

**On the event relativity of modal auxiliaries\***

Valentine Hacquard  
University of Maryland

**Abstract**

Crosslinguistically, the same modal words can be used to express a wide range of interpretations. This crosslinguistic trend supports a Kratzerian analysis, where each modal has a core lexical entry and where the difference between an *epistemic* and a *root* interpretation is contextually determined. A long standing problem for such a unified account is the equally robust crosslinguistic correlation between a modal's interpretation and its syntactic behavior: epistemics scope high (in particular higher than tense and aspect) and roots low, a fact which has led to proposals that hardwire different syntactic positions for epistemics and roots (cf. Cinque's hierarchy). This paper argues that the range of interpretations a modal receives is even more restricted: a modal must be keyed to certain time-individual pairs, but not others. I show that this can be captured straightforwardly by minimally modifying the Kratzerian account: modals are relative to an *event*—rather than a world—of evaluation, which readily provides a time (the event's running time) and (an) individual(s) (the event's participants). I propose that this event relativity of modals can in turn explain the correlation between type of interpretation and syntactic position, without having stipulation of an interpretation-specific height for modals.

---

\* Earlier versions of some of the material presented here can be found in Hacquard (2006, 2008). For very useful comments and discussion, many thanks to P. Anand, J. Anderssen, G. Chierchia, K. von Stechow, A. Kratzer, N. Klinedinst, I. Heim, S. Iatridou, P. Pietroski, P. Portner, A. Williams, audiences at UMass Amherst, UMD, and Sinn und Bedeutung 11, as well as two anonymous NALS reviewers.

**Keywords:** Modality, Cinque's hierarchy, epistemic, event, attitude

## 1. Introduction

It is a rather robust crosslinguistic generalization that the same modal words can be used to express various kinds of possibilities or necessities (cf. Fleishman 1982, Traugott 1988, Bybee et al. 1994, Palmer 2001). Take the English sentence *John may not watch TV, but he may be watching it anyway*. The most natural reading of this sentence is not contradictory, and means that while John is not allowed to watch TV, it is possible, given what we know, that he still does. The first *may* expresses a *deontic* possibility (permission), the second, an *epistemic* possibility (what is possible given what is known). Is this a case of lexical ambiguity (i.e., do English speakers store two different *may*'s in their lexicon), or are we dealing with a single lexical item, whose meaning is partly determined by context? This multiplicity of meanings is quite pervasive across languages: to cite just a few examples, French *pouvoir*, and Italian *potere*, can express an ability, a permission, or an epistemic possibility, among other variants. Similarly, Malay modal *mesti* (Drubig 2001) and the Tamil permission and debitive suffixes can receive both epistemic or deontic interpretations (Palmer 2001). This crosslinguistic pattern argues against a lexical ambiguity account, as it seems highly improbable that the same lexical accident should be found in language after language. Instead, we would like to arrive at a single meaning for each modal, and derive the variety of flavors via some contextual factors. This is exactly what Kratzer proposes in a series of influential papers. According to Kratzer (1981, 1991), there is just one *may*, and just one *must*, respectively an existential and a universal quantifier over a set of accessible worlds. The context dependency of a modal's interpretation arises from a 'conversational background', which determines the set of worlds modals quantify over.

Depending on the context, the same modal will quantify over worlds compatible with certain laws and receive a deontic interpretation, or quantify over worlds compatible with what is known and receive an epistemic interpretation. Going back to our original English sentence, we obtain that *in no world compatible with his parents' orders* does John watch TV, but *in some world compatible with what is known*, John is watching TV. The Kratzerian proposal, then, fits nicely with the crosslinguistic trend to use the same words to express various flavors of modality.

A long-standing problem for this proposal are the systematic syntactic and semantic patterns that set apart epistemic from all non-epistemic—or ‘root’—interpretations of the same modals. First, epistemics and roots differ in the individual the modality is anchored to: epistemics seem to be *speaker-oriented*, roots *subject-oriented* (Bybee et al. 1994), a difference sometimes captured by having roots (but not epistemics) enter in a thematic relation with the subject (Perlmutter 1971, Ross 1969, Jackendoff 1972). Second, roots and epistemics differ in their time of evaluation: epistemics’ time of evaluation is the speech time (or the internal ‘now’ of the attitude, when in the complement of attitude verbs), while that of roots is the time provided by tense (Iatridou 1990, Picallo 1990, Abusch 1997, Stowell 2004). These interactions with time and individuals follow if roots and epistemics occupy different syntactic positions, as in Cinque’s hierarchy, where the ordering of tense, aspect, roots, and epistemics is fixed crosslinguistically (Cinque 1999). However, this ordering is unexpected and unexplainable if the only difference between epistemics and roots is in the kind of worlds the modal quantifies over, as determined by the context. Thus, we are led to the following dilemma: on the one hand, given that the same modal words can express both epistemic and root modality, we want to give them the same lexical entry; on the other, the fact that epistemics and roots differ systematically in height of interpretation suggests that they should be treated as separate elements. I will refer to

this problem as ‘Cinque’s puzzle’. The general trend in resolving it has been to essentially reject a unified account: Drubig (2001), for instance, proposes that epistemics are evidentials, rather than modals, to explain why they take widest scope. Others bite the Cinque bullet and stipulate that modals come in two (UG-supplied) types: epistemics, which take IP complements, and roots, which take VP complements (Jackendoff 1972, Zubizarreta 1982, Picallo 1990, Butler 2003). However, these solutions, which rely on separate, interpretation-specific entries, ultimately fail to explain *why*, crosslinguistically, epistemics and roots are expressed by the same lexical items.

With this paper, I would like to show that the individual relativity and the time relativity of modals go hand in hand. Modals are anchored both to an individual and a time, but crucially, not all time/individual pairings are attested. To illustrate this point, I use the semi-modal *have to*, which, unlike modal auxiliaries *may* or *must*, can be fully declined, and hence shows the full range of possible interpretations. In (1), *have to*’s most natural interpretation is epistemic: it expresses a necessity, given what *the speaker* knows, that John was home. Crucially, the modal’s time of evaluation is the speech time; we are talking about what the speaker *currently* knows (about a past event): *given what I know now, it is necessary that John was home.*

(1) John had to be home (at the time of the crime).

When we embed an epistemic modal under an attitude verb, as in (2), the modal is no longer relative to the *speaker*’s evidence, but rather to that of the *attitude holder*, Mary (Speas 2004, Stephenson 2007). Interestingly, the time of evaluation of the modal has to be the attitude time: (2) expresses a necessity given *Mary*’s evidence *at her thinking time.*

(2) Mary thought that John had to be home (at the time of the crime).

Finally, with a root interpretation, as in (3), the modal expresses a necessity for the subject to take the train, given certain circumstances of the base world, namely *his* circumstances. And crucially, the time of evaluation of a root modal has to be the time provided by tense: we are talking about circumstances of *Mary at the time provided by tense* (a past time).<sup>1</sup>

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

The generalization that emerges is that when a modal is speaker-oriented, it is keyed to the speech time (and gets an epistemic interpretation). When it is attitude holder-oriented, it is keyed to the attitude time (and gets an epistemic interpretation). When it is subject-oriented, it is keyed to the time provided by tense (and gets a root interpretation). This pattern is unexpected, given our current assumptions. We are thus faced with a new puzzle, independently of the root/epistemic distinction: why couldn't a modal express possibilities or necessities (epistemic or other) for the subject at the speech time, or for the speaker at a time prior to the speech time? There is nothing conceptually odd with the latter. In fact, we can express such a necessity with an overt restriction: *given what I knew **at the time**, John had to be home*. Why, then, isn't such an interpretation available in the absence of an overt restriction?<sup>1</sup> I will argue that these constraints follow naturally once we assume that modals are relative to an *event* (rather than a world) of evaluation, as events come naturally with time/individual pairs, namely their running time and

---

<sup>1</sup> As we will see, roots are not always tied to the subject, but rather to the participants of the VP event.

<sup>1</sup> An overt restriction may go missing in cases of free indirect discourse, cf. Section 3.2.

participants. We will see that there are three types of events that a modal can be relativized to: the speech event (which anchors the modality to the speaker and the speech time), an attitude event (which anchors it to the attitude holder and the attitude time), and the VP event (which anchors it to the subject and the time provided by tense).

This event relativity will ultimately allow us to keep a unified Kratzerian account, while deriving the correlation between modal flavor and height of interpretation. I will maintain from Kratzer that each modal has a single lexical entry, not specified for a particular flavor, but I will let this modal freely appear above or below tense. The restriction on the interpretation of the modal will come from the very structural position in which the modal appears. This is how: A modal is relative to an *event of evaluation*; specifically, it has an event variable in its restriction that needs to be bound locally. There are two positions within a clause in which a modal can appear (type-wise): right above VP (the ‘low’ position) or right above tense (the ‘high’ position). In the low position shown in (4), the closest event binder is aspect: aspect binds the modal’s event variable, and thus relativizes the modal to the event it quantifies over (the VP event), and hence to that event’s participants (e.g., the subject) and running time (the time provided by tense). In the high position, the modal cannot be bound by its clausemate aspect. In matrix contexts, as shown in (4), its event variable gets bound by a default topmost speech event binder (cf. Percus 2000). This anchors the modal to the speech event, and thus relativizes it to the speaker and the speech time. In embedded contexts, shown in (4), the modal’s event variable gets bound by the aspect of the *matrix*, which quantifies over the *attitude* event. This anchors the modal to the attitude event, and hence the attitude holder and attitude time. We thus derive the correct time/individual constraints.

- (4) a.  $[_{CP} \lambda e_0 \quad [_{TP} T \quad \mathbf{Asp}_1 \quad \text{Mod } e_1 \quad [_{VP} V \quad e_1 ] ] ]$
- b.  $[_{CP} \lambda e_0 \quad \text{Mod } e_0 \quad [_{TP} T \quad \text{Asp}_1 \quad [_{VP} V \quad e_1 ] ] ]$
- c.  $[_{CP} \lambda e_0 \quad T \quad \mathbf{Asp}_2 \quad \mathbf{Att } e_2 \quad [_{CP} \text{Mod } e_2 \quad [_{TP} T \quad \text{Asp}_1 \quad [_{VP} V \quad e_1 ] ] ] ]$

This event relativity allows for the following reformulation of Cinque’s puzzle: why is it that attitude or speech event-relative (i.e., *high*) modals get an epistemic interpretation, while VP event-relative (i.e., *low*) modals get a root one? I will sketch a solution to this problem which exploits a crucial difference between speech and attitude events, on the one hand, and regular VP events, on the other, namely the fact that only the former have associated *propositional content* (i.e., a set of propositions, such as a set of beliefs for *believe*), which allows them to embed propositions. I will take this propositional content to be essential to the licensing of modals’ epistemic interpretations because it provides the very information state epistemics quantify over. Epistemics, I propose, express compatibility with an information state. This is often what looks like a state of knowledge (hence the name *epistemic*, from Greek *episteme* ‘knowledge’), but not always: a closer look at ‘epistemics’ in embedded contexts shows that it is possible (and even desirable) to view these modals as expressing compatibility with the set of propositions that make up the embedding attitude *directly* (i.e., unmediated by a state of knowledge). Thus a sentence like *John believes that it might be raining* would mean ‘It is raining in some world compatible with John’s beliefs’, rather than ‘It is raining in some world compatible with what John knows in his belief worlds’. If this is correct, an epistemic interpretation will only be available when the modal is relative to an event associated with propositional content (i.e., attitude or speech events), which, in turn, is only possible when the modal is in the ‘high’ position. In contrast, ‘low’ modals, being relative to an event that lacks such propositional

content, only receive root interpretations.

This paper is organized as follows: Section 2 reviews Kratzer's unifying account; Section 3 focuses on the interaction of modals with tense and aspect to show that root and epistemic interpretations correlate with two distinct syntactic positions. Section 4 illustrates modals' sensitivity to time/individual constraints and argues for their event relativity. Section 5 proposes an event-relative semantics for modals; Section 6 sketches a proposal for the association of epistemic and root interpretations with high and low positions.

## 2. A unified account for modals: Kratzer (1981, 1991)

Modals come in various flavors often classified into two classes: *epistemics* and *roots*.<sup>2</sup> The French possibility and necessity modals *pouvoir* and *devoir* help to illustrate this, as they can take the full range of meanings available to modals. (5) provides a context that supports an *epistemic* interpretation. Roughly, (5) expresses the possibility/necessity, given what is *known* (in particular that it is 6 p.m., and that she is not at the office), that Mary is home.

