

CLASS 3: EPISTEMIC AND ROOT MODALS

Cross-linguistically, the same modals can express epistemic and root modality:

- (1) a. John is not in his office. He *must* be home. [epistemic]  
 b. John parked illegally. He *must* pay a fine. [root: deontic]  
 c. John wants a PhD. He *must* write a thesis. [root: teleological]

**French**

- (2) a. ... Il *doit* être chez lui. [epistemic]  
 b. ... Il *doit* payer une contravention. [deontic]  
 c. ... Il *doit* écrire une thèse. [teleological]

**Malay**

- (3) a. Dia *mesti* belajar. [epistemic]  
 He must be studying  
 b. Dia *mesti* belajar. [root: deontic]  
 He must study

Systematic ambiguity?

- Unlikely that the same ambiguity would be found in language after language.
- Ideally: Only one *must* (one *devoir*, one *mesti*...), which expresses necessity. Different ‘flavors’ come from context (Kratzer 1981, 1991)
- However, systematic scope differences between epistemics and ‘roots’ (i.e., all *non epistemic*) favor a polysemy account.

CROSS-LINGUISTIC GENERALIZATION 1: Same modals express root and epistemic modality.  
 CROSS-LINGUISTIC GENERALIZATION 2: Epistemics scope high; Roots scope low.

Today: The *epistemic* vs. *root* distinction:

- Meaning: A difference in modal reasoning (Kratzer 1981, 1991)
- Structure: scope differences
- Deriving structure & meaning correlations (Event-relative modality, Hacquard 2006)

**1. Roots vs. Epistemics: a difference in modal reasoning (Kratzer 1981, 1991)**

Recall our analysis far: Modals are quantifiers over possible worlds, where the set of worlds quantified over is provided by a conversational background<sup>1</sup>:

- (4)  $[[\mathbf{must}]]^w = \lambda f_{\langle s, st \rangle} \lambda q_{\langle st \rangle} . \forall w' \in \cap f(w) : q(w') = 1$   
 (5)  $[[\mathbf{can}]]^w = \lambda f_{\langle s, st \rangle} \lambda q_{\langle st \rangle} . \exists w' \in \cap f(w) : q(w') = 1$

<sup>1</sup> We assume conversational backgrounds are arguments of the modal, following Fintel&Heim (2005). In Kratzer’s original papers, they are parameters of the interpretation function.

Kratzer (1981, 1991) shows that this analysis makes the wrong predictions.

### 1.1. Double relativity of modals

#### 1.1.1. Inconsistent laws and deontic paradoxes

How can we talk about worlds where the law is obeyed, when the law has been broken?

(6) [John just committed a crime.] He must go to jail.

(6) should say that in all of the worlds in which the law is obeyed, John goes to jail. But surely, in all of the worlds in which the law is obeyed, there is no crime!

The problem with the semantics outlined above for sentences like (6) is that it treats the cold fact that John committed murder and the content of the law on a par.

Toy example: The law consists of 3 judgments (all judgments equal).

Judge A: murder is a crime

Judge B: owners of goats responsible for damage caused by goats

Judge C: owners of goats not responsible for damage caused by goats

The above analysis predicts that (7a-b) should be true, and (8a,b) false:

- (7) In view of what the judgments provide
- a. It must be the case that murder is a crime.
  - b. It must be the case that murder is not a crime.
- (8) In view of what the judgments provide
- a. It may be the case that goat owners are liable for damage caused by their goats.
  - b. It may be the case that goat owners are not liable for damage caused by their goats.

Our semantics makes the wrong predictions whenever laws are inconsistent. Why? Because the set of accessible worlds is empty!

