

On the Event Relativity of Modal Auxiliaries¹

Abstract

Cross-linguistically, the same modal words can be used to express a wide range of interpretations. This cross-linguistic trend supports a Kratzerian analysis (Kratzer 1981, 1991), 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 cross-linguistic 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 hard-wire different syntactic positions for epistemics and roots (Cinque 1999). 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 fact 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 to stipulate an interpretation-specific height for modals.

1. Introduction

It is a rather robust cross-linguistic generalization that the same modal words can be used to express various kinds of possibilities or necessities (cf. Bybee et al 1994, Palmer 2001). Take the

¹ Acknowledgements to be provided later.

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), while the second expresses 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 *mays* 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*, or Italian *potere*, can express an ability, a permission or an epistemic possibility, among other variants. Similarly the Malay modal *mesti* (Drubig 2001) or the Tamil permission and debitive suffixes can receive both epistemic or deontic interpretations (Palmer 2001). This cross-linguistic pattern argues against a lexical ambiguity account: It seems highly improbable that the same lexical accident should be found in language after language. Rather, 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 contextual dependency of a modal’s interpretation arises from a ‘conversational background’, which determines the set of worlds the modals quantify over. Depending on the context, the same modal will quantify over worlds compatible with certain laws—and receive a deontic interpretation, or 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, accounts nicely for the cross-linguistic 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. Modals with an epistemic interpretation scope higher than tense and aspect (and tend to do so with negation and quantifiers), while modals with root interpretations scope lower than these elements (cf. Brennan 1993, Drubig 2001, Stowell 2004, a.o.). To illustrate, consider the following examples from Italian, which show the interaction of modals with tense in a morphologically transparent way. In (a), tense and aspect morphology appears on the embedded verb (hence, below the modal), while in (b), it appears on the modal itself (hence above it):

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

The same modal **potere** (*can*) can receive an epistemic or a root interpretation. However, when it appears above tense/aspect, only an epistemic interpretation is available: (a) expresses an epistemic possibility at the speech time of a past event of talking to Maria. When the modal is below tense/aspect, it can only receive a root interpretation: (b) expresses a past permission (or ability) of a concurrent event of talking to Maria. Languages do not always show this interaction

overtly, and may allow past morphology on the modal, regardless of its interpretation (e.g., French *pouvoir*). Yet, the generalization still holds at the interpretation level: with an epistemic interpretation, the modal seems to scope over tense (its evaluation time is the speech time), and with a root interpretation, it scopes below (its evaluation time is past).

Modals with epistemic interpretations thus seem to be interpreted higher than modals with non epistemic interpretations, a fact which a purely contextual account in the Kratzerian tradition cannot explain. In particular, the ordering of tense, aspect, and root and epistemic modals seems to be fixed cross-linguistically (Cinque 1999). 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 modals 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 systematically differ 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 this puzzle has been to essentially reject a unified account: Drubig (2001) and Westmoreland (1998) for instance propose that epistemics are evidentials, rather than modals. Others bite the Cinque bullet and stipulate that modals come in two (UG supplied) types: epistemics, which take IP complements, and root modals, which take VP complements (cf. Jackendoff 1972, Zubizarreta 1982, Picallo 1990, Butler, 2003). This kind of solution, which relies on interpretation-specific separate entries, ultimately leads to a reformulation of the problem, and do not explain *why*, cross-linguistically, epistemics and roots take different sorts of complements even when they are expressed by the same lexical items.

With this paper, I would like to show that Cinque's puzzle is in fact just one side of the coin. Modals are relative not only to a (interpretation-specific) time of evaluation (the speech time for epistemics, the time provided by tense for roots), but to an individual as well: according to Bybee 1995 and Palmer 2001 (among others) epistemics are speaker-oriented, while roots are subject-oriented. It thus appears that, generally, modals are anchored both to an individual and a time. Crucially, not all time/individual pairings are attested. To illustrate this point in English, I use the semi-modal *have to*, which, unlike modal auxiliaries *may* or *must*, can be fully declined, and hence show the full range of possible interpretations. In (2), *have to*'s most natural interpretation is epistemic: it expresses a necessity, given what *the speaker* knows, that John was the murderer. And, as we saw, a modal with an epistemic interpretation is evaluated at 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 the murderer.*

(2) John had to be the murderer.

When we embed an epistemic modal under an attitude verb, as in (3), 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: (3) expresses a necessity given *Mary's* evidence *at her thinking time.*

(3) Mary thought that John had to be the murderer.

Finally, with a root interpretation, as in (4), the modal expresses a necessity for the subject², to take the train, given certain circumstances of the base world, namely *his* circumstances. And crucially, as we saw, the time of evaluation of a root modal has to be the time provided by tense: we are talking about circumstances of *John at the time provided by tense* (a past time).

(4) John 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). Finally, 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: 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 the murderer*. Why, then, isn't such an interpretation available in the absence of an overt restriction? 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 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), attitude events (which anchor 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).

² This is a simplification: root modality doesn't always co-occur with an individual in subject position. We will see that in these cases, the circumstances are tied to a participant of the VP event. The correct generalization will be that when the modality is tied to a VP participant, it is keyed to the time of the VP event (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, and let this modal freely appear above or below tense. However, the particular structural position in which the modal appears will itself restrict the interpretation the modal can receive. 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 (for type reasons): right above VP (the ‘low’ position (5)a)) or, right above tense (the ‘high’ position (5)b-c)). In the low position ((5)a), the closest event binder is aspect: the modal’s event variable gets bound by *aspect*, thereby relativizing the modality to the event aspect quantifies over (i.e., the VP event). This anchors the modal to the VP event participants (e.g., the subject), at the running time of the event (the time provided by tense). In the high position, the modal is anchored to the speech event in matrix contexts, as in ((5)b): the modal’s event variable gets bound by a default, topmost, speech event binder, in the spirit of Percus 2000. This forces the modal to be relative to the speaker, at the speech time. In embedded contexts, as in ((5)c), it is anchored to the attitude event: the modal’s event variable gets bound by the *matrix* aspect, which quantifies over the *attitude* event. This forces the modal to be relative to the attitude holder at the attitude time. We will thus derive the correct time/individual constraints:

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

We will then be in a position to reformulate Cinque’s puzzle as follows: why is it that attitude or speech event-relative (i.e., *high*) modals get an epistemic interpretation, while VP-relative (i.e., *low*) modals get a root one? I will sketch a solution to this problem that 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 licensing a modal with an epistemic interpretation because it provides the very information state epistemic modals quantify over. Epistemics, I propose, express compatibility with an information state. This is often a state of knowledge (hence the name *epistemic*), but not always: a closer look at ‘epistemics’ in embedded contexts will show that it is possible (and perhaps 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, can only receive a root interpretation.

This paper is organized as follows: in section 2, I review Kratzer’s unified account for epistemic and root modals. Section 3 goes over Cinque’s puzzle and shows how epistemics and roots correlate with two different syntactic positions. Section 4 discusses the empirical observation that modals are sensitive to time/individual constraints, which, I argue, reflects the

event-relativity of modals. Section 5 proposes an event-relative implementation for modals. Section 6 revisits Cinque's puzzle and sketches a proposal of how to derive the association of the epistemic and the root interpretation with the high and the low position, respectively.

2. A unified account for epistemic and root modals: Kratzer (1981, 1991)

Modals come in various flavors. The French possibility and necessity modals *pouvoir* and *devoir* will help us illustrate this variety, as they allow the full range of meanings available to modals. The literature highlights two main classes of modals: *epistemics* and *roots*.³ The following example provides a context that supports an *epistemic* interpretation. As their name indicates, epistemics are usually taken to invoke knowledge (from the Greek *episteme*: knowledge). As a first approximation, we take these modals to express the possibility/necessity, given what is known (in particular that it is 6pm, and that Mary is not in the office), that Mary is home.

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

It's 6:00pm. Mary is not in the office. She may/must be home.

The term *root* applies to modals with non epistemic interpretations. (7)a) illustrates *deontic* modality: the modals express a permission or an obligation given Mary's father's orders. *Pouvoir* can also express an *ability*, that of lifting a table in (7)b). (7)c) illustrates *goal-oriented* (or *teleological*) modality (cf. von Stechow and Iatridou 2003), which expresses possibilities and necessities given a particular goal of the subject, and (7)d) a purely circumstantial reading: here, *devoir* expresses a necessity given certain relevant facts or circumstances:

³ 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 in this paper and refer the interested reader to Heim (1982), Brennan (1993), Portner (2008).