(5) *Il est 18 heures. Mary n'est pas au bureau. Elle peut/doit être chez elle.*

'It's 6:00 p.m. Mary is not at the office. She may/must be home.'

The term *root* applies to modals with non-epistemic interpretations. (6) illustrates some of these: (6) is an instance of *deontic* modality: the modals express a permission or an obligation given Mary's father's orders. (6) shows the '*ability*' interpretation of *pouvoir*. (6) illustrates *teleological* modality, which expresses possibilities and necessities given a particular goal.

---

<sup>2</sup> A third class of modals are the so-called quantificational modals as in *Texans can be tall*, where the modal seems to act as a quantifier over the indefinite ('Some Texans are tall'). I put aside this kind of modality here and refer the interested reader to Heim (1982), Brennan (1993), Portner (2009).



- (6) a. *Le père de Mary lui impose un régime très strict. Elle peut/doit manger du brocoli.*  
‘Mary’s father imposes a strict diet on her. She can/must eat broccoli.’
- b. *Mary est très forte. Elle peut soulever cette table.*  
‘Mary is very strong. She can lift this table.’
- c. *Mary veut être à Paris à 17 heures. Elle peut/doit prendre le train pour aller à Paris.*  
‘Mary wants to be in Paris at 5 p.m. She can/must take the train to go to Paris.’

The class of root modals is not simply defined in terms of complementarity to the class of epistemics. In fact, they share an important meaning component: all root modals describe possibilities and necessities given particular circumstances of the world of evaluation, usually centered around the subject (Mary’s diet, her physical condition, her schedule). Kratzer’s account aims at capturing these meaning differences without postulating massive homonymy.

In the modal logic tradition, modals are treated as quantifiers over possible worlds (cf. Carnap 1957, Hintikka 1962, Kripke 1963, Lewis 1968), restricted by an accessibility relation which determines the set of worlds modals quantify over, and hence the particular modal flavor. Kratzer’s (1981) crucial insight is to have the *context* determine the accessibility relation, via what she calls *conversational backgrounds*, brought about by phrases like *in view of what is known*, or covertly by the context. Instead of having as many modals with their accessibility relation wired in as there are modal flavors, the restriction is contextually provided, thereby permitting a single lexical entry for each modal. The contextual nature of the restriction provides a viable account of why (i) each modal can come in a variety of flavors (epistemic, deontic,...), and why (ii) each flavor itself can come in a variety of subflavors (Kratzer 1977): e.g., deontic

*must* in *Mary must pay a fine* can be interpreted as an obligation in view of various kinds of laws (the laws of Boston, the regulations of the IRS,...).

Kratzer proposes two kinds of conversational backgrounds: (i) the *modal base*, which determines a set of accessible worlds; (ii) the *ordering source*, which imposes an ordering on the worlds of the modal base. In the following I depart from Kratzer's formalization, in keeping with the extensional framework I adopt in my proposal, and assume that conversational backgrounds are syntactically represented as arguments of the modal, following von Stechow and Heim (2001). There are two kinds of modal bases: the *epistemic* modal base, which picks out a set of worlds compatible with what is known in the base world; and the *circumstantial* modal base, which picks out a set of worlds compatible with certain circumstances of the base world. Formally, a modal base  $f$  is a function from worlds to sets of propositions: for each world, the set of propositions 'accessible' from it.  $\cap f(w)$  denotes the *intersection* of the set of propositions accessible from  $w$ ; it picks out the set of worlds in which all propositions of the modal base hold: worlds compatible with what is known in  $w$  for (7), worlds compatible with the circumstances in  $w$  for (8):

$$(7) \quad \cap f_{\text{epist}}(w) = \{w' : w' \text{ is compatible with what is known in } w\}$$

$$(8) \quad \cap f_{\text{circ}}(w) = \{w' : w' \text{ is compatible with certain circumstances in } w\}$$

In Kratzer's system, the main difference between epistemics and roots comes from the modal base.<sup>3</sup> Epistemic interpretations arise from epistemic modal bases, roots from circumstantial ones: recall that the examples in (6) all express possibilities or necessities given

---

<sup>3</sup> "Root modality comprises all occurrences of modals with a circumstantial modal base." (Kratzer 1991: 650)

certain circumstances in the base world. Further meaning differences among roots are captured by the *ordering source*, which orders the set of worlds provided by the modal base according to a particular (deontic, bouletic,...) ideal. Not all ordering sources are compatible with all modal bases: a circumstantial modal base is compatible with deontic (laws), bouletic (wishes), or teleological (aims) ordering sources; an epistemic modal base typically combines with a stereotypical ordering source (which orders worlds based on how well they fit stereotypical expectations). Formally, the ordering source  $g$  is also a function from worlds to sets of propositions (sets of laws for deontics, sets of desires for bouletics,...). The set of worlds modals quantify over is derived by a function  $max$  which selects the worlds of the modal base in which a superset of the propositions of the ordering source hold: a modal quantifies over the most ideal worlds (in terms of laws, desires...) of the modal base. Both  $f$  and  $g$  are contextually determined (when not overt). This allows for a single entry each for *must* and for *can*, which only differ in force of quantification:

- (9) a.  $[[\text{must}]] = \lambda f_{\langle s, \text{stt} \rangle} \lambda g_{\langle s, \text{stt} \rangle} \lambda q_{\langle \text{st} \rangle} \lambda w. \forall w' \in \max_{g(w)}(\cap f(w)): q(w') = 1$   
 b.  $[[\text{can}]] = \lambda f_{\langle s, \text{stt} \rangle} \lambda g_{\langle s, \text{stt} \rangle} \lambda q_{\langle \text{st} \rangle} \lambda w. \exists w' \in \max_{g(w)}(\cap f(w)): q(w') = 1$

A crucial advantage of a Kratzerian approach is that it provides a unified treatment of modals: there is just one *can* and one *must*.<sup>4</sup> This accounts nicely for the crosslinguistic use of the same lexical items to express various kinds of modality. It further generates a wide range of possible modal meanings (various subflavors of deontics, bouletics,...) and allows for the contextual nature of this variety.

---

<sup>4</sup> Idiosyncrasies of particular modals (e.g., *might*) are due to selectional properties hardwired into their lexical entries.

In the next section, we will look at structural differences between roots and epistemics that go beyond what a difference in conversational backgrounds can explain. Kratzer (1991) already suggests that her two modal bases could correlate with a difference in argument structure advocated for in the syntactic literature. We will see how to derive such structural differences between roots and epistemics, without losing the unifying strength of her account. From this point on, I will ignore ordering sources, which I take to work exactly as described above: they are selected by context, among those compatible with the modal base they combine with (e.g., a bouletic ordering source requires a circumstantial modal base). The focus of the paper will be the extent to which the selection of the modal base depends, above and beyond contextual factors, on the grammatical environment in which a modal appears.

### **3. Cinque's puzzle**

Crosslinguistically, the same words can be used to express epistemic and root modality, a fact which the Kratzerian account naturally captures. However, just as systematically crosslinguistically, epistemics and roots differ in ways that go beyond meaning differences derivable by context. Epistemics have been argued to scope higher than tense (Picallo 1990, Iatridou 1990, Stowell 2004), aspect (Hacquard 2006, 2009), subjects (Jackendoff 1972, Zubizarreta 1982, von Stechow and Iatridou 2003, Lee 2008), and negation (Coates 1983, Drubig 2001), and roots to scope below all these elements (for an overview, see Hacquard, *to appear*). These scopal facts receive a natural explanation if modals occupy (at least) two different positions—a 'high' position for epistemics and a 'low' position for roots—in accordance with Cinque's hierarchy:

(10) *Cinque's hierarchy* (irrelevant projections omitted)<sup>5</sup>

Mod<sub>epis</sub> > Tense > Aspect > Mod<sub>root</sub>

The correlation between modal flavor and structural position, however, doesn't follow from a purely contextual account. The problem is not merely to let modals appear in two positions, but to ensure that the low position is reserved for roots and the high one for epistemics, a problem often resolved by giving epistemics and roots separate lexical entries that encode structural position. This section focuses on the interaction of modals with tense (Sect. 3.1) and aspect (Sect. 3.2), and reviews proposals that derive structural differences between roots and epistemics (Sect. 3.3).

### 3.1. Interaction with tense

This section argues that modals with a root interpretation are always interpreted within the scope of tense, whereas modals with an epistemic interpretation take scope over tense. While the Kratzerian account (and most of its predecessors) ignores the relation between modals and tense, it is by now widely recognized that modals are relative not 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 last year may not be one today, and vice versa. Importantly, what the relevant time is seems to depend on the particular interpretation of the modal.

The following examples illustrate. When *have to* takes a root (teleological) interpretation, the circumstances and the goal of the subject are evaluated at the time provided by tense (past).

(11) expresses a past necessity, given Mary's circumstances *then*, to take the train *then*. It cannot

---

<sup>5</sup> Cinque's hierarchy further differentiates positions for various roots below aspect. My proposal only depends on whether a modal scopes above tense or below aspect. I thus remain agnostic as to the need for further differentiation.

express a present necessity, given her circumstances *now*, to have taken the train *then*:

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

The evaluation time of a modal with an epistemic interpretation, on the other hand, is always set to the speech time in matrix contexts,<sup>6</sup> or, when embedded, to the internal now of the embedding verb. In other words, the evaluation time of an epistemic can never be back (or future) shifted (Groenendijk and Stokhof 1975, Picallo 1990, Iatridou 1990, Abusch 1997, Condoravdi 2002, Stowell 2004). With an epistemic interpretation, (12) expresses the necessity given what is known *now* that Mary was home at some past (crime) time:

(12) Mary had to be home (at the time of the crime).

✓ $\text{mod}_{\text{epi}} > \text{past}$ : *It is possible, given what is known now, that Mary was home then.*

\* $\text{past} > \text{mod}_{\text{epi}}$ : *It was possible, given what was known then, that Mary was home.*

Imagine that the evidence gathered at the beginning of a murder investigation (a week ago) pointed to Mary being home at the time of the murder: both Mary and her roommate testified that they were having lunch together there. Yesterday however, Poirot established that Mary's roommate had lied, as she was seen by several eyewitnesses elsewhere at that time, debunking Mary's alibi. In this scenario, (12) is false: it cannot describe an epistemic state that held at the

---

<sup>6</sup> Von Stechow and Gillies (2008) argue that in (i), B's utterance can express a *past epistemic* (not just a *counterfactual*) possibility. However, this reading only seems available in answers to *why* questions, where the epistemic's temporal shifting could be due to a covert *because*, able to shift evaluation parameters (Stephenson 2007). For skepticism that (i) involves a past tense scoping over the modal, see Portner (2009).

- (i) 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.)

time when the evidence still pointed to Mary being home.

Nothing conceptual should rule this reading out. In fact, we can report such a past state, but it requires either an embedding attitude verb (as in (13)), an indirect discourse past tense (as in ((13); Boogart 2007), or an *overt* conversational background (as in (13)):

- (13) a. Two days ago, Poirot thought that Mary had to be home.  
b. This didn't make sense, thought Poirot... Mary had to be home.  
c. Given what Poirot knew then, Mary had to be home.

In all these cases, a past morpheme appears on the modal. However, it lacks the characteristic backshifting of a true semantic past tense: the modal's time of evaluation must be Poirot's thinking time; it cannot precede it. This past morpheme seems to reflect instead a 'sequence of tense' rule, according to which the embedded tense morphologically agrees with the higher past.

Note that some speakers may find the past epistemic interpretation of the unadorned (12) acceptable in narrative contexts. These, however, arguably involve some perspectival shift (via temporal subordination or free indirect discourse), which is able to change both the time and the evaluator of the modal. I leave the issue of how and whether context can silently induce such perspectival shifts for future research (cf. footnote 7).

We see that while roots must be evaluated at the time provided by tense, the same modals with epistemic interpretations must be evaluated at the local now. Italian shows this in a morphologically transparent way: when, as in (18b) below, a modal is forced to be in the scope of tense (as indicated by the tense/aspect morphology on the modal itself), the modal cannot get an epistemic interpretation; whereas when, as in (18a), the time of evaluation of the complement

is backshifted with respect to that of the modal, it is the only interpretation available.