#### 1.1.2. Two conversational backgrounds

Kratzer's separation of facts and ideals: modals relative to **two** conversational backgrounds:

- a *modal base*  $f$ , which provides a consistent set of **facts** (e.g., the fact that John committed a crime).
- an *ordering source*  $g$ , which provides a set of **ideals**, moral or other (which may or may not be consistent), which imposes an ordering on the worlds of the modal base

(9)  $g_{deontic}$ :  $\lambda w.\lambda p$ .  $p$  is one of the propositions given by the law in  $w$

**Ordering  $\leq_{g(w)}$ :**

- (10) For all  $u, z \in W$ , for any  $g(w) \subseteq \wp(\wp(W))$ :  
 $u \leq_A z$  iff  $\{p: p \in g(w) \text{ and } z \in p\} \subseteq \{p: p \in g(w) \text{ and } u \in p\}$

*For any pair of worlds  $u, z$ ,  $u$  is at least as close as  $z$  to the ideal set by  $g(w)$  iff the set of propositions that are true in  $z$  is a subset of the set of propositions that are true in  $u$ .*

Modals quantify over *best* worlds of modal base, given ideal set by ordering source.

- (11)  $[[\mathbf{must}]]^w = \lambda f_{\langle s, stt \rangle} \lambda g_{\langle s, stt \rangle} \lambda q_{\langle st \rangle} . \forall w' \in \text{Best}_{g(w)}(\cap f(w)) : q(w') = 1.$   
 (12)  $[[\mathbf{can}]]^w = \lambda f_{\langle s, stt \rangle} \lambda g_{\langle s, stt \rangle} \lambda q_{\langle st \rangle} . \exists w' \in \text{Best}_{g(w)}(\cap f(w)) : q(w') = 1.$   
*where  $\text{Best}_{g(w)}(X)$  selects the most ideal worlds from  $X$ , given the ordering given by  $g(w)$*

*Solving the problem of inconsistent laws (the toy example)*

Assume **modal base** is empty (i.e., the set of accessible worlds is the set of all possible worlds).

**Ordering source**  $g_{\text{deontic}}(w) = \{\text{murder is a crime; goat owners liable; goat owners not liable}\}$

There are no worlds in which *all* of the laws are true. Rather, 3 types of worlds:

- Type 1 worlds in which murder is not a crime
- Type 2 worlds in which murder is a crime and goat-owners are liable
- Type 3 worlds in which murder is a crime and goat-owners are not liable

Type 2 and 3 worlds are the closest worlds to the ideal given  $g$ .

We can derive that:

- It must be that murder is a crime* is true: in *all* of the best worlds, murder is a crime.
- It may be that goat-owners are liable* is true: in *some* of the best worlds, goat-owners liable.
- It may be that goat-owners are not liable* is true: in *some* best worlds goat-owners not liable.
- It must be that murder is not a crime* is false: in *none* of the best worlds is murder not a crime.

## 1.2. Modal bases and ordering sources

### 1.2.1. Two kinds of modal reasoning: circumstantial vs. epistemic

Kratzer argues that there are two different kinds of premises we consider when doing modal reasoning: one is based on known facts, the other on circumstances. The former is at the source of epistemic interpretations, the other at the source of root interpretations.

There is ‘a difference in the kind of premises from which we reason. If we use an epistemic modal, we are interested in what may or must be the case in our world given everything we know already. And if we use a circumstantial modal, we are interested in what can or must happen, given circumstances of a certain kind [...]: facts concerning the outside world, our bodies or our mind, for example’. (Kratzer 1981)

‘Epistemic modality is the modality of curious people like historians, detectives, and futurologists. Circumstantial modality is the modality of rational agents like gardeners, architects and engineers. A historian asks what might have been the case, given all the available facts. An engineer asks what can be done given certain relevant facts.’ (Kratzer 1991)

- (13) a. Hydrangeas can grow here  
b. Hydrangeas might be growing here

**Circumstantial modal base:** picks out worlds in which certain relevant facts or circumstances hold, *e.g.*, climate, soil, special properties of hydrangeas...

**Epistemic modal base:** picks out worlds compatible with available evidence.

If we know for a fact that hydrangeas do not grow here, (13)a) can be true, but (13)b) false.

Epistemic modal bases yield epistemic interpretations; circumstantial modal bases, root ones:  
**Root vs. epistemic = circumstantial vs. epistemic MB**

### 1.2.2. Combinations of modal bases and ordering sources

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

**Circumstantial modal bases** combine with ordering sources related to *laws, aims, wishes*, yielding *deontic, teleological, bouletic* interpretations.