- (7) a. *Le père de Mary lui impose un régime très strict. Elle peut/doit manger du brocoli.*
Mary's father imposes on her a strict diet. 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 doit être à Paris à 17 heures. Elle peut/doit prendre le train pour aller à Paris.*
Mary must be in Paris at 5pm. She can/must take the train to go to Paris.
- d. *Le nez de Mary la chatouille. Elle doit éternuer.*
Mary's nose is tickling her. She must sneeze.

The class of root modals is not simply defined in terms of complementarity to the class of epistemics. They, in fact, share an important meaning component: all 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, the state of her nose). Epistemic and root modals thus seem to differ in the kind of information involved when doing modal reasoning. According to Kratzer (1981, 1991), this is the sole difference between the two classes. Modals are treated as quantifiers over possible worlds (cf. Carnap 1957, Hintikka 1962, Kripke 1963, Lewis 1968, a.o.), restricted by an accessibility relation, which determines the set of worlds the modal quantifies over. Kratzer's crucial insight is to have the *context* determine the accessibility relation, via what she calls *conversational backgrounds*. Such conversational backgrounds can be brought about overtly by phrases such as *in view of what is known*, or covertly by the context. Thus, 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. This is highly welcomed, given the fact that (i) each modal comes in a variety of flavors (epistemic, deontic, teleological, etc.), and that (ii) each flavor itself comes in a variety of subflavors (Kratzer 1977): for instance, deontic *must*, as in ‘*Mary must pay a fine*’ can be interpreted as an obligation in view of various kinds of laws (the laws of Cambridge, the regulations of the IRS, etc.). Kratzer proposes two kinds of conversational backgrounds: (i) the *modal base* which picks a set of accessible worlds and differentiates between epistemics and roots; (ii) the *ordering source*, which picks out a subset of the worlds of the modal base (the best ones given a particular ordering), and further differentiates various kinds of epistemics and roots. In the following I depart from Kratzer’s formalization, to keep in 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; the *circumstantial* modal base, which picks out a set of worlds compatible with certain circumstances of the base world, and is at the source of all root interpretations. Formally, modal bases are functions from worlds to sets of propositions. For short, I use the function $f(w)$ to denote the *intersection* of the set of propositions accessible from w ($f(w)$ picks out the set of worlds in which all the propositions of the modal base hold). f takes a world argument and returns a set of worlds: worlds compatible with what is known in w for (8) and worlds compatible with the circumstances in w for (9):

(8) $f_{\text{epis}}(w) = \lambda w'. w'$ is compatible with what is known in w .

(9) $f_{\text{circ}}(w) = \lambda w'. w'$ is compatible with certain facts/circumstances in w .

In Kratzer's system, the main difference between epistemics and roots comes from the modal base. Epistemic interpretations arise from an epistemic modal base, roots from a circumstantial one. Recall the examples in (7): they all express possibilities or necessities given certain circumstances of the subject in the base world: Mary's diet (a), her physical condition (b), her schedule and goals (c), the state of her nose and the environment that surrounds her (d). Further meaning differences between these four are captured by Kratzer's second conversational background, the *ordering source*, which orders the set of worlds provided by the modal base according to a particular (deontic, bouletic...) ideal. The modal ends up quantifying over the most (legally, bouletically...) ideal worlds of the modal base. Formally, the ordering source is also a function from worlds to sets of propositions (sets of laws for deontics, sets of desires for bouletics...). The most ideal worlds of the modal base are those in which the greatest number of propositions of the ordering source holds.

A modal is thus a quantifier over possible worlds. It is first restricted by a modal base $f(w)$ (*circumstantial* or *epistemic*), which returns a set of accessible worlds that can then be ordered by an ordering source $g(w)$, yielding a subset of these worlds, namely those are most ideal given the ordering source. Not all ordering sources are compatible with all modal bases: a circumstantial modal base is compatible with deontic (laws), bouletic (wishes), teleological (aims) ordering sources, while an epistemic modal base typically combines with a stereotypical ordering source. Both $f(w)$ and $g(w)$ are contextually determined (when not overt). This allows for a single entry for *must* and for *can*, which only differ in force of quantification:

(10) a. For any world w , conversational backgrounds f , g , and proposition q :

$$[[\text{must}]](f)(g)(q) = 1 \text{ iff } \forall w' \in \max_{g(w)} (f(w)): q(w') = 1.$$

- b. For any world w , conversational backgrounds f , g , and proposition q :

$$[[\text{can}]](f)(g)(q) = 1 \text{ iff } \exists w' \in \max_{g(w)}(f(w)): q(w') = 1.$$

where $\max_{g(w)}$ selects the most ideal worlds given the ordering imposed by $g(w)$.

A crucial advantage of a Kratzerian approach is that it provides a unified treatment of modals: there is just one *can* and one *must*⁴. This accounts nicely for the cross-linguistic 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.

In the next section, we look at structural differences that go beyond what a difference in modal bases can explain. From this point on, I will ignore ordering sources, which I take to work exactly as 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

We now turn to the evidence for a correlation between modal flavor and height of interpretation, which argues against a unified account. We will see that epistemics scope over tense (section 3.1) and aspect (section 3.2), while roots scope below these elements, and review proposals for separate entries for roots and epistemics (section 3.3).

⁴ Idiosyncracies of particular modals (e.g., *might*) are due to selectional properties hard-wired in their lexical entries.

3.1. Interaction with tense

In this section, I show that modals with a *root* interpretation are always interpreted in the scope of tense, while modals with an *epistemic* interpretation take scope over 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 this time is seems to depend on the particular interpretation of the modal: the speech time for epistemics, the time provided by tense for roots. The following examples, using English *have to*, illustrate (cf. also the Italian examples in the introduction). With a *root* (goal-oriented) interpretation, the circumstances and goal of the subject are evaluated at the time provided by tense (past). (11) expresses a necessity, given Mary's circumstances *then*, to take the train *then*. It cannot express a necessity given her circumstances *now* to have taken the train *then*.

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

The evaluation time of an *epistemic* modal, on the other hand, is always set to the speech time in matrix contexts, or, when embedded, to the internal now of the embedding verb. In other words, the evaluation time of an epistemic modal can never be back (or future) shifted. This is captured formally by not allowing epistemics to scope below tense (cf. Iatridou 1990, Abusch 1997, Stowell 2004), for instance, by hard-wiring the position for epistemic modals above the tense projection (Cinque 1999, Abraham 2001). With an epistemic interpretation, (12) expresses the necessity given what is known *now* that Mary was the murderer at some past (murder) time:

(12) Mary had to be the murderer. $\checkmark \text{mod}_{\text{epis}} > \text{past}$, $*\text{past} > \text{mod}_{\text{epis}}$

To see this, imagine that the evidence gathered at the beginning of the investigation (a week ago) all pointed to Mary being the murderer: she had no alibi, but many a motive. Yesterday, however, Poirot established that the murder had been committed one hour earlier than originally thought. This fact immediately clears Mary, who was seen by several eyewitnesses elsewhere at that time. In this scenario, (12) is false: it is not possible to have (12) refer to the time when the evidence pointed to Mary. One can force an epistemic interpretation evaluated at a past time with an *overt* conversational background: *given what we knew **then**, Mary had to be the murderer*. However, in the absence of such overt clauses, this past shifting is crucially not available.

That modals interpreted in the past cannot receive an epistemic interpretation is further corroborated by Condoravdi's (2002) examples below. When *might* expresses a past possibility (facilitated by the adverb *still* in (b)), it cannot get an epistemic interpretation, but only a metaphysical/counterfactual reading: there was a possibility, at some past time, that they would win the game (with the further inference that they in fact didn't). In (a), the modal expresses an epistemic possibility *at the speech time* about a past time (facilitated by *already*): it is possible, as far as the speaker knows (*right now*), that (*at some past time*) they won the game⁵:

- (13) a. They might (already) have won the game.
b. They might (still) have won the game.

⁵One may wonder about the nature of the counterfactual reading. Is it triggered by a third type of modal base (metaphysical, according to Condoravdi) and, if so what are its licensing conditions? One alternative is that it is a kind of circumstantial modality (cf. Tancredi 2007, Abusch 2007). Another is that the counterfactual reading reflects the presence of an actual counterfactual modal (Stalnaker 1968) scoping under tense, itself scoping under an epistemic *might*. For discussion, see Hacquard (2006).