- (14) a. Gianni può aver parlato a Maria. (epistemic, \*deontic/ability)  
Gianni can have talked to Maria  
'Gianni may have talked to Maria.'
- b. Giani ha potuto parlare a Maria. (deontic/ability, \*epistemic)  
Gianni has could talk to Maria  
'Gianni was able to talk to Maria.'

These temporal constraints follow if roots scope below tense and epistemics scope above it (and, hence, pick up the local 'now' as their time of evaluation); cf. Picallo (1990), Iatridou (1990), Abusch (1997), Cinque (1999), Abraham (2001), Stowell (2004).<sup>7</sup>

### 3.2. Interaction with aspect

The interaction of modals with aspect seems to corroborate the two positions for modals: one above tense and aspect for epistemics, one below for roots. Aspect is the category of meaning that relates the running time of an event (the VP event) to a time of reference (provided by tense and time adverbials). We focus here on perfective aspect. Formally, perfective is treated as an existential quantifier over the VP event, which places its running time within a reference time interval (cf. Kratzer's 1998 formalization of Klein 1994). The French example below illustrates:

---

<sup>7</sup> Alternatively, we could derive this temporal constraint by having *all* modals (epistemics included) scope below tense, and stipulate that tense must be a present with epistemics (Condoravdi 2002). I choose the other option for two reasons: (i) the two positions are corroborated by the interaction of modals with aspect; (ii) we will see that modals' temporal restrictions go hand in hand with individual restrictions. Assuming a structural source for both will provide a natural explanation for the systematic, and otherwise arbitrary, time/individual anchorings of modals.



(15) Hier matin, Mary lut un livre.

Yesterday morning, Mary read-past-pfv a book

‘Yesterday morning, Mary read a book.’

$\exists e_1[t(e_1) \subset \text{YESTERDAY} \ \& \ \text{Mary read a book}(e_1)]$

Bhatt (1999) shows that in languages with a morphologically overt aspectual distinction (like French), the *perfective* on an ability modal yields what he calls an ‘actuality entailment’, that is, an uncancelable implication that the proposition expressed by the complement was actualized. This is why the continuation in (16) comes out as a contradiction. Actuality entailments are not forced with the *imperfective*. The imperfective comes with its own layer of modality (it reflects the presence of a generic operator; Bhatt 1999), which only forces the proposition expressed by the complement to occur in the worlds provided by the imperfective itself (‘generic’ worlds), but not necessarily the actual world. Languages like English that do not differentiate perfective and imperfective overtly hence lack an unambiguous form that has an uncancelable actuality implication.

(16) Mary put soulever cette table, #mais elle ne la souleva pas.

Mary could-pfv lift this table, #but she NE it lifted not

‘Mary was able to lift this table, #but she didn’t lift it.’

Actuality entailments arise with all root modals, but crucially not with epistemics (Hacquard 2006, 2009): aspect does not affect the non-implicative behavior of modals with an epistemic interpretation. This is shown in (17). Unlike Italian (and as with English *have to*), French allows

tense and aspect morphology to appear on a modal even when it receives an epistemic interpretation, in which case tense and aspect are still obligatorily interpreted below the modal. (17) is thus truly ambiguous between an epistemic ('John may have taken the train') and a root (ability) interpretation ('John was able to take the train'). Importantly, while the latter yields a contradiction (#'John managed to take the train but it's possible he didn't'), no contradiction arises with an epistemic interpretation: the complement need not have taken place in the actual world, but only in some world compatible with the speaker's knowledge.

- (17) John put prendre le train, bien qu'il soit possible qu'il ne l'ait pas pris.  
 John can-past-pfv take the train, even though it is-SUBJ possible he didn't  
 'John may have taken the train, even though it's possible he didn't.'

Hacquard (2006, 2009) argues that actuality entailments result from having aspect take scope over the modal, which happens with roots, but crucially not with epistemics. When aspect takes scope *under* a modal (as happens with epistemics), its world of evaluation is the one provided by the modal, anchoring the event to the modal worlds. When, however, aspect takes scope *over* the modal (as happens with roots), its world of evaluation has to be the matrix one, thereby forcing the event to occur in the actual world. This is sketched below:

- (18) a. Mary put prendre le train. epistemic>aspect  
 'Mary may have taken the train.'  
 b. [<sub>ModP</sub> can [<sub>TP</sub> past [<sub>AspP</sub> perf<sub>I</sub> [<sub>VP</sub> Mary take the train e<sub>1</sub> ] ] ] ]  
 c. *'There is a world w compatible with what is known in the actual world, such that*

*there is a past event in  $w$  which is a train taking event by Mary.'*

- (19) a. Mary put prendre le train. aspect>root  
'Mary was able to take the train.'
- b. [TP past [AspP perf<sub>1</sub> [ModP can [VP Mary take the train  $e_1$  ] ] ] ]
- c. *'There is a past event  $e_1$  in the **actual world**, which in some world compatible with the circumstances in the actual world is a train taking event by Mary.'*

Example (19) is true if there is an actual event  $e_1$  which in some circumstantial world is a train-taking by Mary. The actuality entailment arises via two assumptions: (i) the same event  $e_1$  can occur in several worlds (here, the actual world and the world provided by the root modal); (ii) it keeps its description across worlds via a default principle holding constant across worlds whatever property is part of the VP description (here,  $e_1$  is a train-taking by Mary in all worlds in which it occurs). Note that with the imperfective, the LF will differ from (19) only in that  $e_1$  will be anchored to the worlds provided by the imperfective itself, rather than the actual world: 'There is a past event  $e_1$  in **all 'generic' worlds  $w$** , which in some world  $w'$  compatible with the circumstances in  $w$  is a train-taking event by Mary.' This forces a train-taking by Mary only in generic worlds.

To sum up, epistemics and roots differ in the way they interact with aspect: roots yield actuality entailments with the perfective, epistemics do not. What seems to shield epistemics from this effect is the fact that, unlike roots, they are interpreted above aspect.

### 3.3. Epistemics vs. roots: a difference in height of interpretation

Epistemics and roots differ in meaning: the former talk about possibilities and necessities given

what is known, the latter, given certain circumstances. For Kratzer, this difference results from two types of contextually provided restrictions. Epistemics and roots also differ in their interaction with tense and aspect, a fact which cannot be due to a difference in meaning alone: epistemics are interpreted higher, and roots lower, than tense and aspect. This follows if there are two positions for modals: one above tense (S-level) and one below aspect (VP-level).

Various proposals derive a difference in height of interpretation by having epistemics and roots select different types of arguments (Zubizarreta 1982, Jackendoff 1972, Butler 2003, Hacquard 2009), project in different domains (Picallo 1990), or have different featural makeups (Zagona 2007). More radical accounts deny any connection between them. Westmoreland (1998) and Drubig (2001), for instance, take epistemics not to be modals, but evidentials. Similarly, to explain actuality entailments, Bhatt (1999) proposes that, despite appearances, the ability modal is not a modal, but an implicative predicate (like *manage*). All these proposals ultimately lead to separate, interpretation-specific lexical entries for each modal, and leave open the question of why epistemic and root modality are expressed by the same lexical items crosslinguistically. One could appeal to a diachronic (or metaphoric) explanation: epistemic interpretations tend to develop crosslinguistically from root ones (cf. Sweetser 1982, Papafragou 1998 for discussion). However, while this may account for why roots and epistemics share a common form, it leaves unexplained why roots and epistemics systematically end up with the scopal properties they do.

Again, the analytic goal is not merely to let modals appear in two different positions, but to ensure that the low position is reserved for roots and the high one for epistemics. The above accounts resolve both issues in one fell swoop, by associating modal flavor and modal syntax directly in the lexicon, via two separate lexical entries. Alternatively, we could let modals in general (i.e., not specified for flavor) come in two syntactic types, and then provide independent

reasons for the association between modal flavor and modal syntax (cf. Brennan 1993).<sup>8</sup> This would avoid Cinque's puzzle: there would be just one flavor-independent modal (corresponding to one lexical entry), which would come in two syntactic types (or one flexible/underspecified type), and some independent (i.e., not lexically specified) correlation between modal syntax and modal flavor. This is the path that I choose: I will argue that each modal has a single lexical entry unspecified for flavor (à la Kratzer), which allows it to appear in two different positions. I will propose that independent factors connected with these positions are responsible for the association between modal flavor and height of interpretation.

I first want to show that the interpretation a modal can receive is even more restricted than what we have seen: modals are anchored not only to a time, but to an individual as well, and crucially, not all possible time-individual pairs are attested.

#### **4. Modals' event relativity**

##### **4.1. Individual relativity**

Roots and epistemics are often distinguished in the typology literature by the individual the modality is anchored to. Roots are *subject-oriented*: they express possibilities given the subject's circumstances. Epistemics are *speaker-oriented*: they express possibilities given the evidence available to the speaker (Bybee et al. 1994). In this section, I refine this notion of individual orientation slightly. First, Speas (2004) and Stephenson (2007) show that epistemics are not always relative to the speaker, but rather to the local knowledge bearer: the *speaker* in matrix contexts,<sup>9</sup> the *attitude holder* in embedded contexts. While *must* reports the speaker's epistemic

---

<sup>8</sup> While I follow Brennan in not associating modal syntax with modal flavor in the lexicon, I depart from her explanation for the association, which, for her, is due to arbitrary ontological commitments made by the community of language users. I believe the association is not arbitrary, and this is why it is found in language after language.

<sup>9</sup> This claim may need even further refinement, as it has been argued that the relevant individual for epistemics in

state in (20), it can only report that of Bill in (20). Neither can report John's epistemic state:

(20) a. John must have won.

b. Bill thinks that John must have won.

'Bill thinks that to the best of his/\*my/\*John's knowledge, John must have won.'

Second, while the subject orientation of root modals is traditionally taken to follow from a thematic relation between the subject and the modal (cf. Perlmutter 1971, Ross 1969, Jackendoff 1972), examples like (21) argue that even the most root-like of modals cannot be 'control' predicates (cf. Bhatt 1998, Hackl 1998, Wurmbrand 1999): an ability modal allows for 'weather-it' subjects, as in (21a), and does not necessarily ascribe an ability to its subject, but to some other argument (*the pool*, rather than *a lot of people*, in (21b)). The correct generalization seems to be that root modality is centered around the event described by the VP and its participants. 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:

(21) a. It can rain hard here.

(Hackl 1998)

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

We thus see that roots and epistemics differ in the individual they seem to be anchored to: the local knowledge bearer for epistemics, some participant of the VP event for roots.

---

matrix contexts is not just the speaker, but the speaker and her relevant community (cf. DeRose 1991, von Stechow and Gillies 2008).

## 4.2. Modals are keyed to individual-time pairs

In Section 3.1, we saw that modals are relative to a time of evaluation. For epistemics, this time is the speech time (in matrix contexts) or the internal now of the attitude (in embedded contexts). For roots, it is the time provided by tense. We just saw that modals are also relative to an individual: for epistemics, this individual is the speaker in matrix contexts<sup>10</sup> and the attitude holder when embedded (Stephenson 2007). For roots, this individual is the subject (Brennan 1993, Bybee 1995) or another participant of the VP event. Modals never express an epistemic possibility for the subject nor a circumstantial possibility for the speaker.

Importantly, not all combinations of times and individuals are possible. When a modal is relativized to the subject, as in example (22) below, where *have to* takes a teleological interpretation, it must be relative to the time given by tense (a past time), not the speech time.<sup>11</sup>

- (22) Mary had to take the train.
- a. ‘Given **Mary**’s circumstances **then**, she had to take the train then.’
  - b. \*‘Given **Mary**’s circumstances **now** she had to take the train.’

When a modal is relativized to the speaker, as in the example (23) below, where *have to* is interpreted epistemically, it has to be relativized to the speech time, not a time prior to it.

---

<sup>10</sup> It has recently been argued that epistemics can in fact be evaluated with respect not to the speaker’s evidence but that of any ‘assessor’ hearing the sentence (e.g., MacFarlane 2003, Egan et al. 2004). Such accounts propose that epistemics are relative to an additional parameter in the index, ‘the assessor’. I am not convinced that the evidence that these proposals rest on warrants the addition of such a parameter (for a critical review, see von Stechow and Gillies 2008). However, such proposals could be accommodated within the current account: if the modal is relative to the assessor, it also has to be relative to the assessing time. This could be captured via an ‘assessing’ event.

<sup>11</sup> When there is no clear agentive subject, as in *It could rain hard here in those days*, a root modal is still relativized to a participant of the VP event (e.g., the location *here*) and, crucially, to the time of that VP event (‘in those past days’).