*To sum up:*

**Modal force:** necessity ( $\forall$ ) and possibility ( $\exists$ ), lexically provided.

**Modal base:** **epistemic** or **circumstantial**

**Ordering source:** stereotypical (with epis MB); deontic, bouletic, teleological (with circ MB)

In this system, *epistemics* and *roots* involve a fundamental difference in meaning.

As Kratzer (1991) notes, modal base selection correlates with syntacticians’ epistemic vs. root distinction (Perlmutter 1971, Ross 1969, Jackendoff 1972), according to which epistemic is S-modality, and root VP-modality.

The goal of the next sections is to explain why modal base selection should correlate with syntactic height.

## 2. Epistemics vs. roots: Structural differences

The root/epistemic distinction correlates with scope differences: As we saw last time, **epistemics** tend to scope high (> other modals, negation, tense, subjects), **roots** lower than all of these.

But, there were exceptions. We’ll focus on the tense/subjects interactions.

## 2.1. Interaction with individuals

We saw that **epistemics** can scope over quantifier subject, **root** modals cannot (Brennan 1993):

- (14) a. Every radio may get Chicago stations and no radio may get Chicago stations.  
*It's possible that every radio gets C. stations, it's also possible that none of them do.*  
 b. #Every radio can get Chicago stations and no radio can get Chicago stations.  
*Every radio is such that it gets C. stations, and no radio is s.t. it gets C stations.*

The problem with (14b) can be blamed on a thematic relation between the subject and a **root modal** (Jackendoff 1972, Brennan 1993, a.o.).

However, while there does seem to be a connection between a root modal and its subject, this connection cannot be due to control configuration (Bhatt 1998, Wurmbrand 1999, Hackl 1999):

- (15) a. There have to be fifty chairs in this room. (Bhatt 1998)  
 b. It can rain hard here. (Hackl 1999)  
 c. A lot of people can jump in this pool. (Hackl 1999)

Hence, while root modals are traditionally taken to be *subject-oriented* (Bybee *et al* 1994), a more accurate characterization is that **root modals** are centered around **VP-event participants** (subject or other).

**Epistemics** on the other hand are traditionally *speaker-oriented* (Bybee *et al.* 1994): they report possibilities given the *speaker's* (and his community) knowledge.

In attitude contexts, however, the perspective gets *shifted* to that of the attitude holder (Speas 2004, Stephenson 2007):

- (16) Every boy<sub>1</sub> thinks he<sub>1</sub> must<sub>1</sub> be stupid. (Stephenson 2007)

⇒ Epistemics are speaker/attitude holder oriented, roots are VP-event participant oriented.

## 2.2. Interaction with Tense

**Roots:** time of evaluation is the VP event time, i.e., the time provided by tense:

John wanted to get a PhD...

- (17) He had to write a thesis.  
*Given John's situation & goals **then**, it was necessary he write a thesis.*  
*\*Given John's **current** situation & goals, it is necessary he wrote a thesis then.*

**Epistemics:** the time of evaluation cannot be *back-shifted*. (Groenendijk & Stockhof 1975, Iatridou 1990, Abusch 1997, Stowell 2004...)

In *matrix contexts*, the modal's time of evaluation has to be the speech time:

- (18) John had to be in his office at the time of the crime.  
*Given what we know now, it is necessary that John was in his office then.*  
*\*Given what we knew then, it was necessary that John was in his office.*

In *attitude contexts*, it has to be the internal 'now' of the attitude time:

- (19) Mary thought that John had to be in his office at the time of the crime.  
*It was necessary for M. at her thinking time that J. was in his office at some prior time*  
*\*It is necessary for M. now/at the crime time that J. was in his office at some prior time*

This is often taken as evidence that **epistemics scope above tense** (Iatridou 1990, Abusch 1997, Piccolo 1990, Abraham 2001, Stowell 2004)

However, we saw some counterexamples to the claim that epistemics can't be backshifted (see Homer 2009 for similar examples in French):

- (20) A: [after seeing B looking in a drawer]. Why did you look in there? [Fintel&Gillies 2008]  
 B: My keys might have been in the drawer.  
*It WAS possible that my keys were in there*

Here the epistemic's time of evaluation seems *backshifted* to a past searching time.