Note that it is possible to report a past epistemic state. However, this requires either embedding the modal under an attitude verb (e.g., *two days ago, Poirot thought that Mary had to be the murderer*), or an indirect discourse past tense (cf. Boogart 2007): *this didn't make sense, thought Poirot.... Mary had to be the murderer*. In these cases, while past morphology may appear on the modal, this past morpheme results from a kind of morphological agreement with a higher tense (sequence of tense), and lacks the characteristic backshifting of a semantic past tense. Epistemics, then, cannot take scope below a semantic tense within their clause.

Before moving on, I want to mention a counterexample to this generalization put forth by von Stechow and Gillies (2007) who argue that, in the following exchange, B's utterance can, in fact, express a *past epistemic possibility* (not just a counterfactual possibility)⁶:

(14) 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)

It should be noted that this reading only arises in very limited contexts, namely in answers to *why* questions. B's utterance is a fragment of the full answer '*I looked in the drawer because my keys might have been there*'. What seems to be responsible for this reading is the *because* context, as argued by Stephenson (2007): *because* itself induces a shifting of the evaluation parameters. Thus the sentence '*John looked in the drawer because his keys might be there*' reports a possibility for *John at his search time* (not for the speaker at the speech time).

To sum up, the same modals interact differently with tense depending on their interpretation. This fact cannot be explained in the standard Kratzerian account: there is no reason why modal bases should force tense to be interpreted above or below the modal⁷.

⁶For further skepticism that this example involves a past tense scoping over the modal, see Portner (2008).

3.2. Interaction with aspect

The interaction of modals with aspect seems to corroborate these 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 event described by the VP) to a time of reference (provided by tense and time adverbials). We focus on perfective aspect here⁸. 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:

(15) Hier matin, Mary a lu un livre.

Yesterday morning, Mary read-past-pfv a book.

Yesterday morning, Mary read a book.

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

Bhatt (1999) shows that in languages that have a morphologically overt aspectual distinction (such as French), 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. Hence the continuation in (16) comes out as a contradiction.

⁷Pinon (2001) proposes that two classes of modals can be distinguished on the basis of their interaction with tense, where those that scope below tense receive a historical possibility/necessity interpretation. It isn't entirely clear, however, that historical necessity can generalize to all kinds of root interpretations.

⁸ The implicative effect discussed in this section only happens with perfective. Imperfective comes with its own layer of modality responsible for removing the effect (cf. Bhatt 1999, Hacquard 2006, 2008b).

(16) Mary a pu soulever cette table, #mais elle ne l'a pas soulevée.

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

'*Mary was able to lift this table*', #but she didn't lift it.'

This effect extends to all root modals. Crucially however, the same modals with an *epistemic* interpretation do not yield actuality entailments (Hacquard 2006, 2008a). In the example below, aspect does not affect the non-implicative behavior of a modal with an *epistemic* interpretation. Unlike Italian, French allows tense and aspect *morphology* to appear on a modal, even when it receives an epistemic interpretation, and despite the fact that in this case, tense and aspect are still obligatorily interpreted in the complement. Thus, the sentence in (17) is truly ambiguous between an interpretation where the modal is epistemic (*John may have taken the train*) and one where it is circumstantial (*John was able to take the train*). Importantly, while a circumstantial interpretation 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 a pu 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, 2008b) argues that actuality entailments result from having aspect take scope over the modal, which happens with a root, but crucially not with an epistemic modal. When aspect takes scope *under* a modal (as happens with epistemics), its world of evaluation is the one

provided by the modal: the event only has to occur in 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 a pu courir. epistemic>aspect
Mary may have run.
- b. [ModP can [TP past [AspP perf [VP Mary run]]]]
- 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 running event by Mary.’
- (19) a. Mary a pu courir. aspect>root
Mary was able to run
- b. [TP past [AspP perf [ModP can [VP Mary run]]]]
- c. ‘There is a past event in the **actual world**, and there is a world compatible with the circumstances in the actual world where that event is a run by Mary.’

To sum up, epistemics and roots differ in the way they interact with aspect: roots yield actuality entailments with perfective, epistemics do not. What seems to shield epistemics from this effect is the fact that, contra 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 different types of contextually-provided restrictions. We saw that 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 have to be interpreted higher, and roots lower, than tense/aspect. This falls out naturally if we have two positions for modals: one above tense, and one below aspect (above VP). This relative positioning of modals with respect to tense and aspect is corroborated by Cinque (1999), who argues for the following universal hierarchy, based on a cross-linguistic survey involving various functional heads and adverbs:

(20) Cinque’s Hierarchy of Functional Projections (irrelevant projections omitted):

... MOD_{EPIS} > T > ASP > MOD_{ROOT} > VP

Various proposals derive this relative positioning by having modals come in two types: S-level for epistemics, VP-level for roots. Let’s now briefly turn to the interaction of modals with subjects (for more details, see Brennan 1993, von Stechow and Iatridou 2003, Lee 2008), as it has been used in support for such proposals⁹. Roots, unlike epistemics, seem to enter in a thematic relation with the subject, as evidenced by their interaction with expletive subjects or idioms, suggesting a difference in argument structure between roots and epistemics. As idiom chunks lose their idiomatic meaning in control constructions (*e.g.*, *#the shit wants to hit the fan*), the example in (21) suggests that root ‘*can*’ takes an individual and a property as arguments (VP-level modal), while epistemic ‘*might*’ takes a proposition (S-level modal) (Brennan 1993):

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

⁹ Epistemics also tend to be interpreted above negation and roots below it (cf. Coates 1983, Drubig 2001), but there are many counterexamples (cf. Cormack and Smith 2002, Palmer 2001), which we cannot address here.

However, while there does seem to be a connection between the subject and a root modal, Bhatt (1998), Hackl (1998) and Wurmbrand (1999) have shown that this connection cannot be due to a control configuration: roots should be treated as raising predicates, just like their epistemic counterparts, as they do seem to allow expletive (a) and weather *it* (b) subjects, and do not always express a property of their *subject* (capacity of *the pool* rather than *people* in (c)):

- (22) a. There have to be fifty chairs in this room. [Bhatt 1998]
b. It can rain hard here. [Hackl 1998]
c. A lot of people can jump in this pool. [Hackl 1998]

I take these examples to show that root modality is centered around the event described by the main predicate and its participants, but not necessarily its subject. In most cases, the main participant is indeed the subject, and hence properties of the subject are highlighted. In other cases, however, the location or properties of other participants of the event are more relevant (*here* or *the pool*). The fact that modal statements involving idiom chunks improve when a location is added corroborates the intuition that the relevant factor is not argument structure, but rather whether the modality can be anchored to one of the VP event's participants¹⁰:

- (23) The shit can really hit the fan in this part of the world.

To sum up, modals with an epistemic interpretation scope higher than modals with a root interpretation. The interaction of modals with the subject shows some differences between roots and epistemics, although the evidence doesn't force a control vs. raising analysis. Purported

¹⁰ See Portner (2008) for an alternative view according to which these kinds of modals are really quantificational.

thematic relation between a root modal and the subject instead seem to highlight the fact that root modals are centered around the circumstances of the event described by the main predicate, and especially, but not necessarily, those of its agent. The interaction of modals with tense and aspect, however, clearly supports Cinque's hierarchy and argue for two different positions for modals: one above tense and aspect, and one below.

Various proposals have offered to cash out this difference in height of interpretation by having epistemics and roots select for different types of arguments (e.g., Brennan 1993¹¹, Butler 2003, Zubizarreta 1982, Jackendoff 1972, Hacquard 2008b), by projecting in different domains (Picallo 1990), or by having a different featural makeup (Zagona 2007). More radical accounts deny any connection between them. This is common in the epistemic literature: Westmoreland (1998) and Drubig (2001), for instance, take epistemics not to be modals, but evidentials. Similarly, in order to account for the peculiar actuality entailments that arise from the interaction of root modals and aspect, Bhatt (1999) proposes that, despite appearances, the ability modal (which yields actuality entailments with perfective) is not a modal, but an implicative predicate (like *manage*). Note that, even those S- vs. VP-level proposals that assume some commonality to all modals (quantification over worlds), do also end up with two interpretation specific lexical entries for each modal, and leave open the question of why epistemic and root modals are expressed by the same lexical items cross-linguistically. One could appeal to a diachronic (or metaphoric) explanation: epistemic interpretations tend to develop cross-linguistically from root ones (cf. Sweester 1982; Papafragou 1998 for discussion). However, the problem still remains of

¹¹ Brennan doesn't encode the association between modal flavor and syntactic type directly in the semantics/lexical meaning. Rather, 'the connection between the category of the modal operator's argument and the interpretative class of the resulting sentence is *not* grammatical, but results from what range of things can reasonably be construed as modal properties [and] modal propositions (...), given our world view'. The idea is that it makes sense to use VP-modals to talk about abilities and dispositions because the community of language users recognizes such things. However, there is no independent reason for why the community of speakers couldn't construe epistemics as modal properties, and the connection between syntactic height and modal flavor remains arbitrary.

how the child learns that epistemic interpretations are tied to a high position, and root interpretations to a low one. The solution I choose to explore in this paper is that there really is just one lexical entry for each modal, which can appear in two different positions. We then need independent factors connected with these positions to explain why a high modal must receive an epistemic and a low modal, a root interpretation. This is what I propose in sections 5 and 6. To do so, I first want to show that the interpretation a modal can receive is even more restricted than what we have seen: modals are anchored to particular time-individual pairs, but not others. To derive this generalization I will argue that modals are relative to an event of evaluation.

