

The Character of Epistemic Modality: Evidential indexicals

1. Introduction¹

The influential work on modality in natural language due to Angelika Kratzer has contributed a thesis important for our understanding of Epistemic Modal Auxiliaries (EMAs) like *must* and *might* and other related epistemic expressions:

THE MODAL BASE PRESUPPOSITION: Natural language expressions that contain a modal component in their meaning, including all English modal auxiliaries and EMAs in particular, presuppose a modal base, a function that draws from context a relevant set of propositions which contribute to a premise-semantics for the modal.

We might say that the modal base contextually determines the *premise set* for the truth conditional contribution of the modal (cf. also Veltmann's 1985 premise semantics): truth of EMA ϕ will require that the EMA's prejacent ϕ be appropriately related to this premise set—e.g., follow from or be consistent with that set, or be probabilistically related to it in an appropriate way, etc.

Accepting this thesis for EMAs leaves open (at least) the following two questions about the meaning of English EMAs like *must* and *might*:

- i. What constraints, if any, are there on the character of the premise set for an EMA?
- ii. What is the nature of the relationship between premises and conclusion that is required for truth of the EMA statement?

¹ I am indebted to the members of the OSU project on Perspectival Expressions—Jefferson Barlew, Greg Kierstead, and Eric Snyder—for discussions of this material over many months and their own stimulating exploration of related ideas; to Nate Charlow, Guillermo Del Pinal, Janice Dowell, Kai von Fintel, Hans-Martin Gärdner, Thony Gillies, Leila Glass (see Glass 2013), Jack Hoeksma, Marcus Kneer, Phil Kremer, Ernie LePore, Emar Maier, Friederike Moltmann, Carl Pollard, Jennifer Spenader, Rich Thomason, Judith Tonhauser, and Brandon Waldon for stimulating discussions and comments, and especially to anonymous reviewers for the journal *Semantics and Pragmatics* for tough and extremely useful comments. I am also grateful to audiences at the Rutgers University ErnieFest in 2014, the MASZAT group at the Research Institute for Linguistics of the Hungarian Institute of Sciences in spring 2015, ReDRAW '15 at the University of Groningen, the 2018 meeting of the Society for Exact Philosophy at the University of Connecticut, a University of Toronto Philosophy Colloquium in 2018, the NY Philosophy of Language Workshop in spring, 2019, and the Philosophy of Language and Mind meeting in St. Andrews in August, 2019. The perspective project received invaluable support in 2013-14 from a Targeted Investment in Excellence grant from OSU, a Research Enhancement Grant from the OSU Colleges of the Arts and Sciences, and a supplement to NSF Grant #0952571, the latter originally awarded to Beaver, Roberts, Simons & Tonhauser; and from NSF Grant #1452674, 2015-2018, also to Beaver, Roberts, Simons & Tonhauser. An earlier version of this paper was completed while I was a Senior Fellow in 2014-15 at the Institute for Advanced Studies at Central European University, Budapest, Hungary, sponsored by Budapesti Közép-Európai Egyetem Alapítvány, and again, I am deeply grateful for their support, and for the assistance of OSU, without which I could not have accepted the fellowship. The theses promoted herein are the author's own, and do not necessarily reflect the opinion of the CEU IAS or any of the other sponsoring organizations or individuals cited here.

These questions are in principle independent. More generally, to accept the Modal Base presupposition, you don't have to agree with Kratzer's characterization of (human) necessity or possibility—with what logical relation between premises and “conclusion” (prejacent of the EMA) is required for truth.

Framed in the general Kratzerian approach in which the Modal Base presupposition is a central component, both these questions are presently very fruitful avenues of inquiry. I would argue that as usual for context sensitive expressions, in order to successfully address question (ii), you have to control for the factors that an answer to (i) should illuminate. In other words, you cannot get the semantics right until you properly control for context. Accordingly, here I will argue for at least a partial answer to (i).

That is, I'm interested in the proper constraints on the modal base for an EMA. I offer two hypotheses about those constraints:

INDEXICALITY: EMAs, unlike some other types of modals, are indexical: They are anchored to one of a limited set of agents whose doxastic (belief-based) perspectives are relevant in the discourse at the time of utterance. This is the set of **doxastic centers**.

