Logic and the lexicon: insights from modality[*]

Valentine Hacquard

University of Maryland

0. Introduction

Formal semantics is the study of linguistic meaning using tools from modern logic. Logic is the study of inference, not language, and abstracting away from natural languages was critical to its development: "work in logic just is, to a large extent, a struggle with the logical defects of language," wrote Frege in 1915. Yet the struggle produced tools that proved useful to linguists.

This was due in part to the realization that the perceived "logical defects" arise not only from language itself, but also from the complexity of its use. Speakers do more than assert premises, and in various ways they exploit context to mean more than their words do (Frege 1918, Austin 1962, Grice 1989). If these "pragmatic" aspects of meaning are stripped away, the result is generally more amenable to logical treatment. For instance, if I say "Jo got angry and left," I probably mean that Jo got angry *before* leaving. But this implication of chronology need not come from "and" itself. It might arise instead from a shared presumption that speakers will often narrate events in their natural order (Grice 1975). If that is right, it might be that the best semantics for "and" treats it as the Boolean conjunction, a meaning drawn from the logical toolkit. In this way the logician's struggle has yielded not only formal models for certain lexical meanings (those that participate in general patterns of inference), but also important lessons on their tortuous relation to meaning in language use.

[*] Many thanks to Alexander Williams, Anouk Dieleuveut, Anna Papafragou, and one anonymous reviewer for helpful feedback.

My focus in this chapter is the modal vocabulary: words about possibility and necessity, like "might," "can" and "must." These offer an especially rich study in the linguistic representation of 'logical vocabulary.'[1] On the one hand, logic provides a highly developed formalization of modal inferences, involving operators that correspond somehow to our modal words (Carnap 1947, Prior 1957, Kripke 1959, Hintikka 1963). On the other, what speakers do in using modal words seems to go beyond what the words themselves can contribute. What lexical meaning do speakers actually assign to these words, and what guides the child to this result? Similar questions arise throughout the lexicon (e.g., quantifiers, content-word analyticities, conditionals), so we can apply the lessons of this discussion widely.

The modal vocabulary instantiates a remarkable human property. We can talk about states of affairs beyond the here and now: how the world would be if dinosaurs had not mostly died out, or if all of our needs or desires were to be realized. We can distinguish the merely possible from the necessary, the impossible from the unlikely. Modality is the category of linguistic meaning that enables such talk. Many linguistic expressions encode notions of possibility or necessity, including nouns (*possibility*), adjectives (*possible*), adverbs (*maybe*), verbs (*require*), modal auxiliaries (*must*), and semi-modals (*have to*). This chapter focuses on modal auxiliaries and semi-modals, around which theories of modality have been developed, but contrasts their behavior with modal expressions from other lexical categories.

By using modals, speakers can express certainty or uncertainty about different possibilities (as with (1a) and (1b)), or orders and permissions (as with (2a) and (2b)). But what exactly do words like *must* and *may* contribute, and what is instead contributed by their grammatical context, or by speaker intentions and the circumstances of their use? As we will see, our modal

[1] This chapter focuses on modals, as an instance of logical vocabulary. See Jackendoff, this volume, for an overview of *lexical semantics*.

statements involve a complex interplay of morphology, syntax, semantics, and pragmatics, which make it challenging to pinpoint the exact lexical contributions of the modal words themselves.

(1) a. Jo must have done it.

b. Jo may have done it.

(2) a. You must go.

b. You may go.

First, it is not always easy to tease apart the *semantic* and *pragmatic* contributions involved in a modal statement. A natural use of (2a), for instance, seems to issue a command, similar to the one issued with the imperative "*Go!*". But is the command part of the *literal* meaning of (2a), or might it instead arise indirectly, in virtue of the speaker merely *describing* a necessity? The other complicating factor is the fact that modals can be used to express different *kinds* of possibilities and necessities. The sentences with *must* and *may* in (1) most naturally express 'epistemic' modality: what is required or allowed by the available evidence; those in (2) most naturally express obligations and permissions (so-called 'deontic' modality). How many *musts* and *mays* do we have in our lexicon? Are there distinct epistemic and deontic versions of the modals, or is there just one *must,* and just one *may*, as Angelika Kratzer has influentially argued (1977, 1981)?

This chapter starts with a survey of how possibilities and necessities are encoded in natural language, with an eye toward cross-linguistic similarity and variation. Section 2 introduces the framework that formal semantics inherited from modal logic to analyze modal statements.

Section 3 turns to the division of labor between semantics and pragmatics for modal statements, and section 4 zooms in on the lexical contribution of the modals themselves.

1. Expressing possibility and necessity in natural language

1.1. Notional vs. grammatical modality

I will use the term '*grammatical modality'* (Traugott 2011) to refer to words or morphemes from a dedicated lexical category that express possibility or necessity: in English, modal auxiliaries (*can*, *must, may*, *might*, *could*, *should, will, would*) and semi-modals (*have to, ought to*). This is in contrast with *'notional modality*' (Kratzer 1981), a term that applies to words or morphemes from any lexical category that encode notions of possibility or necessity.

Syntactically, modal auxiliaries belong to the functional domain of the clause; they are in complementary distribution with tense and other auxiliaries like *do*; they do not bear agreement morphology (4c), nor do they allow *do*-support (4a-b), unlike regular verbs (3a-c). Semi-modals are often grouped with modal auxiliaries because they can be used to express the same range of meanings, but they behave syntactically more like verbs: they require *do* support (5a-b), and bear tense/agreement morphology (5c).

(3)	a.	*Did* you <u>leave</u>?	a'.	*<u>Left</u> you?
	b.	You *did* not <u>leave</u>.	b'.	*You <u>left</u> not.
	c.	He <u>leave</u>*s*.	c'.	*He <u>leave</u>.
(4)	a.	**Did* you <u>can</u> leave?	a'.	<u>Can</u> you leave?
	b.	*You *did* not <u>can</u> leave.	b'.	You <u>can</u>'t leave.
	c.	*He <u>can</u>*s* leave.	c'.	He <u>can</u> leave.

(5) a. *Did* you have to leave? a'. *Have you to leave?

b. You *did* not have to leave. b'. *You haven't to leave.

c. He has to leave. c'. *He have to leave.