4. Modals are keyed to individual/time pairs

This section shows that Cinque's puzzle is just one side of the coin: modals exhibit some time/individual constraints, which do not follow, as things stand, from their meaning, but will, once we make our modals relative to an event of evaluation. Zagana (2007) observes that epistemic modals are relative to the speech event (and thus to the speaker at the speech time). In the same vein, I will argue that all modals can be viewed as being relative to *some* event (and thus to its time and participants).

In section 3.2, 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. Modals are also relative to an individual (Bybee 1995, Palmer 2001), though this relativity cannot be encoded in argument structure (cf. section 3.3). For epistemics, this individual is the speaker (and perhaps a larger community, cf. DeRose 1991, von Stechow and Gillies 2007) in matrix contexts¹², and the attitude holder, when

¹²It 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

embedded (Stephenson 2007). For roots, this individual is the subject (Brennan 1993, Bybee 1995). Modals cannot 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 the example below where *have to* takes a circumstantial interpretation, it must be relative to *the time given by tense* (a past time), not the speech time:

- (24) 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.’

Note that 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*).

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

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

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 2007). 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.

When a modal is relativized to the *attitude holder* (every contestant), as in (26), where the modal receives an epistemic interpretation, it has to be relativized to the *attitude time* (cf. Speas 2004, Stephenson 2007), and not the speech time or a time prior to the thinking time:

- (26) Every contestant₁ thought he₁ might be the winner. [adapted from Speas (2004)]
- a. ‘For every contestant x, given what x knew **at x’s thinking time**, it was possible x was the winner.’
- b. *‘For every contestant x, given what x knows **now**, it was possible x was the winner.’

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 interpretation¹³. These restrictions are puzzling. Nothing conceptually should prevent, for instance, a modal to be keyed to the speaker at a past time. In fact, such modality can be expressed with an overt restriction:

- (27) In view of what I knew then, Mary had to be the murderer.

Why is it that, in the absence of an explicit restriction, such a reading unavailable? What is it about the semantics of modals and their grammatical environment that forces these time/individual constraints? I would like to propose that they arise from some event anchoring: by

¹³ 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 2008).

anchoring a modal to an event, we anchor it to its running time and participants. The reason why the overt restriction in (27) allows a past evaluation time is that the restriction 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 possible anchoring events: 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 (e.g., theme), and a time: the time given by tense. In order to make modals sensitive to these events, I will propose in the next section that modals (or more precisely their modal base) take an event argument (instead of a world) in the form of a variable, which needs to be bound.

In the rest of this section, I want to show that the event-relativity of modals is strictly local: modals are relative to the closest event. Consider the following sentence:

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

This sentence 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). Let's now see what happens when we insert a modal right below each of these three events. A modal right below the speech event, as in (29), is relative to the speaker at the speech time: *given what I know now, it is possible that John thought that Mary had taken the train*.

(29) 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, as in (30), 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.*

(30) 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, it is relative to the circumstances of the subject at the time of the taking the train: *given Mary's circumstances the day before yesterday, it was necessary that she take the train* (according to John):

(31) 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 its participants and running time. It further appears that, in fact, it *must*. In (30), *might* has to be keyed the attitude event. It cannot be keyed to the speech event, and thus (30) can never get the following interpretation: *John thought yesterday that it was possible, **given what I know**, that Mary took the train the day before.* Similarly, in (31), the modal has to be keyed to the VP event. The sentence 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 relativized 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 cross-linguistic generalizations: on the one hand, both epistemic and root modals can be expressed by the same lexical items, favoring a unified Kratzerian account. On the other, their systematically different distribution seems to provide evidence to the contrary: 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, VP-subject/VP time, which 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 relative to speech or attitude events, the low ones, those relative to VP-events. This event relativity will lead us to recast our original problem as follows: why is it that only speech or attitude event-relative modals get an epistemic interpretation, while VP event-relative modals get a circumstantial one. We address this question 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 propose to minimally change the lexical entries in (10), such that the modal base now takes an event, rather than a world argument, as in (32). A preliminary event-relative version of our modal bases is given in (33):

(32) For any world w , conversational backgrounds f, g , and proposition q :

- a. $[[\text{must}]](f)(g)(q) = 1$ iff $\forall w' \in \max_{g(w)}(f(e))$: $q(w') = 1$.
- b. $[[\text{can}]](f)(g)(q) = 1$ iff $\exists w' \in \max_{g(w)}(f(e))$: $q(w') = 1$.

Where $\max_{g(w)}$ is the function that selects the set of $\leq_{g(w)}$ -best worlds.

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

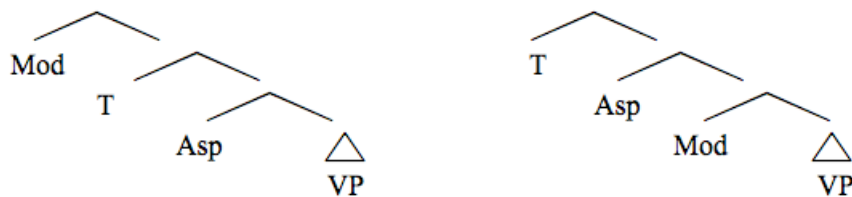
- a. $f_{\text{epis}}(e) = \lambda w'. w'$ is compatible with what is known in e (by $\text{Ag}(e)$ at $\text{time}(e)$).
- b. $f_{\text{circ}}(e) = \lambda w'. w'$ is compatible with certain facts/circumstances of e .

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, which needs to be bound by the closest event binder. We will look at these binders in section 5.2. Before we do so, we have to solve one mechanical problem (independent of event relativity), to allow the same modal, which takes a proposition as its complement to appear in two positions: one above TP and one above VP. I present a possible implementation in section 5.1, 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:

(34)



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. There are several ways around this problem. The solution I adopt here is couched within an extensional framework, where worlds and tenses are explicitly represented in the syntax (cf. Partee 1973, Kratzer 1998, Cresswell 1990, Percus 2000, a.o.). This move, along with a particular view of aspect, will yield two nodes of the right propositional types, with which a

modal can combine: TP and VP. For reasons of space, I won't motivate this framework here, and refer the interested reader to the references cited (see also section 5.2).

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):

- (35) 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 (λ_0), which maps to the actual world in matrix contexts, or by modals or attitude verbs. In the Davidsonian tradition, I take **verbs** to be predicates of events, which have to combine with an event argument. **Aspects** are quantifiers over events: they take a predicate of events (VP) and return a predicate of times, which in turn combines with tense. It is assumed traditionally that aspect is 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 verb¹⁴. Being of the higher

¹⁴ If the syntax doesn't force aspect to be base-generated under tense, could it be merged in the event argument of the modal? Maybe nothing prevents it to merge, but something prevents it to move out (K. von Stechow, p.c.). A modal's restriction is in an island for extraction: neither covert (1a)) nor overt movement (1b)) is allowed:

- (1) a. If every boy_j comes, his_{i/*j} mother will be happy.
b. *Who_i if t_i comes, will Mary be happy?

If aspect were to merge in the modal's restriction, it couldn't get out. Thus only an event pronoun can appear in this position. It gets bound the way *every boy* binds *his* in '*every boy_i will be happy if his_i mother comes*'.

