

The interaction between modals and negation is also suggestive, though a clear pattern doesn't yet emerge. Cross-linguistically, epistemics tend to be interpreted above negation, and roots below it (Coates 1983, Drubig 2001). The Malay examples below illustrate:

(32) a. Dia mesti tidak belajar. (epistemic) Drubig (2001)
he must not study

b. Dia tidak mesti belajar. (deontic)
he not must study

Mesti (*must*) is ambiguous between an epistemic and a deontic interpretation. However, when it appears structurally above negation (*tidak*), the modal only gets an epistemic interpretation, and when below, it only gets a deontic one. Parallel cases can be seen in English, though the modal and negation appear in a fixed order on the surface. The sentence in (33) is ambiguous between an epistemic and a deontic interpretation. However, when the modal takes scope over negation, it must be interpreted epistemically, while, if it takes scope below, it must receive a deontic interpretation:

(33) John may not watch TV…

a. … he never knows any celebrity gossip. epistemic: may>not

b. … his dad is very strict. deontic: not>may

There are, however, counterexamples (Cormack & Smith 2002, Palmer 2001). The following examples contain modals with epistemic interpretations, which seem to scope under negation:

(34) a. Jane doesn't have to be at home.

b. Jane need not be home.

c. Jane can't be home.

Several factors, beyond the epistemic/root distinction, seem to conspire to make the interaction of modals with negation a complex matter, such as the possibility/necessity distinction (Cormack & Smith 2002), the position of negation, which varies cross-linguistically, and idiosyncrasies of various modal auxiliaries (for a thorough typological overview, see de Haan 1997). At best, we find the following weak, but suggestive, generalization (R. Bhatt and A. Rubinstein, p.c.): when a modal is ambiguous between a root and an epistemic interpretation, it is never the case that the modal scopes above negation when it receives a deontic interpretation and below it when it receives an epistemic one, though, all other cases are attested (negation takes scope over the modal no matter the interpretation, negation takes scope in between the two, or below both).

5.4. Interaction with tense

Traditional accounts of modality in general (and Kratzer's specifically) usually ignore the relationship between modals and tense. However, it has been shown that modals cannot be relative just to a world, but to a time as well (cf. Thomason 1984, Ippolito 2002): circumstances or evidence change through time; what was a possibility in the past may not be one in the future, and vice versa. Importantly, what this time *is* seems to depend on the

particular interpretation of the modal: with a root interpretation, the modal's time of evaluation has to be the time provided by tense. With an epistemic interpretation, it has to be the local 'now': the speech time in matrix contexts, the attitude internal 'now' when in the complement of attitude verbs (cf. Iatridou 1990, Picallo 1990, Abusch 1997, Stowell 2004).

Consider the following example, where *have to* gets a *root* (teleological) interpretation; the circumstances and goal of the subject are evaluated at the time provided by tense (past). (35) expresses a necessity, given Mary's circumstances *then*, to take the train *then*. It cannot express a necessity given her circumstances *now* to have taken the train *then*.

(35) Mary had to take the train to go to Paris.

The evaluation time of an *epistemic* modal can never be future-shifted. The only interpretation for (36) is that it may **now** be the case that Marikos will be dead tomorrow, but not that **tomorrow**, it will be possible that Marikos is dead:

(36) Marikos may be dead tomorrow. Groenendijk & Stokhof (1975)

Nor can it be back-shifted. Consider the following example:

(37) Mary had to be the murderer. ✓mod_{epis}>past, *past>mod_{epis}

Imagine that the evidence gathered at the beginning of the investigation, a week ago, all pointed to Mary being the murderer: she had no alibi, but many a motive. Yesterday, however, Poirot established that the murder had been committed one hour earlier than originally thought. This fact immediately cleared Mary, who was seen by several eyewitnesses elsewhere at that time. In this scenario, (37) is judged false: it cannot describe the epistemic state that held at the time when the evidence pointed to Mary. In order for us to report such a past state, we need additionally an embedding attitude verb (as in (38a), an indirect discourse past tense (as in (38b); Boogart 2007), or an *overt* conversational background (as in (38c):

(38) a. Two days ago, Poirot thought that Mary had to be the murderer.

b. This didn't make sense, thought Poirot… Mary had to be the murderer.

c. Given what we knew then, Mary had to be the murderer.

In all these cases, a past morpheme appears on the modal. However, it lacks the characteristic backshifting of a true semantic past tense. For instance in (a), the modal's time of evaluation must be Poirot's thinking time; it cannot precede it. The past morpheme on the modal reflects instead a 'sequence of tense' rule, where the embedded tense morphologically agrees with the higher past tense on 'think', cf. article 57 (Ogihara) *Tense*. (Note that there are some situations in which some speakers find the past epistemic interpretation of (37) acceptable. However, these situations all seem to be narrative contexts, which also involve some kind of temporal subordination or free indirect discourse. Similar complications occur with the 'assessor' of epistemic claims, cf. section 6.2.)

One counterexample to this generalization is put forth by von Fintel & Gillies (2008a), who argue that, in the following exchange, B's utterance expresses a *past epistemic possibility*:

(39) A: Why did you look in the drawer?

B: My keys might have been in there. (=It was possible that my keys were in there)

However, this reading seems to only arise in answers to a *why* question, where the temporal shifting of the epistemic could be due to a covert *because*, able to shift the evaluation parameters (Stephenson 2007).

Finally, the temporal interpretation of modals seems to further differentiate *epistemic* from '*metaphysical'* modality (the modality involved in counterfactuals). Consider the following examples from Condoravdi (2002), who argues that the following contrast results from different scope configurations between the modal *might* and the Perfect (cf. article 49 (Portner) *Perfect and progressive*) along with a felicity condition on the selection of a modal base:

(40) a. They might (already) have won the game.

b. They might (still) have won the game.