Notional modality, however, can be found across all lexical categories, including nouns (*possibility*, *necessity*…), adjectives (*possible*, *necessary*…), adverbs (*maybe*, *necessarily*…), verbs (*require*, *allow*…). Grammatical modals have received the most attention in the literature, and form the empirical basis for formal accounts of modality. However, these modals sometimes show quirks; it is thus useful to contrast their behavior to that of adjectives or verbs that express similar meanings, to tease apart peculiarities that stem from their lexical category from properties essential to notional modality more broadly.

1.2. Modal force and modal flavor

The meanings of modals vary along two main dimensions: "force" and "flavor". We will consider each dimension in turn, and briefly survey variations that we find across languages.

Modal logic distinguishes two main modal *forces*, *possibility* and *necessity*: possibilities leave other possibilities open, necessities do not. The modal auxiliaries of English can be split into these two categories, as shown in (6).

(6) a. Possibility modals: *can, could, may, might*

b. Necessity modals: *must, should, have to, ought to*

However, modal expressions in natural language, however, can express finer-grained distinctions than this dichotomy suggests. The example in (7), for instance, illustrates a difference between the modals *must* and *should*, which both express necessity, but a seemingly "weaker" one for *should*: *must* is associated with a mandatory requirement, *should* with more of a recommendation.

(7) Employees must wash their hands. Everyone else should. Fintel & Iatridou (2008)

And while grammatical modals seem restricted to possibility and (weak and strong) necessity, other lexical categories, like nouns or adjectives, can encode even finer shades of possibility:

(8) a. It's more *likely* that Jo did it than Al.

b. There's a slight *possibility* that Jo did it.

Modal statements express possibilities and necessities allowed by various sorts of consideration, leading to different "flavors" of modality. *Epistemic* modality (from Greek *episteme* 'knowledge') expresses what is possible or necessary given what is known (the available evidence), and *circumstantial* modality, given certain circumstances, with *abilitive* modality a special subcase focused on the subject's physical abilities. Modals can further express what is possible or necessary given different *priorities* (Portner 2009), such as *rules* for deontic modality (Greek *deon* 'obligation'), *desires* for bouletic modality (Greek *boule* 'wish'), or *goals* for teleological modality (Greek *telos* 'goal'). Finally, *metaphysical* modality expresses what is possible given a certain history; this is the modality involved in *counterfactuality*. The following examples illustrate:

(9)	a.	Jo {*might/must*} be the murderer (given what we know).	epistemic
	b.	Jo *can* lift 200lbs (she is very strong).	ability
	c.	I *have to* sneeze (my nose is tickling).	circumstantial
	d.	Participants {*may/have to*} register online (according to the rules).	deontic
	e.	You {*could/should}* try the bisque! (I'd love it if you did!)	bouletic
	f.	You {*can*/*have to*} take a cab (to get to the conference venue.)	teleological
	g.	I {*could/would*} have won, if I hadn't twisted my ankle.	metaphysical

As we will see, epistemic modality tends to pattern differently in its interactions with elements like tense from the other flavors, which themselves tend to pattern together, and are often subsumed under the label "root" modality (Hoffmann 1966).

Modality is often distinguished from two other categories that express related notions: *attitude* predicates (*think, want*…), which express mental states (*belief*, *desire*…), and *evidentials*, which encode *source* of evidence in languages like Korean or Quechua. It can be difficult to know in what category a particular lexical item falls (certain words or morphemes encode more than one category), and some analyses even merge these categories. Here we will assume that modality is distinct from both: attitudes express *mental states*, while modals express possibilities and necessities relative to such mental states; evidentials indicate the *source* of evidence, while epistemic modality expresses certainty *based* on this evidence.

A striking fact about modals in a language like English is that they can be used to express different modal flavors. *Must* and *have to*, for instance, can be used to express epistemic, circumstantial, deontic, bouletic, and teleological necessity. This feature tends to be restricted to

grammatical modals (e.g., English modal auxiliaries), and seems rather common across the world's languages: about half have at least one modal that can be used to express epistemic and deontic flavors (van der Auwera & Ammann 2011).

In some languages, the same word can be used in situations where English speakers would either use a possibility modal or a necessity modal. Such "variable force" modals have been documented in Nez Perce (Deal 2011), illustrated in (10), St'at'imc'ets (Rullmann *et al.* 2008), Washo (Bochnak 2015), and even Old English (Yanovich 2016).

(10) 'inehne-**no'qa**' ee kii lepit ciickan. Nez Perce (Deal 2011)
take-**MOD** you these two blankets
*'you **can** take these two blankets.'*
*'you **must** take these two blankets.'*

Finally, some languages only have modals that can express a single flavor in a single force (e.g. Javanese, Vander Klok 2013).

The rest of this chapter largely focuses on languages with a modal system like English. Summing up its key features, we find a dedicated grammatical category for modals, though modal meanings can be found throughout the lexicon. Grammatical modals are restricted to possibility and necessity in terms of *force*; they however can be used to express different *flavors*.

Words from other lexical categories can express finer shades of possibility, but they typically express a single flavor.[2]

1.3. Possibility and necessity modals: logical inferences

Under all flavors, possibility and necessity modals show the same patterns of entailments and logical equivalences as the quantifiers *some* and *every* in the nominal domain. This is nicely illustrated in Fintel & Heim (2011) with a pair of antonyms like *leave* and *stay* (*leave=not stay*, *stay=not leave*):

(11)	a.	You must leave.	c.	It's not the case that you may stay.
	b.	You may leave.	d.	It's not the case that you must stay.
(12)	a.	Everyone left.	c.	It's not the case that someone stayed.
	b.	Someone left.	d.	It's not the case that everyone stayed.

(11a) entails (11b), just like (12a) entails (12b):[3] if the (a) sentence is true, the (b) sentence has to be true as well. Thus, following (a) with the negation of (b) results in a contradiction, indicated with '#': *#You must leave, but you may not leave*. (11a) and (11c) are logically equivalent (they entail each other), just like (12a) and (12c). (11b) and (11d) are too, as are (12b) and (12d).

