Bounded Modality*

Matthew Mandelkern
All Souls College, Oxford

September 25, 2018

Penultimate draft; to appear in *The Philosophical Review*

Abstract

What does 'might' mean? One hypothesis is that 'It might be raining' is essentially an avowal of ignorance like 'For all I know, it's raining'. But it turns out these two constructions embed in different ways—in particular as parts of larger constructions like Wittgenstein (1953)'s 'It might be raining and it's not' and Moorean sentences like 'It's raining and for all I know, it's not' (Moore, 1942). A variety of approaches have been developed to account for those differences. All approaches agree that both Moore sentences and Wittgenstein sentences are classically consistent. In this paper I argue against this consensus. I adduce a variety of data which I argue can best be accounted for if we treat Wittgenstein sentences as being classically inconsistent. This creates a puzzle, since there is decisive reason to think that ⌜Might p⌝ is consistent with ⌜Not p⌝. How can it also be that ⌜Might p and not p⌝ and ⌜Not p and might p⌝ are inconsistent? To make sense of this situation, I propose a new theory of epistemic modals which aims to account for their subtle embedding behavior and shed new light on the dynamics of information in natural language.

Keywords: *epistemic modality; semantics; local contexts; connectives; attitudes; conditionals; quantifiers*

1 Introduction

Moore (1942) observed that there is something wrong with asserting a sentence like (1):[1]

(1) #It's raining but I don't know it's raining.

*I am grateful to audiences at MIT, the 2016 Rutgers NASSLLI, Northwestern University, Western Michigan University, the Institut Jean-Nicod, PALLMYR 11 at UCL, the University of Pennsylvania, and NYU; to two anonymous referees for this journal, Amir Anvari, Ashley Atkins, David Boylan, Fabrizio Cariani, Sam Carter, Cleo Condoravdi, Kevin Dorst, Daniel Drucker, Vera Flocke, Simon Goldstein, Ben Holguín, Matthias Jenny, Angelika Kratzer, Maša Močnik, Dilip Ninan, Jonathan Phillips, Milo Phillips-Brown, Jim Pryor, Jacopo Romoli, Ginger Schultheis, Daniel Skibra, Shane Steinert-Threlkeld, and Seth Yalcin for very helpful discussion. I thank Cian Dorr, Daniel Rothschild, Paolo Santorio, and Philippe Schlenker for their extensive and invaluable feedback. This paper develops the first chapter of my dissertation (Mandelkern, 2017), and I am particularly grateful to my committee members—Kai von Fintel, Irene Heim, Justin Khoo, Robert Stalnaker, and Stephen Yablo—for their support and feedback.

[1]Moore's sentences were formulated in terms of belief; I stick with knowledge here to make the parallel to Wittgenstein sentences closer, and avoid issues involving neg-raising. The factivity of 'knows' does not play an essential role here (and indeed 'knows' will be necessarily interpreted in a non-factive way in the key examples below): all my points can be made with non-factive variants, provided that care is taken to avoid neg-raising readings. I move freely between using 'and' and 'but' as a conjunction, assuming they have relevantly similar semantics; readers can confirm that alternating between them does not substantially change felicity judgments. I use '#' to indicate pre-theoretic judgments of infelicity.

Generally speaking, *Moore sentences* with the form ⌜p, but I don't know p⌝ or ⌜I don't know that p, but p⌝ tend to be unassertable. In an early discussion of Moore sentences, Wittgenstein (1953, II.x.109) observed a similar phenomenon involving epistemic 'might': the infelicity of sentences like (2), with the form ⌜Might p and not p⌝ or ⌜p and might not p⌝ (which I'll call *Wittgenstein sentences*).[2]

(2) #It might be raining but it's not raining.

Moore and Wittgenstein sentences have played a central role in recent inquiry into the meaning of epistemic modals (words like 'might' and 'must'—and their analogues in other languages—on a broadly epistemic interpretation). On the face of it, an epistemic modal construction like 'It might be raining' seems to mean roughly the same thing as 'For all we know, it's raining', where 'we' refers to some contextually salient group or individual (or more generally, body of evidence). How far does this parallel go? A broad range of recent work has shown striking differences between the ways Wittgenstein sentences and Moore sentences embed. Researchers have made a variety of proposals to account for these differences. Nearly all those proposals, however, agree on one point: namely, that Wittgenstein sentences like (2) are, like Moore sentences, classically consistent.

In this paper, I argue that this consensus is wrong. I adduce a variety of new data which I argue can best be accounted for if we treat Wittgenstein sentences as genuine classical contradictions. But this, in turn, creates a puzzle. If ⌜Might p and not p⌝ (or ⌜Not p and might p⌝) is a classical contradiction, then, assuming 'and' and 'not' have classical semantics, it will follow that ⌜Might p⌝ entails p; but that is clearly false. I propose a solution to this puzzle by giving a new theory of the meaning of epistemic modals and their interaction with embedding operators. The basic idea is that epistemic modals are quantifiers over accessible worlds, as the standard theory has it; but, crucially, their domain of quantification is limited by their *local contexts*. Together with an appropriate theory of local contexts, this theory predicts that Wittgenstein sentences are classical contradictions without also predicting that ⌜Might p⌝ entails p. I argue that this theory accounts for the subtle and surprising embedding behavior of epistemic modals across the board, and does so in a uniquely principled way, shedding new light on the meaning of epistemic modals and on the dynamics of information in natural language.

[2] 'Es dürfte regnen; aber es regnet nicht'; German speakers inform me that 'dürfte' is slightly stronger than 'might' (but 'should' is not a good translation, since 'It should be raining but it's not' can be felicitous, whereas Wittgenstein's sentence can't). I prescind from the complexities of the German modal system here, resting content with 'might'. Yalcin (2007) calls sentences with the form ⌜p and might not p⌝ *epistemic contradictions*; the class of Wittgenstein sentences comprises both Yalcin's epistemic contradictions and variants which permute the order of the conjuncts.

2 The classical strategy

Let's begin by exploring *unembedded* Moore and Wittgenstein sentences. Suppose Ruth asserts the Moore sentence (1):

(1) #It's raining but I don't know it's raining.

There is strong reason to think that the content of Ruth's assertion is *classically consistent*, in the sense of being true at some points of evaluation. There are obviously circumstances in which (1) is true: namely those where it's raining out, but Ruth doesn't know that it's raining out. (1) is nonetheless unassertable. It is fairly easy to see, in broad terms, why this should be so.[3] Asserting something generally involves a commitment to knowing its content. If Ruth asserts (1), she expresses that she knows that it's raining and she doesn't know that it's raining. Assuming knowledge distributes over conjuncts, Ruth thus expresses that she knows that it's raining and that she knows that she doesn't know that it's raining. But such a state could never obtain: if she knows that she doesn't know that it's raining, then she doesn't know that it's raining. Moore sentences, then, are classically consistent, but unassertable on broadly pragmatic grounds.

Now suppose Ruth asserts the Wittgenstein sentence (2):

(2) #It might be raining but it's not raining.

How far should our explanation of the infelicity of a Wittgenstein sentence mirror our explanation of the infelicity of Moore sentences? As we'll see in detail below, there are important differences between Wittgenstein and Moore sentences with respect to how they embed. But there is a simple and persuasive argument that, like Moore sentences, Wittgenstein sentences are classically consistent.[4] If ⌜Might p and not p⌝ is a classical contradiction, and 'and' and 'not' have classical Boolean semantics, then it follows that ⌜Might p⌝ entails p. But that can't be right. ⌜Might p and might not p⌝ is often true (as in 'It might be raining and it might not be raining—I have no idea!'). But if ⌜Might p⌝ entailed p, then ⌜Might p and might not p⌝ would be inconsistent, since it would entail ⌜p and not p⌝. Thus, assuming that 'and' and 'not' have their classical meanings, Wittgenstein sentences must be classically consistent.[5] That, in turn, suggests that the

[3] The details do not matter for our purposes. For a classic formulation, see Hintikka 1962.

[4] See Yalcin 2007 for a clear statement of this argument.

[5] A different reason to think that ⌜Might p⌝ and ⌜Not p⌝ are classically consistent is that many speakers judge that both can be true in the same situation (even though they are not jointly assertable); see Knobe and Yalcin (2014, §3) for experimental evidence to this effect (though this point is controversial; see Egan et al. 2005 for contrary judgments).

situation for unembedded Wittgenstein sentences is much like that for unembedded Moore sentences: they are classically consistent but still unassertable.

And indeed, it looks like we can explain their unassertability using roughly the same strategy that is used to explain the infelicity of unembedded Moore sentences. The general idea is that 'It might be raining' amounts to a proposal to leave open the possibility that it's raining; while 'It's not raining' amounts to a proposal to rule this possibility out. Asserting their conjunction will amount to making both of these proposals simultaneously. And there is something wrong with this, since the two proposals cannot be simultaneously carried out.

Just as for Moore sentences, there are many different ways of spelling out the details of this strategy. Again, these differences don't matter for our purposes. Some approaches, like Veltman (1996), Groenendijk et al. (1996)'s, or Yalcin (2007)'s, spell out this strategy by appeal to a broadly dynamic or informational notion of entailment, a notion relative to which (some) Wittgenstein sentences are inconsistent. Crucially, though, on those approaches, Wittgenstein sentences are still predicted to be *classically* consistent (in the senses relevant to those frameworks)[6]—a fact which, as we will see momentarily, leads to trouble.

Call the broad framework just sketched for accounting for the infelicity of Moore and Wittgenstein sentences the *classical strategy*, since the core of this framework is a commitment to the *classical consistency* of Moore and Wittgenstein sentences. Every extant theory of epistemic modals (with the exception of those I discuss in §7) subscribes to (some variant of) the classical strategy.

3 Wittgenstein disjunctions

The classical strategy yields a striking prediction. On the classical strategy, a Moore or Wittgenstein sentence is infelicitous only because no speaker can coherently be committed to both of its conjuncts. Suppose, however, that someone asserts a *disjunction* of two Moore sentences, or of two Wittgenstein sentences. In asserting a disjunction, the speaker expresses that she knows the disjunction, but *not* that she knows any one disjunct on its own—and thus does *not* express that she knows any one incoherent conjunction. The classical strategy thus predicts that *disjoined* Moore or Wittgenstein sentences will be felicitous.[7]

[6] In the update framework, p is classically consistent just in case there is a context c such that c[p] is non-empty. In the domain framework, p is classically consistent just in case p is true at some point of evaluation.

[7] I know of no discussion of cases with the form of those discussed here prior to Mandelkern 2017, but Dorr and Hawthorne 2013 discusses closely related cases, namely sentences with the form ⌜p and I don't know that p, or q⌝ and ⌜p and might not p, or q⌝ (as well as order variants), and thus is an important precedent for this discussion. These cases illustrate the same basic point I make in this section, albeit in a different way: as Dorr and Hawthorne discuss, the issue with sentences like ⌜p and might not p, or q⌝ is

Is this prediction borne out? Consider first Moore sentences. Suppose I'm party to a court case which has not been decided. I don't know whether I'll win or lose. In this situation, the Moore sentences (3) and (4) are, as always, unassertable:

(3) a. #I'll win but I don't know it.

b. #p and I don't know p.

(4) a. #I'll lose but I don't know it.

b. #q and I don't know q.

But now consider the disjunction of (3) and (4):

(5) a. Either I'll win but I don't know it, or I'll lose but I don't know it.

b. p and I don't know p, or q and I don't know q.

Exactly as the classical strategy predicts, (5) is perfectly felicitous: disjoining Moore sentences bleaches them of their infelicity.

In more detail, the reason the classical strategy predicts that an assertion of a sentence like (3-a) is infelicitous is that the speaker is committing herself to knowing two things: first, that she will win; second, that she doesn't know that she will win. The speaker could know either of these things on its own, but she could not know both together. But now consider the disjunction in (5-a). Each of its disjuncts is such that no speaker could know it to obtain. But a speaker *could* know that *either* the first disjunct is true *or* that the second is true, since to know a disjunction does not require knowing either disjunct. And so when a speaker asserts (5-a), thus expressing that she knows it, she does not commit herself to any kind of inconsistency. On the contrary, for the speaker to know that this disjunction obtains is a fairly straightforward matter: it essentially just amounts to knowing that she is ignorant of the outcome of the court case.

The felicity of disjoined Moore sentences provides striking confirmation of the classical approach to Moore sentences. But matters are different when we turn to Wittgenstein sentences. In the same scenario, the Wittgenstein sentences (6) and (7) are—again, unsurprisingly—unassertable:

(6) a. #I might win but I won't.

b. #Might p and not p.

that the first disjunct is somehow felt to be redundant. The account I give of Wittgenstein sentences, together with any theory of redundancy, accounts for Dorr and Hawthorne's observation.

(7) a. #I might lose but I won't.

b. #Might q and not q.

But—in striking contrast to Moore sentences—their disjunction, in (8), remains just as infelicitous:

(8) a. #Either I might win, but I won't, or I might lose, but I won't.

b. #Might p and not p, or might q and not q.

This pattern is robust across permutations in order as well as variations in our case. Consider first sentences which permute the order of the conjuncts in (5) and (8):

(9) a. Either I don't know I'll win, but I will win; or I don't know I'll lose, but I will lose.

b. I don't know p, and p, or I don't know q, and q.

(10) a. #Either I won't win, but I might; or I won't lose, but I might.

b. #Not p and might p, or not q and might q.

Consider next some further cases. Suppose that either Sue or Mary is my TA, but I can't remember which. Compare (11) and (12):

(11) a. Either Sue is my TA but I don't know it, or Mary is my TA but I don't know it.

b. Either I don't know it, but Sue is my TA; or I don't know it, but Mary is my TA.

(12) a. #Either Sue might be my TA but she isn't, or Mary might be my TA but she isn't.

b. #Either Sue isn't my TA but she might be, or Mary isn't my TA but she might be.

Both variants in (11) sound like roundabout, but coherent, ways of saying that I don't know who my TA is. By contrast, both variants in (12) sound incoherent. Note, further, two important features of these variants. First, unlike our original examples, these do not contain 'will', showing that nothing about the examples turns on the semantics of 'will' (which some, for instance Cariani and Santorio (2018), have argued is itself a modal). Second, these use present tense statives in the complement of 'might'; this selects clearly for an epistemic reading of 'might', ensuring that the modals in question are read epistemically, as intended (Moore, 1959; Condoravdi, 2002).

For a final case, suppose that one of my students, either Mary or Sue, was sick last week, but I can't remember which. Then I can say (13):

(13) a. Either Sue was sick but I don't know it, or Mary was sick but I don't know it.

b. Either I don't know it, but Sue was sick; or I don't know it, but Mary was sick.

By contrast, (14) sounds quite odd on its (default) epistemic reading:

(14) a. #Either Sue might have been sick but wasn't; or Mary might have been sick but wasn't.

b. #Either Sue wasn't sick but might have been; or Mary wasn't sick but might have been.

While there *is* a felicitous reading of (14), it is a metaphysical reading, which communicates something different from (13): it communicates, not something about my knowledge, but rather something about nearby possibilities (that either it was the case that Sue wasn't sick, but there was a close non-actual situation where she was sick; or else that the same held of Mary).

The pattern in question is thus robust: disjoined Moore sentences are generally felt to be coherent, while disjoined Wittgenstein sentences are, by contrast, generally felt to be incoherent.[8]

The fact that disjoined Wittgenstein sentences (call them *Wittgenstein disjunctions*) are infelicitous poses a serious challenge to the classical strategy sketched in the last section. The problem is that the classical strategy predicts that the infelicity of Wittgenstein sentences is due to simultaneously committing to two conjuncts which are incoherent but still *jointly consistent in a classical sense*. If this were right, then, just as for Moore sentences, the infelicity of Wittgenstein disjunctions should wash out when they are disjoined, since when you assert a disjunction, you are not committed to any disjunct. As we have seen, this is exactly the right prediction when it comes to Moore sentences. But it is, apparently, not the right prediction for Wittgenstein sentences.

I will ultimately argue that, contrary to nearly every theory of epistemic modals,[9] Wittgenstein sentences themselves are classical contradictions, just as their behavior in disjunction suggests. For the present, I'll focus on Wittgenstein disjunctions, which make this case very clearly; then I will turn to a broad range of other data which support this hypothesis. In the rest of this section, I introduce the three most prominent

[8]Intuitions in some of these cases will certainly be graded—intuitions about natural language are rarely binary; but every informant I have discussed these data with reports a very strong *contrast* between Wittgenstein disjunctions and Moore disjunctions, and it is that contrast which is a puzzle for standard theories.

[9]This includes the relational theory of Kratzer 1977, 1981 (and its recent elaborations in for instance Ninan 2010, 2016); the update semantics of Veltman 1996; Groenendijk et al. 1996; Beaver 1992, 2001; Aloni 2000; Yalcin 2012a, 2015; Willer 2013; Ninan 2018, the domain semantics of Hacquard 2006; Yalcin 2007; MacFarlane 2011; Klinedinst and Rothschild 2012. Also vulnerable to my objection are the relativist proposals of for instance Egan et al. 2005; Stephenson 2007a,b; Lasersohn 2009, which make essentially the same predictions as the standard relational theory about embedding data (outside of attitude ascriptions, at least); and the constraint semantics of Swanson 2015.

semantic theories of epistemic modals and spell out in more detail the problem that Wittgenstein disjunctions pose for them (in §7 I discuss the existing theories which I think have the best chance of accounting for the data).