(a) gets an epistemic interpretation, facilitated by 'already': *it is possible, as far as we know right now, that at some past time they won the game*; (b) gets a metaphysical interpretation, facilitated by 'still': *there was a possibility at some past time, that they would win the game* (with the further inference that they in fact didn't). Here again, with an epistemic interpretation, the modal's time of evaluation seems unable to get backshifted, even in the presence of a potential backshifter (perfect).

The lack of forward or backshifting of epistemics' time of evaluation is often captured formally by not allowing epistemics to be in the scope of tense (cf. Iatridou 1990, Abusch 1997, Picallo 1990, Abraham 2001, Stowell 2004), either by encoding in the lexical entry of epistemics that they be evaluated at the local time of evaluation or by hard-wiring their position above the tense projection, in line, again, with Cinque's hierarchy.

5.5. Reconciling Kratzer and Cinque

We see that modals interact differently with tense, negation, and quantifiers depending on their interpretation: modals with epistemic interpretations scope high, modals with root interpretations scope low, as in Cinque's hierarchy, where epistemics and roots occupy different fixed positions. This pattern challenges Kratzer's unifying account, according to which epistemics and roots are two contextual variants of the same modal words. Indeed, if modals must appear in predetermined positions, based on their interpretation, then something beyond a contextual parameter must be specified in each of their lexical entries to derive their structural properties. The behavior of epistemic and root modals leads us to two conflicting cross-linguistic generalizations. On the one hand, the same words seem to systematically be used to express both root and epistemic modality, in line with a Kratzerian account. On the other, epistemics and roots seem to systematically differ, notably in the positions in which they appear.

Diachronic and structural approaches

There are several lines one can take to give our syntax and semantics enough freedom to handle idiosyncrasies of roots and epistemics, and still explain why the same words are used cross-linguistically to express root and epistemic modality. One type of explanation for why modals share the same form while having a semantic life of their own is to appeal to a diachronic (or metaphoric) process. Epistemic interpretations tend to develop cross-linguistically from root ones, and interestingly, this historic trend is matched by children's acquisition of modals, with root modals being acquired first (Sweetser 1990, Papafragou 1998). Thus, one could argue that modals are polysemous, but not accidentally so: their various senses are related. For Sweetser (1990), modals encode 'force dynamics' of potential barriers and driving forces. These forces operate in the concrete, external world for root modals, but can be metaphorically extended to the realm of the mental or the abstract, to yield epistemic modality. However, while such diachronic accounts seem to be empirically rooted, they cannot be the full story. It is unclear why each of these senses should inherit the set of scopal (and other) properties it does.

A common way to derive these scopal properties is to assume two different positions (VP-level vs. S-level) for roots and epistemics, by essentially giving them separate lexical entries (cf. Jackendoff 1972, Picallo 1990, Butler 2003, a.o.). This postulation of different lexical entries for roots and epistemics unfortunately leaves unexplained why both types of modality are expressed by the same lexical items cross-linguistically. Brennan (1993) presents an interesting variant, in which modals come in different types, VP-level and S-level modals, but where the root/epistemic distinction is not directly encoded in these two types. The reason most roots correspond to VP-modals and epistemics to S-modals is not a grammatical fact, but the result of certain ontological commitments made and reinforced by the community of language users, which could have been different. This allows Brennan to successfully derive two positions for roots and epistemics without encoding the flavor distinction in the modals' lexical entries (hence resolving the tension between the two

conflicting cross-linguistic generalizations). Yet, the arbitrariness of the correlation between modal type (VP vs. S) and modal flavor is questionable, given that this correlation does not hold only in a single language, or language family, but across languages of different pedigrees. Why should different communities of speakers converge on the same ontological commitments?

Event-relativity approach

Let's review the time and individual restrictions that seem to constrain the interpretation of modals. We saw that modals are generally relative to a time. For epistemics in main clauses, this time is the speech time; for epistemics in complements of attitude verbs, it is the attitude 'now'; and for roots, it is the time provided by tense. Modals are also generally relative to an individual. For epistemics in main clauses, the individual is the speaker, for epistemics in attitude contexts, it is the attitude holder, and for roots, it is often the subject, and sometimes, another participant of the VP event. Putting aside the flavor difference for a moment, one way to recast these generalizations is to say that modals are relative to time/individual pairs, and that crucially, not all time/individual combinations are attested. A modal is either anchored to the speaker at the speech time (*may* in (41a) describes an epistemic possibility for the *speaker* at the *speech time*), the attitude holder at the attitude time (*may* in (41b) describes an epistemic possibility for *Mary* at her *thinking time*), or a participant of the VP event at the time of the VP event, provided by tense (*have to* in (41c) describes a circumstantial necessity for *John* at the *fleeing time*):

(41) a. John may have seen the murderer.

b. Mary thought that John may have seen the murderer.

c. John had to flee the scene.

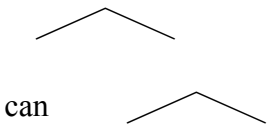
What we do not find are modals anchored to the speaker at the time provided by tense, or to the subject at the speech time (unless, of course, these two times coincide, i.e., with present tense). Why does the interaction of modals with tense and with individuals to go hand in

hand, rather than being independent of each other? One way to derive these time/individual constraints is to make a modal relative to an *event* rather than a world of evaluation (Hacquard 2006, to appear). Doing so will restrict the modal's interpretation by anchoring it to particular time/individual pairs, namely the running time and participants of the event it is relative to. Under this view, the meaning of a modal is not only constrained by the context and the idiosyncrasies of its lexical entry, but by its grammatical environment as well.