- (23) a. Mary had to be home.  
b. ‘Given what **I** know **now**, it must be the case that Mary was home then.’  
c. \*‘Given what **I** knew **then** Mary had to be home.’

When a modal is relativized to the attitude holder (John), as in (24), where the modal receives an epistemic interpretation, it has to be relativized to the attitude time, and not the speech time or a time prior to the thinking time:

- (24) John thought that Mary might be home.  
a. ‘Given what **John** knew at **his thinking time**, it was possible M. was home.’  
b. \*‘Given what **John** knows **now**, it was possible M. was home.’

The empirical generalization that emerges is as follows: when the modal is speaker-oriented, it is keyed to the speech time and gets an epistemic interpretation. When it is attitude holder-oriented, it is keyed to the attitude time and gets an epistemic interpretation. When it is VP participant-oriented, it is keyed to the time provided by tense and gets a root interpretation.<sup>12</sup> These restrictions are puzzling. There is nothing conceptually odd about a modal keyed to the speaker at a past time. As we saw, such modality can be expressed with an overt restriction: *In view of what I knew then, Mary had to be home*. Why is it that, in the absence of an explicit restriction, such a reading is unavailable? What is it about the semantics of modals and their grammatical

---

<sup>12</sup> It has been argued that deontics split into two classes: those that put an obligation on the subject, and those that put it on the addressee (cf. Brennan 1993, Bhatt 1999, Hacquard 2006). Both kinds fit the pattern: when subject-oriented, the modal is VP-time oriented, when addressee-oriented, it is speech time oriented. Addressee-oriented deontics can be viewed as another kind of speech event-relative modality (see Hacquard 2006 and Portner 2009).



environment that forces these time/individual constraints? I propose that they arise from some ‘event anchoring’: by anchoring a modal to an event, we anchor it to its running participants and running time. The reason why an overt restriction allows a past evaluation time is that it itself performs the event anchoring, and relativizes the modal to *what I knew at some past time*.

What are the anchoring events in the absence of an overt restriction? There seem to be three candidates: the *speech event*, which provides an individual—the speaker (its agent)—and a time—the speech time; *attitude events*, which provide an individual—the attitude holder (their experiencer)—and a time—the attitude time; *VP events*, which provide an individual—the subject (their agent) or sometimes other arguments—and a time—the time given by tense.

The event relativity of modals seems to be strictly local: modals are relative to the *closest* event. Take the following sentence, which involves three events: the speech event (an asserting event by the speaker at the speech time); a thinking event (by John, yesterday); a train-taking event (by Mary, the day before yesterday):

(25) John thought yesterday that Mary had taken the train the day before.

Let’s now see what happens when we insert a modal right below the parts of the clause that express each of these three events. The speech event, if represented syntactically, should be in topmost position. And having a modal in the highest possible position (above tense, and presumably right below the speech event), as in (26), makes it relative to the speaker at the speech time: ‘Given what I know now, it is possible that John thought Mary had taken the train.’

(26) John might have thought yesterday that Mary had taken the train the day before.

When the modal is in the high position of an embedded clause (above tense), as in (27), the modal is relative to the attitude holder at the attitude time: ‘Given what John thought yesterday, it was possible that Mary had taken the train.’

(27) John thought yesterday that Mary might have taken the train the day before.

Finally, when the modal is in the low position of the embedded clause (below tense), as in (28), it is relative to the circumstances of the subject at the time of the train taking: ‘Given Mary’s circumstances the day before yesterday, it was necessary that she take the train’ (according to John).

(28) John thought yesterday that Mary *had* to take the train the day before.

A modal can thus be relative to the most local event, and hence to its participants and running time. It further appears that, in fact, it *must*. In (27), *might* has to be keyed to the attitude event. It cannot be keyed to the speech event. (27) can never mean ‘John thought yesterday that it was possible, **given what I know**, that Mary took the train the day before’. Similarly, in (28), the modal has to be keyed to the VP event. (28) could never be interpreted as ‘John thought that it was necessary **given what I know** that Mary took the train’.

The upshot of this section is that a modal seems to be relative to an individual and a time, but that not all individual-time pairs are attested. Instead, the modal has to be keyed to the participants and running time of the most local event.

## 5. Proposal: an event-relative semantics for modals

We started out with two conflicting crosslinguistic generalizations: on the one hand, both epistemic and root modals can be expressed by the same lexical items, favoring a unified account. On the other, their systematically different distribution argues they should get separate lexical entries. We then saw that beyond a root/epistemic distinction, modals seem to be relative to one of three kinds of individual/time pairs: speaker/speech time, attitude holder/attitude time, and VP-event participant/VP-event time. These, I argued, could be obtained by anchoring the modal to speech, attitude, and VP-events, respectively. In this section, I show how to implement this by modifying the Kratzerian account so as to make modals relative to an event rather than a world of evaluation. In our new system, the syntactically high modals are those that are relative to speech or attitude events; the low ones, those that are relative to VP-events. This event relativity will lead us to recast our original puzzle as follows: why is it that only modals relative to speech or attitude events get epistemic interpretations (epistemic modal base), while those relative to VP-events get root ones (circumstantial modal base)? This question is addressed in Section 6.

The first step is to modify our semantics for modals, so that they are relative to an event, rather than a world of evaluation. To do so, I minimally change the lexical entries in (9), such that the modal base is now a function from events, rather than worlds, as in (29). A preliminary event-relative version of our modal bases is given in (30) (type  $\varepsilon$  is for eventualities):

- (29) a.  $[[\text{must}]] = \lambda f_{\langle \varepsilon, \text{st} \rangle} \lambda g_{\langle \varepsilon, \text{st} \rangle} \lambda q_{\langle \text{st} \rangle} \lambda e. \forall w' \in \max_{g(e)}(\cap f(e)): q(w') = 1$   
 b.  $[[\text{can}]] = \lambda f_{\langle \varepsilon, \text{st} \rangle} \lambda g_{\langle \varepsilon, \text{st} \rangle} \lambda q_{\langle \text{st} \rangle} \lambda e. \exists w' \in \max_{g(e)}(\cap f(e)): q(w') = 1$

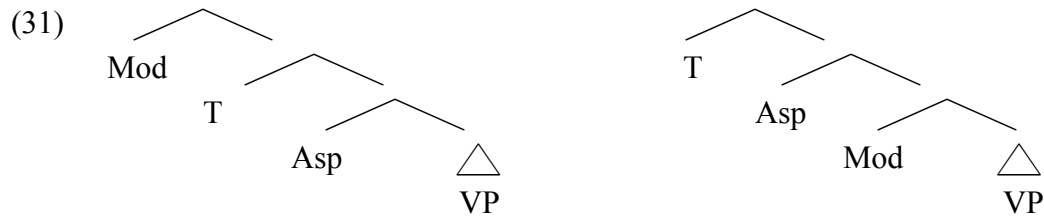
(30) *Recasting the modal bases in event talk* (to be revised)

- a.  $\cap f_{\text{epis}}(e) = \{w' : w' \text{ is compatible with what is known in } e \text{ (by Ag}(e) \text{ at time}(e))\}$
- b.  $\cap f_{\text{circ}}(e) = \{w' : w' \text{ is compatible with the circumstances of } e\}$

To capture the strictly local relation between the modal and the event it is relative to, I will treat the modal’s event argument as a variable and assume that this type of variable needs to be bound by the closest event binder. But first, we need to solve a mechanical problem (independent of event relativity), to allow the same modal to appear in two positions: one above TP and one above VP. I present a possible implementation below, using a fully extensional framework.

### 5.1. One modal, two positions

The upshot of Section 3 is that there are two positions for modals — one above T, one above VP — as depicted in (31).



This section shows how to allow the same modal to appear syntactically in either of these positions. The first stumbling block arises from semantic type considerations: under standard assumptions, if the modal is to combine with a proposition, it should only be able to appear at the TP level. The solution I adopt here is couched within an extensional framework, where worlds and tenses are explicitly represented in the syntax (cf. Partee 1973, Cresswell 1990, Kratzer

1998, Percus 2000, a.o.). This, along with a particular view of aspect, will yield two nodes of the right propositional types for a modal to combine with: TP and VP.

I assume a referential analysis of tense, where tenses are pronouns (cf. Partee 1973, Abusch 1994, Heim 1994, Kratzer 1998). They combine with predicates of times, the way an individual pronoun combines with a predicate of individuals. I borrow the following lexical entries from Kratzer (1998), where the two main tenses are free indexicals (present and past):

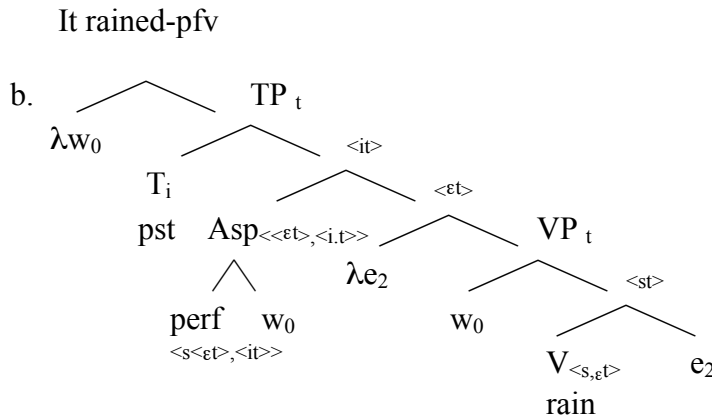
- (32) a.  $[[\text{pres}]]^c$  only defined if  $c$  provides an interval  $t \approx t_0$ . If defined  $[[\text{pres}]]^c = t$   
b.  $[[\text{past}]]^c$  only defined if  $c$  provides an interval  $t < t_0$ . If defined  $[[\text{past}]]^c = t$

Following Percus (2000), I take worlds to be explicitly represented in the syntax as pronouns (situation pronouns in his framework) that need to be bound either by a topmost default world binder ( $\lambda w_0$ ), or by modals or attitude verbs. In the Davidsonian tradition, verbs are predicates of events, which have to combine with an event argument (Davidson 1967). Aspects are quantifiers over events: they take predicates of events (VP) and return predicates of times, which in turn combine with tense. Aspect is traditionally assumed to be base-generated under tense: a verb has a free event variable that gets bound by the aspect above it. I adopt instead an alternative view (suggested in lecture notes by von Stechow 2001), according to which aspect is merged as an argument of the verb. To resolve a type mismatch, aspect (being of a higher quantifier type) moves out of this position, leaving a trace of event type  $\epsilon$ , which it binds from its target position (in a way analogous to the QR operation proposed for quantifiers over individuals in object position). I assume the following lexical entry for the perfective (adapted from Kratzer 1998):

(33)  $[[\text{PERFECTIVE}]] = \lambda w. \lambda P_{\langle \epsilon t \rangle}. \lambda t. \exists e[e \text{ is in } w \ \& \ \tau(e) \subset t \ \& \ P(e) = 1]$

A simple example is presented in (34) below.  $\lambda w_0$  here binds the world argument of aspect and of the embedded VP. Aspect provides existential quantification over the event described by the VP and locates its running time with respect to the time provided by tense. It moves from the VP to a position right below T, creating an abstraction over events:

(34) a. Il plut.



c.  $[[a]]$  only defined if  $t < t_0$ . If defined  $[[a]] = 1$  iff  $\exists e[e \text{ in } w_0 \ \& \ \tau(e) \subset t \ \& \ \text{rain}(e, w_0)]$

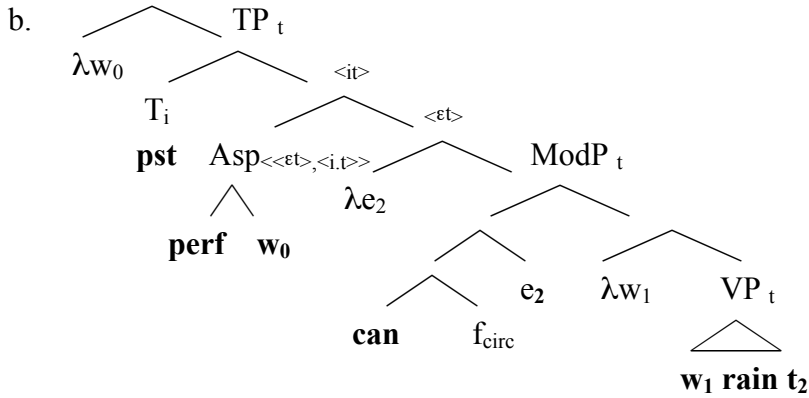
In most cases, aspect movement and base generation are equivalent. However, when a modal intervenes between tense and the VP, as in (38), the movement approach lets the modal have a complement of propositional type and still occur within the scope of aspect.<sup>13</sup>

(35) a. Il put (enfin) pleuvoir.