[2]As a potential counterexample, the adjective *possible* can be used to express epistemic and root possibility. Note, however, that the flavor depends on the finiteness of its complement: *it's possible that Jo is the murderer* (epistemic) vs. *it's possible for Jo to register online* (root).

[3]Assuming the domain is not empty, i.e., that there is at least one individual.

Standard semantic approaches to modals derive these equivalences via a quantificational analysis: while *every* and *some* involve quantification over *individuals*, *must* and *may* involve quantification over "*possible worlds*". Necessity modals are **universal quantifiers**: they quantify over *all* (relevant) worlds, just like *everyone* quantifies over all (relevant) individuals. Possibility modals are **existential quantifiers**: they quantify over *some* (relevant) world, just like *someone* quantifies over some (relevant) individual. We go over this analysis in the next section.

2. Capturing force and flavor in modal logic

The most standard analyses of modals in formal semantics derive from modal logic, where they are treated as quantifiers over *possible worlds* (Carnap 1947, von Wright 1951, Kanger 1957, Hintikka 1961, Kripke 1963). Possible worlds can be viewed as possible '*ways things could have been*' (Lewis 1973). There are countless ways things could have been, each of which represents a different possible world. We can imagine a world just like ours, but where my doctor's appointment started on time. We can imagine another one where it started one minute late, another where it started two minutes late, or one where the doctor didn't show up at all, etc. Each of these possibilities can be viewed as a different world. *Possible worlds* raise fundamental issues as to their metaphysical and psychological plausibility, which semanticists acknowledge, but typically put aside. They assume that human languages have the capacity to represent alternative states of affairs, and use possible worlds as a mere formal tool to represent such alternative states of affairs in the language of the semantic theory.

Possible worlds allow us to formally model the displacing role of modals into states of affairs beyond the here and now. The truth of any statement is evaluated relative to our world, which serves as the world of evaluation. For instance, the sentence "Jo is home" is true in our world if

Jo is home in our world. Modals (and other so-called intensional operators, such as attitude verbs) are special in *shifting* the world of evaluation: the truth of a modal statement in *our* world depends on the truth of the proposition expressed by the modal's complement (its "prejacent") in some *other* world(s). For instance, regardless of whether Jo actually is home, the sentence "Jo might be home" is true in our world, call it *w*, just as long as there is a relevant world, call it *w'*, where Jo is home.

Necessity modals are treated as *universal* quantifiers: they quantify over *all* relevant worlds. In all of these worlds, the proposition expressed by the modal's prejacent is true (logicians use the symbol □ "box" for necessity). Possibility modals are treated as *existential* quantifiers: they quantify over *some* relevant world; in *some* such world, the proposition expressed by the prejacent is true (logicians use the symbol ◊ "diamond" for possibility). (13) illustrates this with truth conditions for the possibility and necessity statements in (11):

(13) a. "You must leave" *is true in w if in* ***all*** *(relevant) worlds w' you leave*

b. "You may leave" *is true in w if in* ***some*** *(relevant) world w' you leave*

This quantificational treatment captures the logical inferences above. If you leave in *all* worlds, then there is *some* world in which you leave: (11a) entails (11b). If you leave in *all* worlds, then there is *no* world in which you stay: (11a) entails (11c) and *vice versa*. If you leave in *some* world, then you do not stay in *all* worlds: (11b) entails (11d) and *vice versa*. This analysis thus derives the relevant relations between possibility and necessity modal statements. It however doesn't allow for further graded notions of possibility or necessity, as quantification is either over *some* or *all* relevant worlds.

The quantificational analysis captures the *flavor* dimension of modality by restricting the domain of quantification to various sets of worlds. With *deontic* modality, the relevant worlds are those compatible with relevant rules in our world; with *bouletic* modality, those compatible with relevant desires; with *epistemic* modality, those compatible with what is known...

What does it mean for a world to be compatible with relevant rules? Imagine that the rules in our household are that children go to bed before 8pm, that they brush their teeth twice a day, that they do not watch TV, and do not whine. A deontically ideal world from our world's perspective then is one where all children go to bed before 8pm, brush their teeth twice a day, and never watch TV, or whine. The deontic necessity statement in (14a) is true in our world *w* if in all such deontically ideal worlds, Jo goes to bed. The deontic possibility statement in (14b) is true if Jo reads a story in at least one of these worlds, even if she doesn't necessarily do so in all of them.

Epistemic modality involves worlds in which all of the known facts in our world hold. These may include the fact that Jo is not in the living room, that it's 8pm, that Jo likes to go to bed after dinner, etc. The epistemic necessity statement in (14c) is true in our world *w* if in all worlds in which all of these facts hold, Jo is in bed.

(14) a. "Jo must go to bed."$_{\text{deontic}}$

is true in w if in all w' compatible with the rules in w, Jo goes to bed.

b. "Jo may read a story."$_{\text{deontic}}$

is true in w if in some w' compatible with the rules in w, Jo reads a story.

c. "Jo must be in bed."$_{\text{epistemic}}$

is true in w if in all w' compatible with what is known in w, Jo is in bed.

Note that (14) gives truth conditions for the entire modal statements, without addressing yet the exact contribution of the modal words themselves. This question will be the focus of section 4. Note further that truth conditions is *all* that is provided in (14): what these truth conditions amount to are *descriptions* of worlds that are ideal from our world's perspective. But a speaker uttering a sentence like (11a) or (14a) often seems to do more than merely *describe* ideal states of affairs: she seems to *demand* that you leave or that Jo go to bed. Similarly, the use of (11b) or (14b) seem to *grant* a permission for you to leave or Jo to read a story. The next section addresses how such demands and permission come about: are they part of the conventional meaning of a *must* or *may* statement, or a pragmatic by-product of an act of describing a necessity or possibility?

3. Modal statements: sentence meaning vs. utterance meaning

Speakers often mean more than the words they utter. We thus need to distinguish the *literal meaning* of sentences from the *speaker meaning* that speakers convey in using them. The former falls within the purview of semantics, the latter that of pragmatics (see Schwarz & Zehr, this volume). Modals are often used in textbook examples of speaker meaning going beyond literal meaning. Here we discuss two cases: the first involves so-called "scalar implicatures", the second "indirect speech acts".