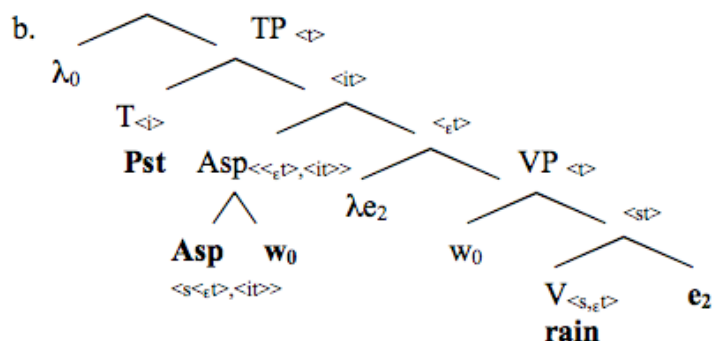
quantifier type, aspect moves out of this position and leaves a trace of type ε (for eventualities), which it binds from its target position (the same way a quantifier over individuals in object position has been argued to move). I assume the following lexical entry for perfective (adapted from Kratzer 1998, modulo the order of arguments and the addition of a world argument in the restriction to derive actuality entailments, cf. Hacquard 2006):

$$(36) \quad [[\text{PERFECTIVE}]] = \lambda w. \lambda P_{\langle \varepsilon, t \rangle}. \lambda t. \exists e[e \text{ is in } w \ \& \ \tau(e) \subset t \ \& \ P(e) = 1]$$

A simple example is presented below. Worlds and tenses are represented in the syntax. λ_0 is Percus's topmost world binder, which 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 argument of the verb to a position right below T, and creates an abstraction over events:

(37) a. Il a plu.

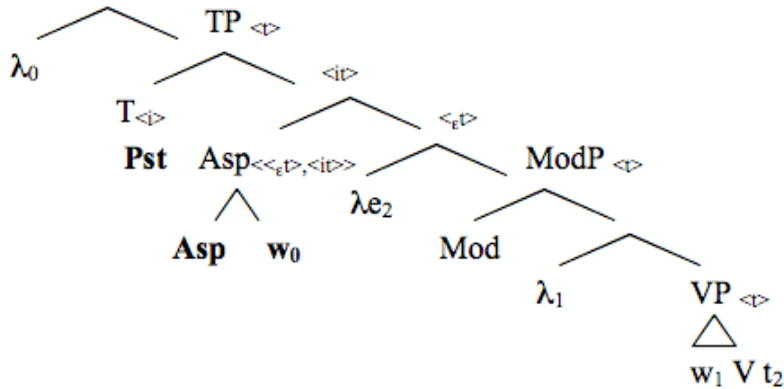
It rained-pfv



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

In most cases, aspect movement and base generation are equivalent. However, when a modal intervenes between tense and the VP, the movement approach lets the modal be of the right proposition type and still scope below aspect:¹⁵

(38)



We may ask at this point which assumptions are crucial for the account. Recall that we want to explain the fact that there are two positions for modals, which each correlates with a particular interpretation—epistemic for the high (above tense/aspect), circumstantial for the low position (below tense/aspect), and yet avoid hard-wiring a modal flavor for each syntactic position. In this proposal, I maintain that there really is just one lexical entry per modal by making two theoretical assumptions: (i) worlds are represented in the object language; (ii) aspect is base-generated as an argument of the verb and moves to T for type reasons¹⁶. An alternative

¹⁵ 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. One more step would have modals merge as arguments of tense, and also move for type reasons. For reasons of space and time, I won't pursue this option further (cf. von Stechow 2001, Shimada 2008).

¹⁶ Italian presents a morphological argument for aspect movement, as perfective is expressed by auxiliary *be* or *have* + participle, and where the particular auxiliary is determined by the embedded verb (unaccusatives select *be*):

- (1) 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 makes sense 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.

would be to invoke flexible types. A modal could combine either with propositions (and return propositions) and thus appear above tense, or with predicates of events (and return predicates of events), and thus appear below aspect. This would also yield two positions for modals, without necessarily specifying a particular interpretation. I chose the current system for three reasons: (i) it allows for a single entry for each modal; (ii) the growing body of evidence that indicates that worlds should be part of the object language is rather convincing (cf. Percus 2000); (iii) aspect movement parallels the type-driven movement of quantifiers over individuals in object position.

5.2. Note on worlds, events, locality and binding

The main motivation for representing worlds as pronouns in the object language (and 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):

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

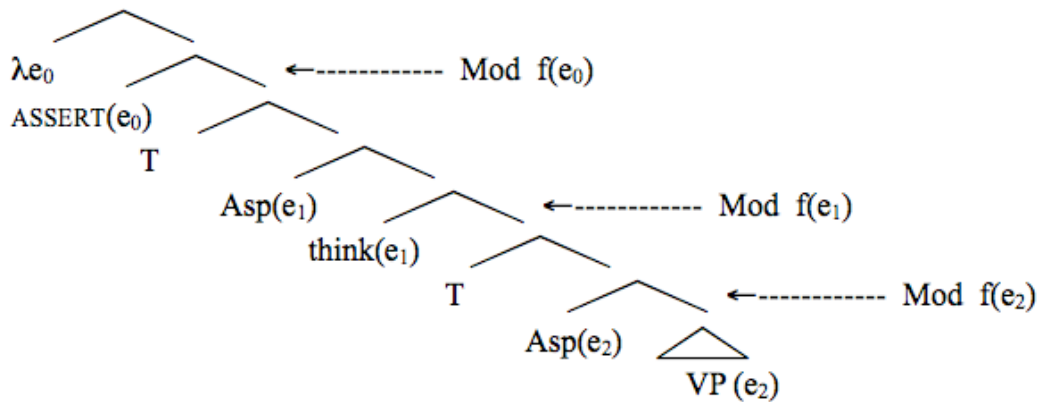
The sentence in (39) is ambiguous between a ‘transparent’ reading which asserts that all worlds in which every *current/actual* semanticist owns a villa in Tuscany are happy worlds (different world indices), and an ‘opaque’ reading, which states that all worlds where every semanticist in *those* worlds owns a villa are happy worlds (same world indices). To capture this, Percus argues that there are explicit world variables, which need to be bound either by a default matrix binder that maps to the actual world, or a binder provided by e.g., a modal. His binding theory differentiates between a world variable in the restriction of a quantifier like *every*, and in its scope. While the former can be bound by either the matrix binder (yielding actual semanticists),

or the binder provided by *would* (yielding counterfactual semanticists), the latter must be bound by the closest binder (*would*), thus preventing unattested readings of (39) about (counterf)actual semanticists owning villas in the **actual world**. There is thus an asymmetry in the locality conditions of world pronouns in the restriction of quantifiers over individuals and those in their scope: only the latter must be bound by the most local world binder. Interestingly, our modals are always relative to the closest *event*. We could reformulate Percus' binding principles (which are stated on a construction by construction basis in Percus 2000) in terms of a more general condition, whereby any world or event argument on the 'spine' of the tree (T, A, M, V) needs to be bound by the closest binder (K. von Stechow, p.c.).

5.3. Event binders

We saw that there are two positions within a clause where a modal can appear: above TP or above VP. 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 above VP, it is relative to the VP event's participants/running time. To derive this event relativity, we 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 λe_0 (à la Percus 2000). 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):

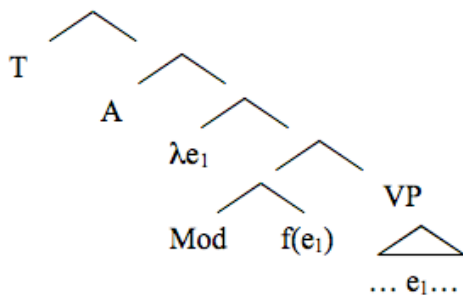
(40)



5.3.1. Aspect

Within a clause, a modal can either appear above TP ('high position') or above VP ('low position'). In the low position, its event argument gets bound by aspect (the closest event binder). That aspect can bind events is uncontroversial: aspect is, by definition, a quantifier over events. It will bind the modal's event variable the way *every* binds a pronoun (e.g., '*every boy₁ loves his₁ mother*'). We need a further syntactic constraint to rule out free event variables. In syntactic jargon, the derivation 'crashes' if the event variable in the modal's restriction is not coindexed with the closest event binder:

(41)



This binding by aspect will occur whenever the modal is in the low position, both in matrix and embedded clauses. What happens to modals in the high position will depend 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 *matrix* aspect, which happens to quantify over the matrix (attitude) verb’s event argument. This binding will anchor 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:

$$(42) \quad [[\text{believe}]] = \lambda p. \lambda x. \lambda w. \forall w' \in \text{DOX}(x, w) [p(w')=1]$$