<sup>13</sup> We could take this proposal one step further by having tenses be quantifiers merged as arguments of aspect, which in turn move for type reasons. Yet another extension would have modals merge as arguments of tense, and also move for type reasons (cf. von Stechow 2001, Shimada 2008).

It could-pfv (finally) rain

'It (finally) managed to rain.'



c.  $[[a]]$  defined if  $t < t_0$ . If defined  $[[a]] = 1$  iff

$\exists e_2 [e_2 \text{ in } w_0 \ \& \ \tau(e_2) \subset t \ \& \ \exists w_1 \text{ compatible with the circumstances of } e_2, \text{ rain}(e_2, w_1)]$

One may wonder whether aspect movement and the representation of worlds in the object language are crucial to my account. The answer is No: these assumptions in tandem only provide one possible implementation for having the same modal appear in two positions. Having worlds in the object language lets both VPs and TPs be of the right propositional type; aspect movement allows aspect to scope over a modal while still having modals come in a single type (i.e., always combine with propositions and return propositions). The same effects could be achieved without either assumption, by allowing the modal to raise its type when it appears in the low position. I find the QR approach plausible, however (and perhaps empirically motivated<sup>14</sup>), given the growing body of evidence for world variables in the object language. What *is* crucial for my

<sup>14</sup> Italian presents a morphological argument for aspect movement, as perfective is expressed by auxiliary *be* or *have* + participle, the particular auxiliary being determined by the embedded verb (unaccusatives select *be*):

- (i) Gianni è potuto andare all cinema.  
Gianni is can-pst-pfv go to the movies.

This is puzzling if aspect is base-generated above the modal, but is motivated under the movement approach: aspect (realized as auxiliary *be*) is selected as an argument of the verb and then moves above the modal for type reasons.

account is to let the same (flavor-independent) modal appear in two positions, as a first step towards explaining the fact that there are two positions for modals, each of which correlates with a particular interpretation (epistemic for the high, circumstantial for the low position), *without* hardwiring flavor into each syntactic position.

## 5.2. Local binding of worlds and events

The main motivation for representing worlds as pronouns in the object language (and for the associated binding theory) in Percus (2000) comes from an otherwise unexpected pattern of attested—and unattested—readings for sentences like the following (see also Farkas 1997):

(36) If every semanticist owned a villa in Tuscany, what a joy it would be. (Percus 2000)

The sentence in (36) is ambiguous between a ‘transparent’ reading, which asserts that all worlds in which every *actual* semanticist owns a villa in Tuscany are happy worlds (different world indices), and an ‘opaque’ reading, which states that all worlds where every semanticist in *those* worlds owns a villa are happy worlds (same world indices).

To capture this, Percus argues that there are explicit world variables, which cannot be free. They need to be bound either by a default matrix binder that maps to the actual world or by a binder provided by, e.g., a modal. His binding theory differentiates between a world variable in the restriction, versus in the scope, of a quantifier like *every*. While world variables in the restriction of a quantifier can be bound by either the matrix binder (yielding actual semanticists) or the binder provided by *would* (yielding counterfactual semanticists), world variables in the scope must be bound by the closest binder (*would*), thus preventing unattested readings of (36)



about (counterf)actual semanticists owning villas in the actual world. Thus only variables in the scope of quantifiers must be bound by the most local world binder.

I propose to generalize Percus' binding principles, which in Percus (2000) are stated on a construction-by-construction-basis as in (37), and extend them to event variables.<sup>15</sup> This will account for the observed interpretation of modals: modals are always relative to the closest *event*.

(37) *Locality of world and event binding*

(i) All world and event variables need to be bound.

(ii) All world and event variables on the 'spine' of the tree (T,A,M,V) need to be bound by the closest binder.

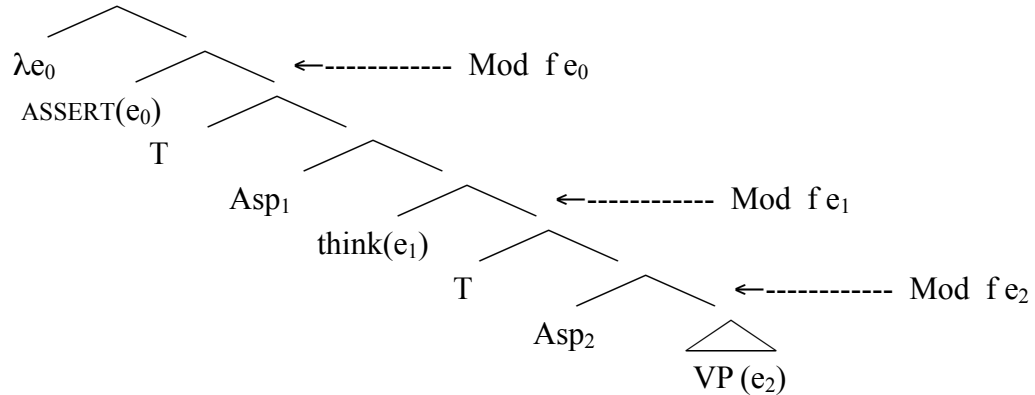
### 5.3. Event binders

We saw that there are two positions within a clause where a modal can appear: above TP ('high' position) or above VP ('low' position). When above TP, the modal is relative to the speech event (speaker/speech time), unless the clause that contains the modal is embedded under an attitude verb, in which case the modal is relative to the attitude holder/attitude time. When the modal is right above the VP, it is relative to the VP-event's participants/running time. To derive this event relativity, I will have the modal's event variable be obligatorily bound by the closest event binder. I propose that there are *two* event binders: aspect, and a default speech event binder  $\lambda e_0$ . This will allow modals to be relativized to *three* kinds of events: the speech event  $e_0$ , which in declaratives is an assertion, attitude events ( $e_1$ ), and VP events ( $e_2$ ):

---

<sup>15</sup> I'm grateful to Kai von Stechow (p.c.) for this generalization of Percus's binding principles.

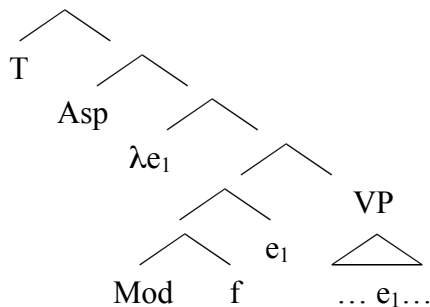
(38)



### 5.3.1. Aspect

When a modal appears in the low position (above VP), its event argument gets bound by aspect (the closest event binder). That aspect can bind events is uncontroversial: aspect is, by the definition presumed here, a quantifier over events. It will bind the modal's event variable the way *every* binds a pronoun (e.g., *Every boy<sub>i</sub> loves his<sub>i</sub> mother*). Recall from the binding condition in (37) that event variables cannot be free. The event variable in the modal's restriction has no choice but to be coindexed with the closest event binder, here aspect:<sup>16</sup>

(39)



This binding by aspect will occur whenever the modal is in the low position (of a matrix or an

<sup>16</sup> If aspect is not base-generated under tense, could it be merged as the event argument of the modal? Maybe nothing prevents it from merging, but something prevents it from moving out (K. von Stechow, p.c.). A modal's restriction is in an island for extraction: neither covert ((ia)) nor overt ((ib)) movement is allowed. Hence, if aspect were to merge in the modal's restriction, it couldn't get out. Only an event pronoun can appear in this position.

- (i) a. If every boy<sub>j</sub> comes, his<sub>i/\*j</sub> mother will be happy.
- b. \*Who<sub>i</sub> if t<sub>i</sub> comes, will Mary be happy?

embedded clause). What happens to modals in the high position depends on whether the modal is in a matrix or an embedded clause (complement of an attitude verb).

We first turn to embedded clauses. When the modal is in the high position of an embedded clause, its event variable cannot be bound by the embedded aspect: the closest binder is the aspect in the matrix, which happens to quantify over the matrix (attitude) verb's event argument. This binding anchors the modal to the attitude event. To see this, we first need to recast our semantics of attitudes in event terms.

In the Hintikka tradition, attitude verbs are treated as quantifiers over worlds, e.g., *believe* quantifies over worlds compatible with the subject's doxastic alternatives:

$$(40) \quad [[\textit{believe}]] = \lambda p. \lambda x. \lambda w. \forall w' \in \text{DOX}(x, w): p(w')=1$$

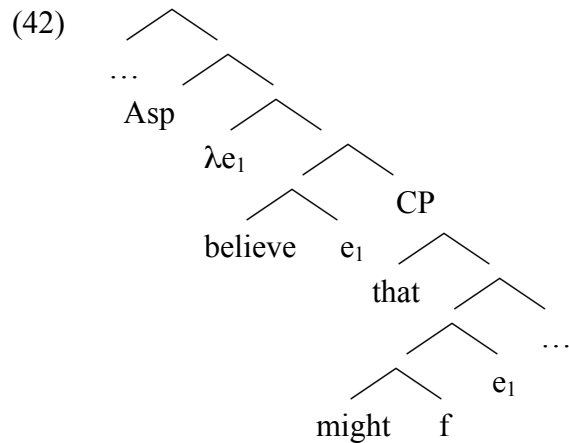
The literature on attitudes usually abstracts away from their aspectual properties and hence ignores their Davidsonian argument (but see Moltmann 2003, Kratzer 2006). However, attitude verbs are verbs, that is, predicates of events. They do differ from regular verbs, like *run* or *sleep*, in that they express quantification over a particular set of worlds. But this quantification can be rendered as an aspect of an event predicate. Let's say that an attitude, besides the proposition that is its 'object', has a set of propositions that is its 'content'. The intersection of the propositions in the content is the set of worlds over which the attitude quantifies. Thus if *John believes it rained*, the *object* of his belief is the proposition that it rained, while the *content* is the set of all propositions that John believes; the intersection of these is the set of worlds compatible with what he believes, his doxastic alternatives. Now we can render the verb *believe* as an event predicate in terms of the experiencer, object, and content of the event, as in (41), where CON(e)

denotes the content of  $e$ .  $\text{CON}(e)$  is defined when  $e$  has propositional content (when  $e$  is an attitude event like *believe*) and undefined when  $e$  lacks such content (as with ordinary verbs like *run*).

$$(41) \quad [[\text{believe}]] = \lambda e. \lambda p. \lambda x. \lambda w. \text{Exp}(e,x) \ \& \ \text{belief}'(e,w) \ \& \ \forall w' \in \cap \text{CON}(e): p(w')=1$$

$$\text{where } \cap \text{CON}(e) = \text{DOX}(\iota x \text{Exp}(e,x), w)$$

Now, when a modal is in the high position of the complement clause of an attitude verb, its closest event binder is the matrix aspect that quantifies over the attitude's event argument. The modal is thus relativized to the attitude event, as shown in (42).



Thus, a modal can be anchored to the VP-event of its own clause, or to a higher attitude event via aspect. We now need to handle the case where a modal is in high position in a matrix clause.

### 5.3.2. Speech event

I take the extensional framework one step further by assuming that the speech event itself is represented syntactically. While this assumption is still controversial, a growing body of literature suggests that a syntactic representation of the speech act is needed to derive a series of independent phenomena (cf. Ross 1970, Rizzi 1997, Ambar 1999, Ginzburg and Sag 2001, Krifka 2001, Tenny and Speas 2004).<sup>17</sup> To explain the modal facts, I will adopt an event version of the speech act representation. There is only one speech event per utterance, in topmost position, whose role is to determine the nature of the utterance: if it is an assertion, it will be an asserting event; if it is a question, a questioning event; and if it is an order (as for imperatives), an ordering event. Semantically, I take illocutionary force to arise from intensional operators which determine different sets of worlds in which the expressed proposition holds. I only consider assertions in the following discussion, for which I borrow Alonso-Ovalle and Menendez-Benito's (2003, forthcoming) ASSERT operator. ASSERT is an implicit universal modal operator, which quantifies over the speaker's epistemic/doxastic alternatives and combines with a proposition:

$$(43) \quad [[\text{ASSERT}]]^c = \lambda p. \lambda w. \forall w' \in \text{DOX}_{\text{speaker of } c}(w): p(w') = 1$$

Under this approach, assertions are viewed as expressions of what the speaker knows/believes.<sup>18</sup>

We can recast this operator in event terms, by using the content function from Section 5.3.1, which here picks out the content of the speech event (ASSERT). Again, 'content' refers to the set

---

<sup>17</sup> Here again, syntactic representation may not be essential. We could assume that free event pronouns (i.e., those that cannot be bound by aspect) must refer to the speech event (which may not need to be represented syntactically). However, one of the goals of this paper is to see how much mileage we can get from a syntactic account based on simple locality and binding principles.