3.1. Scalar implicatures

Recall the logical equivalences in (11), and reconsidered in (15) and (16) below. As we saw, the (a) and (c) sentences entail each other. These entailments are a *semantic* matter: they follow from the *literal* meaning of the sentences. The (a) sentences also seem to imply the (d) sentences

(which correspond to the negation of the (b) sentences). However these inferences are not *entailments*: (a) can be true, but (d) false. "*Someone left, in fact everyone did"* is not a contradiction, and neither is "*You may leave, in fact you have to*".

(15) a. You may leave. c. It's not the case that you must stay.
b. You must leave. d. It's not the case that you must leave.

(16) a. Someone left. c. It's not the case that everyone stayed.
b. Everyone left. d. It's not the case that everyone left.

Grice (1975) coined the term 'implicature' for speaker meanings that give rise to such cancelable inferences. Hearers grasp a speaker's implicatures by considering what they literally said, and explaining this choice in relation to what they could have said, but chose not to. A speaker utters (15a). She could have said something stronger, namely, the *necessity* statement in (15b). By using the *weaker* statement in a context where the *stronger* statement would have been relevant, she might intend to convey that the stronger statement does not hold: if it had, she would have said it instead. The same reasoning holds for (16): by uttering (16a) in a context where (16b) would have been relevant, the speaker can imply that that stronger statement doesn't hold. These types of implicatures are called 'scalar',[4] because they involve "scales" (Horn 1972),

[4]Scalar implicatures are a subtype of 'quantity implicatures', where Grice's maxim of **quantity** ('*Make your contribution as informative as is required*') comes in apparent conflict with the maxim of **quality** ('*Don't say what you believe to be false or lack adequate evidence for*'). The implicature is understood by reasoning that the speaker is not in a position to assert a more

which are conventionalized associations of lexical items, ordered in terms of informativity (<*may, must*>, <*some*, *every*>…). Because of their conventional nature, scalar implicatures tend to be fairly routinely triggered.

3.2. Indirect speech acts

Modals are often featured in *indirect speech acts*, with which a speaker performs an *illocutionary act* by way of performing another, direct, one (Searle 1969). A classic example is shown in (17). The direct act performed with (17)—the one that aligns with its literal meaning—is a *question* about whether the addressee has a certain ability. But (17) is most naturally used as an (indirect) *request* to pass the salt (Searle 1975). This request arises by reasoning that the ability in question is so trivial that the speaker couldn't possibly just want to know about it; instead she must want that ability instantiated. Speakers often use such circumlocutory ways to soften requests and avoid issuing direct commands.

(17) Can you pass the salt?

In examples like (17), it is rather clear that the two illocutionary acts (the question and the request) are distinct. But with some modal statements, the line between direct and indirect act is often blurred. Because modal statements typically express possibilities and necessities relative to the speaker's beliefs or desires, it is sometimes difficult to tell whether a modal statement merely describes an epistemic or deontic possibility consistent with her beliefs or desires, or whether the

---

informative statement, and thus, given compliance with quality, must either not have enough evidence for it, or believe it to be false.

speaker *directly* proffers some degree of confidence or issues a command. Both positions are actively debated in the literature.

A speaker can use a *deontic* modal claim to grant a permission or issue an obligation. Are these acts part of the *conventional* meaning of the modal statement, or do they arise indirectly? A natural use of (18a), for instance, issues an order not unlike the imperative "*Stay!*". Is this order part of the *literal* meaning of (18a), or does (18a) merely *describe* an obligation? To get a feel for the difference, consider (18b). (18b) is not a direct order: it describes an ideal state of affairs for me, the speaker. However, by using it, I can indirectly urge my addressee to stay, by letting her know that it would bring me happiness. We know that (18b) does not directly encode an order to stay, because it can be followed by an imperative ordering the addressee to go ("It'd make me happy if you stayed, but go! I know you need to"). The situation is however less clear with (18a), as following it with the same imperative seems infelicitous (#"You must stay, but go!"). This kind of infelicity has led to proposals where modals like *must* have a 'performative dimension', that is, that part of their conventional meaning is the issue of a command to the addressee (see Ninan 2005, Portner 2009). Alternatively, under the view that (18a) merely *describes* a necessity, the infelicity could be due to the described necessity being relative to rules endorsed by the speaker.

(18) a. You must stay.

b. It'd make me happy if you stayed.

A speaker can use an *epistemic* modal claim to express her certainty about some state of affairs. Because of this, epistemic modals have sometimes been treated as not directly

contributing to the literal, truth conditional, content of the sentence in which they appear, but as mere indicators of the speaker's degree of certainty (Halliday 1970, Palmer 2001, Swanson 2006, a.o.). Alternatively, under the view that modal statements merely describe possibilities or necessities, the expression of certainty could arise indirectly as a pragmatic side effect. For instance, by uttering (19a), the speaker would describe a necessity relative to facts that she believes, and in virtue of this, she would indicate her certainty about Jo's guilt.

(19) a. Jo must be guilty.

b. Jo is guilty.

A complication arises with epistemic necessity statements. A sentence like (19a) seems to make a weaker claim than its unmodalized counterpart (19b), that is the use of (19b) seems to require more confidence or better evidence than the use of (19a) (Karttunen 1972). This fact is puzzling given a standard account of epistemic necessity like the one sketched in section 1. (19a) should entail (19b): if Jo is guilty in all worlds compatible with the relevant known facts, then she should be guilty in our world, which is consistent with these facts. This apparent weakness of *must* statements requires amendments to the standard view (see section 4.2.3), but it can be straightforwardly captured under a performative account, if epistemic modals encode as part of their conventional meaning a weaker commitment from the speaker than a bare assertion.

4. Decomposing modal sentences: modals in the lexicon

So far we have focused on the meaning contribution made by entire modal statements. This section focuses on the contribution of the modal words themselves. Recall that modal sentences

vary along two dimensions: force and flavor. In a language like English, *force* remains constant, but the same sentence can be used to express different modal *flavors*. How does flavor multiplicity arise? Is this a case of *generality* (one general possibility meaning unspecified for flavor) or *ambiguity* (the same string of words corresponds to distinct senses)? And what do the modal words themselves contribute?