*However*, note that this type of counterexamples seems restricted to responses to a *why* question:

Stephenson (2007) argues that the temporal shifting in this case could be due to a covert *because*, able to shift evaluation parameters.

We'll leave aside the semantics of *because* and its shifting properties, and assume that epistemics can't be backshifted by *tense* (and hence scope above it).

### 2.3. Summing up interactions with times and individuals

**Epistemics:** speaker-oriented, speech time-oriented  
 (attitude holder/attitude time in attitude contexts)

**Roots:** subject (or other VP participant)-oriented, VP-event time oriented

Epistemic = TP modality; Root = VP modality

CINQUE'S UNIVERSAL HIERARCHY OF FUNCTIONAL PROJECTIONS:  
 ... MOD<sub>EPIS</sub> > T > ASP > MOD<sub>ROOT</sub> > VP

- This structural difference is unexpected if the epistemic/root distinction boils down to a difference in modal base.
  - We need a way to force epistemics to scope high, and roots to scope low.
- This can be done by giving roots and epistemics **separate lexical entries**, which encode both *flavor* and *syntactic position* (cf. Jackendoff 1972, Picallo 1990, Butler 2003): e.g., epistemics take a proposition, roots a property of times argument.

$$\begin{aligned} [[\text{must}_{\text{epis}}]]^w &= \lambda f_{\text{epis}\langle s, \text{st} \rangle} \lambda g_{\langle s, \text{st} \rangle} \lambda q_{\langle \text{st} \rangle} \cdot \forall w' \in \text{Best}_{g(w)} (\cap f_{\text{epis}}(w)): q(w') = 1 \\ [[\text{must}_{\text{root}}]]^w &= \lambda f_{\text{circ}\langle s, \text{st} \rangle} \lambda g_{\langle s, \text{st} \rangle} \lambda Q_{\langle i, \text{st} \rangle} \lambda t_i \cdot \forall \langle w', t' \rangle \in \text{Best}_{g(w)} (\cap f_{\text{circ}}(w, t)): Q(w', t') = 1 \end{aligned}$$

- The fact that the same words are used to express root and epistemic modality would be an accident in all languages in which it happens...
- But maybe, not such a terrible accident: roots and epistemics are still both modals, and still play similar roles (unlike treating epistemics like evidentials).
- And perhaps, not an accidental accident:

### 3. A motivated polysemy account (Sweetser 1990): diachronic metaphoric process.

- Cross-linguistically, epistemic interpretations tend to develop from root ones.
- Historic development matched by children's acquisition of modals (roots appear around age 2, epistemics not until age 3 ½; cf. Sweetser 1990, Papafragou 1998)

*Sweetser's proposal:*

- Modals encode 'force dynamics' of potential barriers and driving forces. These forces operate in the concrete, external world for roots, but can be metaphorically extended to the realm of the mental/abstract, to yield epistemic modality.
- Modals are polysemous, but this polysemy is motivated. Hence, it's not surprising it should be found in language after language.

**Still**, why is it that epistemic interpretations are tied to a *high* position, roots to a *low* one? Why aren't there languages where *roots* scope high, and *epistemics* scope low?

### 4. Event-relative modality (Hacquard 2006, 2010)

**MAIN CLAIM:** Modals are anchored to an *event*. From that event, we can recover the relevant factors responsible for modals' interpretative constraints.

- Maintains a Kratzerian account, where each modal comes in a single lexical entry, and different 'flavors' come from contextually-given conversational backgrounds.
- But, with two differences:

1. Modals are relative to an *event of evaluation*, rather than a world.
  - o More specifically, conversational backgrounds pick out a set of propositions from an event, rather than a world of evaluation.

*World-relative modal bases:*

- (21) a.  $\cap f_{epistemic}(w): \lambda w'. w'$  is compatible with what we know in  $w$   
 b.  $\cap f_{circumstantial}(w): \lambda w'. w'$  is compatible with the circumstances in  $w$

*Event-relative modal bases:*

- (22) a.  $\cap f_{epistemic}(e): \lambda w'. w'$  is compatible with what we know in  $e$  [to be revised]  
 b.  $\cap f_{circumstantial}(e): \lambda w'. w'$  is compatible with the circumstances of  $e$

2. A modal can appear in two positions (above TP or above VP), but *without* flavor specification. For compositional details of how the same modal can combine with either TP or VP, see Hacquard (2010).