EVIDENTIALITY: The modal base for an EMA is evidential and doxastic, not truly epistemic. I.e., in the terminology current in the literature, weak, not strong, so that *must p* does not entail *p*.

These constraints are modeled as presuppositions triggered by the EMA. That is, they are like selectional restrictions on predicates more generally.

As presuppositions regularly associated with the meanings of the EMAs, the constraints bear on contextual felicity, and are not directly reflected in the proffered content of the EMAs—their regular conventional contribution to the truth conditions for utterances in which they occur. Nonetheless, because they bear on modal accessibility relations assumed by the speaker, thereby restricting the domain of the modal operator, both presuppositions crucially contribute to the proposition expressed in any given context.

In keeping with the general Kratzerian approach to natural language modal auxiliaries, their proffered content can be very simple. Here, I'll assume universal (*must*) or existential (*might*) quantification over the restricted domain of possible worlds. Much of the interesting work is done by other aspects of their semantic character: how they conventionally appeal to the context of utterance to retrieve crucial features of their meaning in that context. Then because discourse is rich and complex, so that different contexts may vary along many distinct parameters, the resulting attested patterns of interpretation accordingly display complex patterns of variation.

Crucially, the pragmatic tools I will use to construct this account are all independently motivated.

- Premise semantics, as originally developed by Kratzer and by Frank Veltman, has long been used to develop sophisticated accounts of epistemic modals in natural language, including in recent work by von Stechow & Gillies (2007a, 2007b, 2008, 2010), Dowell (2011), and Lassiter (2016). von Stechow & Gillies (2010) in particular argue that the premises presupposed by an EMA in a Kratzerian semantics are evidential.

- The notion of a doxastic center in discourse, which will be used here to constrain which doxastic agents are available to anchor EMAS, is motivated both by my own work on indexicals and shifting indexicals (Roberts 2015), and by Jefferson Barlew's (2014, 2017) work on deictic motion verbs, like *come*, across languages.
- The Question Under Discussion (QUD), which will play a role here in explaining the widely observed weakness of an EMA relative to assertion of its prejacent alone, has been motivated as one of the foundations of pragmatic theory in an increasingly large body of work since Ginzburg (1994) and Roberts (1996/2012).
- The notion of a Reference Time, originally due to Reichenbach (1947), has been developed in detailed formal work by Hinrichs (1986), Partee (1984) and Dowty (1986). Partee in particular models it in an update pragmatics using Discourse Representation Theory, and the literature shows how it plays a central role in the interpretation of tense and aspect across languages. Klein (1994) refers to this as the Topic Time, a term picked up by others since. Here, the notion so modeled plays an important role in explaining the full range of data pertaining to Yalcin's (2007) puzzle.

What's novel here is bringing these tools together to contribute to a better understanding of the pragmatics of EMAS, as part of an answer to question (i) above.

In §2, I offer some background on the notions of a doxastic perspective and doxastic centers. In §3 we consider the proposed formal pragmatics and semantics for EMAS. Then in §4, in order to support the proposed analysis, I consider several puzzles about the interpretation of EMAS, including:

- the puzzle over why utterance of *must p* generally sounds like a weaker commitment on the part of the speaker to the truth of *p* than assertion of the simple prejacent *p* alone, and the debate over whether this means that the semantics of the EMAS itself is weak or strong (§4.1);
- the widely observed variability of anchoring of EMAS, and how to predict what the anchor of a given EMA uttered will be understood to be (§4.2);
- purported arguments for modal relativism (Egan, Hawthorne & Weatherson 2005, and others since) (§4.3.1), and
- Yalcin's (2007) puzzle, a version of Moore's paradox for embedded epistemic modals (§4.3.2).

In each case, I argue that the present account provides a simpler, more satisfying solution, both empirically and theoretically, than other proposals intended to deal with one or more of these puzzles.

Thus, though the presuppositional content proposed here is compatible with what Willer (2015) calls a "descriptive account" of epistemic *might* and *must*, such as the simple truth conditions assumed here, it avoids all the problems he poses there for earlier descriptive accounts. A number of non-descriptive accounts have recently been promoted as ways of dealing with these problems, including relativist proposals (like that of Egan et al. 2005), and expressivist accounts (including that of Yalcin (2007), and the dynamic expressivism of Willer (2015) and his other work cited there). But these in turn encounter their own significant difficulties, which are avoided on the present approach.