Before we turn to these questions, let's briefly discuss languages in which the same modal can be used in situations where English speakers would use either a possibility or a necessity modal. Are such modals indeterminate between a possibility and a necessity meaning, or do they represent genuine cases of ambiguity? Neither option seems likely. While proposals for such "variable force" modals differ for various languages, all converge on providing the underlying modal a stable meaning (either as possibility or necessity, or something more complicated), and where speakers make do, using their modal(s) in a wider range of situations than speakers of languages with modals of both forces.

4.1. Meaning indeterminacy

Some linguistic expressions can be used in very different situations. This is due sometimes to genuine *ambiguity*, where two distinct thoughts are expressed by the same sounds, and sometimes to *generality*, where a single, general, thought is expressed, which is compatible with multiple situations. The words *pen* (writing instrument) and *pen* (animal enclosure) present a classic example of lexical ambiguity (two distinct lexical entries), the word *teacher* one of generality (one single lexical entry). Evidence for this distinction comes from Zwicky & Sadock's (1975) Identity of Sense tests, which rely on conjoined structures, as in (20a) and (21a), and elided structures, as in (20b) and (21b).

(20) a. Jo and Al have beautiful pens.

b. Jo has a beautiful pen. Al does too.

(21) a. Jo and Al are teachers.

b. Jo is a teacher. Al is too.

The sentences in (20) require that Jo and Al have the same sorts of things: either they both have beautiful writing implements or they both have beautiful animal enclosures. A 'mixed reading' is impossible. This indicates that *pen* does not express a single general sense that applies equally to Bic pens and pig pens, but is rather ambiguous between two senses. Meanwhile, while the sentences in (21) require that both Jo and Al be teachers, they permit them to teach different subjects. Thus *teacher* is not ambiguous, but expresses a single sense which is simply neutral as to the subject of teaching.

Polysemy shares characteristics with both generality and ambiguity. A *polysemous* term has several senses, somehow related to each other. For instance, we can use "book" to refer to a physical object or to its content. Is this because "book" has a single general sense that applies equally to both? If it were, why wouldn't two copies of Moby Dick count for three books, two paperbacks plus their shared content? Perhaps we should instead say that "book" has multiple senses, differing in their extensions. Unfortunately it is tricky to demonstrate that a term is polysemous, rather than ambiguous or general, using Zwicky & Sadock's tests. Often a plausibly polysemous term, like an ambiguous term, will forbid mixed readings: "Jo and Al have beautiful books" can't seem to describe a situation where Al has a beautiful physical object, and Jo one

with beautiful content. But sometimes a mixed reading is permitted, even with the very same word: "The book is heavy but informative" (Liebesman & Magidor 2017)[5].

When we apply these tests to a modal statement, we see that a mixed reading seems disallowed, suggesting that its meaning is not general. In (22) for instance, the modality is either epistemic or deontic for both individuals: either Jo and Al are both required to eat meat, or they are both likely meat-eaters. The sentences cannot be used to mean that one is a likely meat-eater and the other is required to eat meat.

(22) a. Jo and Al must eat meat.

b. Jo must eat meat. Al too.

This suggests that *must* statements are ambiguous between modal flavors, and do not merely describe a general, unspecified, necessity. But what is the *source* of the ambiguity? Is the word *must itself* ambiguous, is it polysemous, or is the source of the ambiguity external to the word? Linguists from a functionalist tradition[6] tend to assume that modals are polysemous, and thus come in multiple senses (e.g., Sweetser 1990). But in her seminal account of modality, the formal semanticist Angelika Kratzer argues that there is just one *must* and just one *may*, and that the ambiguity arises from additional elements involved in modal statements, either other parts of the sentence, or interactions with the context of speech.

[5]For different views on polysemy, see Carston (2002), Pietroski (2005), Asher (2011), Vicente (2018).

[6]Functional linguistics contrast with formal approaches in aiming to explain grammatical facts from the way language is used.

4.2. Lexical ambiguity?

4.2.1. Against lexical ambiguity

A modal like *must* can be used to express different flavors of modality, including epistemic and deontic ones. Does this mean that we should have as many lexical entries (or senses) for *must* as there are flavors? Kratzer (1977) argues that the problem is made worse by the fact that each flavor comes in many subflavors. The sentence in (23), for instance, can describe obligations relative to different rules: family rules, rules governing school, cities or entire countries... How many distinct *musts* should we have?

(23) Children must wear seatbelts.

Kratzer argues that we further need an extra neutral *must*, for cases like (24), where the flavor seems to be determined by an overt phrase. But once we have a neutral *must*, Kratzer contends, why not assume that there is just this one neutral *must*, which gets its flavor from another expression in its grammatical context, either explicitly, as in (24), or implicitly, as in (23)?

(24) According to DC law, children must wear seatbelts.

A further argument against lexical ambiguity is the fact that multiplicity of flavors is not just a quirk of English, or even Indo-European, as it can be found in many unrelated languages. If this is a lexical accident, why should it occur in language after language? A polysemist might

retort that unlike in the case of *pen*, the different senses a modal can express are related: all express some kind of necessity. There may thus be natural reasons for this accident to keep on occurring (functionalists invoke notions like metaphorical extension, for instance). All else equal, however, the Kratzerian view is more parsimonious. But languages may not abide by the Kratzerian ideal, and parsimony may not be maintainable against the full empirical picture. We turn to the Kratzerian account and empirical arguments that threaten its viability next.

4.2.2. Kratzerian theory

According to the classical Kratzerian theory (Kratzer 1981, 1991), there is just one *must*, unspecified for flavor. Flavor gets determined by a restriction, called "conversational background", which provides the set of worlds the modal quantifies over, with different restrictions giving rise to different flavors. This conversational background can be supplied by context when not overt. An epistemic conversational background, for instance, provides a set of worlds compatible with what is known in our world.

The slightly simplified lexical entries in (25) illustrate. Here *p* is the proposition expressed by the modal's prejacent, and *f(w)* picks out the set of worlds determined by a conversational background *f* at a world *w*. The meanings differ just in the force of quantification over this set: universal quantification for *must*, and existential quantification for *may*. We can see how in this system, there is no lexical ambiguity: we have just one *must* and just one *may*, neither of which specified for flavor.