On the relational semantics—due in particular to Kratzer (1977, 1981), building on earlier work in modal logic—epistemic modals denote quantifiers over a set of worlds, generally assumed to represent a relevant epistemic state or body of evidence. The set of worlds is supplied by a variable which, in turn, is assigned to a function from worlds to sets of worlds (a *modal base* or *accessibility relation*) by a contextual variable assignment g:[10]

(15) *Relational 'might':* $[\![\text{Might}_i \text{ p}]\!]^{g,w} = 1$ iff $\exists w' \in g(i)(w) : [\![\text{p}]\!]^{g,w'} = 1$

Informally: ⌜Might p⌝ says that p is true at some epistemically accessible world.

Given this gloss, it should already be clear why this approach will fail to predict the infelicity of Wittgenstein disjunctions. On this approach, ⌜Might p⌝ means roughly: ⌜For all we [the relevant agents] know, p⌝. But then the meaning of Wittgenstein sentences and Moore sentences will be essentially the same, and so they will be predicted to behave in the same way in disjunctions.

To spell this out more formally, we need a semantics for connectives; I will assume, as is standard in this framework, a Boolean semantics for conjunction, disjunction, and negation:

(16) *Relational conjunction*: $[\![\text{p and q}]\!]^{g,w} = 1$ iff $[\![\text{p}]\!]^{g,w} = 1$ and $[\![\text{q}]\!]^{g,w} = 1$

(17) *Relational disjunction*: $[\![\text{p or q}]\!]^{g,w} = 1$ iff $[\![\text{p}]\!]^{g,w} = 1$ or $[\![\text{q}]\!]^{g,w} = 1$

(18) *Relational negation*: $[\![\text{Not p}]\!]^{g,w} = 1$ iff $[\![\text{p}]\!]^{g,w} = 0$

Now consider a Wittgenstein disjunction with the form ⌜Might$_i$ p and not p, or might$_j$ q and not q⌝ (assume p and q are non-modal; to map this onto our example from above, let p='I win' and q='I lose'). Consider a context which contains two worlds, w and w'. At w, p is true and q false. At w', p is false and q true. Let g take i and j to the accessibility relation which takes each world in the context to the context itself, so $g(i)(w) = g(i)(w') = g(j)(w) = g(j)(w') = \{w, w'\}$. Then our Wittgenstein disjunction will be true at w, thanks to the truth at w of the second disjunct, ⌜Might$_j$ q and not q⌝. This is true at w since w can access

[10] I follow von Fintel (1994)'s implementation here. I simplify Kratzer's theory by ignoring ordering sources and lowering the type of accessibility relations, neither of which matters for us. I use roman letters to stand for sentences and italic letters to stand for the corresponding propositions (which for simplicity I assume are functions from worlds to truth values, or equivalently sets of worlds), implicitly relativized to an index of evaluation and a context. Where p is a proposition or property, $\overline{p}$ is its complement.

a q-world under $g(j)$ (namely w'), and since q is false at w. Thus our Wittgenstein disjunction is classically consistent.

Note moreover that our Wittgenstein disjunction is true *throughout* this context: parallel reasoning shows that w' makes the first disjunct true. If we take a context to represent the conversation's common ground (as in Stalnaker 1974; more below), this shows that Wittgenstein disjunctions are predicted to be not only consistent, but (unlike Moore sentences) can be commonly accepted in some contexts. And I can see nothing untoward about the context as described, or about the way we selected the accessibility relation—it seems perfectly reasonable to think that the accessibility relation will at least sometimes track the context itself (as in for instance Stalnaker 2014).

This shows that the relational framework predicts that Wittgenstein disjunctions are classically consistent, and, worse, coherent in a broader, pragmatic sense. An important question to ask at this juncture is whether there is a pragmatic explanation of their infelicity within the relational framework. It is hard for me to see how this kind of a story would go. Any pragmatic response to this problem would have to acknowledge that there *are* accessibility relations which render Wittgenstein disjunctions consistent, but maintain that such accessibility relations are never in fact chosen in conversation. I cannot see how we would get this result on the basis of the relational theory, plus general facts about rationality of the kind that figure in pragmatic explanations. The question, again, is why we would not simply choose an accessibility relation which allows us to interpret the speaker as saying something coherent—that is, as saying essentially what Moore disjunctions say. There *are* such accessibility relations, and they are, at least apparently, quite natural ones: for instance, an accessibility relation tracking the group's knowledge/beliefs, or the speaker's knowledge/beliefs, would render a Wittgenstein disjunction essentially equivalent to the corresponding Moore disjunction, and thus felicitous. So why don't we interpret Wittgenstein disjunctions this way? In other words, why don't we interpret Wittgenstein disjunctions as (somewhat periphrastic) expressions of ignorance about the relevant facts—just as we interpret Moore disjunctions? I don't see how to spell out a pragmatic answer to this question (see §6.2 for more discussion of pragmatic responses).

It may not be all that surprising that the relational theory cannot make sense of Wittgenstein disjunctions; after all, as we will see in §5 below, it is already well known that the relational theory runs into trouble when it comes to embedded modals. Somewhat more surprising, however, is that this problem extends to the domain and update semantics, both of which have been motivated by the embedding behavior of epistemic modals. At a high level, we have already seen why these accounts run into trouble: both theories predict

that at least some Wittgenstein sentences are classically consistent, and thus that their disjunctions will be coherent. Thus, first, in the domain semantics of Yalcin 2007; MacFarlane 2011, epistemic modals denote quantifiers over a set of worlds s which is supplied, not by an accessibility relation, but rather as a world-independent parameter of the index:

(19) *Domain 'might':* $[\![\text{Might p}]\!]^{s,w} = 1$ iff $\exists w' \in s : [\![\text{p}]\!]^{s,w'} = 1$

This semantics is like the relational semantics, but with a set of worlds substituted where the relational semantics has a function from worlds to sets of worlds. Domain semantics is standardly coupled with Boolean semantics for the connectives, generalized to the domain framework:[11]

(20) *Domain conjunction*: $[\![\text{p and q}]\!]^{s,w} = 1$ iff $[\![\text{p}]\!]^{s,w} = 1$ and $[\![\text{q}]\!]^{s,w} = 1$

(21) *Domain disjunction*: $[\![\text{p or q}]\!]^{s,w} = 1$ iff $[\![\text{p}]\!]^{s,w} = 1$ or $[\![\text{q}]\!]^{s,w} = 1$

(22) *Domain negation*: $[\![\text{Not p}]\!]^{s,w} = 1$ iff $[\![\text{p}]\!]^{s,w} = 0$

In the domain semantics, an assertion of p is a proposal to make the context *accept* p, where a set of worlds s accepts a sentence p just in case $\forall w \in s : [\![\text{p}]\!]^{s,w} = 1$. This gives us a precise notion of pragmatic coherence: we can say that a sentence is coherent just in case it is accepted by a non-empty context. Consider the same context from above; call it s. Recall that $s = \{w, w'\}$; at w, p is true and q false, and at w', p is false and q true. Then $[\![\text{Might p and not p, or might q and not q}]\!]^{s,w} = 1$, since $[\![\text{Might q and not q}]\!]^{s,w} = 1$, since $[\![\text{Might q}]\!]^{s,w} = 1$ (since s contains q-worlds) and $[\![\text{Not q}]\!]^{s,w} = 1$ (since q is false at w). By parallel reasoning on the other disjunct, $[\![\text{Might p and not p, or might q and not q}]\!]^{s,w'} = 1$. And thus s accepts ⌜Might p and not p, or might q and not q⌝. (The order of the conjuncts doesn't matter here; s also accepts ⌜Not p and might p, or not q and might q⌝). Thus in the domain framework, Wittgenstein disjunctions are classically consistent, since they are true at some points of evaluation. They are also pragmatically coherent (or, in Yalcin (2007)'s terminology, *informationally consistent*), since they are accepted at some non-empty contexts.

Let's turn, finally, to the update semantics for epistemic modals, due to Veltman (1996). That theory is given within the dynamic framework of Heim (1982, 1983), in which the semantic value of a sentence p is a *context change potential* (*CCP*), denoted [p]: a function from contexts to contexts. The semantics of atoms, negation, and conjunction are given as follows in Heim (1983)'s system (using standard postfix notation:

[11] Klinedinst and Rothschild 2012 diverges from this assumption in a way I discuss in §6.2.

c[p] is [p] applied to c).[12] For any context c:

(23) *Update atoms*: For atomic p, $c[\text{p}] = \{w \in c : \text{p is true in } w\}$

(24) *Update conjunction*: $c[\text{p and q}] = c[\text{p}][\text{q}]$

(25) *Update negation*: $c[\text{Not p}] = c \setminus c[\text{p}]$

We augment Heim's semantics with the entries for disjunction and 'might' from Veltman 1996:

(26) *Update disjunction*: $c[\text{p or q}] = c[\text{p}] \cup c[\text{q}]$

(27) *Update 'might'*: $c[\text{Might p}] = \{w \in c : c[\text{p}] \neq \varnothing\}$

On this semantics, ⌜Might p⌝ is a test on the context: it leaves the context unchanged if it is compatible with p, and otherwise induces a crash to the empty set. In a dynamic framework, following Veltman 1996, the question of whether a sentence p is pragmatically coherent is the question of whether there is a non-empty context which remains unchanged when updated with p: that is, whether there is a non-empty context c such that $c[p] = c$ (if there is, we say that c accepts p). It turns out that the same context s from above accepts ⌜Might p and not p, or might q and not q⌝. Recall again that $s = \{w, w'\}$, with p true and q false at w, and p false and q true at w'. To update s with ⌜Might p and not p, or might q and not q⌝, we first update s with ⌜Might p⌝, which leaves s unchanged since it contains a p-world. Then we update it with ⌜Not p⌝, removing the p-world, so we have just $\{w'\}$. Likewise, we update s with ⌜Might q⌝, which leaves s unchanged, since it contains a q-world. Then we update it with ⌜Not q⌝, removing the q world, so we have just $\{w\}$. Then we take the union of $\{w\}$ with $\{w'\}$, giving us s back again. Thus, again, the update approach predicts that Wittgenstein sentences are consistent (in the dynamic sense that they do not take every context to the empty set), and, moreover, that they are accepted by some non-empty contexts, and thus are pragmatically coherent.

One interesting feature of the update semantics is that—unlike for the relational or domain semantics—order matters: on the update semantics, no non-empty context accepts a Wittgenstein disjunction with right-embedded modals, like ⌜p and might not p, or q and might not q⌝ (at least as long as p and q are not

[12] Veltman (1996) gives a different entry for conjunction, on which $c[\text{p and q}] = c[\text{p}] \cap c[\text{q}]$. I follow most of the subsequent literature in using Heim's entry; substituting in Veltman's does not help, and indeed deprives update semantics of its success in predicting the infelicity of most Wittgenstein disjunctions with right-embedded modals.

themselves modal).[13] This is a success, and the underlying idea—that p is somehow taken into account when we process ⌜Might not p⌝—will influence my own proposal below. But it is a limited success. As we saw above, the infelicity of Wittgenstein disjunctions is order invariant; but in the update framework, it is only left conjuncts which influence the interpretation of modals in right conjuncts, and not *vice versa*. So this cannot be the end of the story.[14]

In sum, the relational, domain, and update semantics all predict that Wittgenstein disjunctions will be coherent. These frameworks predict that disjoined Wittgenstein sentences and disjoined Moore sentences mean roughly the same thing, and thus should feel equally felicitious; but, as we have seen, this is the wrong prediction.

4 Bounded modality

How can we explain the infelicity of Wittgenstein disjunctions? One simple explanation would be that, *pace* the received view, Wittgenstein sentences are classical contradictions after all. This would immediately explain the infelicity of Wittgenstein disjunctions, since the disjunction of two classical contradictions is itself a classical contradiction.

This reaction may seem too quick, given the argument we have seen that ⌜Might p⌝ and ⌜Not p⌝ are classically consistent. But I will argue that this reaction is the correct one, based on a wide range of data that show that Wittgenstein sentences *always* seem to embed like classical contradictions. Existing theories capture some of these facts by giving suitably fine-tuned semantics for embedding operators which predict limited equivalences of Wittgenstein sentences with classical contradictions *within certain embedding environments*. But these theories inevitably fall short of the full range of data. Disjoined Wittgenstein disjunctions provide a particularly dramatic illustration of this failure, but they constitute only one of many we will explore. I will argue, again, that the reason for this is that these theories fail to make the right

[13] ⌜p and might not p⌝ can still be consistent in the update semantics when p itself contains a modal, contrary to a number of claims in the literature; see Mandelkern 2018a for discussion.

[14] The picture for dynamic semantics changes slightly if, instead of Veltman (1996)'s, we adopt the disjunction from Groenendijk et al. 1996, on which $c[\text{p or q}] = c[\text{p}] \cup c[\text{Not p}][\text{q}]$. Then Wittgenstein disjunctions will no longer be coherent. They will, however, remain *consistent*—in many cases they will take a context to a non-empty set. It is not entirely clear what judgment is predicted in this framework for a sentence which is consistent but not coherent. What is clear, however, is that there is predicted to be a *difference* between consistent and inconsistent sentences. Indeed, a central argument of Groenendijk et al. (1996) (discussed in §6) for the domain semantics concerns exactly this difference. By contrast, however, there seems to be no detectable difference in status between left-modal and right-modal Wittgenstein disjunctions. So this approach does not solve the problem for the update semantics. Moreover, changing the entry for disjunction doesn't help with the other problems I raise for the update semantics in §5, which do not involve disjunction.

generalization: namely, that Wittgenstein sentences are classical contradictions.

If we are to accept this generalization, however, we must find a way to do so without also predicting that ⌜Might p⌝ and ⌜Not p⌝ are inconsistent, which we know to be false. In this section I sketch a semantics which makes Wittgenstein sentences inconsistent, without predicting that ⌜Might p⌝ and ⌜Not p⌝ are jointly inconsistent. My theory, which I call the *bounded theory*, comprises two parts. The first concerns the architecture of embedding operators in natural language: following literature on presupposition projection, I propose that embedding operators systematically make available a quantity of information (a *local context*) which can influence the interpretation of embedded material. The second part of the theory concerns the meaning of epistemic modals: I propose that epistemic modals carry a semantic constraint which ensures that their domain of quantification is limited by their local contexts.

In the rest of this section, I will flesh out these two ideas. I begin by giving a brief intuitive characterization of local contexts; then spell out how my theory of epistemic modals will exploit local contexts; and then come back to local contexts, spelling out the intuitive idea more precisely.[15]

4.1 Local contexts: The basics

It is standard in theorizing about natural language to posit that the interpretation of an assertion depends on the assertion's *context*. There are different ways to spell out precisely what a context amounts to. For our purposes, the most useful approach treats a context as a set of possible worlds which in some sense represents the common commitments of the conversation (see Stalnaker 1974, 2002, 2014 for standard characterizations and discussion). Starting in the 1970s, a variety of theorists proposed that there is some sense in which contexts can *shift* within a sentence. The motivation for this was the phenomenon of presupposition projection. 'Susie stopped smoking' presupposes that Susie used to smoke. Now consider (28) and (29), which embed this sentence in different ways:

(28) If Susie used to smoke, then she stopped smoking.

(29) If Susie started exercising, then she stopped smoking.

Note that (29) still presupposes that Susie used to smoke (the presupposition *projects* to (29)), whereas (28)

[15] Let me note that my theory concerns the *semantics* of epistemic modals, and remains agnostic about their pragmatics/post-semantics/assertoric content (see for instance Ninan (2010); MacFarlane (2014) on the distinction), and thus is compatible with various accounts of the latter, including contextualist, relativist, and expressivist ones (though it may turn out that my semantics lends itself better to one of these post-semantic theories than to another).

does not. The main line of thinking about presupposition proposes to make sense of this roughly as follows.[16] Presuppositions must be entailed by their contexts. When presupposition triggers are unembedded, that means the presupposition in question has to be entailed by the conversation's context (the *global* context). When they are embedded, however, this requirement is interpreted relative to a *local* context. In (28), the local context for the consequent entails that Susie used to smoke. Thus this presupposition is entailed by its local context, and so the sentence as a whole puts no further constraints on the global context. But things are otherwise in (29), where the local context for the consequent does not entail that Susie used to smoke, and thus the presupposition of the consequent still puts constraints on the global context.

This suffices to give a sense of the original motivation for countenancing local contexts. In a moment we will return to the question of what exactly they amount to; what is important for now is simply that a local context for an expression is a quantity of information derived in a systematic way from the global context plus the expression's linguistic environment.

4.2 The bounded theory

My theory of epistemic modals builds on the relational theory. But it augments the relational truth conditions with a constraint which rules out certain accessibility relations, relative to the modal's local context. In particular, the constraint limits admissible accessibility relations to those which quantify over a subset of the local context: that is, to those under which only local context worlds are accessible. Formally, representing local contexts with κ, our semantics runs:[17]

(30) *Bounded 'might'*: $[\![\text{Might}_i\ \text{p}]\!]^{g,\kappa,w}$

a. defined only if $\forall w' : g(i)(w') \subseteq \kappa$; [*locality constraint*]

b. if defined, true iff $\exists w' \in g(i)(w) : [\![\text{p}]\!]^{g,\kappa,w'} = 1$. [*relational truth conditions*]

(30-b) encodes the ordinary relational truth conditions for epistemic modals. The innovation in this semantics is the *locality constraint* in (30-a). The locality constraint ensures that an epistemic modal claim is only

[16] See especially Stalnaker 1974; Karttunen 1974; Heim 1982, 1983.