Finally, in §5, I offer some conclusions.

2. Background: Perspective and doxastic centers

Take a doxastic perspective to be the set of worlds where all the propositions are true that a given agent a believes to be true at a given time t in a given world w — a 's belief set at t in w . This is captured in classical modal logic by a modal accessibility relation Dox such that $\text{DOX}(\langle\langle a, t \rangle, w \rangle)$ yields the doxastic perspective of a at t in w . A central claim in the present account is that in any given discourse at any given time, there is a limited set of doxastic perspectives that are relevant and salient at that time. For example, as argued by Ninan (2010) and Stalnaker (2014), the Common Ground is such a purported belief state—the purported common beliefs of the interlocutors in the discourse. Accordingly, the purported beliefs of the speaker and addressee(s), and their joint perspective reflected in the Common Ground at the time of utterance are always relevant. Similarly, most attitude predicates report on the beliefs and other belief-based attitudes (knowledge, hopes, fears, expectations, etc.) of the agent of the attitude at the time at which the attitude state is reported to obtain, making that agent's doxastic perspective relevant and salient in the course of interpreting the propositional object of the attitude. We will say that the agent of a salient perspective in a given discourse at a given time is one of the available **discourse centers** at that time.

One argument for the pervasive role of doxastic point of view in natural language comes from Jefferson Barlew's (2017) cross-linguistic work on deictic motion verbs, like *come*, with special reference to English and the Bantu language Bulu. He shows quite clearly that *come* and Bulu *zu* are alike in presupposing that the location of the intended destination, or reference point, of the movement from one location to another is where the relevant agent *believes* (or fears, or hopes, or imagines. . .) *herself to be*, and not necessarily where she actually is. Just to give the flavor of this, here's just a bit of the relevant data from English:²

- (1) [Context: Abondo (speaker) and Bella (addressee) are in Avebe (destination). Guy is in Ebolowa (source). Guy leaves Ebolowa, traveling to Avebe. As he is traveling, Abondo tells Bella:]
Guy is coming to Avebe.
- (2) [Context: Abondo (speaker) and Bella (addressee) are in Avebe. Guy is in Ebolowa (source). Guy leaves Ebolowa, traveling to Kribi (destination). As he is traveling, Abondo tells Bella:]
#Guy is coming to Kribi.

² Regarding the symbol '#' used to indicate judgements about examples: one common usage in linguistic semantics and pragmatics is to indicate that the example (or the marked part of it) is infelicitous in the context in which it is uttered. This usually implicates that the expression might be felicitous in another context. When this kind of judgment is made, I generally make a practice of giving the relevant context. But in case none is given, that means that the example is uttered "out of the blue"—meaning that there is no specific relevant issue (explicit or implicit) in the context of utterance. Note that the "out of the blue" context *is* a context, albeit a special one. It's ok out of the blue to say *Please help me!*, but not *Is she hungry?* The latter could be said by one stranger to the other if they are watching a child cry, even if they hadn't previously spoken, but that is not an out of the blue context. It has been brought to my attention that at least some philosophers use the symbol with a different sense, meaning something like "infelicitous in all contexts", but that is not its sense here. If we cannot find any context in which an example may be felicitously uttered, that would seem to indicate that it is semantically anomalous in itself, which would be marked it with '!'.

These would be consistent with the claim that the reference point for *come*, the intended destination, is that of the speaker at utterance time. But Barlew argues that, *pace* all the proceeding literature, what really matters is where the relevant agent takes himself to be, based on examples like the following (evoking a famous example of the *de se* due to Morgan 1970):

- (3) [Context: Last week, Chicago baseball player Ernie Banks was hit on the head. He is now a lucid amnesiac. After the accident, Ernie was transported to Boston to work with an amnesia specialist. For all he knows, he has never been to Chicago. He has been reading about the baseball player Ernie Banks, but does not realize that he is reading about himself. He reads that President Obama was in Chicago 3 weeks ago and while there met Ernie Banks. The doctor later tells her friend:]
- a. #Ernie believes that President Obama came to Chicago.
 - b. Ernie believes that President Obama traveled to Chicago.
- (4) [Context: Identical to (3), except that Ernie regains his memory.]
Ernie believes that President Obama came to Chicago.