(25) a. At a world *w*, "**must**" names a relation between propositions *p* and backgrounds *f*, true just when *p* is true in **all** worlds *w'* compatible with *f(w)*

b. At a world *w*, "**may**" names a relation between propositions *p* and backgrounds *f*, true just when *p* is true in **some** world *w'* compatible with *f(w)*

Kratzer's theory not only explains flavor multiplicity without invoking ambiguity,[7] it also overcomes empirical challenges that the original quantificational analysis faced. Seeing this requires introducing a further complication in Kratzer's system. In the full system, modals are relative not to just one, but *two* conversational backgrounds. This double relativity is motivated in part by puzzles of deontic modality. Imagine that Jo has committed a crime, for which the law requires that she goes to jail. We can report this with the deontic necessity statement in (26), which should be true if Jo goes to jail in all worlds compatible with the law. But how could Jo have committed a crime in these worlds? Such worlds are supposed to be crime free! Kratzer argues that we can solve this conundrum by separating facts from ideals. Jo has committed a crime: this is an irrevocable fact. We can nonetheless invoke worlds that best fit the law amongst those imperfect worlds in which this crime has occurred: in all such worlds, Jo goes to jail.

(26) Jo must go to jail.

[7]Formally, a conversational background is a function from worlds to sets of propositions (e.g. propositions that describe some known facts). One problem with this purely contextual account of flavor determination is that nothing about a set of propositions makes it inherently epistemic vs. deontic. Consequently, whatever meaning differences we find between a modal like *might* (which is typically only associated with epistemic and metaphysical uses) vs. one like *can* (which is typically only associated with root uses) can't be attributed to a difference in value of the conversational background as determined by context (Nauze 2008, Kratzer 2012, Harr 2019).

To implement this separation of facts and ideals, Kratzer proposes to make modals relative to two conversational backgrounds. The first, called the *modal base*, is based on facts: it picks out an initial set of worlds in which certain facts hold (for instance, the fact that Jo committed a crime). The second, called the *ordering source*, is based on various ideals, such as laws, needs, or desires. The modal ends up quantifying over a subset of the initial set of worlds, namely those that best fit the ideal (for instance, those that best obey the law).

This double relativity provides a solution to Karttunen's puzzle, mentioned in section 3.2. Recall that an epistemic necessity statement like (19a) intuitively feels weaker than its unmodalized counterpart (19b). However, a necessity modal should quantify over *all* worlds compatible with the known facts, including our world, thus (19a) should entail (19b). Kratzer argues that the entailment fails to go through because the modal doesn't necessarily quantify over *all* epistemic worlds. The modal takes an epistemic modal base, which does pick out all of the worlds compatible with the known facts, including our world, but it also involves a stereotypical ordering source, which pares down this set of worlds to only those that best fit stereotypical expectations. Our world may not be stereotypical, thus (19a) need not entail (19b).

Double relativity also allows for some graded notions of possibility. Recall that in our initial quantificational analysis, modals either quantified over some or all relevant worlds. We thus couldn't distinguish between slight and regular possibility, or strong and weak necessity. Double relativity allows worlds to be *ranked* according to some ideal, which in turn allows for talk about better or worse possibilities.[8]

[8]Some of the finer-grained gradability that modal adjectives in particular can express may require more than double relativity; for an overview, see Lassiter (*to appear*).

4.2.3. Lexical ambiguity after all?

In the last few decades, new empirical data has been brought to light, arguing that the different flavors expressed by a modal like *must* cannot share the same lexical entry. First, as discussed in section 3.2, modals can be used to perform different speech acts: deontics to give orders or permissions, epistemics to indicate a degree of certainty. The various accounts that derive these speech acts by encoding a performative dimension to the modal words themselves necessarily assume lexical ambiguity, given that the speech act performed differs by flavor. However, this challenge may be easy to skirt, if these speech acts can be derived in a more pragmatic way, as discussed earlier. A speaker using a deontic *must* sentence, for instance, would literally merely describe a deontic necessity, but could indirectly issue a command in virtue of counting on her audience to understand why she is describing what is necessary: namely that she endorses the rules that underwrite and wants her audience to comply with them.

A second and perhaps more worrisome concern for the Kratzerian view is that cross-linguistically, modals tend to interact differently with elements like tense, depending on the flavor they express.[9] To see this in English, we need to turn to the semi-modal *have to*, which, unlike modal auxiliaries, can be tensed. Consider (27). With a deontic interpretation, (27) expresses a *past* obligation: *Jo was required to be home*. With an epistemic interpretation, however, the sentence can express a *current* necessity about a past state of affairs: *it is necessary, in virtue of what we now know, that Jo was home*.

[9]This is part of a more general pattern: epistemic and root modals differ in their interactions with aspect, negation, quantifiers, and other modals. For an overview see Hacquard (2011).

(27) Jo had to be home at the time of the crime.

Not only can *had to* report a present epistemic necessity, it may have to, unlike a past tensed verb like *seemed*. Consider the scenario in (28). (28a) reports a past seeming state, which no longer holds. But if we try to make roughly the same point with (28b), it sounds distinctly odd. This suggests that it can only report a *current* epistemic state, which conflicts with the continuation denying that Jo was home.[10]

(28) Earlier this week, the available evidence pointed us in the wrong direction. For instance,

a. while Jo **seemed to** be home at the time of the crime, we now know that she wasn't.

b. ??while Jo **had to** be home at the time of the crime, we now know that she wasn't.

These facts are usually captured as a matter of *scope*, that is, the position where the modal has to be interpreted relative to elements like tense.[11] Epistemic modals scope above tense; hence

[10]Putting the modal in an adjunct helps sharpen the contrast (A. Williams, p.c.), see van Dooren (*to appear*) for experimental support.