<sup>18</sup> Or possibly, a subset of the speaker's beliefs, namely those he is committed to defending (Gunlogson 2001).

of propositions that provide the worlds quantified over—here, the speaker’s doxastic alternatives:

$$(44) \quad [[\text{ASSERT } e_0]] = \lambda p. \lambda w. \text{Assert}'(e_0, w) \ \& \ \forall w' \in \cap \text{CON}(e_0): p(w') = 1$$

*where*  $\cap \text{CON}(e_0) = \text{DOX}(\iota x \text{ Holder}(x, e_0), w)$ ;  $\iota x \text{ Holder}(x, e_0) = \text{speaker}$

Recall that Percus (2000) introduced a default world (situation) binder in topmost position. This world binder ( $\lambda w_0$ ) will always immediately dominate the speech event, thereby anchoring the speaker’s doxastic alternatives to the actual world.<sup>19</sup>

A declarative will thus have as its speech event an *asserting* event, and the proposition expressed by the sentence will be true in all worlds compatible with the content of that event, that is, in all the doxastic alternatives of the speaker of  $e_0$ . The following sentence illustrates:

- (45) a. John is the murderer.
- b.  $[\lambda w_0 \lambda e_0 [\text{ASSERT } e_0 w_0 \lambda w_1 [\text{John is the murderer } w_1] ] ]$
- c.  $\text{Assert}'(e_0, w_0) \ \& \ \forall w' \in \cap \text{CON}(e_0) [\text{John is the murderer in } w']$
- d. ‘John is the murderer in all of the speaker’s doxastic alternatives.’

We can now straightforwardly derive the speech event relativity of modals in matrix

---

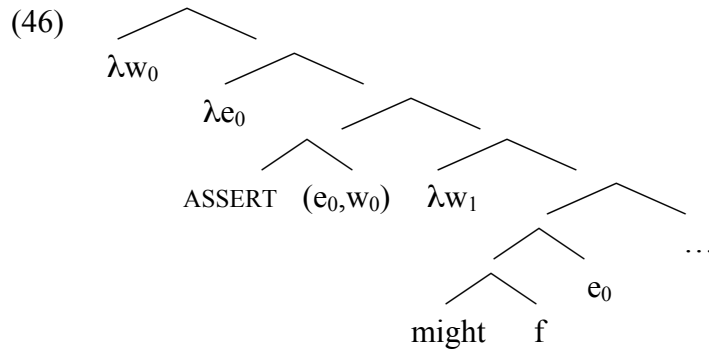
<sup>19</sup> We could simplify the architecture by getting rid of this world binder and having simply the default event binder  $\lambda e_0$ , under the view that events occur single-worldly: the speech event  $e_0$  would single out one particular world ( $w_0$ ), namely *the* world in which  $e_0$  occurs (P. Anand, p.c.). We could thus get rid of the world argument of ASSERT:

$$(i) \quad [[\text{ASSERT } e_0]] = \lambda p. \forall w' \in \cap \text{CON}(e_0): p(w')=1$$

*where*  $\cap \text{CON}(e_0) = \text{DOX}(\iota x \text{ Holder}(x, e_0), w_0)$ ;  $\iota x \text{ Holder}(x, e_0) = \text{speaker}$  ;  $w_0 = \iota w(e_0)$

Similarly, we could get rid of the world argument of attitude and even regular verbs (as suggested by an anonymous reviewer). While this simpler view is attractive, I do not exploit it here, to allow the same event to occur in different worlds and derive actuality entailments; cf. Section 3.2.

contexts: as shown in (46), when the modal is in the high position, its event variable has to be bound by  $\lambda e_0$ .



#### 5.4. Clause architecture

I have proposed that there are two kinds of event binders: aspect (quantifier over events) and a default, topmost event binder. Every sentence contains a single speech event (ASSERT for declaratives), whose modal semantics (in terms of quantification over the doxastic alternatives of the speaker) enables it to embed a proposition. In this framework, we see that embedded and matrix clauses have a similar architecture: both are treated as complements of an attitude event.

$$(47) \quad [\lambda e_0 \text{ ASSERT } (e_0) \quad \text{T Asp}_1 \text{ attitude } (e_1) \quad [\text{CP T Asp}_2 \text{ V}(e_2) ] ]$$

$$\forall w \in \text{CON}(e_0) \qquad \qquad \qquad \forall w' \in \text{CON}(e_1)$$

Let's now summarize the binding possibilities for modals. Recall that, in the Kratzerian tradition, there is just one possibility and one necessity modal, which take propositions as complements. A modal can merge either above or below tense within a clause. When it merges above tense, we have two options: (a) in matrix contexts, the modal's event variable is bound by the default event

binder (and anchored to the speech event); (b) in embedded contexts, it is bound by the matrix aspect that quantifies over the attitude event (and is anchored to the attitude event). When below tense (c), the modal's event variable is bound by aspect (the modal is anchored to the VP event). These three scenarios are summarized in (48).

- (48) a.  $[\lambda e_0 \text{ ASSERT } (e_0) \quad [\text{ModP Mod f } e_0 [\text{TP T } [\text{AspP Asp}_3 [\text{VP V}(e_3) ] ] ] ]$   
 b.  $[\lambda e_0 \text{ ASSERT } (e_0) \text{ T Asp}_1 \text{ attitude } (e_1) [\text{CP } [\text{ModP Mod f } e_1 [\text{TP T } [\text{AspP Asp}_3 [\text{VP V}(e_3)]]]]]]]$   
 c.  $[\lambda e_0 \text{ ASSERT } (e_0) [\text{TP T } [\text{AspP } [\text{Asp}_3 \quad [\text{ModP Mod f } e_3 \quad [\text{VP V}(e_3) ] ] ] ] ]$

We thus have three events a modal can be relative to, as well as two modal bases: epistemic and circumstantial, with the latter serving as the source of all root interpretations. This should yield six kinds of interpretations, listed in (49) below. However, half of these do not seem to be attested (cf. (49b,d,e)). Section 6 revisits each modal base and sketches an account of how to block these unattested readings.

- (49) a. Speech event, epistemic: *given what I know now...*  
 b. \*Speech event, circumstantial: *given the circumstances of the speech event...*  
 c. Attitude event, epistemic: *given what the attitude holder knows at the attitude time...*  
 d. \*Attitude event, circumstantial: *given the circumstances of the attitude event...*  
 e. \*VP event, epistemic: *given what the subject knows at the VP time...*  
 f. VP event, circumstantial: *given the circumstances of the VP event...*



## **6. Solving Cinque’s puzzle: correlating syntactic height with modal flavor**

Cinque’s puzzle can now be reformulated as follows: why do speech or attitude event-relative (high) modals get epistemic interpretations, and VP event-relative (low) modals, root ones? In this section I sketch a proposal for the association of epistemicity and speech and attitude events (in contrast to ordinary VP events). I argue that what sets speech/attitude events apart from ordinary events is (what I am calling) their associated propositional ‘content’, which I take to be crucial for licensing epistemic modal bases. Sections 6.1 and 6.2 discuss epistemic and root interpretations; Section 6.3 sums up the connection between modal base and syntactic position.

### **6.1. Epistemics**

#### **6.1.1. The nature of epistemic modality**

Before we look at how a modal combines with an epistemic modal base, let’s reconsider briefly the nature of epistemic modality. The literature on epistemic modality is vast, and I cannot do it justice here (for an overview, see von Stechow and Gillies 2007 and Portner 2009). The main issue I want to point out is that the traditional encoding of epistemic modality in terms of knowledge runs into problems, which can be avoided if we let its semantics be more flexible.

The shortcomings of standard accounts appear in embedded contexts. Consider the following example, where we embed an epistemic under the attitude *believe* (I use *might* to illustrate, as, for idiosyncratic reasons, it can only be epistemic):

(50) John believes it might be raining.

With a traditional account, we obtain the following truth conditions for (50): ‘In all of John’s

doxastic alternatives, there is a world compatible with what John knows in those doxastic alternatives in which it is raining'. Such accounts must rely on the assumption that the believer doesn't use false or unjustified beliefs as a basis for the proposition expressed by the complement of the modal (Stephenson 2007). But this is not a trivial assumption: it gives John great powers of introspection, whereby he can discern the facts he knows from those he merely believes.<sup>20</sup> Intuitively, however, (50) doesn't require of John such a degree of awareness, and seems to make the more modest claim that rain is compatible with what John believes. Note that even if we were to take epistemics to be belief- rather than knowledge-based ('In all of John's doxastic alternatives, there is a world compatible with what John believes in those doxastic alternatives in which it is raining'), we would still need the unsavory assumption that whatever John *believes* to be compatible with his beliefs *is* actually compatible with his beliefs, to capture our intuition that (50) means that rain is compatible with John's beliefs (cf. Yalcin 2007).<sup>21</sup>

This problem of introspection can be avoided if we let the epistemic modal quantify directly over the worlds provided by the embedding attitude, unmediated by a state of knowledge. Doing so would yield the following truth conditions for (50): 'In some world compatible with what John believes, it is raining'. I would thus like to propose that epistemic modality expresses possibilities/necessities given a particular information state, as determined directly by the embedding attitude, rather than a hardwired knowledge/doxastic state. To do so, I recast the epistemic modal base by using our content function CON(*e*). An epistemic modal base has an event pronoun *e* as an argument and provides a set of accessible worlds, namely those compatible with the content of that event. What this content is depends on the event the modal is

---

<sup>20</sup> For a similar argument in matrix contexts, see Tancredi (2007), who argues that epistemics cannot be knowledge-based, but rather must be doxastic-based.

<sup>21</sup> An anonymous reviewer notes that the desired reading can be obtained in a modal logic that derives the equivalence  $\Box_{\text{epist}} \Diamond_{\text{epist}} p \equiv \Diamond_{\text{epist}} p$ .

relative to:  $\text{CON}(e)$  picks out the propositional content of the event  $e$  that the modal is anchored to:<sup>22</sup>

(51) *Recasting the epistemic modal base in event talk* (final)

$$\cap_{\text{epis}}(e) = \{w' : w' \text{ is compatible with } \text{CON}(e)\}$$

This recasting of the epistemic modal base will avoid the problem of introspection and more importantly for our purposes, provide some constraints on the licensing conditions for the modal base: only those events that have associated content (i.e., speech and attitude events) will be able to license an epistemic modal base. If the event argument of the modal base is one without content, such as a running or screaming, the domain of modal quantification will be undefined and the sentence will be meaningless. In the remainder of this section, we look at epistemic modals in matrix contexts (Sect. 6.1.2) and attitude contexts (Sect. 6.1.3), and discuss why a low modal cannot license an epistemic modal base (Sect. 6.1.4).

### 6.1.2. Matrix contexts

Recall our speech event from Section 5.3.2, whose content we took to be the doxastic alternatives of the speaker in declarative sentences:

$$(52) \quad [[\text{ASSERT } e_0]] = \lambda p. \lambda w. \text{Assert}'(e_0, w) \ \& \ \forall w' \in \cap \text{CON}(e_0): p(w')=1$$

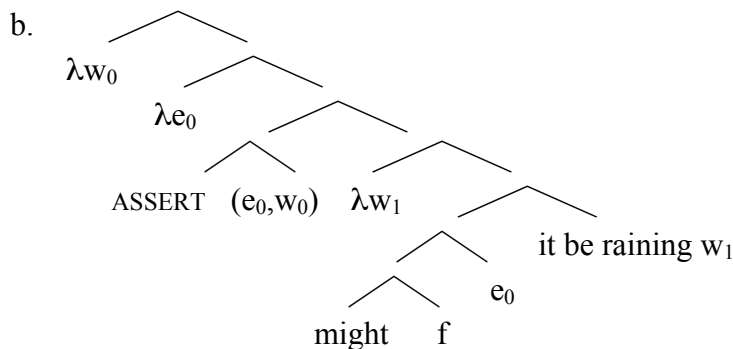
$$\textit{where } \cap \text{CON}(e_0) = \text{DOX}(\iota x \text{ Holder}(x, e_0), w); \iota x \text{ Holder}(x, e_0) = \text{speaker}$$

---

<sup>22</sup> An anonymous reviewer suggests that the attitude verb could contribute the modal base directly, unmediated by an event. Such a proposal is in fact offered by Yalcin (2007) for epistemic modals. However, the event-relative account is more general in that it applies to both roots and epistemic modals and aims at explaining why both types of modals are expressed by the same items crosslinguistically.