[17] 'Must' will be the dual of 'might', as usual. Note that the locality constraint will project, so that the semantics for 'must' will run as follows: $[\![\text{Must}_i\ \text{p}]\!]^{g,\kappa,w}$ a. defined only if $\forall w' \in \kappa : g(i)(w') \subseteq \kappa$; b. if defined, true iff $\forall w' \in g(i)(w) : [\![\text{p}]\!]^{g,\kappa,w'} = 1$. I will continue to focus on 'might' here. I assume that 'might' does not itself shift the local context for its prejacent. As Simon Goldstein has pointed out to me, this assumption may not in fact be correct given our local context algorithm, though the question is subtle. Changing this assumption would affect the interpretation of stacked modals. Since this is not our focus here, I'll stick with this assumption for the sake of simplicity.

well-defined if the modal is associated with an accessibility relation which only quantifies over worlds in the local context. In other words, the locality constraint ensures that accessibility relations can never reach across the boundaries set by local contexts.[18]

The locality constraint is broadly inspired by the update semantics: the locality constraint ties the interpretation of epistemic modals to their local informational environment.[19] But the present approach ends up being very different from the update approach in its empirical reach. In particular, adopting the present approach makes it possible to adopt a *symmetric* theory of local contexts—something which, as I discuss in §6, is difficult to do satisfactorily in the dynamic (or update) framework, and which turns out to be essential for capturing the embedding behavior of epistemic modals.

What kind of definedness condition is the locality constraint? We do not need a precise answer to this question for present purposes. A natural approach is to understand it broadly along the lines on which semanticists analyze gender, person, and number features on pronouns—features which, like the locality constraint, are standardly treated as definedness conditions which constrain admissible variable assignments.[20] All that matters for us is that this definedness condition ensures that modals will be associated with accessibility relations which respect the locality constraint.

4.3 Local contexts, more precisely

To complete the exposition of my account, I turn to a more detailed account of local contexts. There has been much controversy over what local contexts amount to. The controversy is two-fold. First, what are the empirical facts? Second, how can we predict those facts in a systematic and explanatory way?[21] Here I will

[18] The locality constraint bears some similarity to Stalnaker (1975)'s proposal that the selection function for indicative conditionals always selects a world in the context (thanks to Irene Heim for drawing my attention to this connection). I hope to explore the relation between these constraints in future work.

[19] See especially Beaver 2001, which brings out this feature of the update semantics; as well as Klinedinst and Rothschild 2012; Stalnaker 2014 for related discussions in the update and relational semantics.

[20] Thus for instance 'she$_i$' will have the value of $g(i)$, *provided* $g(i)$ *is a single female*, and otherwise will be undefined. A standard treatment of these definedness conditions is as presuppositions, in particular as *expressive presuppositions* in the sense of Soames 1989—conditions a context must meet in order for a sentence to express a proposition in that context. (To be distinguished from semantic presuppositions; if we go this way, then the undefinedness in question will naturally project universally, in a weak Kleene fashion, ensuring that speakers will always choose variable assignments which ensure that all parts of a sentence are well-defined. I leave these projection properties implicit in the semantics that follows.) See Sudo (2012) for references and recent discussion. Why should we encode the role of local contexts as a definedness condition, rather than as a restriction on the domain of the epistemic modal (along the lines of domain restrictions on nominal quantifiers, as in Stanley and Szabó 2000)? An approach along the latter lines is well worth considering. But its logic ends up being far more revisionary than the logic of the bounded theory, and the bounded theory's conservativity gives it better resources for explaining the felt validity of inference patterns like conjunction introduction, which at least on the most straightforward implementation of a domain restriction approach is invalid. Thanks to a referee for this journal for helpful comments on this point.

[21] See especially Schlenker 2008a, 2009; Rothschild 2015 for recent discussion.

follow one variant of Schlenker (2009)'s answer to these questions. The payoff of this choice will be in the empirical coverage it affords, not *a priori* considerations, and so I will not say much for the present to justify it.

Schlenker, building on Stalnaker 1974, proposes that the local context of a given expression in a sentence is whatever information would be felt to be *redundant* at that point in the sentence: that is, whatever information you could add to that sentence while being guaranteed not to change its truth conditions, given what is already taken for granted in the context. This provides a heuristic for determining what information is *already available* at a given point in the sentence, given a truth-conditional semantics for the relevant embedding operator: we treat that already available information as the local context introduced by the embedding operator.

There are two ways to flesh this idea out, depending on how we think about redundancy. Is redundancy a *temporal* matter—a matter of what material is redundant *given how things stand at a particular moment in time* in processing a sentence? If so, then we should take an *asymmetric* approach, and say that the local context for an expression is whatever is redundant at that point in the sentence, *given everything to the left of that point*. Or is redundancy a static, *informational* matter—a matter of what material is redundant, *given the rest of the information in a sentence*? If so, then we should take a *symmetric* approach, and say that the local context for an expression is whatever is redundant at that point in the sentence, *given everything else in the sentence*. It is the second, symmetric notion that will be useful for present purposes, for reasons that may already be clear, given the symmetric nature of the judgments involving Wittgenstein disjunctions, and which I discuss further in §6 below.

Schlenker makes this symmetric approach to local contexts precise as follows:[22]

Definition 4.1. Symmetric Local Contexts:

The *symmetric local context* of expression e in syntactic environment a_b and context c is the strongest $[\![y]\!]$ such that for all d of the same type as e: $\forall w \in c : [\![a(d \wedge y)b]\!]^w = [\![adb]\!]^w$.

This algorithm makes precise the intuition that a local context is whatever content adds no new information at a given point in a sentence, given everything else in that sentence and the background assumptions in

[22] Italic letters here range over strings of the object language. For more formal details, and more formal proofs of the results stated here, see Schlenker 2009. My presentation is a slight variation on Schlenker's, which stipulates that any presupposed material in a and b is ignored; this won't affect present discussion. Since the semantics of epistemic modals will ultimately depend on local contexts, to avoid circularity we restrict the domain of the quantifier in this algorithm to non-modal sentences; for simplicity, we ignore any other sources of context-sensitivity. 'adb' is the sentence obtained by concatenating the values of a, d, and b.

the context. To see how it works, let's go through an example that will be central in what follows. Assume as a starting point the classical Boolean semantics for 'and'. Given that, what is the local context for a left conjunct? That is, what is the local context for q in a sentence ⌜q and p⌝ and context c? To answer this we ignore q, and look at the schema ⌜_ and p⌝. Note that, whatever ends up going in for '_', this sentence is guaranteed to be equivalent to ⌜(p and c and _) and p⌝. In other words, given the rest of the information in the sentence, plus the background information c, adding ⌜p and c⌝ to the left conjunct is guaranteed not to change the truth-value of the sentence, no matter what the left conjunct amounts to. It is easy to show that we cannot add any *more* information while still having this guarantee. And so our local context algorithm predicts that the local context for '_' in this sentence is $c \cap p$; generally, the local context for a left conjunct is the global context plus the information expressed by the right conjunct. By exactly parallel reasoning, the local context for a right conjunct is the global context plus the information expressed by the left conjunct, since our algorithm is perfectly symmetric. I will treat local contexts derived in this way as parameters of the index, accessible for semantic interpretation.[23] Thus, given the considerations just sketched, 'and' will have the semantics in (31):

(31) *Bounded conjunction*: $[\![\text{p and q}]\!]^{g,\kappa,w} = 1$ iff $[\![\text{p}]\!]^{g,\kappa^{\text{q}}_g,w} = 1$ and $[\![\text{q}]\!]^{g,\kappa^{\text{p}}_g,w} = 1$

κ^{p}_g is set intersection relative to g: $\kappa^{\text{p}}_g = \kappa \cap \{w : [\![\text{p}]\!]^{g,\kappa,w} = 1\}$.[24]

5 The data

With this exposition in hand, we can explore how the bounded theory makes sense of the embedding behavior of epistemic modals. I begin by explaining how the bounded theory pulls off the tricky task set out above: predicting that Wittgenstein sentences are classically inconsistent, while also predicting that ⌜Might p⌝ is consistent with ⌜Not p⌝. I will then show how, by making this key move, the bounded theory accounts for the embedding behavior of epistemic modals in a very broad range of environments.

[23] Broadly following Klinedinst and Rothschild 2012.

[24] I will sometimes leave off the variable assignment subscript in what follows, for readability. The local context for *unembedded* modals is predicted by this approach to be the global context. Thus on the bounded theory, unembedded modals will be interpreted as quantifying over a subset of the global context set. I won't devote much discussion to this prediction, but I believe it is plausible; see Stalnaker 2014 for discussion.

5.1 Wittgenstein sentences

Consider a Wittgenstein sentence with the form of (32) as evaluated in a context c:[25]

(32) Might$_i$ p and not p.

Recall that the local context for a conjunct is the local context for the whole conjunction (in this case c) intersected with the content of the other conjunct (in this case $\overline{p}$). Thus the local context for the left conjunct in (32) will be $c^{\overline{\text{p}}}$. The locality constraint in our semantics thus ensures that, in evaluating ⌜Might$_i$ p⌝, we do so relative to an accessibility relation with the property that only $c^{\overline{\text{p}}}$-worlds are in its image, and thus, *a fortiori*, that only $\overline{p}$-worlds are in its image. And so the left conjunct of (32) is guaranteed to be false at any world and variable assignment, if the whole conjunction is well-defined at that world and variable assignment. That is, for any g relative to which the whole conjunction is well-defined and any world w, w will not be able to access a p-world under $g(i)$, and so ⌜Might$_i$ p⌝ will be false. And so the whole conjunction will be false.

This reasoning was perfectly general, and so it shows that any sentence with the form of (32) will be false at any world and variable assignment, provided it is well-defined relative to that variable assignment. Importantly, since our local context algorithm is entirely symmetric, this account will extend immediately to Wittgenstein sentences which permute the order of the conjuncts, that is, those with the form ⌜p and might$_i$ not p⌝. And thus this reasoning shows that, on the bounded theory, any Wittgenstein sentence will be a classical contradiction, in the sense that it is false at any index of evaluation where it is defined (I will use 'contradiction' in this sense—of being false wherever defined—throughout what follows).

Let us take a step back to appreciate how the bounded theory accomplishes this. Recall the puzzle laid out above. On the one hand, ⌜Might p⌝ seems to be consistent with ⌜Not p⌝; otherwise, ⌜Might p⌝ would entail p. On the other hand, the behavior of Wittgenstein disjunctions (and other data we'll review shortly) suggests that ⌜Might p and not p⌝ is inconsistent. The bounded theory makes sense of these facts by predicting that ⌜Might p⌝ is indeed consistent with ⌜Not p⌝—that is, in many cases, both are true at the same world relative to a given variable assignment (namely, at any $\bar{p}$-world, relative to any variable assignment under which that world can access p-worlds). But, if we evaluate their *conjunction* relative to a variable assignment which makes each conjunct true at a given world, the result will be undefined. The conjunction will be well-defined

[25] I will leave subscripts off 'might's in general except where it is helpful for exposition, as in this case.

only if we evaluate it relative to an accessibility relation which renders the conjunction false at any world. And so, despite the joint consistency of the two conjuncts, the conjunction ends up being a contradiction.

The situation here is subtle, but the underlying mechanics are conservative: classical core truth conditions for the connectives, plus the possibility of changes in interpretation arising from the structure of information in discourse. On my theory, again, ⌜Might$_i$ p⌝ is consistent with ⌜Not p⌝. And the underlying logic is essentially classical, in the sense that classical entailments will preserve truth as long as they preserve definedness conditions (we can call this kind of classicality 'Strawson classicality', building on von Fintel (1999)).[26] But the caveat here is crucial. For any variable assignment g, there may be worlds where ⌜Might$_i$ p⌝ and ⌜Not p⌝ are both true, relative to g. But if this is so, then their conjunction—⌜Might$_i$ p and not p⌝ or ⌜Not p and might$_i$ p⌝—will be *undefined*, relative to g, at that world (and so conjunction introduction in this case will not preserve truth because it does not preserve definedness). When we consider a conjunction with this form, if we wish to maintain definedness, we will be forced to evaluate the conjunction relative to a different variable assignment which makes the modal conjunct, and thus the whole conjunction, false at any world. In short: conjoining two sentences can change the way that we interpret those sentences. Interpreted separately, ⌜Might$_i$ p ⌝ and ⌜Not p⌝ can both be true relative to a given world and variable assignment; but their conjunction can never be interpreted in such a way that it is true, relative to any world and variable assignment. Their conjunction will be just what it appears to be: a classical contradiction.

5.2 Wittgenstein disjunctions

With this in hand, let us turn to a careful exploration of the embedding behavior of epistemic modals, and of whether this behavior matches the predictions of our theory. I begin with the data point that I used above to motivate the bounded theory: Wittgenstein disjunctions. The fact that Wittgenstein sentences are classical contradictions on our theory suffices to predict that Wittgenstein disjunctions, too, are classical contradictions on our theory. This is for the simple reason that, on our account of disjunction—as on any plausible account—the disjunction of two classical contradictions will always itself be a classical contradiction: a

[26] More precisely, whenever Φ classically entails p relative to a variable assignment and context, then, in the bounded system, Φ entails p relative to that variable assignment and context as long as p and all the elements of Φ are well-defined relative to that variable assignment and context (i.e., Φ *Strawson entails* p in von Fintel's sense). All this also goes for **K**, the weakest normal modal system, which incorporates classical propositional logic. The proof of these facts turns on the observation that, *apart from definedness conditions*, the bounded theory is exactly like the standard semantics for **K**. Thus any sentence that is true in **K** will be true or undefined on my semantics, and any sentence false in **K** will be false or undefined in my semantics (relative to a world, variable assignment, and local context). And so, if a given inference preserves truth in **K**, it will preserve truth in my semantics as well, as long as all the premises and the conclusion are well-defined.

disjunction cannot be true at any world where both of its disjuncts are false.[27] And so the bounded theory's explanation of the infelicity of Wittgenstein sentences extends straightforwardly to an explanation of the infelicity of their disjunctions: Wittgenstein disjunctions, on the bounded theory, will be classical contradictions. Moreover, if we assume that no parallel locality constraint applies to the interpretation of 'knows', we explain the divergence between Wittgenstein disjunctions from Moore disjunctions.

5.3 Modals and conditionals

The bounded theory thus accounts for Wittgenstein disjunctions. One data point, however, is not enough to motivate a new theory: the test of a semantic theory is its ability to make plausible predictions about embedding behavior in a wide range of environments. In the rest of this section I argue that the bounded theory indeed makes the correct predictions about embedded epistemic modals in a wide range of embedding environments. The data I explore here support both the central hypothesis of the bounded theory—that Wittgenstein sentences are contradictions—as well as the details of the bounded theory. (The rest of this section will remain fairly informal. In Appendix A, I compile all of the semantic entries relevant to evaluating the claims I make here; interested readers can use those entries to verify my claims.)

I began by exploring the behavior of epistemic modals embedded in the scope of modals and conditionals. Consider first Wittgenstein sentences embedded under an epistemic modal:

(34) a. #It might be that (John might be sick and he isn't).

b. #Might (might p and not p).

(35) a. #It might be that (John isn't sick but he might be).

b. #Might (not p and might p).

Sentences like this sound quite bad. Existing theories, since they predict that Wittgenstein sentences are consistent, predict that sentences with this form are consistent and coherent.[28] In particular, on those theories, a

[27] Given our framework for local contexts, disjunction gets the semantics in (33):

(33) *Bounded disjunction*: $[\![\text{p or q}]\!]^{g,\kappa,w} = 1$ iff $[\![\text{p}]\!]^{g,\kappa_g^{\overline{\text{q}}},w} = 1$ or $[\![\text{q}]\!]^{g,\kappa_g^{\overline{\text{p}}},w}$=1

where $\kappa_g^{\overline{\text{p}}} = \kappa \setminus \{w : [\![\text{p}]\!]^{g,\kappa,w} = 1\}$. David Boylan points out to me an unintuitive feature of this semantics for disjunction, together with my semantics for epistemic modals, namely that ⌜p or might$_i$ p⌝ is predicted to be equivalent to p. I'm unsure of how bad this consequence is: disjunctions like this are decidedly odd, and insofar as they are interpretable, I'm inclined to think that we interpret the 'or' as a particle signaling revision, rather than as an ordinary disjunction (on this use of 'or' see for instance Szabolcsi 1997, 2015; Rawlins 2016).

[28] I know of no discussion of cases with this form prior to Mandelkern 2017. Gillies 2018 makes the same observation and proposes a variant on the update 'might' which predicts this data point. On that variant, [Might p] checks whether a non-empty part of the

sentence like (34) or (35) is predicted to be roughly equivalent to (36), which is perfectly coherent:

(36) It might be that (John isn't sick, but, for all we know, he is sick).

By contrast, since the bounded theory predicts that the embedded clause in (34) and (35)—a Wittgenstein sentence—is classically inconsistent, it predicts that sentences with this form are also classically inconsistent (since a sentence consisting of 'might' taking scope over a classical contradiction can never be true). Sentences with the form of (34) and (35) thus provide striking further confirmation of the thesis that Wittgenstein sentences are classical contradictions.[29]

These data also help put to rest what might have seemed a natural response to Wittgenstein disjunctions: that the puzzle they raise is one about *disjunction*. This idea, tempting though it is, is hard to flesh out; I have not been able to find a plausible entry for disjunction which accounts for the infelicity of Wittgenstein disjunctions within the context of standard theories of epistemic modals.[30] But in any case, the present data show that this response misses the central lesson of Wittgenstein disjunctions, since similar puzzles arise in the form of sentences like (34) and (35) (and a great deal more data reviewed momentarily) which do not contain disjunctions.

Let us turn now to conditionals. Yalcin (2007) points out that conditionals with Wittgenstein sentences in their antecedents are infelicitous:[31]

(37) a. #If it's raining and it might not be, then we should still bring an umbrella.

b. #If p and might not p, then q.