[11]To be precise, the generalization is not strictly speaking that *had to* can't report a past epistemic necessity. "*I thought that Jo had to be home*" clearly reports a necessity at a *past* thinking time. In that sentence, the past tense on the main verb, *'thought*,' introduces the relevant past time, and the necessity is simply concurrent with the thought. The question is whether the modal can scope under the tense in its *own* clause (not the tense of a higher clause, like the past tense on *thought*). And this is a matter of debate. Some argue that it must, and question data like (27) and (28) (e.g. Rullmann & Matthewson 2018). Others argue that it cannot, and that

their time of evaluation is not affected by a past tense that appears in the same clause. Deontic (and other root) modals, on the other hand, scope below tense, hence their evaluation time is shifted by a past tense (Picallo 1990, Stowell 2004, Hacquard 2006, a.o.).

This scope difference goes against the Kratzerian view: if deontic and epistemic modals scope in different positions how can they share a lexical entry? We may need (at least) two separate *musts*, each specified for flavor *and* scope. There are ways to maintain the Kratzerian view by complicating how the modal combines with its restriction: there would still just be one neutral *must*, which could freely appear above or below tense, but different *modal bases* would be available in different positions (Hacquard 2006, 2010, Kratzer 2012). But now, parsimony might argue for ambiguity. This apparent advantage however seems to diminish when we consider issues of learnability in the next section. On any analysis of modals, the child needs to recognize the association between flavor and scope. Since this association is not an idiosyncratic fact about each individual modal, it seems wrong to encode it via lexical stipulations. Thus in the end, the Kratzerian view might on balance be the simplest.

5. Learnability issues

Children acquiring the modals of their language presumably face the same challenges as semanticists trying to figure out their meanings. The first challenge we discussed, namely the difficulty in untangling the semantic and pragmatic contributions of modal statements is further amplified for children, who are only ever exposed to speaker meanings. How do they extract the

---

counterexamples involve special contexts, which introduce higher operators responsible for shifting the modal's evaluation time (e.g. Stowell 2004, Boogart 2007, Hacquard 2006).

literal content of speakers' modal claims? Might they have certain expectations about how modal meanings are packaged in natural language?

The second challenge comes from the fact that the same modal word can express different flavors, but that these flavors interact differently with tense. Children not only need to figure out that a word like *have to* can express epistemic and deontic necessity, but also that with an epistemic flavor (but not a deontic one), it outscopes tense. What prevents them from assuming that *have to* uniformly scopes below tense, just like any tensed verb? Positive evidence for epistemics scoping over tense is virtually absent in speech to children: epistemic modals are infrequent, and almost always occur in the present tense.[12] It can't just be a matter of notional meaning (e.g., an incompatibility between epistemic meanings and pastness), since predicates with epistemic meanings like *know, seem*, or *be likely* all happily scope below tense.

These scope facts create problems for both the classical Kratzerian account and for a pure lexical ambiguity account, the former because it cannot tie particular flavors to particular scopal positions, the latter, because if scope is a lexical idiosyncrasy of two homonyms, it is one impossibly hard to detect. There seems to be something special about grammatical modals that goes beyond the meanings they express, which dictates their scopal behavior, and which learners have to somehow be privy too. Cinque (1999) proposes that functional elements (which is roughly to say closed class or grammatical elements), like tense and modals, are rigidly and universally organized in a particular order, with epistemic modals above, and root modals below tense. Because Cinque's hierarchy proposes that different flavors of modals occupy different

[12]Only one epistemic *had to* was found out of 2,400 occurrences of *have to* in maternal speech in the Manchester corpus (339,795 utterances) (Theakston et al. 2001), analyzed in van Dooren et al. 2017.

positions, it fits well with ambiguity proposals where different lexical entries for different flavors can occupy different positions, and a learner equipped with something like Cinque's hierarchy might expect a functional (*i.e.* grammatical) epistemic modal to scope above tense, unlike a verb with an epistemic meaning. Another possibility, alluded to at the end of section 4, is to tie *modal bases* to particular positions, rather than the modals themselves. If that is right, the spirit of the Kratzerian view might be maintainable. Either way, there is something particular to grammatical modals which allows them to express different flavors of modality, but constrains their scopal behavior based on the flavor expressed, and which both children and semanticists need to figure out. (for more on the acquisition of logical vocabulary, see Crain, this volume).

6. Conclusion

Languages have various means of expressing notions of possibility and necessity. A language like English has a dedicated grammatical category of modals, words like *can* or *must*, alongside verbs, adverbs, nouns and adjectives that express similar notions. Because modal statements involve a complex interplay of morphology, syntax, semantics, and pragmatics, isolating the exact lexical contributions of the modal words themselves presents challenges for both semanticists and children acquiring their language. The first is that it isn't always easy to untangle the semantic and the pragmatic contributions of modal statements, given that modals are routinely used to perform illocutionary acts that go beyond mere descriptions of possibilities and necessities. The second is that the same words can be used to express different flavors of modality, raising questions as to whether each of these words comes in different lexical entries in our mental lexicon. However way these questions ultimately get resolved, modals provide a rich terrain to explore how meaning gets packaged in natural language.

References

Asher, Nicholas (2011). *Lexical Meaning in Context*. Cambridge: Cambridge University Press.

Austin, John (1962). *How to do things with words*. Oxford: Clarendon Press.

van der Auwera, Johan and Andreas Ammann (2011). 'Overlap between situational and epistemic modal marking'. In M. Dryer & M. Haspelmath (eds) *The world atlas of language structures online*. Munich: Max Planck Digital Library.

Bochnak, Ryan (2015). 'Variable force modality in Washo'. In T. Bui and D. Özyıldız (eds) *Proceedings of North East Linguistic Society* 45: 105-114.

Boogart, Rony (2007). 'The past and perfect of epistemic modals'. In L. de Saussure, J. Moescher, and G. Puskas (eds) *Recent advances in the syntax and semantics of tense, aspect and modality*. Berlin and New York: Mouton de Gruyter.

Carnap, Rudolph (1947). *Meaning and Necessity*. Chicago: University of Chicago Press.

Carston, Robyn (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell Publishers.

Cinque, Guglielmo (1999). *Adverbs and functional heads: A cross-linguistic perspective*. Oxford and New York: Oxford University Press.