We have already seen intuitions that modals are centered around an event: the main event for root modals, and the speech event for epistemics (cf. Jespersen 1924, Palmer 2001, Zagona 2007). In Hacquard (2006, *to appear*), I propose to cash out this event-relativity by using a Kratzerian semantics, except that modals (and in particular modal bases) are relative to an *event* of evaluation, rather than a *world* of evaluation. There are three kinds of events that modals can be anchored to: speech events, VP-events and attitude events. I argue that by relativizing modals to an event rather than a world of evaluation, one gets all and only the attested time-individual pairs: the running time and participants of the events of evaluation. Modals are thus either relative to the VP event (and hence its participants—e.g., the subject— and its running time—determined by tense), the speech event (and hence the speaker and the speech time), or an attitude event (and hence its attitude holder and attitude time).

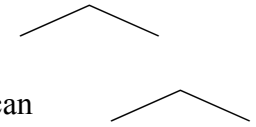
Here is a formal sketch: modals keep their standard lexical entries, but their modal bases take an event pronoun *e*, which needs to be bound locally. Assuming that in a standard world-relative system, modal bases and worlds are represented in the *object language*, and not as parameters, the only difference between a world-relative and an event-relative system is that the argument of the modal base *f* for the latter is an event rather than a world pronoun:

(42) a. *world-relative modal*



f($\mathbf{w_1}$) q

b. *event-relative modal*



can

f($\mathbf{e_1}$) q

In the Davidsonian tradition, verbs (including attitude verbs) are predicates of events, whose event argument is quantified over by Aspect (article 48 (Filip) *Aspectual Class and Aktionsart*). There are two types of event binders: Aspect, and a default speech event e_0, which I assume is represented in the object language. The event-relativity of the modals in (41) is derived as follows; modals can appear (for type reasons) in either one of two positions within a clause, roughly corresponding to Brennan's S-level vs. VP-level modals: above tense or above the VP. In a main clause, a modal located above tense is bound by the speech event e_0 (41)/(43a); in a complement clause, a modal located above the embedded tense is bound by the aspect that quantifies over the embedding attitude event e_2 ((41)/(43b); a modal located below tense and aspect is bound by the aspect that quantifies over the VP event e_1 ((41)/(43c):

(43) a. John may have seen the murderer.

$[_{CP}\ \mathbf{e_0}\ \lambda e_0$ Mod f($\mathbf{e_0}$) $[_{TP}$ T $Asp_1\ \lambda e_1[_{VP}$ V e_1]]]

may *past pfv J. see murderer*

b. M. thought that J. may have seen the murderer.

$[_{CP}\ e_0\ \lambda e_0$ T $\mathbf{Asp_2}\ \lambda e_2$ Att $\mathbf{e_2}$

thought

$[_{CP}$ Mod f($\mathbf{e_2}$) $[_{TP}$ T $Asp_1\ \lambda e_1[_{VP}$ V e_1]]]]

may *J. see murderer*

c. John had to flee the scene.

$[_{CP}\ e_0\ \lambda e_0\ [_{TP}$ T $\mathbf{Asp_1}\ \lambda e_1$ Mod f($\mathbf{e_1}$) $[_{VP}$ V $\mathbf{e_1}$]]]

past pfv *have to* *J. flee*

This derives the right time and individual constraints. The connection between the type of events modals are relative to and the root and epistemic distinction is indirect. I propose that, usually, only modals relative to speech and attitude events can combine with an epistemic modal base because only those events have associated 'propositional content' (i.e., the propositions that make up the attitude, such as a set of beliefs), which provides an information state required by an epistemic modal base. VP-event-relative modals, on the other hand, get a default circumstantial modal base:

(44) a. $f_{epis}(e) = \lambda w'$. w' is compatible with the 'content' of e

b. $f_{circ}(e) = \lambda w'$. w' is compatible with the circumstances of e

Many complications arise, which we cannot discuss here. The main advantage of an event relative system, is that it allows for a single (flavor independent) entry for modals à la Kratzer, but whose meaning is partly constrained by its grammatical environment (i.e., it depends on what the closest event binder is), making sense of the fact that low modals in Cinque's hierarchy (i.e., those that scope *below* aspect) receive a root interpretation, while high modals (i.e., those that scope *above* tense) receive an epistemic interpretation.

6. Questioning the modal semantics of modals

In section 5, we saw how epistemics and roots differ in their interactions with tense, negation and subjects. These interactions showed us that modality cannot be a completely hermetic system, as its interpretation is affected, and perhaps partly determined, by neighboring grammatical elements. The tension there was to reconcile a unifying account with an account that could derive these scope interactions. We now turn to two different and independent challenges to any 'modal' account of modals more generally: in section 6.1 we discuss a puzzle involving *root* modals and their interaction with aspect, where modals seem to lose their modal dimension and force their complement clause to be actualized. In section 6.2 we

discuss claims that *epistemics* are not part of the proposition expressed by the sentence they appear in: they do not involve quantification over possible worlds, but merely make an extra-truth conditional contribution.

6.1. Root modals as implicative predicates?

This section discusses the interaction of root modals and aspect. Consider the following Greek example from Bhatt (1999), which shows that ability modals do not always merely express a possibility, but sometimes force the proposition expressed by their complement to occur in the actual world. Such '*actuality entailments*' (using Bhatt's terminology) happen when modals combine with *perfective*, but not with *imperfective* aspect:

(45) a. Borusa na sikoso afto totrapezi ala ∂en to sikosa. Bhatt (1999)
can-**impf**.1s NA lift.nonpast.pfv.1s this table, but NEG it lift-impf
'(In those days) I could lift this table, but I didn't lift it.'

b. Boresa na tu miliso (#ala ∂en tu milisa).
can-pst.**pfv** na him talk.nonpast.pfv.1s #but neg him talk.past-pfv
'I was able to talk to him (#but I didn't talk to him).'

This is not an idiosyncrasy of the Greek ability modal. Bhatt shows that the effect happens in languages, which, unlike English, have an overt morphological distinction between perfective and imperfective aspect, such as Hindi, French, Italian, Bulgarian, etc. Furthermore, this effect further extends to *all* root interpretations (Hacquard 2006).