(38) a. #If it might not be raining and it is, then we should still bring an umbrella.

input context *accepts* p. As Gillies discusses, assuming that ⌜If p, q⌝ presupposes [Might p], this also accounts for the infelicity of ⌜If might p and not p, q⌝. Paolo Santorio has suggested to me that Gillies's entry may also help with the data discussed below of the form ⌜If might p, if not p, q⌝, though the picture there is subtle and I am not convinced a successful story can be told. Gillies's theory does not as far as I can tell help with Wittgenstein disjunctions, unless we combine it with a treatment of disjunction like that discussed in Footnote 30.

[29] It may be objected that speakers generally aren't happy about stacked modals in any case. But, as Moss 2015 discusses at length, we *are* sometimes willing to countenance epistemic modals in the scope of other epistemic modals, especially with lexical variation in the modals, as in 'It's possible that John might be sick.' But varying lexical items in (34) and (35) does not improve them; for instance, 'It's possible that (John might be sick and he isn't)' sounds just as bad as (34), suggesting that the problem with it is, after all, a semantic, not syntactic, problem.

[30] One option, suggested to me by Jacopo Romoli, would be to adopt the disjunction from Zimmermann 2000; Geurts 2005, on which ⌜p or q⌝ ends up roughly equivalent to ⌜Might p and might q⌝. Wittgenstein disjunctions would then be equivalent to sentences of the form ⌜Might (might p and not p) and might (might q and not q)⌝; the explanation of their infelicity would go by way of whatever explains the infelicity of the latter. But as the present discussion shows, this just pushes the puzzle one step back, since standard approaches don't account for the infelicity of a conjunction of this form. Thanks to a referee for this journal for encouraging me to explore a solution from disjunction.

[31] Yalcin only discusses the order in (37), but the point generalizes to (38).

b. #If might not p and p, then q.

Here, once again, Wittgenstein sentences embed just like classical contradictions: these sentences sound roughly equivalent to sentences like (39):

(39) # If it's raining and it's not raining, then we should still bring an umbrella.

This is, again, immediately predicted by the bounded theory, on which sentences with the form of (37) and (38) are indeed predicted to be equivalent to sentences which embed contradictions in their antecedents, providing the resources to explain their infelicity.[32]

Unlike any of the data we've seen so far, these are in fact captured by some versions of the relational semantics (Ninan, 2016) and domain semantics (Yalcin, 2007). (The standard update semantics conditional (Gillies, 2004) accounts for sentences like (37), but not order permutations like (38).) The domain and relational semantics accomplish this by giving a special-purpose semantics for the conditional to predict these data. By contrast, the interpretation of these data which I advocate in light of the full complement of data is as further confirmation that Wittgenstein sentences are, after all, classical contradictions. One piece of evidence that the bounded theory captures these data in a more plausible way than the domain or relational semantics comes from close variations on (37) and (38) like (40), which is due to Paolo Santorio (p.c.):

(40) a. #If it might be raining, then if it's not raining, then we should still bring an umbrella.

b. #If might p, then if not p, then q.

Sentences with this form sound equivalent to sentences like (37) and (38), and are equally bad. This is not, however, predicted by the relational, domain, or update semantics, which all predict that sentences like (40) will be perfectly consistent and coherent.[33] By contrast, the bounded theory predicts that (40) will be equivalent to ⌜If p, then if not p, then q⌝. Provided we adopt a semantics for the conditional which validates the *Import-Export* inference pattern (which lets us predict the latter to be infelicitous), we also predict (40) to be infelicitous: both will be equivalent to ⌜If p and not p, then q⌝.[34] This provides further support for the

[32] There is more to say about what is wrong with an indicative conditional with a classically inconsistent antecedent. One possibility is that indicative conditionals presuppose their antecedents to be compatible with the global context, a presupposition which could not be met here (see for instance Stalnaker 1975; von Fintel 1998b; Gillies 2009).

[33] Yalcin (2007)'s domain conditional ⌜If p, then q⌝ is true at information state s and world w iff q is accepted at the maximal subset of s that accepts p. An information state s that includes p-worlds and $\overline{p}$-worlds, and where all the p-worlds are also q-worlds, will accept ⌜If might not p, then if p, then q⌝. Likewise for Ninan (2016)'s (essentially equivalent) relational conditional. On Gillies (2004)'s update semantics for the conditional, ⌜If p, then q⌝ is accepted at an information state c just in case $c[\text{p}] = c[\text{p}][\text{q}]$; s will again accept our sentence on this semantics.

[34] *Import-Export* says that ⌜If p, then if q, then r⌝ is equivalent to ⌜If p and q, then r⌝. If *Import-Export* holds, then ⌜If p, then if q,

present approach. And it illustrates an important high-level point (which will also be illustrated by quantifier data below): what is required to account for the full range of data is a thoroughly symmetric system of local contexts, and not just a symmetric entry for conjunction (which doesn't figure in this example).

5.4 Attitudes

Let's turn now to look at epistemic modals in the scope of attitude predicates. Yalcin (2007) points out the infelicity of Wittgenstein sentences under 'Suppose':[35]

(41) a. #Suppose it's raining and it might not be.

b. #Suppose (p and might not p).

(42) a. #Suppose it might not be raining but it is.

b. #Suppose (might not p and p).

Here, again, Wittgenstein sentences seem to embed like contradictions: (41) and (42) sound roughly equivalent to 'Suppose it's raining and it's not'. This is exactly what is predicted by the bounded theory, on which these just are equivalent, since Wittgenstein sentences *are* contradictions. Given suitable semantics for 'suppose', the relational, update and domain theories can also capture this data point; but these data provide further support for our hypothesis that Wittgenstein sentences just are classical contradictions (rather than that their embedding behavior here has to do with the particular semantics we give to 'suppose').

This account extends to variants on (41) and (42) with a high-scope conjunction, as in (43). At a high level, this is because, on our account, ⌜A supposes might p⌝ entails that p is compatible with A's suppositions, provided they are consistent; ⌜A supposes not p⌝ entails that $\overline{p}$ is entailed by A's suppositions; these can both be true only if A's suppositions are inconsistent.[36]

then r⌝ is equivalent to ⌜If p and q, then if q, then r⌝, meaning that, given Schlenker's system of local contexts, the local context for p entails *q*; then ⌜If might p, then if not p, then q⌝ will be equivalent to ⌜If p, then if not p, then q⌝. McGee 1985 argued for *Import-Export* (see Khoo and Mandelkern 2018 for a recent defense), which is validated by a wide range of semantics for the conditional, for instance by Kratzer 1981; McGee 1985; Gillies 2009; Mandelkern 2018b. See the appendix for an extension of McGee (1985)'s semantics to the present framework; that extension accounts for the present data by keeping track of successive antecedents in the index and then conjoining them using the bounded conjunction. A different approach would treat local contexts, not as parameters, but rather as the denotations of dedicated object language variables whose value is calculated using Schlenkers algorithm. These could then be exploited in the same way that local context parameters are in the present theory.

[35] Again, Yalcin focuses on variants with the modal in the right conjunct, but his point extends to order permutations.

[36] Our predictions here depend on what local context is introduced by 'suppose', which depends on subtle choices about the syntax-semantics interface; see Schlenker 2009, §3.1.2 for discussion. I follow Schlenker in taking the local context introduced by an attitude verb to be the set of worlds compatible with that attitude verb: thus for instance where $S_{A,w}$ is the set of worlds compatible with A's suppositions in w, the local context introduced by 'supposes' in context c is $S_{A,w}$. I am not fully satisfied with Schlenker's route to this conclusion, however. A different conclusion that we can arrive at *via* different assumptions about

(43) a. #Suppose it's raining and suppose it might not be raining!

b. #Suppose p and suppose might not p!

Things are similar, *mutatis mutandis*, for Wittgenstein sentences under other attitude predicates.[37] For instance, on the bounded theory, (45) will be equivalent to (46), which, again, matches intuitions:

(45) John believes it might be raining and it's not.

(46) John believes it's raining and it's not.

An interesting final prediction concerns the contrast between (47), which is felt to ascribe contradictory beliefs to the speaker, versus (48) and (49), which are not:[38]

(47) a. #I believe it's raining, but I believe it might not be raining.

b. #I believe p, and I believe might not p.

(48) a. I believe it's raining, but it might not be raining.

b. (I believe p), and (might not p).

(49) a. I believe it's raining, but I know it might not be raining. (Hawthorne et al., 2016)

b. I believe p and I know might not p.

This is predicted in the present system, in which (47) will always ascribe contradictory beliefs to the speaker

syntax/semantics is that an attitude predicate like 'suppose' introduces as its local context the set of worlds that, for all that is accepted in the context, are compatible with what the agent in question supposes. This approach suffices to account for the data discussed here, but may fail in some edge cases not discussed here. More work is required here; thanks to Daniel Rothschild and Philippe Schlenker for very helpful discussion.

[37] Paolo Santorio brings to my attention the question of how to account for sequences like the following:

(44) #Suppose the following. It's raining. It might not be raining.

This is as infelicitous as the single-sentence variants considered above. What are the predictions of Schlenker's algorithm about these sentences? Intuitions about redundancy suggests that the local context for q in ⌜A supposes the following. p. q. . . . ⌝ is the set of worlds compatible with what A supposes, plus p: thus ⌜A supposes the following. p. p.⌝ will generally strike us as redundant. If this is right, then we can account for the infelicity of (44) by saying that when an operator in one sentence takes scope over subsequent sentences, as in (44), the proper input for Schlenker's algorithm is the whole sequence of sentences. This is independently motivated by the redundancy data just mooted, and can also be motivated by data from presupposition projection. Generally speaking, then, we should say that the input for our local context algorithm when we consider the local context for an expression is the whole scopal unit containing that expression, where a scopal unit is the smallest string of sentences containing the expression which is not itself within the semantic scope of an embedding operator. Standardly a scopal unit will thus just be the sentence containing the expression in question; but in cases like these, the scopal unit will be the whole sequence.

[38] Beddor and Goldstein (2018) argue that (47) is coherent. Their arguments are worth serious exploration, but as far as I can tell their conclusion is at odds with most speakers' judgments. A related issue concerns a variety of surprising data from epistemic modals under 'knows', for instance the fact that ⌜S knows might p⌝ is generally inadmissible in a context in which p has been ruled out (Lasersohn, 2009). This is not straightforwardly predicted by our view; more work is required here.

(assuming it is well-defined); while (48) and (49) will not, since the modals are not in the scope of 'believes'. Similar contrasts will be predicted, *mutatis mutandis*, for other attitudes.[39]

5.5 Quantifiers

I turn, finally, to epistemic modals under quantifiers, which provide some of the most puzzling and subtle tests for a theory of epistemic modals.[40] I will show in this section that the bounded theory, together with standard truth-conditional entries for the quantifiers, makes strikingly accurate predictions about epistemic modals in these environments.[41]

Consider first modals under 'some'. Groenendijk et al. (1996); Aloni (2000); Yalcin (2015); Ninan (2018) observe that sentences like (50) and (51) are infelicitous (both modified from Groenendijk et al. 1996):[42]

(50) a. #Someone who is hiding in the closet might not be hiding in the closet.

b. #Some(p)(might$_i$ not p).

(51) a. #Someone who might not be hiding in the closet is hiding in the closet.

b. #Some(might$_i$ not p)(p).

Once more, Wittgenstein sentences seem to embed just like contradictions: (50) and (51) sound roughly equivalent to 'Someone who is hiding in the closet isn't hiding in the closet'. This is surprising from the point of view of the standard relational semantics (together with a standard truth-conditional semantics for 'some'), on which (50-b) or (51-b) will be true just in case there is an individual a which is p, and it is compatible with what the relevant agents know that a is not p. It should be very easy for these conditions to obtain, if, say, we know of two people that one of them is p and one of them isn't, but we don't know

[39] Doesn't (49) entail (47)—since knowledge entails belief—in which case (49) should entail that the speaker has inconsistent beliefs? Not in my system, in which the entailment from knowledge to belief will only be Strawson valid. In particular, if (49) is true, we cannot conclude that (47) is true: and in fact, if we interpret the 'might' relative to the same accessibility relation in (47) as in (49), then, whenever (49) is true, (47) will be undefined.

[40] See Beaver 1994; Groenendijk et al. 1996; Gerbrandy 1998; Aloni 2000, 2001; Yalcin 2015; Rothschild and Klinedinst 2015; Ninan 2018; Moss 2018 for extensive discussion. Roman letters in the discussion that follows stand for open sentences, and italic letters for the corresponding predicates obtained by abstracting over unbound variables.

[41] As Yalcin 2015 discusses, the domain and standard relational semantics do not predict any of these data. As the work just cited shows, update semantics can make sense of at least some of these data; since things get quite complicated here, I will not attempt a detailed comparison.

[42] Groenendijk et al. (1996); Aloni (2000) discuss variants on these sentences involving anaphora. An important question which I will leave open here is how to extend the present theory to those cases. If we adopt an e-type theory of anaphora, the extension is straightforward; things are less clear if we have a dynamic theory of anaphora.

which is which (say, we know that either Bill or Sue (but not both) is hiding in the closet, but we don't know which). But, again, the bounded theory straightforwardly accounts for these facts: open Wittgenstein sentences embed like open contradictions because Wittgenstein sentences *are* classical contradictions. The upshot is that sentences with the form of (50) or (51) will themselves be contradictions, accounting for their infelicity.[43]

A similar point concerns definite descriptions. As Aloni (2000) observed, sentences like (53) are infelicitous:[44]

(53) a. #The biggest flea might be the smallest flea. (Aloni, 2000)

b. #The(p)(might$_i$ not p).

This is, again, puzzling on the standard relational semantics for 'might' (plus any standard account of 'the'), on which (53-a) should mean, roughly, that there is a unique biggest flea a, and, for all we know, a is the smallest flea—in other words, that we don't know which flea is biggest, which is perfectly coherent. But the incoherence of (53) is, again, just what the bounded theory predicts: Wittgenstein sentences embed under 'the' like contradictions, because they are contradictions.[45]

Let me point to a final, intriguing set of predictions of the bounded theory. Yalcin (2015), crediting Declan Smithies, points out that sentences like (54) are felicitous.

---

[43] In more detail, assuming a standard semantics for 'some', our local context algorithm predicts that the local context for p in ⌜Some(p)(q)⌝ entails q (in a type-general sense of entailment; thus when I say in this section that the domain of quantification is limited to q-worlds, this is periphrastic: as the semantics in (52) makes clear, the relevant local context will in fact vary with the witnesses to the quantification). Given that, our semantics for 'some' will be the following (where p and q have the type of functions from indices to functions from individuals to truth values):

(52) *Bounded 'some'*: $[\![\text{Some(p)(q)}]\!]^{g,\kappa,w}=1$ iff $\exists x: [\![\text{p}]\!]^{g,\kappa\cap\{w':[\![\text{q}]\!]^{g,\kappa,w'}(x)=1\},w}(x)=[\![\text{q}]\!]^{g,\kappa\cap\{w':[\![\text{p}]\!]^{g,\kappa,w'}(x)=1\},w}(x)=1$

Now substitute ⌜Might$_i$ not p⌝ for q in (52), so that we have a sentence with the form of (50-b). Given the standard Tarski semantics for meta-language '∃', the right-hand side of this biconditional will then be true iff for some individual a, $[\![\text{p}]\!]^{g,\kappa\cap\{w':[\![\text{Might}_i\text{ not p}]\!]^{g,\kappa,w'}(a)=1\},w}$ holds of a, and $[\![\text{Might}_i\text{ not p}]\!]^{g,\kappa\cap\{w':[\![\text{p}]\!]^{g,\kappa,w'}(a)=1\},w}$ holds of a. Assume these two conditions hold for some a at some w. Then, thanks to locality, under the accessibility relation $g(i)$, w can only access worlds in the local context, and thus w can only access worlds where a is p. It follows from the core relational part of our semantics that the right conjunct is false. So these two conditions cannot both hold for any a at w; thus there is *no* witness to the truth of (50-b) at w, and thus (50-b) is false at w. Since w was chosen arbitrarily, this shows that (50-b) is false at any world where defined. Our truth conditions in (52) are entirely symmetric, so exactly the same reasoning will show that (51-b) is, likewise, a contradiction.

[44] At least in a context where there is more than one flea, so that the same flea cannot be both biggest and smallest.

[45] The local context for q in ⌜The(p)(q)⌝ entails p, according to our algorithm; we can see this by noting the felt equivalence between ⌜The(p)(q)⌝ and ⌜The(p)(p and q)⌝. Then a sentence with the form of (53-b) will be felt to say: there is a unique p-individual a, and there is some accessible world (under $g(i)$) where a is $\overline{p}$. But because the local context for the nuclear scope entails p, the accessibility relation $g(i)$ will only be able to access worlds where a is p. And so no world will be able to access a world under $g(i)$ where a is $\overline{p}$. (53-b) will thus again be a contradiction. See the appendix for a semantic entry for 'the' which bears out this informal reasoning.

(54) a. Not everyone who might be sick is sick.

b. Not every(might$_i$ p)(p).

This is surprising, since, on the face of it, (54-b) looks logically equivalent to sentences with the form of (51-b) (⌜Some(might$_i$ p)(not p)⌝), which, as we saw above, are infelicitous. So why is (54) substantially more felicitous than (51-b)?

It is not immediately clear that our theory helps. If we treat 'every' as the standard universal quantifier from first-order logic, then the local context for p in ⌜Every(p)(q)⌝ will entail $\overline{q}$, since ⌜$\forall x : p(x) \rightarrow q(x)$⌝ is logically equivalent to ⌜$\forall x : (p(x) \wedge \neg q(x)) \rightarrow q(x)$⌝. If we go this way, then ⌜Every(might$_i$ p)(p)⌝ will always be true where defined: since the modal in its restrictor (its first argument) will only be able to access $\overline{p}$-worlds under $g(i)$, the restrictor will be true of no individuals, and thus the sentence will be trivially true. And so its negation in (54-b) is (wrongly) predicted to be a contradiction, just like (51-b).