#### 4.1. Event anchoring

*Recall:* modals are anchored to individuals and to times.

**Epistemics:** <speaker-oriented ; speech time oriented> (in matrix contexts)

- (23) John had to be in his office.  
*Given what I know now, it is necessary that John was in his office.*

<attitude-holder ; attitude time> (in attitude contexts)

- (24) Paul thought that John had to be in his office.  
*Given what Paul knew then, it was necessary that John was in his office.*

**Roots:** <VP-event participant ; VP-time>

- (25) a. John had to write a thesis (in order to get a PhD).  
*Given John's situation & goals then, it was necessary he write a thesis.*  
 b. Many people can fit in this elevator.  
*Given the elevator's capacities now, it is possible that many people fit in it.*

However, *not all* combinations of time/individual pairs are attested (regardless of modal flavor):

- \*<VP-participant-speech time>  
 (26) John had to be in his office.  
*\*Given J's circumstances/evidence now, it is necessary that he was in his office.*
- \*<speaker-VP time>  
 (27) John had to be in his office.  
*\*Given what I knew then, it is necessary that he was in his office.*
- \*<Attitude holder-Attitude time>  
 (28) Mary thought that John might be the murderer.  
*\*Given what M. knows now/at time of crime, it is possible John was the murderer.*

**Empirical claim:** Time and individual constraints

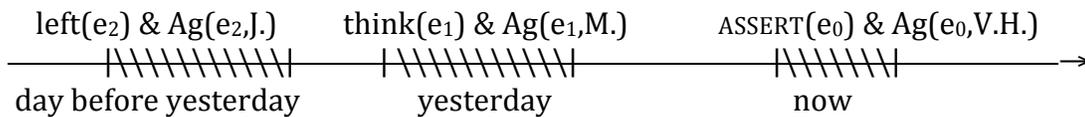
- If a modal is *speaker*-oriented  $\Rightarrow$  keyed to *speech time*  $\Rightarrow$  EPIS
- If a modal is *attitude holder*-oriented  $\Rightarrow$  keyed to *attitude time*  $\Rightarrow$  EPIS
- If a modal is *VP participant*-oriented  $\Rightarrow$  keyed to *VP time*  $\Rightarrow$  CIRC

**Theoretical claim:** Time and individual constraints arise from ‘event’ anchoring:

- From an *event* we can recover information about its participants & running time
- By anchoring a modal to an event, we anchor it to that event’s participants & running time

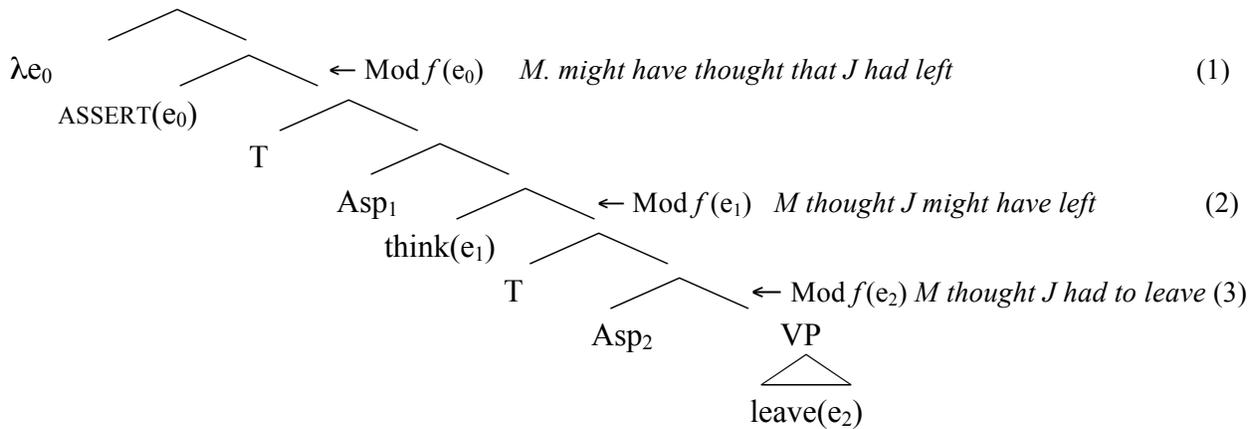
**What are the anchoring events?**

(29) Mary thought yesterday that John had left the day before.