Actuality entailments pose a serious challenge for any modal account, by seemingly eradicating the very property of displacement that defines modals. They further show the dangers of focusing on languages like English or German, which sometimes have idiosyncratic properties that can obscure the bigger picture (such as not distinguishing aspect morphologically). Should we give up standard accounts? Are root modals not modals after

all? Bhatt (1999) in fact proposes an account of the ability modal, which denies altogether that it is a modal, and treats it instead as an *implicative* predicate. Following Karttunen's (1971) analysis of the implicative *manage*, Bhatt argues that with an ability statement, what is asserted is the complement clause, and a further meaning component, that the complement requires some effort, is added as a conventional implicature. The lack of actuality entailment with imperfective arises from an additional layer of modality associated with the imperfective itself.

Hacquard (2006, 2009) proposes a way to derive actuality entailments and keep a relatively standard modal semantics for root modals. Root modals are regular quantifiers over possible worlds, and actuality entailments arise from the configuration of aspect and the modal: while epistemics scope above tense and aspect and are thereby immune to actuality entailments, root modals scope below aspect, and are thus susceptible to them. Given that aspect is what quantifies over the VP event, it locates that event in time and in a *world*. By having (perfective) aspect scope over the modal, that world has to be the actual world (imperfective brings in an additional layer of modality as in Bhatt 1999). Under this account, actuality entailments are ultimately of the same nature as the scope interactions of root and epistemic modals with tense and subjects, and the problem reduces to explaining why epistemics and roots scope in different positions.

6.2. Do epistemics contribute to truth conditional content?

It is often assumed in the descriptive literature that epistemics do not contribute to the truth conditional content of the sentence they appear in, but rather express a speaker's comment about, or commitment to, the proposition expressed by their complement (cf. Halliday 1970, Palmer 2001). This intuition has been formalized recently in various ways, from treating epistemics as *evidentials* (Westmoreland 1998, Drubig 2001), to having them modify or perform a different speech act, such as a kind of 'doxastic advice' (Swanson 2006). There

is, for instance, a tight connection between evidentials and epistemics, as both deal, to a certain extent, with speakers' *evidence* (cf. Aikhenvald 2004). Evidentials are often said *not* to contribute to the truth conditional content of the sentence they combine with, but rather indicate the speaker's grounds for expressing that sentence (cf. Faller 2002); likewise, accounts of epistemics as evidentials take them to lack truth conditional content. The precise nature of the connection between the two categories is under active debate: some argue that epistemics are a kind of evidentials, and thus lack truth conditional content (Drubig 2001), others that *evidentials* are a kind of epistemic modals, and thus make a truth conditional contribution in terms of quantification of possible worlds (McCready & Ogata 2007, Matthewson, Rullman & Davis 2007). There is at least some evidence that epistemic modals have an evidential component, as argued by von Fintel & Gillies (2007), with examples like (46):

(46) It must be raining.

A speaker can utter (46) felicitously in a windowless room after seeing a few people coming in with wet umbrellas, but not when standing outside in pouring rain. This indicates that an epistemic modal's felicity conditions require that the evidence the modal claim is based on be indirect, or involve an inference.

The main supporting evidence for all non-truth-conditional accounts of epistemics is the fact that epistemics are notoriously hard to embed. As we saw, they tend to scope over tense, negation, and quantifiers. It has further been claimed that they cannot occur in questions, in antecedents of conditionals, or in complements of attitude verbs (cf. Jackendoff 1972, Drubig 2001). There are, however, counterexamples to these unembedability claims. We saw, for instance, that some epistemics scope under negation. Furthermore, the ability of epistemics to embed seems to depend on Lyons's (1977) *subjective/objective* distinction. Subjective epistemics are taken to rely on the speaker's personal and subjective evidence, while objective epistemics rely on more objective grounds, which the speaker shares with a

relevant community. A modal in a sentence such as *'it might rain'* is in principle ambiguous between a subjective reading (say, if I utter this sentence not having read any weather report, based solely on the dubious fact that my arthritis is acting up), and an objective one (say, if it is uttered by a meteorologist after consulting various radar maps). Lyons claims that only *subjective* epistemics lack truth conditional content, while *objective* epistemics get an ordinary modal semantics, a claim supported by the fact that while subjective epistemics resist embedding, objective ones embed more freely—they can appear in questions (Papafragou 2006), and in the scope of quantifiers, obviating the 'ECP' (Tancredi 2007).

A subjective/objective split, however, raises a by-now familiar dilemma: if objective and subjective epistemics are truly separate modals, which differ in whether they make a truth conditional contribution, why should they share a lexical entry? A perhaps more promising way to derive a subjective/objective distinction while maintaining a unified account is to have the modal be relativized either to the speaker's information state for the former, or that of the speaker and his community for the latter (cf. Papafragou 2006, Portner 2009; von Fintel & Gillies 2008b propose an interesting pragmatic way of deriving this distinction, which may avoid having to lexicalize this information).

But why should such a distinction correlate with epistemics' ability to embed? Perhaps, subjective epistemics can embed to the same extent than objective ones, but some embeddings are unattested for felicity reasons. Papafragou (2006), for instance, argues that subjective epistemics do not appear in questions, since it would be odd for a speaker to ask his addressee about his own epistemic state. As a matter of fact, we do actually find subjective epistemics in some attitude contexts (Portner 2009); *might* in (47) seems to express a *subjective* epistemic possibility based on John's *subjective* beliefs:

(47) John believes that it might be raining.

The question of whether epistemics can embed — and if not, what in their semantics prevents them to do so — requires further empirical investigation. What is clear at this point

is that, while we do find cases of embedded epistemics, their distribution is still limited: as we saw, epistemics cannot scope under tense, apparently regardless of the subjective/objective distinction. Furthermore, while there does seem to be something to the subjective/objective distinction, both the data supposed to tease them apart and the nature of the distinction are still somewhat controversial.