But if we take a more sophisticated approach to the semantics of 'every', then our theory does *not* predict that the local context for p in ⌜Every(p)(q)⌝ entails $\overline{q}$. To see the motivation for this, note that, contrary to our predictions if we treat 'every' as '$\forall$', ⌜Every(p)(q)⌝ is intuitively *not* equivalent to ⌜Every(p and not q)(q)⌝, as witnessed, for instance, by the contrast between (55) and (56):

(55) Every dog is in the genus *canis*.

(56) #Every dog which isn't in the genus *canis* is in the genus *canis*.

Again, these are predicted equivalent if 'every' is treated as '$\forall$'. We can break this equivalence, however, if we follow many researchers in positing that the meaning of 'every' diverges from that of '$\forall$' in that the former presupposes that the extension of its restrictor is non-empty.[46] This accounts for the felt difference between (55) and (56): the former presupposes (truly) that there is some dog, and asserts (truly) that anything which is a dog is in the genus *canis*; the latter asserts the same thing, but presupposes (falsely) that there is a dog not in the genus *canis*. The total content of each thus ends up being quite different: the former can be (in fact, is) true, whereas the combined presupposed and asserted content of the latter cannot be true.

If we take into account the existence presupposition of 'every' in calculating the local context for its restrictor, then ⌜Every(p)(q)⌝ and ⌜Every(p and not q)(q)⌝ are not invariably equivalent, and thus our algorithm predicts that the local context for p will not entail $\overline{q}$; it turns out just to be the global context.[47]

[46] See Heim and Kratzer (1998, §6.8) for discussion and references.

[47] The proof is identical to a parallel result given in Mandelkern and Romoli 2017. As they discuss, the result holds in general only

Thus, in particular, in ⌜Every(might$_i$ p)(p)⌝, as assessed at context c, the modal's local context will be c. And so we will interpret the embedded modal just as we would interpret an unembedded modal. ⌜Every(might$_i$ p)(p)⌝ will thus have the meaning that we would naively expect it to have: it will simply say that anyone who is p in some accessible world is in fact p. It will thus be neither a contradiction nor a tautology, and so neither will its negation: ⌜Not every(might$_i$ p)(p)⌝ will have the perfectly coherent meaning that someone who is p in an accessible world is in fact $\overline{p}$. Our theory thus accounts in a principled way for the fact that ⌜Not every(might$_i$ p)(p)⌝ strikes us as coherent, in surprising contrast to ⌜Some(might$_i$ p)(not p)⌝.

If the local context for the restrictor of 'every' is blind to the information carried by its nuclear scope (its second argument), then it may be objected that we predict that a sentence with the form of ⌜Every(might$_i$ p)(not p)⌝ is consistent. But this appears to be wrong, as witnessed by the infelicity of (57):

(57) #Everyone who might be sick is not sick.

But we can easily proffer an alternate explanation of the infelicity of sentences like this. 'Every', again, is associated with a non-emptiness presupposition, thanks to which (57) presupposes that someone might be sick. But (57) simultaneously asserts that no one who might be sick is sick. Assuming that the accessibility relation for 'might' generally can only access worlds compatible with what the speakers jointly accept (a standard assumption, which also follows from our semantics), we cannot coherently accept both (57)'s presupposition and its asserted content.[48]

Evidence that this is the right explanation of the infelicity of (57) comes from two observations. The first is that similar infelicity arises for variants of (57) with Moore sentences:

(58) #Everyone who for all we know is sick, is not sick.

The oddness of (58) is readily explained in the same way we've explained the oddness of (57), providing support for our pragmatic explanation of the infelicity of the (57). Second, note that (59) is felicitous and sounds roughly equivalent to (60):

(59) Not everyone who might be sick is well.

(60) Someone who might be sick is sick.

if we distinguish presupposition failure from falsity in Schlenker's algorithm.

[48] A similar point can be made regarding sentences with the form ⌜The (might$_i$ not p)(p)⌝.

(59) and (60) are not tautologous; both communicate the substantive proposition that someone is sick. This is exactly what the bounded theory predicts, but this would not be predicted if (57) were a contradiction: then its negation in (59) would wrongly be predicted to be a tautology.

Let me generalize the reasoning here. It is generally accepted that all (or nearly all) natural language quantifiers are *conservative*, in the sense that any quantificational structure ⌜Q(a)(b)⌝ is equivalent to ⌜Q(a)(a and b)⌝ (see Barwise and Cooper 1981). This means that, for any quantificational structure ⌜Q(a)(b)⌝, on our algorithm, the local context for b will entail a. But the converse does not invariably hold: the local context for a will only entail b when ⌜Q(a)(b)⌝ is invariably equivalent to ⌜Q(a and b)(b)⌝. That equivalence sometimes holds, as in the case of 'some'. But it often fails, for various reasons—for instance, due to a presupposition of 'Q' (as we've just seen for 'every'), or due to the lexical semantics of 'Q' (as for 'most'; see §7.3). When that equivalence is blocked, we expect to find asymmetries like those observed between (51) and (54).

This is a striking prediction; it is borne out in the data we've seen so far, but requires more systematic exploration. I cannot undertake that here, but let me note one upshot. Many take 'weak' (that is, generally non-presuppositional) quantifiers like 'some' to sometimes have presuppositional readings.[49] Depending how we analyze this—and how this interacts with our local context algorithm—this may then lead us to predict that there are readings of weak quantifiers where they pattern like strong (presuppositional) quantifiers with respect to embedded modals. This, in turn, may explain Yalcin (2015)'s observation that ⌜Some(might$_i$ p)(not p)⌝ is sometimes more felicitous than ⌜Some(not p)(might$_i$ p)⌝. If a quantifier like 'some' has presuppositional uses, then the bounded theory will predict that, on its presuppositional uses, ⌜Some(might$_i$ p)(not p)⌝ will be consistent. By contrast, ⌜Some(not p)(might$_i$ p)⌝ will never be consistent. This furnishes a possible explanation for the subtle contrast between these two.

There is much more to explore here, but this discussion should suffice to give a sense of the empirical reach of the bounded theory.[50] The bounded theory makes sense of the surprising and subtle embedding

[49]Thanks to Kai von Fintel for discussion; see von Fintel 1998a.

[50]Let me point out three areas which need further exploration. First, Daniel Rothschild points out to me that quantifiers like 'a few' or 'at least three' look to be conservative in their restrictor; but nonetheless, a sentence like 'A few people who might be sick aren't sick' sounds acceptable. This is indeed puzzling for my theory. One question to pursue is whether, insofar as these are felicitous, it is because we are bringing out a presuppositional reading of the quantifier. Second, Ninan (2018) points out the oddness of conjunctions like 'Any flea might be the biggest, but of course the smallest flea can't be the biggest'. My theory—like one that Ninan is arguing against—predicts sequences like this to be consistent, albeit only when the two epistemic modals are differently indexed. It is not clear to me that this is the wrong prediction: perhaps the *prima facie* oddness of Ninan's conjunction is due to a *prima facie* tendency to co-index the modals. The third issue concerns tense. As von Fintel and Gillies (2011) observed, past 'might's, like ⌜Might have p⌝, are acceptable in contexts in which p has been ruled out. Intriguingly, as Boylan (2018) has

behavior of epistemic modals, not just in Wittgenstein disjunctions, but in a wide range of further embedding environments. Taken together, these data provide striking confirmation of the hypothesis that Wittgenstein sentences are genuine classical contradictions, as well as of the bounded theory's particular approach to predicting this fact.[51]

6 Order, explanatory power, and the dynamics of information

In this section, I will reflect on a few high-level features of the bounded theory which help bring out what is distinctive about it; why we cannot replicate its predictions in a principled way on a purely pragmatic basis, or in an update or domain framework; and what its upshots are for other semantic and pragmatic systems which might operate with an apparatus of local contexts.

6.1 Order

As we saw above, the infelicity of Wittgenstein disjunctions does not depend on the order of the conjuncts. The same holds for the other data points introduced above: for instance, Wittgenstein sentences under epis-

observed, future epistemic modality seems to work differently: ⌜Will be possible that p⌝ is unacceptable when p is ruled out—even if the speaker knows that p will again become compatible with the relevant information. On first blush, my theory predicts the latter data point but not the former: once a tensed sentence p is accepted in a given context, it will be redundant under either future or past tense operators. But this leaves unexplained the first data point. One possibility to explore here is that the past tense in sentences like these acts not just as a temporal past but also as some kind of modal distancer (following Iatridou 2000).

[51] Many of the sentences which I have marked with a '#' can be used in ways that are not terrible, given enough contextual support. Following Kratzer and Phillips 2017, suppose ornithology students are learning to identify birds. Students tend to have trouble distinguishing sparrows from finches. The instructor says (61):

(61) ?Whenever a bird is shown that isn't a sparrow, but might be, the students hesitate.

Given a plausible semantics for 'whenever', (61) will be predicted, on the bounded theory, to be equivalent to (62):

(62) #Whenever a bird is shown that isn't a sparrow and is a sparrow, the students hesitate.

But obviously (61) manages to communicate something different than (62); and while (61) is by no means impeccable, it certainly does not sound as strange as (62). Similar readings can be brought out in other cases; the generalization seems to be that the hash-marks used throughout the paper identify a strong default reading, but that, provided sufficient contextual set-up, improved readings are sometimes available (compare Egan et al. (2005)'s 'exocentric' readings). There are at least three ways to make sense of this within the present framework. The first is to follow the strategy of Stephenson 2007a; MacFarlane 2011 and to maintain that, on the non-default readings, we interpret the epistemic modals in question as embedded under a covert operator which (only) shifts the local context—so that, for instance, (61), on its acceptable reading, is parsed as 'Whenever a bird is shown that isn't a sparrow, but, for all the students know, might be, the students hesitate'. 'For all the students know' will shift the local context relative to which 'might' is embedded to the worlds compatible with the students' knowledge, explaining why (61) communicates what it does, and is not equivalent to (62). The second option is to hold that the algorithm for calculating local contexts is itself only a strong default which can be violated given sufficient contextual set-up. (There is a parallel here to local accommodation in the theory of presupposition projection (Heim, 1983); as well as Schlenker (2011)'s proposed solution to the proviso problem.) A third option is that the locality constraint is only a defeasible default. All these options seem like viable routes for making sense of non-default readings of the sentences under discussion, and I will not try to decide between these here. More systematic exploration, however, is obviously required: the key question is why these rescue mechanisms are not always available to rescue our bad sentences, and instead require substantial contextual set-up.

temic modals are infelicitous in either order; likewise in conditionals.

These data thus help resolve a longstanding empirical controversy about epistemic modals. Dynamic approaches to epistemic modals claim that the interpretation of modals under connectives is sensitive to the intrasentential dynamics of information; but, by contrast to the bounded theory, they claim that that sensitivity is fundamentally asymmetric. (As we discuss in a moment, the same goes for recent pragmatic approaches.) The data we have surveyed suggest that there is something right in this approach and something wrong: the interpretation of epistemic modals *is* sensitive to the intrasentential dynamics of information, but *not* to order. The intrasentential dynamics of information, at least when it comes to epistemic modals, is a symmetric matter.[52]

This does not mean that the dynamics of epistemic modality *across discourses* is also symmetric; that, of course, is false, and is no commitment of the bounded theory. This lets us account for the fact that, as Groenendijk et al. (1996) observe, (63) differs from (64):

(63) It might be raining outside. [. . .] It isn't raining outside.

(64) It isn't raining outside. [. . .] It might be raining outside.

There is a difference between these sequences: (63) looks like an ordinary, monotonic evolution of information; (64) looks like a change of mind. This shows that order matters *across sequences of assertions*. The bounded theory captures that fact straightforwardly: insofar as the first assertion in (64) changes the global context, the local context for the modal in (64) will differ from that for the modal in (63), since in calculating the local context for the modal in (63), the second sentence will not be taken into account (the symmetry in our local context algorithm is symmetry *within sentences*, not symmetry across discourses). But the existence of order effects *across* sequences of the form ⌜Might p. . . . Not p⌝, and *vice versa*, does not show there are order effects *within* sentences which conjoin these; and, I think, the data we have surveyed provide a powerful argument against that hypothesis.[53]

[52] And note, again, that it is not just conjunction which must be symmetric: the data from conditionals and quantifiers show that the symmetry must run deeper than this.

[53] See Yalcin (2012b) for similar conclusions. I am thus committed to giving a broadly pragmatic explanation of the infelicity of a *sequence* like 'It might be raining out. It's not.' But a pragmatic explanation is ready at hand: indeed, the classical approach spelled out in §2 provides just such an explanation. Recall that my objection to the classical approach is not that it fails to account for the infelicity of unembedded Wittgenstein sentences, but rather that, if that approach is the *only* explanation of their infelicity, then we cannot make sense of their behavior in embedded environments. And so there is no reason that the bounded theory cannot appeal to that explanation to account for the infelicity of sequences like this. (This does not open me up to my own objection from Wittgenstein disjunctions, since we cannot embed a series of sentences in a disjunction.) Let me note finally that, while there do not seem to be order effects as far as judgments about embedded Wittgenstein sentences go, this is not to say there are

6.2 Consequences of symmetry

Symmetry is thus an essential feature of the intrasentential dynamics of epistemic modality—and thus of the bounded theory's ability to capture those dynamics. And, as I will now explain, the symmetry of the bounded theory makes it very difficult to reproduce its predictions in a plausible way on a pragmatic basis, or in a framework like the domain or update semantics.

Consider first a pragmatic approach. In introducing Wittgenstein disjunctions in §3, I explained in very general terms the challenge for a pragmatic approach within a standard relational approach: such a pragmatic approach must explain why we cannot resolve the context sensitivity of modals in a way which renders Wittgenstein disjunctions felicitous, even though there are quite natural accessibility relations (for instance those tracking the speaker's or group's evidence) which would indeed render them felicitous, and would make them essentially equivalent to the corresponding disjoined Moore sentences. Similar considerations apply to all the other data we have explored, which, as we saw throughout, would be perfectly felicitous if we interpreted them relative to any number of quite natural accessibility relations.

In spite of these challenges, Dorr and Hawthorne 2013 attempt to spell out a pragmatic treatment of the data from Yalcin 2007 reviewed above. Whether or not they are successful in accounting for those particular data (which, recall, all involve Wittgenstein sentences with a modal in the right, not left, conjunct) is not something I will address here. Without answering that question, however—and without getting into the details of their system—we can point to one important feature of their system which makes it inapt for accounting for the data discussed here: the system is fundamentally asymmetric. As Dorr and Hawthorne write:

> It does seem, in general, to be harder for inheritance [the pragmatic process by which accessibility relations are limited by salient information] to operate from right to left. This is not hard to explain. Other things equal, our interpretation of a modal will be influenced by the facts about what is salient when it is uttered. While we can of course go back and rethink the interpretation of earlier material in the light of later developments, we will expect co-operative speakers to prefer modes of expression that do not require this. (Dorr and Hawthorne, 2013, 886)

Indeed, it seems to me that this asymmetry will be a central feature of *any* plausible pragmatic account.[54] A

not order effects when it comes to resolving the context-sensitivity of epistemic modals in general (for instance, as Muffy Siegel has reminded me, modal subordination is clearly order sensitive). This is something that we could capture by saying that the way in which accessibility relations are assigned is a broadly pragmatic process which depends on order; this is, of course, consistent with adopting a symmetric underlying theory of local contexts (the latter only constrain, and do not fully determine, the modal's accessibility relation).

[54]Compare the account given in Stojnić 2016, which incorporates similar asymmetries by tying the interpretation of epistemic

pragmatic approach thus predicts that right-to-left inheritance in, say, a Wittgenstein disjunction of the form ⌜Might p and not p, or might q and not q⌝ is dispreferred, and thus that there should a *default* reading of such sentences which is felicitous. This is clearly the wrong prediction. Not only is the default readings of such sentences an infelicitous one; moreover, a felicitous reading hardly seems accessible at all (see Footnote 51 for exceptions). Similar considerations go for the other data reviewed above. The lack of any order contrast in those data poses a serious challenge for any pragmatic account of the data, over and above the very general challenge discussed at the outset.[55]

The symmetry of the data in question also complicates any attempt to account for them in a principled way in a domain or update framework. A natural attempt in a domain framework would couple the domain semantics for 'might' with symmetric connectives along the following lines:[56]

(65) *Symmetric domain conjunction*: $[\![\text{p and q}]\!]^{s,w} = 1$ iff $[\![\text{p}]\!]^{s^{\text{q}},w} = 1$ and $[\![\text{q}]\!]^{s^{\text{p}},w} = 1$

(66) *Symmetric domain disjunction*: $[\![\text{p or q}]\!]^{s,w} = 1$ iff $[\![\text{p}]\!]^{s^{\bar{\text{q}}},w} = 1$ or $[\![\text{q}]\!]^{s^{\bar{\text{p}}},w} = 1$

Read $s^{\text{p}} = s \cap \{w : [\![\text{p}]\!]^{s,w} = 1\}$, and $s^{\bar{\text{p}}} = s \setminus \{w : [\![\text{p}]\!]^{s,w} = 1\}$. If we adopt these connectives, then Wittgenstein sentences will be classically inconsistent—accounting for much (albeit not quite all) of the data surveyed above. But this approach runs into serious trouble elsewhere. Consider a disjunction like (67):

(67) a. The keys might be upstairs, or they might be downstairs.

b. Might p or might q.