“In (3), the anchor, Ernie Banks, believes of Ernie Banks that he is located at the destination of the motion event. Nevertheless, (3a) is unacceptable because Ernie does not realize that he believes this about himself. His belief is not *de se*. In the minimally different context in (4) where Ernie believes *de se* that he was in Chicago, the example is acceptable.”

Barlew (pp.31ff, especially 38-43) argues that the possible anchors for *come* are those with salient doxastic perspectives—the same set of potential doxastic centers as Roberts (2015) argues are available for the indexicals. He backs up this claim with corpus evidence from The Corpus of Contemporary American English, COCA (Davies 2008). It seems like the following types of context serve to make a perspective sufficiently salient that it can serve to anchor a deictic motion verb like *come*:

- In discourse, the perspectives of speaker and addressee(s) are always salient. See (5) below.
- In the complement of an attitude predicate the perspective of the agent of the attitude (usually expressed by the embedding subject) is salient. This was first noticed by Fillmore (1975) and Hockett (1990). Oshima (2006a,b) calls this “deictic perspective shift”. See (6) and (7) below.
- In FID (Free Indirect Discourse), a common literary style where the author adopts the point of view of one of her (3rd person) characters—that character’s perspective may anchor ‘come’ (Doron 1991, Sharvit 2008, Eckardt 2014). See (8) below.

Addressee’s perspective:

- (5) [Context: Ann is in Cleveland, and Beth is in New York. They are talking on the phone. Ann says *Where is John these days?* Beth answers:]
John is in Chicago. However, he is coming to Cleveland tomorrow.

Perspective of the agent of an embedding attitude:

- (6) [Context: Ann is in Cleveland, Ben is in New York, and Chris is in Denver. On the phone, Ann asks Ben *Where is John these days?* Ben says:]
Chris {thinks/says} that John is coming to Denver today.
- (7) [Context: The interlocutors are in Columbus.]
- a. Ron says he's in New York, and he says Tom is coming to New York.
 - b. Ron thinks he's in New York, and he thinks Tom is coming to New York.
 - c. Ron is imagining he's in New York, and he's imagining Tom is coming to New York.
 - d. Ron wishes he was in New York, and he wishes Tom was coming to New York.
 - e. Ron is pretending he's in New York, and he's pretending Tom is coming to New York.
 - f. Ron is fears he's in New York, and he fears Tom is coming to New York.
 - g. Ron is wonders if he's in New York, and wonders if Tom is coming to New York.
 - h. Ron hopes to be in New York, and he hopes Tom will come to New York.
 - i. Ron is dreaming he's in New York, and he's dreaming Tom is coming to New York.

According to Barlew: “The key feature in these examples [in (7)] is that the anchoring implication is entailed by the information state relative to which *come* is interpreted.”

He notes (2017:40): “Importantly, other perspectival expressions shift under attitude predicates in exactly the same way (Mitchell, 1986; Sells, 1987; Speas and Tenny, 2003; Oshima, 2006c; Smith, 2009; Roberts, 2015), again providing support for the notion that *come* is a prototypical perspectival expression.”

Barlew also observes that the literary style FID can facilitate deictic perspective shift:

- (8) But with Mr. Ramsay bearing down on her, she could do nothing. Every time he approached - he was walking up and down the terrace - ruin approached, chaos approached. She could not paint She rejected one brush; she chose another. When would those children come? When would they all be off? she fidgeted . . .
[from Virginia Woolf's *To the Lighthouse*, as quoted in Doron (1991:52), where Woolf is describing the inner state of her character Lily Briscoe]

In FID, the author adopts the perspective of one of her characters; still using the 3rd person, she nonetheless may anchor a number of indexical expressions to the 3rd person doxastic agent of the perspective: Here the deictic motion verb *come* is anchored to the agent whose perspective is implicitly adopted, Lily Briscoe, who is also understood to be the source of the judgment that the approach of Mr. Ramsay was “ruin. . .chaos”.