Another challenge to traditional accounts is the claim that any *assessor* of an epistemic modal statement can disagree with its content (even if she is not part of the conversation, but merely eavesdropping), and thus that the truth of an epistemic statement is relative not merely to the speaker, but to the perspective of the sentence's *assessor* (MacFarlane 2003, Egan, Hawthorne & Weatherson 2004, Stephenson 2007, a.o.). While the kind of data used in support of assessor-relativity suggests that the question of whose information state epistemics are relative to is quite complex, there seem to be ways to account for it without invoking the 'assessor' machinery. Recall from our discussion of the subjectivity of epistemics that they seem to sometimes not be relative solely to the knowledge of the speaker, but to that of the speaker and her community. It could be that, sometimes, a sentence's assessor is really part of the speaker's community, whose knowledge state needs to be factored in (cf. von Fintel & Gillies 2007).

The precise nature of epistemic modality is thus still a matter of debate. There are, however, good reasons for wanting to maintain a truth conditional account in terms of quantification over possible worlds. First, epistemics can embed and contribute to truth conditional content, as the following contrast shows (von Fintel & Gillies 2007):

(48) a. If there *might* have been a mistake, the editor will have to reread the manuscript.

b. If there is a mistake, the editor will have to reread the manuscript.

Second, epistemic modals *are* expressed by the same words as other kinds of modals cross-linguistically. How, then, can we derive their idiosyncrasies? With respect to the

embedability facts, we saw several solutions that give epistemics widest scope, and thereby derive their difficulty to embed under certain categories. What about other peculiarities? One way to capture idiosyncrasies of modals, while maintaining a standard semantics in terms of quantification over worlds, is to encode them in a separate dimension of meaning. For instance, to handle the evidential nature of epistemics, von Fintel & Gillies (2007) suggest that epistemic modals needn't be evidentials *per se*: they could contribute an extra speech act of 'proffering' beyond their standard truth conditional contribution as quantifiers over possible worlds. Portner (2009) takes such a line further by proposing that other modals also contribute an extra speech act: for deontics, a command, for roots, an assertion responsible for actuality entailments. However, while encoding such a dimension is a step forward in being able to account for peculiarities of various modalities while maintaining a unified truth conditional semantics, it ultimately faces our original dilemma of explaining why these various modalities are expressed by the same modal words cross-linguistically: even if encoded in a separate dimension, modality-specific information still needs to be lexically specified in separate entries. That is, unless, one could find a (non lexical) way to have the type of performativity somehow fall out from the type of modality.

7. Modality and its kin

Several systems share many similarities with modals, and one may question the extent to which these similarities are due to the limited set of resources these systems can appeal to, or whether they reflect deeper dependencies. To understand the connection between modals and these various systems, it has proven fruitful to look at the way they interact. One such system is *mood*. Mood and modality are often discussed together, as they seem intimately connected, in that both signal the non actual; certain languages do not have modal auxiliaries but a rich mood system, others, like English, have an impoverished mood system, but a rich modal auxiliary system (Palmer 2001). Are the two separate systems, or alternatives to one

another? How do they interact in languages that have both systems (e.g., Romance languages)? There, modals seem to systematically select for subjective mood (when they do not combine with an infinitival complement), suggesting that mood may be more of a morphological reflex, rather than a semantically contentful category. For more on mood, see Farkas (1985), Giannakidou (1997), Portner (1997), and article 50 (Portner) *Verbal mood*.

Another system to consider is *evidentiality*. We saw that epistemic modals share similarities with evidentials. Is it that epistemic modals are really evidentials? Or that evidentials are really modals? Alternatively, could there be two separate systems, standing in some kind of dependency? One possibility is that evidentials impose restrictions on a modal's conversational background, say, by determining the ordering source of epistemic modals, as suggested by Portner (2007). This could explain why epistemic modals seem to have an evidential component, without forcing them to *be* evidentials. Further research on the interaction of the two systems in languages that have both rich evidential and modal systems will shed more light on the connection between the two.

Another case at hand is *imperatives* (article 67 (Han) *Imperatives*). Imperatives share many similarities with modals, and are sometimes argued to employ, in part, the Kratzerian machinery, cf. Han (1999). Portner (2007) shows that imperatives can receive the same flavors of interpretation as root modals (deontic, bouletic, teleological). A sentence such as *Have an apple!*, for instance, could be taken as an order or an invitation, depending on context. Yet, imperatives and modals differ in important ways: while it is fairly uncontroversial that root modals make a truth conditional contribution, this is not so for imperatives. So, whence the similarities? Portner argues that imperatives and root modals are intimately connected, in part because of the way the discourse evolves and affects context-sensitive modals. Just like declaratives are added to the common ground and affect the interpretation of subsequent modals (which is why the following sequence is infelicitous: *It isn't raining. #It might be raining*), imperatives affect the interpretation of subsequent root

modals by helping determine their ordering source. So here again, we find a potential dependency that goes beyond the appeal to the same resources.

Finally, consider attitude verbs. We mentioned that attitude verbs share many similarities with modals: they can have epistemic/doxastic-type interpretations (*know, believe*) or root-like interpretations (bouletics *'want', 'wish'*, commands '*order'*…). Are these two systems analogous but independent, or is there a deeper connection? The connection might simply be that both involve quantification over possible worlds (Hintikka 1962). However, evidence for a deeper connection comes from the fact that epistemic modals cannot appear in the complements of all attitude verbs. Anand & Hacquard (2009) argue that epistemics cannot appear in complements of root-like attitudes (a), but only in that of what Stalnaker (1984) calls 'attitudes of acceptance' (b), and propose that epistemic modals are in fact anaphoric to the *content* of an embedding attitude verb of the right type:

(49) a. John {believes, argues, assumed} that the Earth might be flat.

b. *John {hopes, wishes, commanded} that the Earth might be flat.