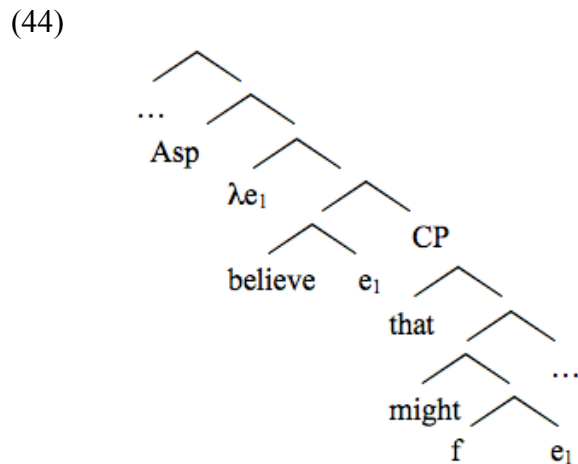
The literature on attitudes usually abstracts away from their aspectual properties and hence ignores their Davidsonian argument (but see Moltmann 2003, Kratzer 2006). However, attitude verbs are verbs, and as such, are *predicates of events*. They do differ from regular verbs, like *run* or *sleep*, in that they have ‘content’, that is, a set of propositions associated with them. To avoid confusion, let me make clear that by ‘content’ of the attitude, I do not mean the proposition that it takes as its complement, but rather the set of propositions that make up that attitude, i.e., that provides the set of worlds quantified over: for *believe*, for instance, this content is the set of doxastic alternatives of the subject. To encode this special lexical property of attitude verbs, while still treating them as predicates of events (i.e., giving them an event argument), I propose to directly link the propositional content of the attitude to its event argument via a CONTENT function (CON(e)). CON(e) picks out, for each attitude, the set of propositions that make up that attitude (e.g., the set of doxastic alternatives of the subject for *believe*). Thus, by ‘content of an event’, I do not mean the metaphysical makeup of that event, but rather refer to the set of

propositions associated with the attitude event, as determined by the attitude's lexical properties. $CON(e)$ is defined when e has propositional content (when e is an attitude event like *believe*), and undefined when e lacks such content, as is the case with ordinary verbs like *run*. The new lexical entry in (43) gives *believe* a Davidsonian argument, and recasts the quantification over the subject's doxastic alternatives via $CON(e)$:

$$(43) \quad [[\text{believe}]] = \lambda e. \lambda p. \lambda x. \lambda w. \text{Exp}(x,e) \ \& \ \text{belief}'(e,w) \ \& \ \forall w' \in \cap CON(e) [p(w')=1]$$

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

Now, when a modal is in the high position of an attitude verb's complement clause, 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:



Thus, a modal can be anchored to a VP or attitude event via aspect. We now need to handle the case where a modal is in high position in a matrix. I will assume a topmost default event binder, which will take on the role of Percus' (2000) world binder, as we will see in the next section.

5.3.2. Speech event

I will 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, Ginsburg and Sag 2001, Ambar 1999, Krifka 2001, Tenny and Speas 2004). For the case at hand, we will adopt an event version of the speech act representation, which will enable us to explain the modal facts in a straightforward way, and further allow for a simple clausal architecture, where embedded and matrix clauses share a similar structure. 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, it will be a questioning event, and if it is an order (as for imperatives), it will be an ordering event. Semantically, illocutionary force can be captured by intensional operators which determine different sets of worlds in which the proposition expressed holds. We only consider assertions in the following, for which I borrow Alonso-Ovalle and Menendez-Benito's (2003, 2008) ASSERT operator. ASSERT is an implicit universal modal operator, which quantifies over the speaker's doxastic alternatives and combines with a proposition:

$$(45) \quad [[\text{ASSERT}]]^c = \lambda p. \lambda w. \forall w' \in \text{Dox}_{\text{speaker of } c}(w)[p(w')]$$

Under this approach, assertions are viewed as expressions of what the speaker believes¹⁷. This is presumably what underlies Moore's Paradox (e.g., *#It is raining but I don't believe it is*). We can recast this operator in event terms, by using the content function from the previous section,

¹⁷As Alonso-Ovalle and Menendez-Benito note, this is an oversimplification. We may need to quantify over not simply what the speaker believes but what he takes to be common ground, or perhaps a subset of the speaker's beliefs, namely those that he is committed to defending (cf. Gunlogson 2001).

which will pick out the content of the speech event (ASSERT). Again, ‘content’ refers to the set of propositions that provide the worlds quantified over, here, the speaker’s doxastic alternatives:

$$(46) \quad [[\text{ASSERT } e_0]] = \lambda p. \lambda w. \forall w' \in \cap \text{CON}(e_0)[p(w')]$$

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

Recall that Percus (2000) introduced a default world (situation) binder λ_0 in topmost position. We can now replace this world binder by a default event binder λ_{e_0} , under the view that events occur single-worldly¹⁸: the speech event e_0 singles out one particular world (w_0), namely the world in which e_0 occurs.

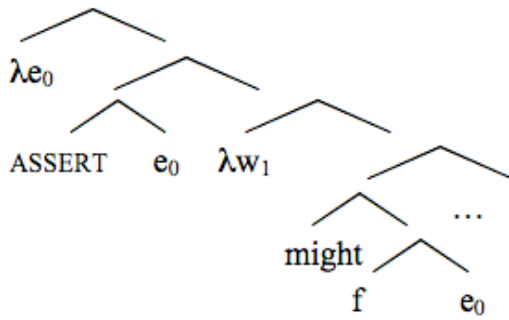
A declarative will thus have as its speech event an *asserting* event, such that the proposition expressed by the sentence will be true in all worlds compatible with the content of that event, that is, in all of the doxastic alternatives of the speaker of e_0 (in the (unique) world and time of e_0).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

- (47) a. John is the murderer.
 b. $[\lambda_{e_0} [\text{ASSERT } e_0 [\text{John is the murderer }]]]$
 c. $\forall w' \in \cap \text{CON}(e_0) \text{ John is the murderer in } w'$
 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 contexts: when the modal is in the high position, its event variable has to be bound by λ_{e_0} :

¹⁸ It would still be possible to have the same event occur in various worlds, by having both a root world and a root event binder. I chose not to do this here for simplicity. I’m grateful to Pranav Anand for the simpler representation.

(48)



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, which serves the traditional role of the index (provides the world and time of the utterance, namely the world and time of the speech event). 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:

$$(49) \quad [\lambda_{e_0} \text{ ASSERT } e_0 \quad \text{T Asp}_1 \text{ attitude } (e_1) \quad [\text{CP T Asp}_2 \text{ V}(e_2)]]$$

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

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 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 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):

- (50) a. $[\lambda_{e_0} \text{ ASSERT } e_0 \quad [\text{ModP Mod } f(e_0) [\text{TP T } [\text{AspP Asp}_1 [\text{VP V}(e_3)]]]]$
 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 and two modal bases: this should yield six kinds of interpretations. However, half of these do not seem to be attested ((b), (d), (e) below):

- (51) 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...*

In the next section, I show how to block these unattested readings and explain why only modals relative to speech and attitude events (i.e., the 'high' modals) receive an epistemic interpretation.

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 receive an epistemic interpretation, and VP-event-relative ('low') modals a root

interpretation? In this section I will sketch a proposal for the association of epistemicity and speech and attitude events (in contrast to ordinary VP events). I will argue that what sets speech and attitude events apart from ordinary events is their associated propositional *content*, and that this *content* is crucial for licensing an epistemic modal base. Section 6.1 discusses epistemic and section 6.2 circumstantial interpretations. Section 6.3 summarizes the connection between modal bases and syntactic positions.

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 possibly give it justice here (for an overview, see von Stechow and Gillies 2007 and Portner 2008). The main issue I want to point out here is that the traditional encoding of epistemic modality in terms of knowledge forces us to make a bizarre assumption, which can be avoided if we let its semantics be more flexible. The shortcomings of knowledge-based accounts get heightened in embedded contexts. Consider the following example where we embed an epistemic under the attitude *believe* (we use *might* as, for idiosyncratic reasons, it can only be epistemic):

(52) John believes it might be raining.

With a traditional account, we should obtain the following truth conditions: '*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*'. The problem with this account, is that it 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 (see Stephenson 2007). But this is not a trivial assumption: it gives John immense powers of introspection, whereby he can discern the facts that he knows from those he merely believes¹⁹. Intuitively, however, (52) doesn't require from John such a degree of awareness, and seems to make the more modest claim that 'rain in LA' is compatible with what John believes.