Deal Amy Rose (2011). 'Modals without scales'. *Language* 87.3, 559-585.

van Dooren, Annemarie, Anouk Dieuleveut, Ailis Cournane, and Valentine Hacquard (2017). 'Learning what *must* and *can* must and can mean'. In *Proceedings of the Amsterdam Colloquium*, 225-235.

van Dooren, Annemarie (to appear). 'The temporal perspective of epistemics in Dutch'. *Proceedings of Sinn und Bedeutung* 24.

Fintel, Kai and Irene Heim (2011), Intensional semantics, unpublished lecture notes.

Fintel, Kai and Sabine Iatridou (2008). How to say ought in foreign: The composition of weak necessity modals. *Time and modality*. Dordrecht: Springer, 115-141.

Frege, Gottlob (1915). My basic logical insights. *The Frege Reader*, 321-324.

Frege, Gottlob (1918). *The thought: A logical inquiry*. *Mind* 65.259, 287-311.

Grice, Herbert P. (1975). 'Logic and conversation'. *Speech acts*. Brill, 41-58.

Grice, Herbert P. (1989). 'Studies in the way of words'. Cambridge: Harvard University Press.

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

Hacquard, Valentine (2010). 'On the event relativity of modal auxiliaries'. *Natural language semantics* 18.1, 79-114.

Hacquard, Valentine (2011). 'Modality'. In C. Maienborn, K. von Heusiger, and P. Portner (eds) *Semantics: An International Handbook of Natural Language Meaning*. HSK 33.2, 1484-1515.

Halliday, Michael AK. (1970). 'Functional diversity in language as seen from a consideration of modality and mood in English'. *Foundations of language*, 322-361.

Harr, Quinn (2019). *Epistemic "might"*: *a non epistemic analysis*. PhD. Dissertation. University of Maryland.

Hintikka, Jakko (1961). 'Modality and quantification'. *Theoria* 27: 119-128.

Hintikka, Jakko (1962). 'Knowledge and belief: An introduction to the logic of the two notions'. Ithaca: Cornell University Press.

Hoffmann, T. Ronald (1966). 'Past tense replacement and the modal system'. In A. Oettinger (ed), *Mathematical Linguistics and Automatic Translation*. Cambridge, MA: Harvard University, Harvard Computational Laboratory.

Horn, Larry (1972). *On the semantic properties of logical operators in English*. Ph.D. Dissertation. UCLA.

Kanger, Stig (1957). 'The morning star paradox'. *Theoria* 23: 1-11.

Karttunen, Lauri (1972). 'Possible and must'. *Syntax and semantics* 1.

Kratzer, Angelika (1977). 'What *must* and *can* must and can mean'. *Linguistics and Philosophy* 1.

Kratzer, Angelika (1981). 'The notional category of modality. In H.-J. Eikmeyer and H. Rieser (eds) *Words, worlds, and contexts. New approaches in word semantics*. Berlin: Mouton de Gruyter.

Kratzer, Angelika (1991). 'Modality'. In: A. von Stechow and D. Wunderlich (eds) *Semantik: Ein Internationales Handbuch Zeitgenoessischer Forschung*. Berlin: Mouton de Gruyter: 639-650.

Kratzer, Angelika (2012). *Modals and conditionals: New and revised perspectives*. Oxford: Oxford University Press.

Kripke, Saul (1973). 'Semantical analysis of modal logic, normal propositional calculi'. *Zeitschrift für mathematishe Logic und Grundlagen der Mathematik* 9: 67-96.

Lassiter, Daniel (*to appear*). 'Graded Modality'. In L. Matthewson, C. Meier, H, Rullmann, and E. Zimmermann (eds) *The Companion to Semantics*. Oxford: Wiley.

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

Liebesman, D. and Magidor, O. (2017), Copredication and Property Inheritance. Philosophical Issues, 27: 131-166.

Nauze, Fabrice (2008). *Modality in typological perspective*. PhD. Dissertation. University of Amsterdam.

Ninan, Dilip (2005). Two puzzles about deontic necessity. In J. Gajewski, V. Hacquard, B. Nickel, and S. Yalcin (eds) *New work on modality*. MITWPL.

Palmer, Frank R. (2001). *Mood and modality*. Cambridge: Cambridge University Press.

Picallo, Carme (1990). 'Modal verbs in Catalan'. *Natural Language and Linguistic Theory* 8.

Pietroski, Paul (2005). 'Meaning Before Truth.' *Contextualism in Philosophy*. Preyer and Peters, (eds.). Oxford: Oxford University Press.

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

Prior, A.N. (1957). *Time and modality*. Oxford: Oxford University Press

Matthewson, Lisa, Hotze Rullman and Henry Davis (2007). 'Evidentials as Epistemic Modals: Evidence from St'at'imcets'. *The Linguistics Variation Yearbook* 7.

Rullmann Hotze and Lisa Matthewson (2018). 'Towards a theory of modal-temporal interaction'. *Language* 94.2, 281-331.

Searle, John R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Searle, John R. (1975). Indirect speech acts. *Speech acts*. Brill, 59-82.

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

Swanson, Eric (2006). *Interaction with context.* Ph.D. Dissertation. Massachusetts Institute of Technology.

Sweetser, Eve (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.

Theakston, Anna, Lieven, Elena, Julian Pine, and Caroline Rowland (2001). 'The role of performance limitations in the acquisition of verb-argument structure: An alternative account'. *Journal of child language*, *28*(1), 127-152.

Traugott, Elizabeth (2011). 'Modality from a historical perspective'. *Language and Linguistics Compass* 5.6, 381-396.

Vander Klok, Jozina (2013). 'Pure possibility and pure necessity modals in Paciran Javanese. *Oceanic Linguistics* 52.2, 341-374.

Vicente, Agustin (2018). 'Polysemy and word meaning: An account of lexical meaning for different kinds of content words'. *Philososphical Studies* 175: 947-968.

von Wright, Georg (1951). *An essay on Modal Logic*. North-Holland, Amsterdam.

Yanovich, Igor. (2016). 'Old English **motan*, variable force modality, and the presupposition of inevitable actualization'. *Language* 92:3, 489-5121.

Zwicky, A. M. and J. M. Sadock (1975). 'Ambiguity tests and how to fail them', *Syntax and Semantics* 4, 1-36.