Comparing modals and various systems such as mood, evidentials, or attitude verbs, and looking at the way they interact, it appears that their similarities may not be completely accidental, but may reflect instead the appeal to the same resources, and sometimes even deeper, perhaps anaphoric, dependencies.

8. Conclusions

We began with the cross-linguistic generalization that the same words express various flavors of modality, a fact supported by Kratzer's unifying account, which captures modals' context-dependency. We saw that this pan-modal generalization was counterposed by several flavor-specific idiosyncrasies, which cast doubt on unification (and more generally, challenge analyses in terms of quantification over possible worlds). Thus, while context undeniably plays an important role in determining the flavor of modals, it cannot by itself

disallow say, an epistemic modal to be evaluated in the past or the future, or to scope below a quantifier or negation; and yet, this is what we find cross-linguistically. We considered new accounts, which reconciled these two generalizations either by making the semantics of modals partly dependent on their grammatical environment, or by exploring additional dimensions of meaning.

For reasons of space, this survey article had to be limited in scope both theoretically and empirically. In the remaining paragraphs, I will mention some of the areas we skipped, and provide references for the interested reader. We focused on modals and their interactions at the sentence level. But modals, being context sensitive, are clearly affected by the way discourse evolves, as we briefly mentioned in section 7. This is illustrated in cases of so-called *modal subordination*, where modals appear in anaphoric relations cross-sententially (cf. Roberts 1989, Geurts 1995, Frank 1997, and (article 75 (Geurts) *Accessibility and Anaphora*). Theory-wise, we focused on static approaches; however, theoretical alternatives are found in dynamic frameworks (cf. Groenendijk, Stokhof & Veltman 1996 and references therein). Beyond framework differences, the way discourse affects a modal's interpretation, or the issues surrounding subjective and objective epistemic modality (in particular, the way they affect and are affected by the knowledge state of discourse participants), both seem like good areas to better understand the division of labor between semantics and pragmatics, and the representation of meaning in static or dynamic terms (cf. von Fintel & Gillies 2007, Yalcin 2007). Empirically, our discussion was limited to modal auxiliaries, ignoring nouns, adjectives, and adverbs. But adverbs like *possibly* or *necessarily* express the same kinds of possibilities and necessities than modal auxiliaries do, and are traditionally given the same semantics. One interesting research question is how modal auxiliaries and modal adverbs interact, and in particular, the puzzling non-additive, but rather 'agreeing' effect that results from so-called *modal concord* (cf. Geurts & Huitink 2006, Zeijlstra 2008). We have also ignored the modal component involved in some non-lexical categories of meaning as well,

such as tense, aspect, or mood (for discussion, see articles 48 (Filip) *Aspectual class and Aktionsart*, 50 (Portner) *Verbal mood*, and 57 (Ogihara) *Tense*), or the covert modality present in infinitivals (Bhatt 1999), as well as the modality involved in conditionals, which played a major role in shaping Kratzer's theory (article 59 (von Fintel) *Conditionals*). While we focused on the interaction of modals with *viewpoint* aspect, we ignored the interaction of modals with the *lexical* aspect of the verb phrase they combine with, and in particular the fact that stative verbs tend to force epistemic interpretations of the modals they combine with (e.g., *John must love Mary*), while eventives tend to force root interpretations (e.g., *John must go to Paris*).

An important lesson we learned from looking at the interactions of modals with elements like tense or aspect is that these elements seem to constrain the interpretation of modals. While earlier work either de-emphasized these constraints or derived them by appeal to syntactic design, we may now be in a position to start explaining these puzzles without recourse to cartographic appeal. For example, as we saw, epistemic interpretations are restricted by whether a modal scopes above tense. Perhaps only high scoping modals get epistemic interpretations because modals obey local dependencies, and epistemic interpretations depend on a 'high' element (such as a speech or attitude event). Perhaps, more generally, the interpretation of modals is constrained by various elements because these elements participate in determining that interpretation. The moral of the past decade is this: now that we have robust accounts of modality, which can handle both the logical properties of modals and their context-dependency, we no longer need to study modality as a hermetic system. The next chapter, it seems, will be to explain exactly how modals affect and are affected by their surrounding environment, both at the sentence level and the discourse level.

Acknowledgments

For helpful comments and discussion, many thanks to Pranav Anand, Kai von Fintel, Nathan Klinedinst, Dave Kush, Terje Londhal, Paul Pietroski, Paul Portner, Alexis Wellwood, and especially Rajesh Bhatt.

9. References

Abraham, Werner 2001. Modals. Toward explaining the 'epistemic non-finiteness gap. In: R. Müller & M. Reis (eds.). *Modalität und Modalverben im Deutschen*. Hamburg: Helmut Buske Verlag, 7-36.

Abusch, Dorit 1997. Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20(1): 1–50.

Aikhenvald, Alessandra 2004. *Evidentiality*. Oxford: Oxford University Press.

Anand, Pranav & Valentine Hacquard 2009. Epistemics with attitudes. In: *Proceedings of Semantics and Linguistics Theory (SALT) XVIII*. Ithaca, NY: CLC Publications. http://hdl.handle.net/1813/13025.

Bhatt, Rajesh 1998. Obligation and possession. In: H. Harley (ed.). *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect, MITWPL 32*. Cambridge, MA, 21-40.

Bhatt, Rajesh 1999. *Covert Modality in Non-Finite Contexts*. Ph.D. Dissertation. University of Pennsylvania.

Boogart, Rony 2007. The past and perfect of epistemic modals. In: L. de Saussure, J.

Moescher, & G. Puskas (eds.). *Recent Advances in the Syntax and Semantics of Tense, Aspect and Modality*. Berlin and New York: Mouton de Gruyter, 47-69.

Brennan, Virginia. 1993. *Root and Epistemic modal auxiliary verbs*. Ph.D. Dissertation. University of Massachusetts, Amherst.

Butler, Jonny 2003. A minimalist treatment of modality. *Lingua* 113: 967–996.