This problem of introspection could be avoided if we let the epistemic modal quantify directly over the worlds provided by the embedding attitude, unmediated by a state of knowledge²⁰. Doing so would yield the following truth conditions for (52): '*In some world compatible with what John believes, it is raining*'. I thus would like to propose that epistemic modality expresses possibilities/necessities given a particular information state, as determined directly by the embedding attitude, as opposed to a hard-wired knowledge state. To do so, I recast the epistemic modal base by using our content function CON(*e*). CON(*e*) will pick out the propositional content of the attitude event *e* that the modal is anchored to. Thus, an epistemic modal base has an event pronoun *e*, which needs to be bound by an event that has 'CONTENT', and returns a set of worlds compatible with the content of that event:

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

$$f_{\text{EPISTEMIC}}(e) = \lambda w'. w' \text{ is compatible with CON}(e)$$

¹⁹ For a similar argument in matrix contexts, see Tancredi (2007) who argues that epistemics cannot be knowledge-based, but rather doxastic-based.

²⁰ Much more needs to be said to see if this claim holds of all attitudes. For evidence that it does, see Anand and Hacquard (2008).

This recasting of the epistemic modal base avoids the problem of introspection, and more importantly for our purposes, provide some constraints on its licensing conditions: only those events that have associated content (i.e., speech and attitude events) will be able to license an epistemic modal base. If the modal base is bound by a contentless event (as is usually the case with VP events), the sentence will be undefined. In the remainder of this section, we look at epistemic modals in matrix (section 6.1.2) and attitude contexts (section 6.1.3). We then show why a low modal (i.e., an aspect-bound) cannot license an epistemic modal base (section 6.1.4).

6.1.2. Matrix contexts

Recall our speech event from section 5.2.3, whose content we took to be the doxastic alternatives of the speaker in declarative sentences:

$$(54) \quad [[\text{ASSERT } e_0]] = \lambda p. \lambda w. \forall w' \in \cap \text{CON}(e_0) [p(w')]$$

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

These doxastic alternatives will serve as the information state which a speech-event bound epistemic modal quantifies over. Let's see how, by turning to an example where a modal is bound by this speech event:

$$(55) \quad \text{It might be raining.}$$

$$\text{ASSERT}(e_0) \ \& \ \forall w' \in \cap \text{CON}(e_0): \exists w'' \in \cap \text{CON}(e_0): \exists e[e \text{ in } w'' \ \& \ \text{rain}(e, w'')]$$

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

Because both the modal and ASSERT quantify over the same set of worlds (those compatible with the content of e_0), we have a vacuous layer of quantification. The LF in (55) is thus equivalent to that in (56). We obtain that it is compatible with my (the speaker's) beliefs that it is raining:

(56) $\text{ASSERT}(e_0) \ \& \ \exists w' \in \cap \text{CON}(e_0): \exists e[e \text{ in } w' \ \& \ \text{rain}(e, w')]$

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

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 doxastic-based. Crucially, this depends on what we take assertions to invoke (cf. footnote 17). An interesting consequence of this proposal is that assertions and matrix epistemics are made of the same cloth: if assertions are knowledge-, doxastic-, or commitment-based, so will matrix epistemics.

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

(57) It must be raining.

$\text{ASSERT}(e_0) \ \& \ \forall w' \in \cap \text{CON}(e_0): \forall w'' \in \cap \text{CON}(e_0): \exists e[e \text{ in } w'' \ \& \ \text{rain}(e, w'')]$

In all worlds compatible with the content of e_0 (the speaker's beliefs) it is raining.'

This raises another potential issue²¹: 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) argued, an *ordering*

²¹ Thanks to Paul Portner (p.c.) for pointing this issue out.

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 our sentence with an epistemic modal embedded under *believe*, repeated in (58). Our intuitions were that the sentence as a whole describes compatibility of rain with John's beliefs:

(58) John believes that 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):

- (59) a. [John believe(e) [CP that [ModP might f(e) [TP it is raining]]]]
 b. $\exists e[e \text{ in } w \ \& \ \text{Exp}(e,J.) \ \& \ \text{belief}^{\circ}(e) \ \& \ \forall w' \in \cap \text{CON}(e):$
 $\qquad \qquad \qquad \exists w'' \in \cap \text{CON}(e): \exists e'[e' \text{ in } w' \ \& \ \text{rain}(e',w'')]]$
 c. $\exists e[e \text{ in } w \ \& \ \text{Exp}(e,J.) \ \& \ \text{belief}^{\circ}(e) \ \& \ \exists w' \in \cap \text{CON}(e): \exists e'[e' \text{ in } w' \ \& \ \text{rain}(e',w'')]]$

Believe describes a belief event (state) of John's. In all worlds w' compatible with the content of that belief state, there is a world w'' compatible with the content of that belief state in which it is raining. Notice that, here again, we have a layer of vacuous quantification. The LF in (59)b is thus equivalent to that in (59)c: '*it is compatible with John's beliefs that it is raining*'. This is precisely the meaning we were after. This reformulation of the epistemic modal base avoids the problem of introspection and further enables us to explain the limited distribution of epistemics.

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 an epistemic modal base. We have seen that both speech and attitude events have content, thus, speech event- and attitude-relative modals (which happen to be the high modals) can receive an epistemic modal base. What about VP-event-relative (low) modals?

Usually, the VP complement of a modal describes a non contentful event. We thus predict that a modal relative to such an event cannot get an epistemic interpretation. This prediction is borne out. The following sentence in which the modal is anchored to the VP event (and its agent *Mary*) cannot mean that it is compatible with what *Mary* believes/knows that she lifts this table. Instead, the sentence receives a circumstantial interpretation to which we will turn to in the next section.

(60) Mary can lift this table.

$\exists e: \exists w' \in \cap \text{CON}(e): \text{lift_this_table_by Mary}(e, w')$

*'#It is possible, given what Mary **believes**, that she lifts this table'*

What if, however, the VP complement of the modal is itself an attitude verb? Our 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 an epistemic possibility/necessity for the subject, that is, a possibility/necessity given the content of the attitude. This prediction seems to be borne out, as the following examples from German and French illustrate:

(61) 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 perf₂ [can f(e₂) [see(e₂) [Mary...

Given that *see* or *notice* 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, will express an 'epistemic' possibility, i.e., a possibility given what John knows. This is indeed a possible interpretation for (61): *Mary being nice was/became compatible with what John knows.*²²

To sum up, an epistemic modal base requires a contentful event. 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 in event terms:

(62) $f_{\text{CIRC}}(e) = \lambda w'. w'$ is compatible with the circumstances of e

What are the circumstances of an event? I propose that, beyond the spatial and temporal particulars surrounding the event, these circumstances include the properties of the event's participants (the agent, the theme...) at the event's location and the event's specific time.

²² Interestingly, with stative attitude verbs like *believe* this type of readings seems unavailable, suggesting that lexical aspectual restrictions may further constraint the range of interpretations. I leave this issue for future research.

Usually, a low modal will take a circumstantial modal base. As we saw in the previous section, the epistemic one is unavailable when the VP event lacks content. Thus, (63) will obligatorily receive a circumstantial modal base:

(63) Mary can lift this table.

$\exists e: \exists w' \in f_{\text{CIRC}}(e): \text{lift_this_table_by_Mary}(e, w')$

There is an event e such that in some world compatible with the circumstances of e (its agent's and theme's physical properties at the time of e), e is a table lifting by Mary.

Here the event is an event of *lifting this table* by Mary. The event participants are Mary and the table, and thus both the properties of Mary (how strong she is, how much training she has had, etc.) and properties of this table (how heavy it is, how voluminous...) at the time of the event, will be relevant circumstances for evaluating the modal.

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 properties of the event participants, not necessarily that of its agent. The circumstances of the event could, for instance, highlight properties of the object, or the location:

(64) A lot of people can jump in this pool. [Hackl 1998]

Note that a circumstantial modal base may require *some* anchoring to a participant: the circumstantial claims with expletive subjects or idioms in (65) are degraded without an explicit location. I leave a precise account of how these circumstances get computed for future research,

and simply note that, while a circumstantial modal may take into consideration general properties of the base world (gravity, laws of physics, etc.), it seems to prioritize properties of its participants: a sentence like ‘*John can lift this table*’ makes a claim about John’s strength and the table’s weight, above and beyond a claim about the impact of gravity in this world (in contrast to ‘*On the moon, John can lift this table*’, where properties of the moon (and thus its gravity) take on a more prominent role).

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

There is no definedness condition for circumstantial modal bases. Thus, 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 at this point, however, prevents *high modals* from taking a circumstantial modal base (though we could stipulate a default preference). We should thus be able to find examples such as the following:

- (66) Mary may take the train.