These doxastic alternatives serve as the information state a speech event-bound epistemic modal quantifies over. Let's use example (53) to see how:

(53) a. It might be raining.



c.  $\text{Assert}'(e_0, w) \ \& \ \forall w' \in \cap \text{CON}(e_0): \exists w'' \in \cap \text{CON}(e_0): [\text{it is raining in } w'']$

Because both *might* and *ASSERT* in (53) quantify over the same set of worlds (those compatible with the content of  $e_0$ ), the higher modal (*ASSERT*) quantifies vacuously (cf. Hacquard 2006). Expression (53) is equivalent to (54). We obtain that it is compatible with the speaker's beliefs that it is raining:

(54)  $\text{Assert}'(e_0, w) \ \& \ \exists w'' \in \cap \text{CON}(e_0): [\text{it is raining in } w'']$

'In some world compatible with the content of  $e_0$  (the speaker's beliefs) it is raining.'

Note that this raises the question of whether *matrix* epistemics are knowledge- or belief-based. Crucially, this depends on what we take assertions to invoke (cf. footnote 19). An interesting consequence of this proposal is that assertions and matrix epistemics are made of the same cloth:

if assertions are knowledge-, belief-, or commitment-based, so will matrix epistemics.

With a necessity modal, we again obtain a vacuous layer of quantification:

(55) It must be raining.

Assert'(e<sub>0</sub>,w) &  $\forall w' \in \cap \text{CON}(e_0): \forall w'' \in \cap \text{CON}(e_0): [\text{it is raining in } w'']$

'In all worlds compatible with the content of e<sub>0</sub> (the speaker's beliefs) it is raining.'

This raises another potential issue:<sup>23</sup> Shouldn't this LF be equivalent to the unmodalized version *It is raining*, despite our intuitions that the modalized version expresses something weaker than the unmodalized one (Karttunen 1972)? I take this difference in meaning to result from the presence of an *ordering source* in the restriction of *must*. As Kratzer (1991) argues, an ordering source further restricts the set of worlds the modal quantifies over, and may thus exclude the actual world, explaining why the version with *must* feels weaker.

### 6.1.3. Attitude contexts

Recall sentence (50) with an epistemic modal embedded under *believe*, repeated in (56). Our intuition was that the sentence as a whole describes compatibility of rain with John's beliefs:

(56) John believes it might be raining.

Let's compute its truth conditions, given our new semantics for attitudes and epistemic modals (for simplicity, I omit the temporal restriction of aspect and the speech event):

---

<sup>23</sup> Thanks to Paul Portner (p.c.) for pointing this issue out.



precisely what my account does.

#### 6.1.4. Unavailability of the epistemic modal base for low modals

The new formulation of the epistemic modal base makes a prediction: content-less events will not be able to license epistemic modal bases. We have seen that both speech and attitude events have content. Thus, speech event and attitude event-relative modals (which happen to be the high modals) can receive an epistemic modal base. But what about VP event-relative (low) modals?

Usually, the VP complement of a modal describes a non-contentful event. We predict that a modal relative to such an event cannot get an epistemic interpretation. This prediction is borne out, as shown with (59): when the modal is anchored to the VP event, and hence its agent *Mary* and its running time *last night*, (59) cannot mean that ‘it was necessary, last night, given what *Mary* **believed/knew**, that she took the train to go to Paris’.

(59) Last night *Mary* had to take the train to go to Paris.

What if, however, the VP complement of the modal is itself an attitude verb? My theory predicts that an epistemic modal base should be licensed, even when the modal is bound by aspect. In that case, the modal should be interpreted as a possibility/necessity for the subject, at the time provided by tense, given the content of the subject’s attitude. This prediction seems to be borne out, as the following examples from German and French illustrate:

- (60) a. John konnte sehen/merken, dass Mary nett war. (I. Heim, p.c.)  
 b. John a pu voir/remarquer que Mary était gentille.  
 ‘John could see/notice that Mary was nice.’  
 [John perf<sub>2</sub> [ can f e<sub>2</sub> [see(e<sub>2</sub>) [ Mary...]

Since *see* or *notice* is contentful and roughly means ‘come to know’, we expect that, with an LF where the modal is in the low position (which we know from the fact that it scopes below tense and yields actuality entailments in French), the modal, which is relativized to the attitude event, can express an ‘epistemic’ possibility, i.e., a possibility given what John came to know. This is indeed a possible interpretation for (60): ‘Mary being nice was compatible with what John came to know.’<sup>25</sup>

To sum up, an epistemic modal base requires a contentful event as its argument. Given that event arguments need to be bound locally, this explains why speech/attitude-relative (‘high’) modals can be epistemic but VP-event-relative (‘low’) modals usually cannot.

## 6.2. Circumstantial modal base

In this section we turn to the circumstantial modal base. Recall our reformulation, according to which it provides a set of worlds compatible with the circumstances of an event of evaluation:

$$(61) \quad \cap f_{\text{circ}}(e) = \{w' : w' \text{ is compatible with the circumstances of } e\}$$

What are the circumstances of an event? This turns out to be a complex matter, and a

---

<sup>25</sup> Interestingly, with stative attitude verbs like *believe* this type of readings seems unavailable, suggesting that lexical aspectual restrictions may further constrain the range of interpretations. I leave this issue for future research.



comprehensive semantics of the event dependence of this modal base will have to await another occasion. As a first stab, I propose that these circumstances include the immediate material surroundings of the event and its participants at the event's time and location.

Usually, a low modal takes a circumstantial modal base. As we saw in Section 6.1, the epistemic one is unavailable when the VP-event lacks content. (62) thus obligatorily takes a circumstantial modal base (as sketched in the simplified LF below, ignoring the speech event):

(62) (Last night) Mary had to take the train (to go to Paris).

[ T [ Asp<sub>1</sub> [ Mod f e<sub>1</sub> [ VP(e<sub>1</sub>) ] ] ] ]

$\exists e_1[e_1 \text{ in } w \ \& \ \forall w' \text{ compatible with the circumstances of } e_1: [\text{Mary take the train}(e_1, w')]]$

Here the event the modal is relative to is a train-taking by Mary. The circumstances of this event will be the immediate circumstances surrounding Mary, last night, as she is about to take the train: her need to go to Paris, the fact that her car broke down, that there is a train that goes to Paris at a particular time, etc.

One may worry about how such circumstances get picked when the event doesn't occur in reality. What if Mary did not, in fact, take the train last night? What is the event whose circumstances determine the worlds of the modal base? Recall that in a configuration with a low modal like (62), the same event  $e$  occurs both in the world of evaluation ( $w$ ) and in the modal worlds ( $w'$ ): in all of these worlds  $e$  is a train-taking by Mary (cf. Section 3.2). Now,  $e$  may not occur in reality: the value of  $w$  depends on what aspect is used. With perfective,  $w$  is the actual world (or, if we do not ignore the speech event, the doxastic alternatives of the speaker): Mary has to have actually taken the train. With imperfective,  $w$  denotes the worlds provided by the

imperfective itself (e.g., ‘generic’ worlds, which may or may not include the actual world). (62) can be true even if Mary doesn’t *actually* take the train. The event whose circumstances determine the worlds of the modal base is a train-taking by Mary in all generic worlds. Its circumstances are those that surround Mary, last night, as she is about to take the train.<sup>26</sup>

Recall that circumstantial interpretations are available even in sentences with no obvious agentive subject. Under the current proposal, this is expected: what matters are the circumstances of the event participants, not necessarily those of its agent. The circumstances of the event could, for instance, highlight properties of the location.<sup>27</sup>

(63) A lot of people can jump in this pool. (Hackl 1998)

There is no definedness condition for circumstantial modal bases. In principle, we should be able to find them both with low and high modals. Yet, the circumstantial modal base seems to be reserved for low modals. That a low modal should take a circumstantial modal base follows from the fact that an epistemic modal base is simply not available when the VP is contentless. Nothing, however, precludes *high modals* from taking a circumstantial modal base. ‘*Mary may be the murderer*’ should be able to mean: ‘In some world compatible with the circumstances of the speech event, Mary is the murderer.’

Ruling out such interpretations is tricky, and I can only offer rough speculations here.

---

<sup>26</sup> Of course, we now need to make sure that the circumstances in the ‘generic’ worlds somehow line up with the circumstances in the actual world. This preserving of the circumstances is presumably due to the semantics of the generic operator, which generalizes over ideal events that are still anchored in reality. I have in mind a generic (imperfect) operator along the lines of Portner (1998) and Ferreira (2005), which takes a circumstantial modal base.

<sup>27</sup> A circumstantial modal base may require *some* anchoring to a participant: the circumstantial claims with expletive subjects or idioms in (i) are degraded without an explicit location, which might explain why it was first thought in the syntactic literature that root modals are control predicates.

- (i) a. It can rain hard ??(here).
- b. The shit can really hit the fan ??(in this part of the world). (Hacquard 2006)

Unless we want to stipulate a default preference for epistemic interpretations whenever available, we need a better understanding of what the ‘circumstances of the speech event’ are. Based on our previous discussion, it should be the immediate circumstances surrounding the speech event. We have two options. Either this is somehow an incoherent meaning (the circumstances of an event should be somehow material), and we thus never consider a circumstantial modal base in such cases. Alternatively, we take the circumstances of the speech event to be the conversational context in which the speech event takes place: what is common ground, what are the speaker’s beliefs, etc. In this case, the circumstantial modal base may look more and more like the epistemic one, to the point where we may not be able to distinguish their meanings.<sup>28</sup>

### **6.3. Correlating height and modal flavor**

Let’s take stock. To capture their individual/time relativity, we have made our modals relative to an event, by modifying the Kratzerian account so that a modal takes an event variable in its restriction rather than a world. We saw that a modal can appear in one of two positions within a clause: either right above TP or right above VP. In the high position, the modal’s event variable has to be bound by the speech event binder, unless it is embedded under an attitude, in which case it is bound by the matrix aspect quantifying over the embedding attitude event. In the low position, the modal has to be bound by the aspect quantifying over the VP event. Our resulting puzzle was why high (attitude or speech event-relative) modals receive epistemic interpretations, and low (VP event-relative) modals circumstantial ones. I proposed to recast the epistemic modal base in terms of *content*: an epistemic modal base needs to be bound by a contentful event, which both attitude and speech events are, but regular VP-events aren’t. This explained why high

---

<sup>28</sup> An anonymous reviewer suggests that the circumstantial and epistemic modal bases be collapsed into one, relying on the event to distinguish the nature of the modal base. I leave the exploration of this option for future research.

modals can be epistemic, but low modals cannot (unless the VP event is itself an attitude). A circumstantial modal base (which picks out worlds compatible with the circumstances of the event binding it), on the other hand, doesn't have any restriction. It is, therefore, the modal base usually associated with low modals. We thus obtain that high modals tend to be epistemic and low modals circumstantial, without having to stipulate two separate entries for each modal.

## **7. Conclusion**

We have seen that the same words can express epistemic and root modality. The challenge was to give these modal words a unified semantics that could still explain why they interact differently with tense and aspect depending on their interpretation. I proposed that modals are always relative to an event, rather than a world of evaluation. By relativizing modals to an event, we were able to derive otherwise unexplained restrictions on a modal's interpretation (i.e., time/individual constraints). By looking further at the type of event binders available at various syntactic positions, we could make sense of the fact that only certain modal bases are available at certain positions. This allowed us to remove some of the arbitrariness behind Cinque's hierarchy: modals scoping above tense and aspect are epistemic because modals in a high position are relative to contentful events—that is, events that invoke an information state, the very stuff epistemic modality is made of. Modals scoping under tense and aspect are root modals because low modals are relative to contentless events, which cannot license an epistemic modal base. Abstracting away from the particular implementation I have pursued here, the general contention of this paper is that modals must be relative to something more fine-grained than a world, or a world-time pair, and that seemingly arbitrary constraints on the meaning of modals may be due to general syntactic principles germane to locality and binding phenomena.