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

Carnap, Rudolph 1957. *Meaning and Necessity*. Chicago: University of Chicago Press.

Cinque, Guglielmo 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford Studies in Comparative Syntax. Oxford and New York: Oxford University Press.

Coates, Jennifer 1983. *The Semantics of the Modal Auxiliaries.* London: Croom Helm.

Condoravdi, Cleo 2002. Temporal interpretations of modals. In: D. Beaver, S. Kaufman, & B. Clark (eds.). *Stanford Papers in Semantics*. Palo Alto, CA: CSLI Publications, 59-88.

Copeland, Jack 2002. The genesis of possible worlds. *Journal of Philosophical Logic* 31: 99–137.

Cormack, Annabel & Neil Smith 2002. Modals and negation in English. In: S. Barbiers, F. Beukema & W. van der Wurff (eds.). *Modality and its Interaction with the Verbal System*. Amsterdam: John Benjamins, 133–163.

Drubig, Hans B. 2001. On the Syntactic Form of Epistemic Modality. Ms. Tübingen, University of Tübingen.

Egan, Andy, John Hawthorne & Brian Weatherson 2004. Epistemic modals in context. In: G. Preyer & G. Peter (eds.). *Contextualism in Philosophy*. Oxford: Oxford University Press, 131-169.

Faller, Martina 2002. *Semantics and Pragmatics of Evidentials in Cuzco Quecha*. Ph.D. Dissertation. Stanford University.

Farkas, Donka 1985. *Intensional Descriptions and the Romance Subjunctive Mood.* New York: Garland.

Feldman, Fred 1986. *Doing the Best We Can.* Dortrecht: Reidel.

von Fintel, Kai 2006. Modality and Language. In: D. Borchert (ed.). *Encyclopedia of Philosophy.* Detroit: MacMillan, 20-27.

von Fintel, Kai & Anthony Gillies 2007. An opinionated guide to epistemic modality. In: T. Gendler & J. Hawthorne (eds.). *Oxford Studies in Epistemology, Vol. 2*. Oxford: Oxford University Press, 32-62.

von Fintel, Kai & Anthony Gillies 2008a. CIA Leaks. *Philosophical Review* 117(1): 77–98.

von Fintel, Kai & Anthony Gillies 2008b. *Might* made right. To appear in: A. Egan & B. Weatherson (eds.). *Epistemic Modality*. Oxford: Oxford University Press.

von Fintel, Kai & Sabine Iatridou 2003. Epistemic Containment. *Linguistic Inquiry* 34: 173-198.

von Fintel, Kai & Sabine Iatridou 2004. Seminar Notes. Ms. Massachusetts Institute of Technology.

Fleischman, Suzanne 1982. *The Future in Thought and Language: Diachronic Evidence from Romance*. Cambridge: Cambridge University Press.

Frank, Anette 1997. *Context Dependence in Modal Constructions.* Doctoral Dissertation, University of Stuttgart.

GAMUT, L.T.F. 1991. *Logic, Language, and Meaning. Vol. 1: Introduction to Logic; Vol. 2: Intensional Logic and Logical Grammar*. Chicago and London: University of Chicago Press.

Geurts, Bart 1995. *Presupposing*. Doctoral Dissertation. University of Stuttgart.

Geurts, Bart & Janneke Huitink 2006. Modal concord. In: P. Dekker & H. Zeijlstra (eds.). *Proceedings of the ESSLLI Workshop Concord Phenomena at the Syntax-Semantics Interface, Malaga*, 15–20.

Giannakidou, Anastasia 1997. *The Landscape of Polarity Items*. Ph.D. Dissertation. University of Groeningen.

Groenendijk, Jeroen & Martin Stokhof 1975. Modality and conversational information. *Theoretical Linguistics* 2: 61–112.

Groenendijk, Jeroen, Martin Stokhof & Frank Veltman 1996. Coreference and modality. In: S. Lappin (ed.). *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell, 179–213.

de Haan, Ferdinand 1997. *The Interaction of Modality and Negation.* New York: Garland.

Hackl, Martin 1998. On the Semantics of '*Ability* Ascriptions'. Ms. Cambridge, MA, Massachusetts Institute of Technology.

Hacquard, Valentine 2006. *Aspects of Modality.* Ph.D. Dissertation. Massachusetts Institute of Technology.

Hacquard, Valentine 2009. On the interaction of aspect and modal auxiliaries. *Linguistics and Philosophy* 32(3): 279–315.

Hacquard, Valentine to appear. On the event relativity of modal auxiliaries. *Natural Language Semantics*.

Hintikka, Jakko 1961. Modality and quantification. *Theoria* 27: 119–128.

Hintikka, Jakko 1962. *Knowledge and Belief*. Ithaca, NY: Cornell University Press.

Hoffmann, T. Ronald 1966. Past tense replacement and the modal system. In: A. Oettinger (ed.). *Mathematical Linguistics and Automatic Translation*. Cambridge, MA: Harvard University, Harvard Computational Laboratory. Report NSF-17.

Horn, Laurence. 1972. *On the Semantic Properties of the Logical Operators in English.* Bloomington, IN: Indiana University Linguistics Club.

Hughes, George & Max Cresswell 1996. *A New Introduction to Modal Logic.* London and New York: Routledge.

Huitink, Janneke 2008. Scoping over epistemics in English and Dutch'. In: *Proceedings of CIL, Seoul*, 2077-2089.

Iatridou, Sabine 1990. The past, the possible and the evident. *Linguistic Inquiry* 21(1): 123–129.

Ippolito, Michela 2002. *The Time of Possibilities*, Ph.D. Dissertation. Massachusetts Institute of Technology.

Jackendoff, Ray 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT Press.

Jespersen, Otto 1924. *A Modern English Grammar*. Heidelberg: Karl Winter; Copenhagen: Einar Munksgaard.