∃w’ compatible with the circumstances of the speech event: Mary takes the train(w’)

What does (66) mean? What are the circumstances of the speech event? They should consist mainly of properties of the speaker. It is unclear that such a meaning would be coherent. If it

were judged incoherent, this configuration would not be considered and only an epistemic interpretation would arise. Alternatively, we might start considering properties of the speaker as a speaker (is the speaker informative? are his beliefs consistent? etc.), in which 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. I leave this issue as an open question.

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 a high (attitude or speech event relative) modal receives an epistemic interpretation, and a low (VP event relative) modal a circumstantial one. 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 verbs aren't. This explained why high 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 thus the modal base usually associated with low modals. We thus obtain that high modals tend to be epistemic and low modals circumstantial, without having two separate entries for each modal.

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.) *Modalität und Modalverben in Deutschen*. Hamburg: Helmut Buske Verlag.

- Abusch, D.: 1997, 'Sequence of Tense and Temporal *De Re*'. *Linguistics and Philosophy*, **20.1**.
- Abusch, D.: 2007, 'Temporal and Circumstantial Dependence in Counterfactual Modals', in *Proceedings of the Amsterdam Colloquium*.
- Alonso-Ovalle, L. and P. Menendez-Benito: 2003, 'Some Epistemic Indefinites', in M. Kadowaki and S. Kawahara (eds.) *Proceedings of NELS 33*. Amherst, MA: GLSA, 1-12.
- Alonso-Ovalle, L. and P. Menendez-Benito: 2008, 'Modal Indefinites', ms.
- Ambar, M.: 1999, 'Aspects of Focus in Portuguese', in L. Tuller and G. Rebuschi (eds.) *The Grammar of Focus*. Amsterdam and Philadelphia: John Benjamins.
- Anand, P. and V. Hacquard: 2008, 'Epistemics with Attitudes', in *Proceedings of SALT XVIII*.
- Bhatt, R.: 1998, 'Obligation and Possession', in H. Harley (ed.) *Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect*, MITWPL 32. Cambridge, MA.
- Bhatt, R.: 1999, *Covert Modality in Non-Finite Contexts*. Ph.D. Thesis, UPenn.
- Brennan, V.: 1993, *Root and Epistemic modal auxiliary verbs*, Ph.D. Thesis, UMass, 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*. 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*. 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: University of Chicago Press.
- Carnap, R.: 1957, *Meaning and Necessity*. Chicago: University of Chicago Press.

- Cinque, G.: 1999, *Adverbs and Functional Heads: A Cross-Linguistic 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*. Palo Alto: CSLI Publications.
- Cormack, A. and N. Smith: 2002, 'Modals and Negation in English', in S. Barbiers, F. Beukema and W. van der Wurff (eds.) *Modality and its Interaction with the Verbal System*. Amsterdam: John Benjamins. 133-163.
- Cresswell, M.: 1990, *Entities and Indices*. Kluwer.
- Davidson, D.: 1967, 'The Logical Form of Action Sentences', in N. Rescher (ed.) *The Logic of Decision and Action*. Pittsburg: University of Pittsburg Press, 81-120.
- 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 Tuebingen.
- Egan, A., J. Hawthorne, and B. Weatherson: 2004, 'Epistemic Modals in Context', in G. Preyer and G. Peter (eds.) *Contextualism in Philosophy*. Oxford: Oxford University Press.
- Farkas, D.: 1997, 'Evaluation Indices and Scope', in A. Szabolcsi (ed.) *Ways of Scope Taking*. Kluwer Academic Publishers, pp. 183-215.
- von Fintel, K.: 2001, *Advanced Semantics Seminar*, lecture notes, MIT.
- von Fintel, K. and T. Gillies: 2007, 'CIA Leaks', in *Philosophical Review* 117.
- 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.

- Ginzburg J. and I. Sag: 2001. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, CA: CSLI.
- Hackl, M.: 1998, 'On the Semantics of "Ability Ascriptions"', ms., MIT.
- Hacquard, V.: 2006, *Aspects of Modality*. Ph.D. Thesis. MIT.
- Hacquard, V.: 2008a, 'Speaker-oriented vs. subject-oriented modality: a split in implicative behavior. *Proceedings of Sinn und Bedeutung 11*.
- Hacquard, V.: 2008b, 'On the Interaction of Aspect and Modal Auxiliaries', ms., University of Maryland.
- Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. Thesis, University of Massachusetts, Amherst.
- Heim, I.: 1994. 'Comments on Abusch's theory of Tense', in H. Kamp (ed.) *Ellipsis, Tense and Questions*. University of Amsterdam.
- Hintikka, J.: 1962, *Knowledge and Belief*. Cornell University Press.
- Hintikka, J.: 1969, 'Semantics for Propositional Attitudes' in J. W. Davis *et al.* (eds.) *Philosophical Logic*. Dordrecht, 21-45.
- Iatridou, S.: 1990, 'The Past, the Possible and the Evident', *Linguistic Inquiry* **21.1**, 123-129.
- Ippolito, M.: 2002, *The Time of Possibilities*, Ph.D. Thesis, MIT.
- Jackendoff, R.: 1972, *Semantic Interpretation in Generative Grammar*. Cambridge, MA: The MIT 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**.
- 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*. Berlin: De Gruyter, 639-650.
- Kratzer, A.: 1998, 'More Structural Analogies Between Pronouns and Tenses', in D. Strolovich and A. Lawson (eds.) *Proceedings of SALT VIII*. Ithaca, NY: CLC Publications, Cornell University.
- Kratzer, A.: 2006, 'Decomposing Attitude Verbs'. Talk in honor of A. Mittwoch. The Hebrew University Jerusalem.
- Kratzer, A., and J. Shimoyama: 2002, 'Indeterminate Pronouns: The View from Japanese, in Y. Otsu (ed.), *Proceedings of the 3rd Tokyo Conference on Psycholinguistics*.
- Krifka, M. 2001, Quantifying into Question Acts. *Natural Language Semantics* 9(1).
- Kripke, S.: 1972, 'Naming and Necessity', in D. Davidson and H. Gilbert (eds.) *Semantics of Natural Language*, Dordrecht: Reidel.
- 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., UC Berkeley.
- Moltmann, Fr.: 2003, 'Propositional Attitudes without Propositions' *Synthese* 135.
- 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.
- Partee, B.: 1973, 'Some Structural Analogies Between Tenses and Pronouns'. *The Journal of Philosophy*.

- Percus, O.: 2000, 'Constraints on some other variables in syntax'. *Natural Language Semantics* **8**, 173-229.
- Picallo, M.: 1990. 'Modal Verbs in Catalan. *Natural Language and Linguistic Theory* **8**.
- Piñón, C.: 2003, 'Being able to' in G. Garding and M. Tsujimura (eds.) *WCCFL 22 Proceedings*. Somerville, MA: Cascadilla Press, 384–397.
- Portner (2008) *Modality*. Oxford: Oxford University Press.
- Rizzi, L.: 1997, 'The Fine Structure of the Left Periphery, in L. Haegeman (ed.) *Elements of Grammar*. Dordrecht: Kluwer.
- Ross, J. R.: 1970, 'On Declarative Sentences', in R. A. Jacobs and P. S. Rosenbaum (eds.) *Readings in English Transformational Grammar*. Waltham, Mass.: Ginn.
- Shimida, 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.
- Stalnaker, R.: 1968, 'A Theory of Conditionals'. *Studies in Logical Theory. American Philosophical Quarterly*.
- von Stechow, A.: 2003, 'Feature Deletion under Semantic Binding: Tense, Person and Mood under Verbal Quantifiers', In *Proceedings of NELS 33*.
- Stephenson, T.: 2007, *Toward a Theory of Subjective Meaning*. Ph.D. Thesis, MIT.
- Stowell, T.: 2004, 'Tense and Modals', in J. Guéron & J. Lecarme (eds.) *The Syntax of Time*. Cambridge, MA: MIT Press.
- Sweester, E.: 1982, 'Root and Epistemic Modals: Causality in two worlds. *Proceedings of BLS 8*.
- 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. Guentner (eds.) *Handbook of Philosophical Logic, Vol. II*, 135-165 D. Reidel Publishing Co.
- Westmoreland, R.: 1998, *Information and Intonation in Natural Language Modality*. Ph.D. Thesis, 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*.
- Zagona, K.: 2007, 'On the syntactic features of epistemic and root modals', in L. Eguren, and O. Fernandez Soriano (eds.) *Coreference, Modality & Focus: Studies on the Syntax/Semantics Interface*, pp. 221-236.
- Zubizarreta, M. L.: 1982, *On the Relationship of the Lexicon to Syntax*. Ph.D. Thesis. MIT.