Intuitively, a sentence like (67-a) should be accepted by an information state s that includes worlds where the keys are upstairs and worlds where the keys are downstairs. But in the present framework, s won't accept (67-a).[57] Worse, s *does* accept the *negation* of (67-a), which is clearly the wrong prediction. Even worse, s

modals to the interpretation of pronouns in general.

[55] Could there be other pragmatic explanations of the infelicity of our data? I don't see any promising candidates. One possibility suggested to me by Danny Fox is a story based on redundancy. But I have not been able to see how such a story would go: first, Wittgenstein disjunctions are not generally equivalent to any of their proper parts, or (given standard theories of modals) locally contradictory or redundant (the two standard criteria for redundancy; see Mayr and Romoli 2016 for a helpful recent overview); second, it is hard to see how any such story would distinguish Wittgenstein disjunctions from Moore disjunctions, which *are* felicitous (even if they do strike us as a bit periphrastic), and are equally wordy. Another thought along these lines would be that there is some semantic or syntactic constraint which generally rules out disjunctions of conjunctions which include a 'might' claim. But disjunctions which have this form are generally acceptable; here's an example: 'Either John won the case and might then get a big payout; or he settled and might then get a smaller payout. Either way, he'll probably get some money out of this.' Moreover, much of our data involved structurally simpler sentences, where a similar response is obviously not plausible.

[56] The semantics for negation remains unchanged from above. Compare Klinedinst and Rothschild 2012, which couples the domain semantics with dynamic, but asymmetric, connectives.

[57] Let s include both p-worlds and q-worlds. The local context for the left disjunct will be $s \setminus \{w : [\![\text{Might q}]\!]^{s,w} = 1\} = \varnothing$. So, in

also accepts 'The keys might be upstairs and they might be downstairs', which means this is predicted to be consistent with, and indeed jointly acceptable with, the negation of (67-a). This is clearly unacceptable. The bounded theory, although it has a symmetric disjunction, does not face this problem, because it does not *identify* the domain of quantification of 'might' with its local context; instead it has the local context *limit* the domain, which, as readers can confirm for themselves, avoids this issue.[58]

One response would be to adopt a *symmetric* conjunction together with an *asymmetric* disjunction, like Klinedinst and Rothschild (2012)'s, or the non-dynamic disjunction assumed in §3 above.[59] This would indeed avoid the present problem. But it leads to other problems. First, this approach still would not account for some of the data discussed above, in particular the infelicity of conditionals with the form ⌜If might p, then if not p, then q⌝, as well as much of the quantifier data. Second, the logic of this approach is quite strange. For example, conjunction elimination is not even Strawson valid on this semantics (by contrast to the bounded theory). To see this, consider an information state s which includes both p- and $\overline{p}$-worlds. ⌜Must p⌝ and ⌜Must not p⌝ are both false at this information state (relative to any world), since it's neither the case that all the worlds in s are p-worlds, nor that all the worlds in s are $\overline{p}$-worlds. But, startlingly, on the present approach, the *conjunction* of these claims, ⌜Must p and must not p⌝, will be true at s (relative to any world), since the local context for each conjunct will be the empty set, making each 'must'-claim trivially true. This is clearly unacceptable. There are various ways we could respond here; as Paolo Santorio points out to me, for instance, we could say that modals must be evaluated relative to non-empty local contexts. But other issues arise in the neighborhood. For instance, on this approach, for non-modal p, ⌜p and must p⌝ is predicted to be Strawson equivalent to p, which means that the inference from ⌜p and must p⌝ to ⌜Must p⌝ will not be semantically valid. The inference *will* still be valid in the logic of acceptance, but its semantic invalidity will have downstream consequences even in the logic of acceptance; for instance, ⌜Might (p and q)⌝ will no longer entail ⌜Might p⌝ in the logic of acceptance.[60] So it seems to me that the resulting logic

evaluating (67-b) at $\langle s, w\rangle$, we will evaluate the left disjunct at $\langle \varnothing, w\rangle$, and thus the left disjunct will be false, since the empty set does not include a p-world. Things are exactly parallel for the right disjunct.

[58] See §7.1 for further discussion, and for discussion of a related problem concerning disjunction. This discussion raises an interesting possibility of a kind of synthesis of the bounded theory and domain semantics: as on the bounded theory, we could let local contexts limit accessibility, rather than identify modal domains with local contexts; but, as on the domain theory, we could eschew accessibility relations, so that modals are instead directly associated with an information state (which must be a subset of the local context, but need not be identical to it). Coupled with our theory of local contexts, this would account for all our data without running into the problems pointed to here. I don't see much motivation for going this way, but, as far as the concerns of this paper go, such an approach is a viable variant on the bounded theory. Thanks to Paolo Santorio for very helpful comments on this point.

[59] Thanks to Paolo Santorio for encouraging me to explore this possibility.

[60] For instance, ⌜Might (must p and p)⌝ can be accepted while ⌜Might (must p)⌝ isn't, and indeed while the latter's negation is.

remains unsatisfactory.

A final and more abstract—but, I think, crucial—worry about the present approach is that it is unsatisfying from an explanatory point of view. As a methodological principle, we do not want to have to stipulate the interaction of embedding operators with epistemic modals on an *ad hoc* basis. Rather, we want to justify our choice in terms of a principled and unified algorithm, like the Schlenkerian algorithm the bounded theory adopts. This point is related to the 'and*' objection to Heim (1983)'s framework (Soames, 1989; Heim, 1990): why should conjunction have the left-to-right asymmetry encoded in her framework, rather than a right-to-left asymmetry (or none at all)?[61] I cannot see how we could justify in a principled way the coupling of a symmetric conjunction with an asymmetric disjunction—in sharp contrast to the principled, uniformly symmetric approach of the bounded theory.[62]

These comments bring out a feature of the bounded theory which has remained in the background so far but is worth emphasizing. Most theories of epistemic modals account for their embedding behavior by stipulating, on a case-by-case basis, how a given embedding operator should interact with embedded epistemic modals, based on particular embedding data.[63] The bounded theory, by contrast, has a single algorithm which makes predictions about the interaction of epistemic modals with embedding operators across the board: once the basic truth-conditional semantics of the embedding operator is specified, its interaction with epistemic modals follows without further stipulation. A theory of this nature is explanatorily preferable to one which specifies the interaction of embedding operators with epistemic modals in their scope on a case-by-case basis.

One way to see this is that a theory like the bounded one is more predictive, and thus more easily falsifiable, than a theory which goes connective-by-connective. Thus (building on analogous discussion in Schlenker 2008a) consider the question of how epistemic modals behave under a connective like 'unless' or 'because'. Without a general algorithm like the one I am proposing here, to answer this question we would have to first specify the core truth-conditional contribution of these connectives and then *separately* stipulate the interaction of these connectives with modals, based on particular data points like (68) and (69):

[61] Things would be even worse for this mix-and-match variant of the domain framework than for Heim's framework, which is at least uniformly asymmetric in a way that can be made precise (see Rothschild 2015).

[62] This approach would be particularly unsatisfying insofar as data from presupposition projection show symmetries more clearly across disjunction than across conjunction.

[63] This is brought out clearly in Ninan (2016)'s discussion of Yalcin 2007. Ninan shows that Yalcin's semantics for epistemic modals and connectives do not account for his data. The key move in Yalcin's system is rather adopting semantics for 'suppose' and 'if' custom-built to account for those data.

(68) a. #Unless John might not be sick and he is, we should get a doctor's opinion.

b. #Unless (might not p and p), q.

(69) a. #Because John might not be sick and he is, we should get a doctor's opinion.

b. #Because (might not p and p), q.

By contrast, the bounded theory, once it specifies the core truth-conditional semantics of 'unless' and 'because', already predicts how they interact with embedded epistemic modals—and, on any plausible such semantics, will predict the infelicity of (68) and (69). This is not to say that the domain, update, or relational semantics *couldn't* account for the data in (68) or (69)—the point, rather, is that they will not do so without case-by-case stipulations that go over and above truth-conditional analyses of these connectives; whereas the bounded theory accounts for these data as soon as we specify the truth conditions of 'unless' and 'because'.

In sum, the primary argument for the bounded theory is empirical: we should accept it because it captures a very wide range of data better than its competitors. But—even if a competitor theory could be found which captured the data as well as the bounded theory—we would then have to ask which theory captured those data in a more principled way. In particular, we should prefer a theory which captures the data in terms of a single predictive generalization, rather than by way of connective-by-connective stipulations.

Matters are similar when we turn to the update semantics. A natural first pass at a symmetric update semantics would go as follows (our semantics for negation, again, remains unchanged):

(70) *Symmetric update conjunction*: $c[\text{p and q}] = c[\text{p}][\text{q}] \cap c[\text{q}][\text{p}]$

(71) *Symmetric update disjunction*: $c[\text{p or q}] = c[\text{Not q}][\text{p}] \cup c[\text{Not p}][\text{q}]$

If we adopt these connectives, then we predict that Wittgenstein sentences will be inconsistent (as long as their modals take non-modal complements), again accounting for the bulk of the data seen above (though again missing some of the quantifier and conditional data). But the symmetric disjunction faces the same objection as above: ⌜Might p or might q⌝ will not be accepted in a context which contains p-worlds and q-worlds, and in fact will always 'crash' such a context to the empty set. And here there is an even more basic problem with symmetric disjunction. (71) is no longer a truth-conditionally respectable disjunction: it is not truth-conditionally equivalent to Boolean (inclusive) disjunction, but rather to exclusive disjunction.[64]

[64] Thanks to Daniel Rothschild for pointing this out to me. Suppose that s contains a world w where p and q are both true, and no other p-worlds or q-worlds. Intuitively, we want an update of s with [p or q] to result in $\{w\}$, since ⌜p or q⌝ is true at w. Instead, on this approach, it would result in $\varnothing$.

The issue is not just that we have picked the wrong symmetric dynamic connectives: as Rothschild (2015) argues, there is just no satisfactory symmetric dynamic semantics for disjunction.

Again, we could adopt a symmetric conjunction with a static or asymmetric disjunction (like Veltman (1996)'s or Groenendijk et al. (1996)'s, respectively). But this mix-and-match theory, while more empirically tenable, would, first, still fail to account for the infelicity of conditionals with the form ⌜If might p, then if not p, then q⌝; and, second, would face the explanatory challenge sketched above for the mix-and-match version of the domain theory: it could not, as far as I can see, be justified on the basis of a general explanatory and predictive algorithm like the one the bounded theory appeals to. I should emphasize that this challenge is not by any means meant to be decisive: it constitutes, I think, a strong *pro tanto* consideration in favor of the bounded theory, though one that could be overridden if other empirical benefits of, say, a domain or update semantics could be found. But, at the present state of research, it seems to me that the bounded theory remains both empirically and explanatorily superior to this version of the update semantics.

It is, finally, worth pointing out that, whatever our entry for disjunction, the symmetric dynamic conjunction sits very poorly with one of the characteristic ambitions of the dynamic framework in which the update semantics is spelled out: to give a unified account of anaphora, presupposition, and epistemic modality with the same system of local contexts. Together with standard dynamic theories of anaphora and presupposition, the symmetric conjunction predicts, for instance, that presuppositions always project out of conjuncts, which is clearly wrong, and that an indefinite in a left conjunct cannot license a definite in the right conjunct. The symmetric conjunction thus will not work with existing theories of these phenomena. This is not an argument against *dynamic semantics* in general (whatever exactly dynamic semantics is;[65] some may wish to describe the bounded theory as a dynamic theory). Indeed, it is possible to rewrite the bounded theory in a more superficially dynamic fragment, on which semantic values are functions from contexts to contexts rather than from indices to truth values. Rather, the present considerations constitute an argument against the particular treatment of epistemic modals in the update semantics—and in particular an argument that we cannot adopt the strategy of the update semantics of exploiting the asymmetric local contexts which dynamic semantics uses to account for anaphora and presupposition in order to also account for epistemic modals.

[65] See Rothschild and Yalcin 2015, 2016.

6.3 The dynamics of information in natural language

Let me close this section by taking a step back from epistemic modality to briefly consider the broader context of the bounded theory.

The bounded theory makes essential use of the system of local contexts developed in Schlenker 2009. The ability of Schlenker's algorithm to capture the subtle data discussed here speaks in support of its claim to capture something fundamental about natural language. But the present implementation of Schlenker's algorithm departs from Schlenker's own implementation in a few important ways. A first departure concerns the fact that we require that local contexts are *semantically* accessible; this is crucial for the locality constraint to be stateable (though local contexts need not be elements of the index, as I have assumed; they could instead be denoted by dedicated object-language variables, as suggested in Footnote 34). This differs from the pragmatic way in which Schlenker (2008a) initially introduced his algorithm (building on earlier pragmatic work in Stalnaker 1974; Karttunen 1974), though is consistent with later discussion in Schlenker 2009.

A more substantial departure concerns order. If the bounded theory is correct, then local contexts must be calculated symmetrically, and only symmetrically, at least insofar as they influence the interpretation of epistemic modals. By contrast, most work on the role of local contexts in presupposition projection, anaphora, and the theory of redundancy assumes that these systems involve genuine asymmetries. Symmetric algorithms for calculating local contexts have indeed been proposed, by Schlenker as well as by Rothschild 2015. But Schlenker proposes that the symmetric algorithm is recruited only when the asymmetric algorithm seems to be going wrong. And Rothschild suggests that a left-to-right parse and a right-to-left parse are typically both available. Neither of these approaches in application to epistemic modals would account for the data we have seen here: if both an asymmetric and symmetric (or left-to-right and right-to-left) algorithm were available, then we would expect that interlocutors would simply choose the algorithm which would render the data felicitous, resulting in prominent felicitous readings for many of the sentences which we have seen are marked.

There are a number of options in light of this divergence. One option with obvious theoretical attractions is to try to spell out a theory on which epistemic modals and presupposition (and perhaps anaphora and redundancy) operate on the same, symmetric system of local contexts. If we take this route, then we will have to attribute observed asymmetries in presupposition projection to the interference of independent

factors having to do with the way that we process sentences across time. There is much work to be done to see if an approach like this is plausible.[66] Another option would be to postulate that epistemic modals and presupposition work on different systems of local contexts, but systems that are intimately and systematically related. For instance, we might find that the best theory of presupposition operates on Schlenker's asymmetric system of local contexts (which is a very close variant of the symmetric system), and find, moreover, a theoretically satisfying explanation of why we calculate symmetrically in the case of epistemic modals and asymmetrically in the case of presupposition projection. A third possibility is that the best theory of presupposition (and perhaps also of anaphora and redundancy) will not make crucial use of Schlenker's system of local contexts at all; perhaps, as trivalent (George, 2008) and DRT-based (van der Sandt, 1992) theories have proposed, the notion is irrelevant to presupposition projection. If this were correct, we might then simply conclude that presupposition and epistemic modals operate on different systems. This would not be surprising *a priori*, and it would not deprive the bounded theory of its empirical accuracy or explanatory strength.

All of these theoretical options remain open; future empirical and theoretical work will hopefully clarify the relationship between epistemic modals and these other systems. The bounded theory advances that work by clarifying the demands on the system of local contexts from the perspective of epistemic modals. The bounded theory itself, however, does not stand or fall with the answers to those questions; my argument for the bounded theory is based on its empirical coverage and explanatory power vis-à-vis epistemic modals, and does not depend on any assumptions about the relation of epistemic modality to other semantic and pragmatic systems.[67]

[66] For a suggestion along those lines, see Rothschild 2008a. For some recent experimental evidence both for and against an approach like this, see Chemla and Schlenker 2012; Schwarz 2015; Mandelkern et al. 2017b. Straightforwardly applying Schlenker's symmetric algorithm to presupposition projection faces a technical problem pointed out by Rothschild (2008b); Beaver (2008); see Schlenker (2008b) for a possible, if not altogether satisfying, response.

[67] A different question to ask about unification concerns the cross-modal picture. If the bounded theory is correct, it to some degree leaves intact the Kratzerian promise of a unified theory of modality across modal flavors, insofar as a broadly relational approach remains the default candidate for treatment of other flavors of modality (though such a treatment has come in for challenges in those cases as well; see for instance Cariani et al. 2013; Cariani 2013; Mandelkern et al. 2017a for recent critiques of the relational approach to deontic and agentive modals, respectively). Having said that, not all modals are subject to the locality constraint characteristic of epistemic modals; while probability modals like 'probably' and 'likely' seem to be (a fact which can easily be captured by building semantics for those modals on top of the architecture of the bounded theory), root modals (deontic, agentive, and circumstantial) and weak necessity modals of different flavors ('ought', 'should') do not seem to be subject to a similar constraint. But this still leaves open the possibility of a broadly unified approach to natural language modals: the idea would be that all modals have a core relational semantics, but that epistemic modals are semantically distinguished by the locality constraint.

7 Alternative solutions

This completes my exposition of the bounded theory. In closing, I compare the bounded theory to what I see as the most promising alternative approaches to the data reviewed here. I first introduce a close variant on the bounded theory which I argue is worth serious consideration; I then explore the three extant theories of epistemic modals which seem best able to account for our data, pointing to problems for each of them.