Kamp, Hans 1975. Reference and Quantification in Tense and Model Logic. In: S. Schmidt (ed.) *Pragmatics II*. München: Finck, 150-197.

Kanger, Stig 1957. The Morning Star paradox. *Theoria* 23: 1–11.

Karttunen, Lauri 1971. Implicatives. *Language* 47(2): 340–358.

Kaufmann, Stephan, Cleo Condoravdi & Valentina Harizanov 2007. Formal approaches to modality. In: W. Frawley (ed.). *The Expression of Modality*. Berlin: Mouton de Gruyter, 72-106.

Kratzer, Angelika 1976. Was *können* und *müssen* bedeuten können müssen. *Linguistische Berichte* 42: 128–160.

Kratzer, Angelika 1977. What *must* and *can* must and can mean. *Linguistics and Philosophy* 1(3): 337–355.

Kratzer, Angelika 1978. *Semantik der Rede: Kontexttheorie, Modalwörter, Konditionalsätze.* Königstein: Scriptor.

Kratzer, Angelika 1981. The notional category of modality. In: H.-J. Eikmeyer & H. Rieser (eds.). *Words, Worlds, and Contexts. New Approaches in Word Semantics*. Berlin: Mouton de Gruyter, 38-74.

Kratzer, Angelika 1991. Modality. In: A. von Stechow & D. Wunderlich (eds.). *Semantik: Ein internationales Handbuch zeitgenössischer Forschung*. Berlin: Mouton de Gruyter, 639–650.

Kripke, Saul 1963. Semantical analysis of Modal Logic I, Normal Propositional Calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9: 67–96.

Lewis, David 1973. *Counterfactuals*. Cambridge, MA: Harvard University Press.

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

Matthewson, Lisa, Hotze Rullman & Henry Davis 2007. Evidentials as epistemic modals: Evidence from St'at'imcets. *The Linguistics Variation Yearbook* 7. Amsterdam: John Benjamins, 201-254

McCready, Eric & Norry Ogata 2007. Evidentiality, modality and probability. *Linguistics and Philosophy* 30(2): 147–206.

MacFarlane, John 2003. Epistemic Modalities and Relative Truth. To appear in: A. Egan & B. Weatherson (eds.). *Epistemic Modality*. Oxford: Oxford University Press.

Nauze, Fabrice 2008. *Modality in Typological Perspective*. Ph.D. Dissertation. Amsterdam University.

Palmer, Frank R. 2001. *Mood and Modality*. 2nd edition. Cambridge: Cambridge University Press.

Papafragou, Anna 1998. The acquisition of modality: implications for theories of semantic representations. *Mind and Language* 13: 370-399.

Papafragou, Anna 2006. Epistemic modality and truth conditions. *Lingua* 116: 1688–1702.

Perlmutter, David 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Picallo, Carme 1990. Modal verbs in Catalan. *Natural Language and Linguistic Theory* 8(2): 285–312.

Portner, Paul 1997. The semantics of mood, complementation, and conversational force. *Natural Language Semantics* 5: 167–212.

Portner, Paul 2007. Imperatives and modals. *Natural Language Semantics* 15(4): 351–383.

Portner, Paul 2009. *Modality*. Oxford: Oxford University Press.

Prior, Arthur 1957. *Time and Modality*. Oxford: Oxford University Press.

Roberts, Craige 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12: 683–721.

Ross, John R. 1969. Auxiliaries as main verbs. In: W. Todd (ed.). *Studies in Philosophical Linguistics (Series 1)*. Evanston, IL: Great Expectations Press, 77–102.

Speas, Peggy 2004. Person (and mood and tense) and indexicality. Paper presented at the Harvard Workshop on Indexicals, Speech Acts, and Logophors, November 20, 2004.

Stalnaker, Robert 1984. *Inquiry*. Cambridge, MA: The MIT Press.

Stephenson, Tamina 2007. *Toward a Theory of Subjective Meaning*. Ph.D. Dissertation. Massachusetts Institute of Technology.

Stowell, Tim 2004. Tense and modals. In: J. Guéron & J. Lecarme (eds.). *The Syntax of Time*. Cambridge, MA: The MIT Press.

Swanson, Eric 2006. *Interaction with Context*. Ph.D. Dissertation. Massachusetts Institute of Technology.

Sweetser, Eve 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Tancredi, Chris 2007. A Multi-Modal Theory of I-Semantics. Ms. Tokyo, University of Tokyo.

Thomason, Richard 1984. Combinations of tense and modality. In: D. Gabbay & F. Guenthner (eds.). *Handbook of Philosophical Logic, Vol. 2*. Dordrecht: Reidel, 135–165.

Traugott, Elizabeth 1988. Pragmatic strengthening and grammaticalization. In: S. Axmaker, A. Jaisser & H. Singmaster (eds.). *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 406–416.

Westmoreland, Robert 1998. *Information and Intonation in Natural Language Modality.* Ph.D. Dissertation. Indiana University.

von Wright, Georg 1951. *An Essay on Modal Logic*. Amsterdam: North-Holland.

Wurmbrand, Susi 1999. Modal verbs must be raising verbs. In: S. Bird, A. Carnie, J. Haugen & P. Norquest (eds.). *Proceedings of West Coast Conference in Formal Linguistics* 18. Somerville, MA: Cascadilla Press, 599-612.

Yalcin, Seth 2007. Epistemic modals. *Mind* 116: 983–1026.

Zagona, Karen 2007. On the syntactic features of epistemic and root modals. In: L. Eguren & O. Fernandez Soriano (eds.). *Coreference, Modality & Focus: Studies on the Syntax/Semantics Interface*, 221–236.

Zeijlstra, Hedde 2008. Modal concord is syntactic agreement. In: M. Gibson & T. Freidman (eds.). *Proceedings of Semantics and Linguistics Theory XVII.* Ithaca, NY: CLC Publications, 317-332.