## References

- Abraham, W.: 2001, Modals: Toward explaining the ‘epistemic non-finiteness gap’, in R. Mueller and M. Reis (eds.) *Modalitat und Modalverben im Deutschen*, 7-36. Hamburg: Helmut Buske Verlag.
- Abusch, D.: 1994. Sequence of tense revisited: Two semantic accounts of tense in intensional contexts, in H. Kamp (ed.) *Ellipsis, tense and questions*, 87-139. DYANA deliverable R2.2.B, University of Amsterdam.
- Abusch, D.: 1997, Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20(1):1-50.
- Alonso-Ovalle, L. and P. Menendez-Benito: 2003, Some epistemic indefinites, in M. Kadowaki and S. Kawahara (eds.) *Proceedings of NELS 33*, 1-12. Amherst, MA: GLSA.
- Alonso-Ovalle, L. and P. Menendez-Benito: forthcoming, Modal indefinites, *Natural Language Semantics*.
- Ambar, M.: 1999, Aspects of focus in Portuguese, in L. Tuller and G. Rebuschi (eds.) *The grammar of focus*, 23-53. Amsterdam and Philadelphia: John Benjamins.
- Anand, P. and V. Hacquard: 2009, Epistemics with attitudes, in *Proceedings of SALT 18*. Ithaca, N.Y.: CLC Publications. <http://hdl.handle.net/1813/13025>.
- Bhatt, R.: 1998, Obligation and possession, in H. Harley (ed.) *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, MITWPL 32, 21-40. Cambridge, Mass.: MIT.
- Bhatt, R.: 1999, *Covert modality in non-finite contexts*. PhD dissertation, University of Pennsylvania.
- Brennan, V.: 1993, *Root and epistemic modal auxiliary verbs*, PhD dissertation, University of Massachusetts at Amherst.

- Boogart, R.: 2007, The past and perfect of epistemic modals, in L. de Saussure, J. Moeschler, G. Puskas (eds.) *Recent advances in the syntax and semantics of tense, aspect and modality*, 47-69. Berlin and New York: Mouton de Gruyter.
- Butler, J.: 2003, A minimalist treatment of modality. *Lingua* 113: 967-996.
- Bybee, J.: 1995, The semantic development of past tense modals in English, in J. Bybee and S. Fleischman (eds.) *Modality and grammar in discourse*, 503-517. Amsterdam and Philadelphia: John Benjamins.
- Bybee, J., R. Perkins, and W. Pagliuca: 1994, *The evolution of grammar: tense, aspect and modality in the languages of the world*. Chicago: The University of Chicago Press.
- Carnap, R.: 1957, *Meaning and necessity*. Chicago: The University of Chicago Press.
- Cinque, G.: 1999, *Adverbs and functional heads: a crosslinguistic perspective*, Oxford Studies in Comparative Syntax. Oxford and New York: Oxford University Press.
- Coates, J.: 1983, *The semantics of the modal auxiliaries*. London: Croom Helm.
- Condoravdi, C.: 2002, Temporal interpretations of modals, in D. Beaver, S. Kaufman, B. Clark (eds.) *Stanford Papers in Semantics*, 59-88. Stanford, Cal.: CSLI Publications.
- Cresswell, M.: 1990, *Entities and indices*. Dordrecht: Kluwer.
- Davidson, D.: 1967, The logical form of action sentences, in N. Rescher (ed.) *The logic of decision and action*, 81-120. Pittsburgh: University of Pittsburgh Press,.
- DeRose, K.: 1991, Epistemic possibilities. *The Philosophical Review*, Vol. C, No. 4.
- Drubig, H. B.: 2001, On the syntactic form of epistemic modality, ms., University of Tübingen.
- Egan, A., J. Hawthorne, and B. Weatherson: 2004, Epistemic modals in context, in G. Preyer and G. Peter (eds.) *Contextualism in philosophy*, 131-169. Oxford: Oxford University Press.
- Farkas, D.: 1997, Evaluation indices and scope, in A. Szabolcsi (ed.) *Ways of scope taking*, 183-

215. Dordrecht: Kluwer.
- Ferreira, M.: 2005, *Event quantification and plurality*. PhD dissertation, MIT.
- von Fintel, K.: 2001, *Advanced semantics seminar*, lecture notes, Fall 2001, MIT.
- von Fintel, K. and A. Gillies: 2007, An opinionated guide to epistemic modality. In T. Gendler and J. Hawthorne (eds.), *Oxford studies in epistemology*, Vol. 2, 32-62. Oxford: Oxford University Press.
- von Fintel, K. and A. Gillies: 2008, CIA leaks, in *Philosophical Review* 117(1):77–98.
- von Fintel, K. and I. Heim: 2001, Notes on intensional semantics, ms., MIT.
- von Fintel, K. and S. Iatridou: 2003, Epistemic containment. *Linguistic Inquiry* 34:173-198.
- Fleischman, S. 1982. *The future in thought and language: diachronic evidence from Romance*. Cambridge: Cambridge University Press.
- Ginzburg J. and I. Sag: 2001. *Interrogative investigations: the form, meaning, and use of English interrogatives*. Stanford, Cal.: CSLI Publications.
- Groenendijk, J. and M. Stokhof: 1975, Modality and conversational information. *Theoretical Linguistics* 2:61-112.
- Gunlogson, C.: 2001, *True to form: rising and falling declaratives as questions in English*. PhD dissertation, University of California at Santa Cruz.
- Hackl, M.: 1998, On the semantics of ability ascriptions, ms., MIT.
- Hacquard, V.: 2006, *Aspects of Modality*. PhD dissertation, MIT.
- Hacquard, V.: 2008, Speaker-oriented vs. subject-oriented modality: a split in implicative behavior. *Proceedings of Sinn und Bedeutung 11*, 305-319. Barcelona: Universitat Pompeu Fabra.
- Hacquard, V.: 2009, On the interaction of aspect and modal auxiliaries. *Linguistics and*

- Philosophy* 32: 279-315.
- Hacquard, V.: to appear, Modality, in C. Maienborn, K. Heusiger and P. Portner (eds.), *Semantics: an international handbook of natural language meaning*. Berlin: Mouton de Gruyter.
- Heim, I.: 1982, *The semantics of definite and indefinite noun phrases*. PhD dissertation, University of Massachusetts at Amherst.
- Heim, I.: 1994. Comments on Abusch's theory of tense, in H. Kamp (ed.) *Ellipsis, tense and questions*, 143-170. University of Amsterdam.
- Hintikka, J.: 1962, *Knowledge and belief*. Ithaca, N.Y.: Cornell University Press.
- Iatridou, S.: 1990, The past, the possible and the evident, *Linguistic Inquiry* 21(1):123-129.
- Ippolito, M.: 2002, *The time of possibilities*, PhD dissertation, MIT.
- Jackendoff, R.: 1972, *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
- Karttunen, L.: 1972: *Possible and must*. In J. Kimball (ed.) *Syntax and Semantics Vol. 1*, 1-20. New York: Academic Press.
- Klein, W.: 1994, *Time in language*. London/New York: Routledge.
- Kratzer, A.: 1977, What *must* and *can* must and can mean. *Linguistics and Philosophy* 1:337-355.
- Kratzer, A.: 1981, The notional category of modality, in H.-J. Eikmeyer and H. Rieser (eds.), *Words, worlds, and contexts. new approaches in word semantics*. Berlin: de Gruyter.
- Kratzer, A.: 1991, Modality, in A. von Stechow and D. Wunderlich (eds.) *Semantik: Ein internationales Handbuch zeitgenoessischer Forschung*, 639-650. Berlin: de Gruyter.
- Kratzer, A.: 1998, More structural analogies between pronouns and tenses, in D. Strolovich and



- A. Lawson (eds.) *Proceedings of SALT 8*. Ithaca, N.Y.: CLC Publications, Cornell University.
- Kratzer, A.: 2006, Decomposing attitude verbs. Talk in honor of A. Mittwoch. July 4, 2006, The Hebrew University, Jerusalem.
- Krifka, M.: 2001, Quantifying into question acts. *Natural Language Semantics* 9(1):1-40.
- Kripke, S.: 1963, Semantical analysis of modal logic I, Normal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9: 67–96.
- Lee, F.: 2008, Deriving subject scope in modal constructions: the case of modal *have*, ms., UCLA.
- Lewis, D.: 1968, *On the plurality of worlds*. Oxford: Basil Blackwell.
- MacFarlane, J.: 2003, Epistemic modalities and relative truth, ms., University of California at Berkeley.
- Moltmann, F.: 2003, Propositional attitudes without propositions. *Synthese* 135:77-118.
- Palmer, F. R.: 2001, *Mood and modality*. Cambridge: Cambridge University Press.
- Papafragou, A.: 1998, The acquisition of modality: implications for theories of semantic representations. *Mind and Language* 13:370-399.
- Partee, B.: 1973, Some structural analogies between tenses and pronouns. *The Journal of Philosophy* 70: 601-609.
- Percus, O.: 2000, Constraints on some other variables in syntax. *Natural Language Semantics* 8, 173-229.
- Perlmutter, D.: 1971, *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.
- Picallo, M.: 1990, Modal verbs in Catalan. *Natural Language and Linguistic Theory* 8: 285-312.

- Portner, P.: 1998, The progressive in modal semantics. *Language* 74(4): 760-787.
- Portner, P.: 2009, *Modality*. Oxford: Oxford University Press.
- Rizzi, L.: 1997, The fine structure of the left periphery, in L. Haegeman (ed.) *Elements of grammar*, 281-337. Dordrecht: Kluwer.
- Ross, J. R.: 1969, Auxiliaries as main verbs. In W. Todd (ed.), *Studies in philosophical linguistics* (Series 1), 77-102. Evanston, Ill.: Great Expectations Press.
- Ross, J. R.: 1970, On declarative sentences, in R. A. Jacobs and P. S. Rosenbaum (eds.) *Readings in English transformational grammar*, 222-272. Waltham, Mass.: Ginn.
- Shimada, J.: 2008, Head movement, binding theory and phrase structure, ms., MIT.
- Speas, P.: 2004, Person (and mood and tense) and indexicality. Paper presented at the Harvard Workshop on Indexicals, Speech Acts, and Logophors, November 20, 2004, Cambridge, Mass.
- Stephenson, T.: 2007, *Toward a theory of subjective meaning*. PhD dissertation, MIT.
- Stowell, T.: 2004, Tense and modals, in J. Guéron and J. Lecarme (eds.) *The syntax of time*, 495-537. Cambridge, Mass.: MIT Press.
- Sweetser, E.: 1982, Root and epistemic modals: causality in two worlds. *Proceedings of BLS* 8, 484-507. Berkeley: Berkeley Linguistics Society.
- Tancredi, C.: 2007, A multi-modal theory of I-semantics, ms., University of Tokyo.
- Tenny, C. and P. Speas: 2004, Configurational properties of point of view roles, in A. M. Di Scullio (ed.) *Proceedings of UQAM Asymmetry Workshop*.
- Thomason, R. 1984: Combinations of tense and modality, in D. Gabbay and F. Guenther (eds.) *Handbook of Philosophical Logic, Vol. 2*, 135-165. Dordrecht: Reidel.
- Traugott, E.: 1988, Pragmatic strengthening and grammaticalization. In S. Axmaker, A. Jaisser

- and H. Singmaster (eds.), *Proceedings of the 14<sup>th</sup> Annual Meeting of the Berkeley Linguistics Society*, 406-416. Berkeley: Berkeley Linguistics Society.
- Westmoreland, R.: 1998, *Information and intonation in natural language modality*. PhD dissertation, Indiana University.
- Wurmbrand, S.: 1999, Modal verbs must be raising verbs, in S. Bird, A. Carnie, J. Haugen and P. Norquest (eds.) *Proceedings of WCCFL 18*, 599-612. Somerville, Mass.: Cascadilla Press.
- Yalcin, S.: 2007, Epistemic modals. *Mind* 116: 983–1026.
- Zagona, K.: 2007, On the syntactic features of epistemic and root modals, in L. Eguren, and O. Fernandez Soriano (eds.) *Coreference, modality and focus: studies on the syntax/ semantics interface*, 221-236. Amsterdam: Benjamins.
- Zubizarreta, M. L.: 1982, *On the relationship of the lexicon to syntax*. PhD dissertation, MIT.