Crucially, Barlew argues that “non-interlocutors [who aren’t the agent of an embedding attitude or the perspectival anchor of FID] are not generally acceptable anchors for *come*” even when they are contextually highly salient. See (9), where “despite the fact that the interlocutors are talking about Peyton Manning, he is not an acceptable anchor.”

- (9) [Context: Ann and Beth are in New York. Beth says:]
 Sarah is in Chicago right now. However, she is a Denver Broncos fan and loves
 quarterback Peyton Manning who has a game there tomorrow. #Therefore, she is coming
 to Denver tomorrow.

Note that Peyton Manning the individual is salient at the time of utterance of the last sentence; for example, he could felicitously serve as the antecedent for the pronoun *he* in a different final sentence: *He always puts on a great show*. Furthermore, Manning presumably self-ascribes being located in Denver. However, his perspective *per se* is not under discussion, hence neither salient nor relevant, and therefore he cannot be the anchor of *come*.

While space precludes discussion here, in Roberts (2015) I argue there that indexicals are always anchored to one of the salient discourse centers in this sense: the speaker for *I*, the addressee for *you*, their join for *we*, and in languages with shifted indexicals (e.g., where 1st person pronouns can be used to refer to the agent of an embedding attitude) the agent of the embedding attitude. In language after language, the same limited set of potential doxastic centers—a set we’ll examine just below—are the only potential referents of indexicals. It is striking to find this evidence across domains that are *prima facie* unrelated, and have heretofore been treated with unrelated semantic tools. I will argue below that the same inventory is crucial to understanding the possible doxastic anchors for EMAS.

In order to see how this doxastic anchoring plays out in the semantics of EMAS, we first need to briefly consider the kinds of pragmatic tools we will need to model the sense in which EMAS are perspectival, anchored to some salient doxastic perspective in discourse.

2.1 Doxastic states and doxastic perspectives

We can model the notions we need for the account of EMAS in terms of standard notions of what a doxastic state is and how it reflects an agent’s perspective, from the literature on the *de se*. Think of it intuitively as follows: A doxastic state is that of a given agent. Since our beliefs may change over time, we model the agent as an ordered pair: an agent at a time; call this a **doxastic center**. Of course, when we use this notion to determine a doxastic accessibility relation over worlds, we’ll consider the doxastic center’s beliefs in a given world. So call the ordered pair of a doxastic center at a world the **doxastic origin** for the accessibility relation. Then we assume a function **DOX** from doxastic origins to sets of worlds, yielding the set of worlds accessible from a given origin; call this the **doxastic perspective** of the given center at that world. Summarizing, we have the following:

A **proposition** is a set of centered worlds.

A **doxastic center** is an ordered pair consisting of a doxastic agent *a* and a time *t*: $\langle a, t \rangle$.

A **doxastic origin** is an ordered pair of a doxastic center and a world: $\langle \langle a, t \rangle, w \rangle$.

A **doxastic point of view** is a doxastic origin and its associated doxastic relation DOX.

A **doxastic perspective** is the information accessible from the point of view of a given doxastic origin $\langle \langle a, t \rangle, w \rangle$: $\text{DOX}(\langle \langle a, t \rangle, w \rangle)$.

Sometimes I get lazy and talk about a doxastic center as an agent, but technically it's an agent at a time. A doxastic origin is, of course, a centered world.

Why does the value of a doxastic accessibility relation amount to a perspective? There's a formal analogy: A point of view in actual space, the latter modeled as a three-dimensional Cartesian coordinate structure, can be characterized as a point in that space together with a vector for which the point serves as origin—a location plus an orientation (Barlew 2015). Together, the coordinate structure, origin and vector amount to a way of modelling an agent's perspective in physical space, with the agent at the origin. The vector suggests an accessibility relation from the vector's origin to other points ("locations") in the space—those perceptually accessible from the adopted point of view plus orientation. Summarizing:

Locative space is a set of points in a 3-dimensional Cartesian coordinate structure.

A **locative origin** is a point in locative space.