7.1 The weak bounded theory

I consider, first, an alternative of my own which I argue we should take seriously. Recall that the locality constraint in the bounded theory says that all worlds can access only local context worlds. This constraint is stronger than what is strictly merited by the data we've seen. We can explain all those data with a slightly weaker constraint which concerns only *local context worlds*, as follows (call this variant the *weak bounded theory*):

(72) *Weak bounded 'might'*: $[\![\text{Might}_i\ \text{p}]\!]^{g,\kappa,w}$

a. defined only if $\forall w' \in \kappa : g(i)(w') \subseteq \kappa$; [*weak locality constraint*]

b. if defined, true iff $\exists w' \in g(i)(w) : [\![\text{p}]\!]^{g,\kappa,w'} = 1$. [*relational truth conditions*]

Coupled with the system of local contexts outlined above, this theory is identical to the bounded theory except that it puts no constraints on accessibility for worlds *outside* the modal's local context: where the (strong) bounded theory says that *all* worlds can access only local context worlds, the weak bounded theory says that *local context worlds* can access only local context worlds. Wherever the bounded theory predicted that a sentence was a contradiction, the weak bounded theory predicts that it is a *contextual* contradiction (that is, that it is false in every world in any context in which it is assessed), and thus that it embeds in a way which is contextually equivalent to a contradiction. A sentence which is guaranteed to be a contextual contradiction will generally sound just as infelicitous as a full-blown contradiction, and so the weak bounded theory can account for all the data that we have seen so far in essentially the same way as the strong bounded theory.

The weak bounded theory has a number of features which recommend it over the strong bounded theory. The weak bounded theory is *strictly weaker* than the strong bounded theory: it puts fewer constraints on

accessibility relations than the strong bounded theory. This has important empirical upshots. Consider (73):

(73) It might be raining, or it might not be.

Suppose that the 'might's in (73) are both evaluated relative to an accessibility relation which can access both rain- and not-rain-worlds from every context world. Then, on the strong bounded theory, (73) will be undefined, whereas on the weak bounded theory, it will be defined and true. Now, even on the strong bounded theory, there are many *other* accessibility relations relative to which (73) will be both defined and true, and so this point does not seem decisive to me; but this is an unattractive feature of the strong bounded theory avoided by the weak bounded theory.

A second empirical consideration in favor of the weak bounded theory comes from considerations about reflexivity, which were brought to my attention by Cian Dorr. The strong bounded theory is incompatible with the assumption that accessibility relations for epistemic modals are always reflexive, and thus cannot appeal to the most obvious explanation of the infelicity of a sequence like ⌜Must p. Not p⌝. The strong bounded theory can still assume that accessibility relations for epistemic modals are always *locally reflexive*: that is, that local context worlds can always access themselves. That, in turn, suffices to predict that ⌜Must p. Not p⌝ will be incoherent. But other cases suggest that local reflexivity is not enough. For instance, (74), suggested to me by Cian Dorr, seems to entail (75):

(74) Either John is here and must be upstairs, or Susie is here and must be upstairs.

(75) It's not the case that: John is here and Susie is here and both are downstairs.

But, curiously, this entailment is not valid on the strong bounded theory, even if we assume local reflexivity. On the other hand, the weak bounded theory is compatible with full reflexivity. And, if we assume full reflexivity, the weak bounded theory validates the desired inference from (74) to (75).

The strong bounded theory, in turn, has one consideration in its favor: it fits more naturally with a restrictor theory of conditionals than the weak bounded theory. For instance, if we assume that conditional antecedents add their content to the local context, then on a restrictor approach, the strong bounded theory predicts that ⌜If p, [must/might] q⌝ will be interpreted essentially in the way Kratzer 1981, 1986, 1991 argues it should be. By contrast, the weak bounded theory does not fit as well with a restrictor approach. How much this consideration weighs will depend on whether we want to adopt a restrictor theory of conditionals, and

if so, what its details amount to.[68]

Further discussion of these issues will have to await future work. I have elected to focus on the strong version of the bounded theory in this paper mainly for presentational reasons. The weight of the present considerations seem to me to tell in favor of the weaker version over the stronger version,[69] but both versions strike me as being worth further exploration.

7.2 Idempotent update semantics

I turn now to discuss three alternate approaches to epistemic modals from the existing literature which promise an interesting alternate account of the data I discussed above.[70] I begin by examining two enrichments to the update semantics. The first, given in Yalcin (2015), augments a standard update semantics (relevantly the same as the one given in §3) with the following constraint:[71]

(76) a. *Consecutive Update Idempotence*: For any sentence p with sentential constituents q and r, if $[p]$ is defined anywhere in part in terms of the composite update function $[q][r]$, that composite update function must be idempotent.

b. *Definition of Idempotence*: $[q][r]$ is idempotent iff $\forall c : c[q][r] = c[q][r][q][r]$

This constraint rules out as undefined sentences embedding a Wittgenstein sentence, since in the update semantics, Wittgenstein sentences will either be inconsistent or fail to be idempotent, accounting for most (but not all) of the data we saw above. The problem is that this constraint overgenerates. Consider a sentence like (77):

(77) a. John might be sick and Sue might be sick, but either John isn't sick or Sue isn't sick.

b. $(\Diamond p \wedge \Diamond q) \wedge (\neg p \vee \neg q)$.

[68] As Cian Dorr points out to me, since other kinds of modal (for instance, deontic) seem not to be sensitive to local contexts in general in the same way as epistemic modals, it's not clear that the resulting restrictor theory would retain the original attractions of the restrictor approach in any case.

[69] Cian Dorr has pointed out a variety of further considerations in this direction concerning elided and quantified modals which I find convincing, but don't have space to discuss here.

[70] Let me, note just to set aside, two other systems I will not discuss. One is developed by Hawke and Steinert-Threlkeld (2016). That system has problems with negation, predicting that ⌜Not must p⌝ is equivalent to ⌜Not p⌝, but recent bilateral developments in Aloni 2016, Steinert-Threlkeld 2017, Chapter 3 avoid this problem. Another system, sketched in Kratzer and Phillips 2017, uses situation semantics to make sense of some of the data discussed above. Both systems seem well worth serious further exploration.

[71] In the following sections I abbreviate natural language sentences with a standard modal propositional language for readability. A similar constraint is given in Klinedinst and Rothschild 2014. That constraint, interestingly, does not predict that a sentence like (77-a) is uniformly infelicitous, but faces the same objections raised against the view discussed in §7.3.

(77-a) is perfectly felicitous. But it violates *Consecutive Update Idempotence*. The CCP of a sentence with the form of (77-a) will be defined in terms of the sequence $[\Diamond p \wedge \Diamond q][\neg p \vee \neg q]$. But this sequence is not idempotent. *Consecutive Update Idempotence* thus wrongly predicts that sentences with the form of (77) will be infelicitous. The bounded theory, by contrast, correctly predicts that sentences like (77-a) are perfectly consistent and coherent.[72]

7.3 Accept update semantics

A second enrichment to the update semantics is given by Rothschild and Klinedinst (2015), who augment a standard update semantics with a further constraint, as follows:

(78) *Successive Update Rule (SUR)*: Let $[\cdot]$ be the update semantics denotation function from §3. Let $\mathbf{[\cdot]}$ be a new function defined recursively from $[\cdot]$ as follows: if there is no non-empty c such that $c[p] = c[q] = c$, then $c\mathbf{[}p\mathbf{][}q\mathbf{]} = \varnothing$. In all other cases, $c\mathbf{[\cdot]} = c[\cdot]$.

This stipulates that, unless there is some context which *accepts* each of two successive updates, then those successive updates yield the empty set. If we treat $\mathbf{[\cdot]}$ rather than $[\cdot]$ as our semantic value function for natural language, then this approach predicts that Wittgenstein sentences of all kinds are inconsistent, since no context accepts both $\ulcorner\Diamond p\urcorner$ and $\ulcorner\neg p\urcorner$, for any p, once again capturing the bulk of the data above. And this approach avoids my objection to Yalcin's idempotence rule, since, while $[\Diamond p \wedge \Diamond q][\neg p \vee \neg q]$ is not idempotent, there are contexts which accept both $\ulcorner(\Diamond p \wedge \Diamond q)\urcorner$ and $\ulcorner(\neg p \vee \neg q)\urcorner$.

I think this approach is promising. Let me bring out two observations, however, which speak against it. First, consider again (59):

(59) Not everyone who might be sick is well.

On the bounded theory, (59) communicates that someone is sick. This prediction, again, seems correct. By contrast, *SUR* predicts that (59) is true whenever its non-emptiness presupposition is met, and thus communicates nothing more than that someone might be sick. This seems incorrect: (59) tells us, not only that someone might be sick, but that someone *is* sick.

[72] Yalcin floats the idea that the *Consecutive Update Idempotence* rule applies only as a general preference, which is suspended when doing so would allow us to avoid disaster. If we took that tack, then we could explain the felicity of (77-a), but we would no longer explain the infelicity of Wittgenstein disjunctions, since presumably the constraint would just be suspended for left-modal Wittgenstein disjunctions, to avoid disaster. The difficulty is finding a principled way to explain why the constraint would apply to Wittgenstein disjunctions but not to (77-a).

A second problem is that *SUR* is sensitive to syntactic structure in a way which is not obviously correct. Compare (79) and (80), which differ only in where negation is placed—a difference which does not generally effect the interpretation of such sentences:

(79) a. Most people who might be sick aren't sick.

b. Most(might p)(not p).

(80) a. It's not the case that most people who might be sick are sick.

b. Not most(might p)(p).

SUR predicts a difference between sentences with the form of (79) and those with the form of (80). Given plausible semantic assumptions (as in Rothschild and Klinedinst 2015), updating with (79) goes by way of a sequence with the form $[\Diamond p][\neg p]$ and thus will be predicted to be inconsistent by *SUR*. By contrast, since negation takes high scope in (80), updating with (80) will not go by way of any objectionable successive updates, and thus (80) is predicted to be consistent.

Contrary to these predictions, there seems to be no contrast in felicity between (79) and (80). It also seems to me that both are consistent. Suppose that 100 people tested positive for a given sickness, but we know that the test has a rate of false positives above .5. It seems that in this scenario I can felicitously assert (79-a) or (80-a). That both sentences have the same status, and that both are felicitous, is the prediction of the bounded theory.[73] A similar point can be extended to other quantifiers, like 'few'. Broadly speaking, the bounded theory predicts invariance under certain syntactic permutations which *SUR* predicts will affect interpretation, but which do not seem to.

Let me close with a more abstract point against *SUR*. There is a sense in which *SUR* bleaches asymmetry out of the update framework. But there is something strange about constructing an asymmetric framework in the first place, only to bleach the asymmetry out later on. Why not move directly to a symmetric framework? There may be high level arguments for this kind of approach, but, at least at first blush, it looks unnecessarily roundabout.

[73] The local context for p in ⌜Most(p)(q)⌝ entails neither q nor $\overline{q}$: ⌜Most(p)(q)⌝ is not in general equivalent to ⌜Most (p and q)(q)⌝ or to ⌜Most(p and not q)(q)⌝.

7.4 Probabilistic contents: Moss 2015

I address, finally, the semantic system given in Moss 2015, 2018. I will gloss over some complexity in Moss's system which will be irrelevant for us. The essentials are the following. First, every sentence denotes a set of probability measures. 'Might' takes as its complement a set of probability measures. Connectives come in two flavors: one in which they they take as arguments sentences that denote sets of probability measures, one in which they take as arguments clauses that denote ordinary possible-worlds propositions (the latter have their ordinary Boolean semantics). To let modal and non-modal clauses combine, and to ensure that every sentence denotes a set of probability measures, Moss introduces a type-shifter $\mathcal{C}$ which takes a non-modal clause p to the set of probability measures built on sample spaces that entail p. In a Wittgenstein sentence with the surface form $\ulcorner\Diamond p \wedge \neg p\urcorner$or $\ulcorner\neg p \wedge \Diamond p\urcorner$ (order doesn't matter in this system), a modal clause is conjoined with a non-modal one. We therefore need to insert type-shifters for well-formedness. Here are two possibilities:

(81) $\Diamond\mathcal{C}p \wedge \mathcal{C}\neg p$.

(82) $\Diamond\mathcal{C}p \wedge \neg\mathcal{C}p$.

Given the type-flexibility of connectives in Moss's system, these are both well-formed. (81) will say (very roughly) ⌜Might p, and definitely not p⌝. In Moss's system, no probability measure satisfies both these conjuncts (to do so, it would have to have $Pr(p) > 0$ and $Pr(\overline{p}) = 1$). And so (81) will have as its semantic value the empty set. By contrast, (82) says (very roughly) ⌜Might p, and not definitely p⌝. There *are* probability measures that satisfy both conjuncts of (82) (to do so, they only have to have $Pr(p) > 0$ and $Pr(p) < 1$), and so (82) will have a non-empty denotation. If we stipulate that Wittgenstein sentences are always parsed as in (81), rather than as in (82) (for instance, by stipulating that type-shifters are always added as high as possible) then we predict that Wittgenstein sentences denote the empty set, accounting for much of the data above.

But there is trouble in the neighborhood. Compare the following:

(83) a. It's not the case that (it must be snowing or it must be raining), but it might be raining.

b. $\neg(\Box\mathcal{C}s \vee \Box\mathcal{C}r) \wedge \Diamond\mathcal{C}r)$

(84) a. #It's not the case that (it must be snowing or it is raining), but it might be raining.

b. $\neg(\Box\mathcal{C}s \vee \mathcal{C}r) \wedge \Diamond\mathcal{C}r)$

(83-a) and (84-a) are strikingly different: (83-a) is coherent—it says that we don't know whether it's raining, but it's not certain to be snowing—while (84-a) is incoherent (it is just an elaborated Wittgenstein sentence). In order to be interpretable in Moss's system, however, (83-a) and (84-a) *must* be parsed as in (83-b) and (84-b), respectively. But a little reflection shows that (83-b) and (84-b) are semantically equivalent in Moss's framework, since $\ulcorner\mathcal{C}r\urcorner$ and $\ulcorner\Box\mathcal{C}r\urcorner$ are equivalent in Moss's framework (they both denote the set of probability measures whose sample space entails r). This means that—whatever stipulations we might add about when type-shifting happens—Moss's system will predict that (83-a) and (84-a) are semantically equivalent; that both are interpreted like (83-a); and thus that both will be felt to be coherent. All three of these predictions are clearly wrong: (83-a) and (84-a) are strikingly inequivalent, and (84-a) is not felt to be coherent.

There is an element in Moss's system that I have ignored so far, namely a semantic role for partitions of logical space. It is true that, *relative to some choices of partition*, a sentence like (84-a) will denote the empty set. But it is hard to see how this helps, because this also holds for (83-a). A defender of Moss's view might try to argue that we always evaluate sentences like (84-a) relative to such partitions, accounting for their infelicity. But given that (83-a) and (84-a) are semantically equivalent, why would we always choose partitions which render the latter incoherent and the former coherent?[74] Short of an answer to this question, the present problem (which can be easily multiplied in a variety of ways) seems like a serious one.

8 Conclusion

There is strong reason to believe that ⌜Might p⌝ is classically consistent with ⌜Not p⌝; otherwise ⌜Might p⌝ would entail p. And if two sentences are jointly consistent, it follows in classical logic that their conjunction is consistent. These facts have led nearly everyone studying epistemic modals to conclude that Wittgenstein sentences are classically consistent. I have argued that this consensus is wrong. I have adduced a wide range of data—some new, involving epistemic modals embedded under disjunctions, modals, and in conditionals; some already known, involving epistemic modals embedded under attitude verbs, in condition-

[74]Thanks to Daniel Drucker and Sarah Moss for discussion on these points. Moss (2018), discussing a different but related problem, proposes a background presumption that a relevant proposition either must be the case or can't be the case. Adding this presumption would, of course, strengthen these problematic sentences to contradictions. But this leaves open the question, again, of why this background presumption obtains in the case of (84-a) but *not* in the case of the semantically equivalent (83-a). (One possibility is that this assumption holds for a given sentence just in case that sentence is nowhere embedded under an overt modal. But this can't be right, since embedding (84-a) under a modal does not improve it.)

als, and in quantified environments—which together show that Wittgenstein sentences always embed like classical contradictions. From these data I have drawn what seems to me the most natural conclusion: that, in spite of the fact that ⌜Might p⌝ is classically consistent with ⌜Not p⌝, Wittgenstein sentences nonetheless are classical contradictions.[75] I have proposed a theory of epistemic modals and their interaction with embedding operators which predicts just this, and which makes accurate predictions about the subtle embedding behavior of epistemic modals across the board. My theory makes these predictions in a logically conservative and principled way: given a truth-conditional semantics for an embedding operator, the bounded theory predicts how modals will behave under that operator, without *ad hoc* stipulations about how embedding operators shift the interpretation of epistemic modals. This suggests that the bounded theory is not only an advance in extensional adequacy, but also in explanatory accuracy.

The bounded theory shows that there is something right in two of the main approaches to epistemic modals. On the one hand, as the standard relational view has it, the truth conditions of epistemic modals are close to those of operators in modal logic: ⌜Might p⌝ is true only if p is true in some accessible world. On the other hand, as dynamic semanticists have argued, the interpretation of epistemic modals is sensitive to what information is locally available when we process them—though in a way which turns out to be starkly at odds with the standard dynamic approach. Thus, while the core meanings of ⌜Might p⌝ and ⌜For all we know, p⌝ turn out to be closely related, the two interact with their local information in very different ways. Interpretation of the former, but not the latter, is constrained by its local informational environment: epistemic modals contain something like an indexical reference to their local information.

Why are epistemic modals constrained like this—and what is the connection to presupposition, which seems to be dependent on its local information in a similar way (although not necessarily with the same symmetry properties)? These are big questions which require much further inquiry, but I suspect the answer will build on the insight that epistemic modals are, generally speaking, used to coordinate on information in conversation; the locality constraint makes epistemic modals a remarkably subtle tool for coordinating information in a local, and not just global, fashion.