**3 different times, 3 agents  $\Rightarrow$  3 different events:**

- **speech event** (assertion): <speaker, speech time>
- **attitude event:** <thinker, yesterday>
- **VP event:** <leaver, day before yesterday>



**(1) Modal above matrix TP  $\Rightarrow$  Speech event relative**

*Mary might have thought that John had left*

$\exists w'$  comp. w/ what  $Ag(e_1)$  (=speaker) knows at  $t(e_1)$  (=speech time) s.t. M thought J left in  $w'$

**(2) Modal above embedded TP  $\Rightarrow$  Attitude event relative**

*Mary thought that John might have left*

$\dots \exists w'$  comp. w/ what  $Ag(e_2)$  (= Mary) knows at  $t(e_2)$  (= think time): J left.

**(3) Modal above VP  $\Rightarrow$  VP event relative**

*Mary thought that John had to leave*

$\dots \forall w'$  compatible with  $Ag(e_3)$  (= J.)’s obligations at  $t(e_2)$  (= leaving time): J left.

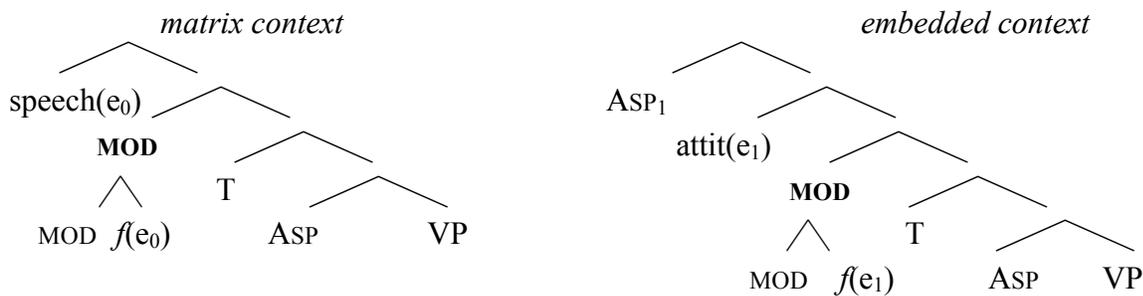
**Modal’s event argument anchored to the closest event:**

- (30) Mary thought that John might have left.  
 \*Mary thought that, given what **I** know **now**, it was possible John left

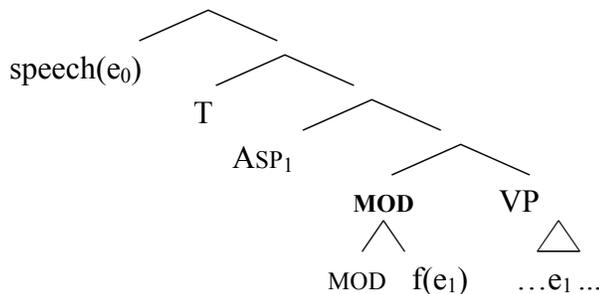
**4.2. Event-relative modality**

- Modals have an event (rather than a world) argument that needs to be bound locally.
- 2 positions for modals (‘high’ = above TP, or ‘low’ = above VP).
- This yields 3 binding possibilities, and hence 3 possible event anchorings:

**‘High’ modal: relative to speech event (matrix), or attitude event (embedded)**



**‘Low’ modal: relative to VP event**



- **Event relativity** derives all and only attested individual-time pairs: modals are either <VP-event participant/time>, <attitude holder/time> or <speaker/speech time> oriented.
- But it doesn’t yet explain the association between modal height and modal interpretation (*high* modals = epistemics; *low* modals = roots).
- It does allow for a reformulation of the question, however:
  - Why do only modals relative to speech or attitude events (i.e., *high* modals) get epistemic interpretations (i.e. take an epistemic modal base)?

*Next time:* on why epistemic modality is associated with ‘high’ modals.