A Appendix: Semantics

Here I compile the semantic entries of the bounded theory for ease of reference. With g a variable assignment and κ a context (set of worlds), the bounded theory is as follows:

[75] Or at least classical *contextual* contradictions, as on the variant given in §7.1.

(85) $[\![\text{Might}_i\ \text{p}]\!]^{g,\kappa,w}$

a. defined only if $\forall w' : g(i)(w') \subseteq \kappa$;

b. if defined, true iff $\exists w' \in g(i)(w) : [\![\text{p}]\!]^{g,\kappa,w'} = 1$

The weak bounded theory given in §7.1 is as follows:

(86) $[\![\text{Might}_i\ \text{p}]\!]^{g,\kappa,w}$

a. defined only if $\forall w' \in \kappa : g(i)(w') \subseteq \kappa$;

b. if defined, true iff $\exists w' \in g(i)(w) : [\![\text{p}]\!]^{g,\kappa,w'} = 1$

The semantics for the connectives and 'supposes' that follow from our local context algorithm, plus standard semantic assumptions about the underlying truth-conditional semantics, are as follows:

(87) $[\![\text{p and q}]\!]^{g,\kappa,w} = 1$ iff $[\![\text{p}]\!]^{g,\kappa^{\text{q}}_g,w} = 1$ and $[\![\text{q}]\!]^{g,\kappa^{\text{p}}_g,w} = 1$, where $\kappa^{\text{p}}_g = \kappa \cap \{w : [\![\text{p}]\!]^{g,\kappa,w} = 1\}$

(88) $[\![\text{p or q}]\!]^{g,\kappa,w} = 1$ iff $[\![\text{p}]\!]^{g,\kappa^{\overline{\text{q}}}_g,w} = 1$ or $[\![\text{q}]\!]^{g,\kappa^{\overline{\text{p}}}_g,w} = 1$, where $\kappa^{\overline{\text{p}}}_g = \kappa \setminus \{w : [\![\text{p}]\!]^{g,\kappa,w} = 1\}$

(89) $[\![\text{Not p}]\!]^{g,\kappa,w} = 1$ iff $[\![\text{p}]\!]^{g,\kappa,w} = 0$

(90) $[\![\text{A supposes q}]\!]^{g,\kappa,w} = 1$ iff $\forall w' \in S_{A,w} : [\![\text{q}]\!]^{g,S_{A,w},w'} = 1$, where $S_{A,w}$ is the set of worlds compatible with A's suppositions in w

For quantifiers, let 'p' and 'q' denote one-place predicates; then:

(91) $[\![\text{Some(p)(q)}]\!]^{g,\kappa,w} = 1$ iff $\exists x : [\![\text{p}]\!]^{g,\kappa\cap\{w' : [\![\text{q}]\!]^{g,\kappa,w'}(x)=1\},w}(x) = [\![\text{q}]\!]^{g,\kappa\cap\{w' : [\![\text{p}]\!]^{g,\kappa,w'}(x)=1\},w}(x) = 1$

(92) $[\![\text{The(p)(q)}]\!]^{g,\kappa,w} = 1$

a. presupposes $\exists! x : [\![\text{p}]\!]^{g,\kappa,w}(x) = 1$.

b. when the presupposition is satisfied, true iff $\exists x : [\![\text{p}]\!]^{g,\kappa,w}(x) = [\![\text{q}]\!]^{g,\kappa\cap\{w' : [\![\text{p}]\!]^{g,\kappa,w'}(x)=1\},w}(x) = 1$

(93) $[\![\text{Every(p)(q)}]\!]^{g,\kappa,w}$

a. presupposes $\exists x : [\![\text{p}]\!]^{g,\kappa,w}(x) = 1$

b. when the presupposition is satisfied, true iff $\forall x : [\![\text{p}]\!]^{g,\kappa,w}(x) = 1 \rightarrow [\![\text{q}]\!]^{g,\kappa\cap\{w' : [\![\text{p}]\!]^{g,\kappa,w'}(x)=1\},w}(x) = 1$

Finally, for conditionals, we generalize McGee (1985)'s variant of Stalnaker (1968)'s theory to our framework. We now include in our indices a pair comprising a Stalnakerian selection function f and a set of sentences Φ (we assume that in a matrix context, Φ is empty). Where A is an atomic sentence, let $[\![\text{A}]\!]^{g,\kappa,\langle f,\Phi\rangle,w} = 1$ iff $f([\![\text{and}(\Phi)]\!]^{g,\kappa,\langle f,\varnothing\rangle}, w)$ makes A true (according to the relevant atomic valuation), where 'and(Φ)' is the n-ary bounded conjunction defined in the obvious way by generalizing the binary bounded conjunction given above. Then our semantics for the conditional is the following:

(94) $[\![\text{If p, q}]\!]^{g,\kappa,\langle f,\Phi\rangle,w} = [\![\text{q}]\!]^{g,\kappa^{\text{p}}_g,\langle f,\Phi\cup\{\text{p}\}\rangle,w}$

References

Aloni, M. (2000). Conceptual covers in dynamic semantics. In Cavedon, L., Blackburn, P., Braisby, N., and Shimojima, A., editors, *Logic, Language and Computation*, volume III.

Aloni, M. (2016). FC disjunction in state-based semantics. Slides for Logical Aspects of Computational Linguistics (*LACL*), Nancy, France.

Aloni, M. D. (2001). *Quantification Under Conceptual Covers*. PhD thesis, University of Amsterdam, Amsterdam.

Barwise, J. and Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2):159–219.

Beaver, D. (1994). When variables don't vary enough. In Harvey, M. and Santelmann, L., editors, *Semantics and Linguistic Theory (SALT)*, volume 4, pages 35–60.

Beaver, D. (2001). *Presupposition and Assertion in Dynamic Semantics*. CSLI Publications: Stanford, CA.

Beaver, D. (2008). As brief as possible (but no briefer): Commentary on 'Be articulate'. *Theoretical Linguistics*, 3(38):213–228.

Beaver, D. I. (1992). The kinematics of presupposition. In *ITLI Prepublication Series for Logic, Semantics and Philosophy of Language*. University of Amsterdam.

Beddor, B. and Goldstein, S. (2018). Believing epistemic contradictions. *Review of Symbolic Logic*, 11(1):87–114.

Boylan, D. (2018). What the future 'might' brings. Manuscript, MIT.

Cariani, F. (2013). 'Ought' and resolution semantics. *Noûs*, 47(3):534–558.

Cariani, F., Kaufmann, M., and Kaufmann, S. (2013). Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy*, 36(3):225–259.

Cariani, F. and Santorio, P. (2018). *Will* done better: Selection semantics, future credence, and indeterminacy. *Mind*, 127(505):129–165.

Chemla, E. and Schlenker, P. (2012). Incremental vs. symmetric accounts of presupposition projection: an experimental approach. *Natural Language Semantics*, 20(2):177–226.

Condoravdi, C. (2002). Temporal interpretation of modals: Modals for the present and for the past. In Beaver, D., Kaufmann, S., Clark, B., and Casillas, L., editors, *The Construction of Meaning*, pages 59–88. CSLI Publications: Stanford, CA.

Dorr, C. and Hawthorne, J. (2013). Embedding epistemic modals. *Mind*, 122(488):867–913.

Egan, A., Hawthorne, J., and Weatherson, B. (2005). Epistemic modals in context. In Preyer, G. and Peter, G., editors, *Contextualism in Philosophy: Knowledge, Meaning and Truth*, chapter 6, pages 131–169. Oxford University Press.

von Fintel, K. (1994). *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts at Amherst.

von Fintel, K. (1998a). Evidence for presuppositional indefinites. Manuscript, MIT.

von Fintel, K. (1998b). The presupposition of subjunctive conditionals. In Percus, O. and Sauerland, U., editors, *The Interpretive Tract*, volume 25, pages 29–44. MIT Working Papers in Linguistics, Cambridge, Massachusetts.

von Fintel, K. (1999). NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics*, 16(2):97–148.

von Fintel, K. and Gillies, A. (2011). 'Might' made right. In Egan, A. and Weatherson, B., editors, *Epistemic Modality*. Oxford University Press.

George, B. R. (2008). Presupposition repairs: A static, trivalent approach to predicting projection. Master's thesis, University of California, Los Angeles.

Gerbrandy, J. (1998). Identity in epistemic semantics. In *Third Conference on Information Theoretic Approaches to Logic, Language and Computation*.

Geurts, B. (2005). Entertaining alternatives: Disjunctions as modals. *Natural Language Semantics*, 13(4):383–410.

Gillies, A. (2004). Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616.

Gillies, A. (2009). On truth conditions for *If* (but not quite only *If*). *Philosophical Review*, 118(3):325–349.

Gillies, A. S. (2018). Updating data semantics. To appear in *Mind*

Groenendijk, J., Stokhof, M., and Veltman, F. (1996). Coreference and modality. In *Handbook of Contemporary Semantic Theory*, pages 179–216. Oxford: Blackwell.

Hacquard, V. (2006). *Aspects of Modality*. PhD thesis, MIT.

Hawke, P. and Steinert-Threlkeld, S. (2016). Informational dynamics of epistemic possibility modals. *Synthese*.

Hawthorne, J., Rothschild, D., and Spectre, L. (2016). Belief is weak. *Philosophical Studies*, 173(5):1393–1404.

Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst.

Heim, I. (1983). On the projection problem for presuppositions. In Barlow, M., Flickinger, D. P., and Wiegand, N., editors, *WCCFL*, volume 2, pages 114–125. Stanford, Stanford University Press.

Heim, I. (1990). Presupposition projection. In van der Sandt, R., editor, *Presupposition, Lexical Meaning, and Discourse Processes: Workshop Reader*. University of Nijmegen.

Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell.

Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of Two Notions*. Cornell University Press, Ithaca, NY.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31(2):231–270.

Karttunen, L. (1974). Presupposition and linguistic context. *Theoretical Linguistics*, 1(1-3):181–194.

Khoo, J. and Mandelkern, M. (2018). Triviality results and the relationship between logical and natural languages. *Mind*.

Klinedinst, N. and Rothschild, D. (2012). Connectives without truth tables. *Natural Language Semantics*, 20(2):137–175.

Klinedinst, N. and Rothschild, D. (2014). Epistemic contradictions: why idempotence is hygienic. Handout.

Knobe, J. and Yalcin, S. (2014). Epistemic modals and context: Experimental data. *Semantics and Pragmatics*, 7(10):1–21.

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

Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H. and Rieser, H., editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter.

Kratzer, A. (1986). Conditionals. *Chicago Linguistics Society*, 22(2):1–15.

Kratzer, A. (1991). Modality. In von Stechow, A. and Wunderlich, D., editors, *Semantics: An International Handbook of Contemporary Research*, pages 639–650. de Gruyter, Berlin.

Kratzer, A. and Phillips, J. (2017). Epistemic modals—are there any? David Lewis Lecture, Princeton University.

Lasersohn, P. (2009). Relative truth, speaker commitment, and control of implicit arguments. *Synthese*, 166(2):359–374.

MacFarlane, J. (2011). Epistemic modals are assessment sensitive. In Egan, A. and Weatherson, B., editors, *Epistemic Modality*. Oxford University Press.

MacFarlane, J. (2014). *Assessment Sensitivity: Relative Truth and Its Applications*. Oxford University Press, Oxford.

Mandelkern, M. (2017). *Coordination in Conversation*. PhD thesis, Massachusetts Institute of Technology.

Mandelkern, M. (2018a). Dynamic non-classicality. Manuscript, All Souls College.

Mandelkern, M. (2018b). Import-export and 'and'. *Philosophy and Phenomenological Research*.

Mandelkern, M. and Romoli, J. (2017). Parsing and presuppositions in the calculation of local contexts. *Semantics and Pragmatics*, 10(17).

Mandelkern, M., Schultheis, G., and Boylan, D. (2017a). Agentive modals. *The Philosophical Review*, 126(3):301–343.

Mandelkern, M., Zehr, J., Romoli, J., and Schwarz, F. (2017b). Asymmetry in presupposition projection: The case of conjunction. In Burgdorf, D., Collard, J., Maspong, S., and Stefánsdóttir, B., editors, *Semantics and Linguistic Theory (SALT)*, volume 27, pages 504–524.

Mayr, C. and Romoli, J. (2016). A puzzle for theories of redundancy: Exhaustification, incrementality, and the notion of local context. *Semantics and Pragmatics*, 9(7):1–48.

McGee, V. (1985). A counterexample to modus ponens. *The Journal of Philosophy*, 82(9):462–471.

Moore, G. (1942). A reply to my critics. In Schilpp, P., editor, *The Philosophy of G.E. Moore*. Northwestern University, Evanston, IL.

Moore, G. (1959). Certainty. In *Philosophical Papers*. Allen and Unwin.

Moss, S. (2015). On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics*, 8(5):1–81.

Moss, S. (2018). *Probabilistic Knowledge*. Oxford University Press.

Ninan, D. (2010). Semantics and the objects of assertion. *Linguistics and Philosophy*, 33(5):355–380.

Ninan, D. (2016). Relational semantics and domain semantics for epistemic modals. *Journal of Philosophical Logic*.

Ninan, D. (2018). Quantification and epistemic modality. *The Philosophical Review*.

Rawlins, K. (2016). Particle 'or'. Handout, New York Philosophy of Language Workshop.

Rothschild, D. (2008a). Presupposition projection and logical equivalence. *Philosophical Perspectives, 22, Philosophy of Language*, pages 473–497.

Rothschild, D. (2008b). Transparency theory and its dynamic alternatives: Commentary on 'Be articulate'. *Theoretical Linguistics*, 34(3):261–268.

Rothschild, D. (2011/2015). Explaining presupposition projection with dynamic semantics. *Semantics and Pragmatics*, 4(3):1–43.

Rothschild, D. and Klinedinst, N. (2015). Quantified epistemic modality. Handout for talk at Birmingham.

Rothschild, D. and Yalcin, S. (2015). On the dynamics of conversation. *Noûs*, 51(1):24–48.

Rothschild, D. and Yalcin, S. (2016). Three notions of dynamicness in language. *Linguistics and Philosophy*, 39(4):333–355.

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(3):333–377.

Schlenker, P. (2008a). Be articulate: a pragmatic theory of presupposition projection. *Theoretical Linguistics*, 34(3):157–212.

Schlenker, P. (2008b). Presupposition projection: Explanatory strategies. *Theoretical Linguistics*, 38(3):287–316.
Schlenker, P. (2009). Local contexts. *Semantics and Pragmatics*, 2(3):1–78.
Schlenker, P. (2011). The proviso problem: a note. *Natural Language Semantics*, 19(4):395–422.
Schwarz, F. (2015). Symmetry and incrementality in conditionals. In Schwarz, F., editor, *Experimental Perspectives on Presuppositions*, volume 45 of *Studies in Theoretical Psycholinguistics*, pages 195–213. Springer.
Soames, S. (1989). Presupposition. In Gabbay, D. and Guenthner, F., editors, *Handbook of Philosophical Logic, Volume IV: Topics in the Philosophy of Language*, pages 553–616. Kluwer Academic Publishers.
Stalnaker, R. (1968). A theory of conditionals. In Rescher, N., editor, *Studies in Logical Theory*, pages 98–112. Oxford: Blackwell.
Stalnaker, R. (1974). Pragmatic presuppositions. In Munitz, M. K. and Unger, P., editors, *Semantics and Philosophy*, pages 197–213. New York University Press, New York.
Stalnaker, R. (1975). Indicative conditionals. *Philosophia*, 5(3):269–86.
Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25(5-6):701–721.
Stalnaker, R. (2014). *Context*. Oxford University Press.
Stanley, J. and Szabó, Z. G. (2000). On quantifier domain restriction. *Mind and Language*, 15(2-3):219–261.
Steinert-Threlkeld, S. (2017). *Communication and Computation: New Questions About Compositionality*. PhD thesis, Stanford University.
Stephenson, T. (2007a). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy*, 30(4):487–525.
Stephenson, T. (2007b). A parallel account of epistemic modals and predicates of personal taste. In Puig-Waldmueller, editor, *Sinn und Bedeutung*, volume 11, pages 583–97.
Stojnić, U. (2016). One's *Modus Ponens*: Modality, coherence, and logic. *Philosophy and Phenomenological Research*, 93(3):1–48.
Sudo, Y. (2012). *On the Semantics of Phi Features on Pronouns*. PhD thesis, Massachusetts Institute of Technology.
Swanson, E. (2015). The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic*, 45(121):121–146.
Szabolcsi, A. (1997). Quantifiers in pair-list readings. In Szabolcsi, A., editor, *Ways of Scope-Taking*, pages 311–347. Springer Netherlands.
Szabolcsi, A. (2015). What do quantifier particles do? *Linguistics and Philosophy*, 38(2):159–204.
Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261.
Willer, M. (2013). Dynamics of epistemic modality. *Philosophical Review*, 122(1):45–92.
Wittgenstein, L. (2001/1953). *Philosophical Investigations*. Wiley-Blackwell, 3rd edition.
Yalcin, S. (2007). Epistemic modals. *Mind*, 116(464):983–1026.
Yalcin, S. (2012a). Context probabilism. In Aloni, M., Kimmelman, V., Roelofsen, F., Sassoon, G. W., Schulz, K., and Westera, M., editors, *The 18th Amsterdam Colloquium*, pages 12–21.
Yalcin, S. (2012b). Dynamic semantics. In Russell, G. and Fara, D. G., editors, *Routledge Companion to Philosophy of Language*. Routledge.
Yalcin, S. (2015). Epistemic modality *de re*. *Ergo*, 2(19):475–527.
Zimmermann, T. E. (2000). Free choice disjunction and epistemic possibility. *Natural Language Semantics*, 8(4):255–290.