A **locative point of view** consists of a locative origin O and a vector V through the locative space, anchored in that origin: $\langle O, V_O \rangle$

A **locative perspective** is the set of points perceptually accessible from a given locative point of view: $\text{Percep}(\langle O, V_O \rangle)$

It is an old idea (Lyons 1977, and many other references) to extend the locative notion metaphorically to model a more abstract notion of point of view and perspective. Here is a way of formalizing that intuition: For *de se* interpretation, we can metaphorically generalize this locative model in order to characterize an agent's doxastic point of view—how they see the world as characterized in their belief state, yielding their doxastic perspective on a matter. In the locative structure, the agent takes a "stance" at the origin, oriented along the vector, and from that stance so oriented the agent has perceptual access to a region of physical space. Just so, in the doxastic case, the "space" is that of possibility, as usual in formal semantics modeled as the set of possible worlds. In this space, the perspective is that of an agent-at-a-time, an ordered pair we call a (doxastic) center, located in one of the possible worlds. Hence, the origin of a doxastic point of view is the ordered pair of a doxastic center and the possible world in which the agent has her beliefs at that time; this is a centered world: $\langle \langle a, t \rangle, w \rangle$, a the agent, t the time, w the world, in keeping with the literature on the *de se* (Lewis 1979). Parallel with Barlew's locative point of view, a doxastic point of view is a doxastic origin (the base centered world) plus a doxastic relation DOX, the latter reflecting the agent's orientation in the space of possibilities. $\text{DOX}(\langle \langle a, t \rangle, w \rangle)$ yields a set of centered worlds, the agent a 's **belief set** at t in w . But now each "point" in the space is a centered world $\langle \langle a', t' \rangle, w' \rangle$, a' at t' being "where" a at t in w takes herself to be in the doxastically live possibility w' , in keeping with the information to which she is privy. Thus, a doxastic origin serves as the base centered world in a *de se* belief relation, and the belief set of an agent at a time in a world is her perspective on the way things are from that point of view.

I adopt Stalnaker's (2008) constraints on the doxastic accessibility relation DOX. His **belief state** is a pair consisting of a centered world and its DOX-related belief set:

- the **base (centered) world**: a doxastic origin in the terms defined above.

- the **belief set**: the value of DOX applied to the base centered world, the doxastic perspective $\text{DOX}(\langle\langle a, t \rangle, w \rangle)$ in the terms defined above, a set of centered worlds. In each pair of a doxastic center and a world $\langle c, w' \rangle$ in the belief set, the c represents the individual that the base subject takes herself to be in w' , a world which, for all she believes, may be actual. The worlds in these pairs are those which would be accessible from the base world under Hintikka's doxastic accessibility relation relativized to the base center.

Stalnaker modifies Lewis' characterization of the accessibility relation over centered worlds. Chief among his reasons for doing so is the problem of *calibration*, crucial to comparing cognitive states. Hintikka's approach to propositional attitudes via modal accessibility relations makes possible comparison of the content of the objects of such attitudes across times and across persons: Two individuals A and B believe the same proposition p at t in w just in case both $\text{DOX}(A, t, w) \subseteq p$ and $\text{DOX}(B, t, w) \subseteq p$; similarly one individual believes p at two distinct times: $\text{DOX}(A, t, w) \subseteq p$ and $\text{DOX}(A, t', w) \subseteq p$, where $t \neq t'$. But in Lewis' account of centered worlds, an agent's centered belief set may include two distinct centered worlds which differ only in the center, so that they have the same world argument (and the same time); such a belief set indicates confusion about who the agent is within a single possible world. This means that two agents' belief states are not in general comparable—the centered worlds in their belief sets can never be the same unless they are mutually deluded in believing they are one and the same individual, since the centers in general will be distinct from one agent's set to the other's; and we cannot just compare the worlds in their belief states, since that might collapse possible distinctions in what one of the agents believes about who s/he is (in the example above, failing to reflect the agent's confusion about who she is in a single world). So in defining his doxastic accessibility relation, Stalnaker rules out such Lewisian accessibility relations with the constraint (*):

DOX is a binary relation on the set of all centered worlds that is transitive, Euclidean and serial, and also satisfies condition (*):

- (*) For any centers, c^* , c' and c'' , and worlds w and w' : if $\langle c^*, w \rangle \text{DOX} \langle c', w' \rangle$ and $\langle c^*, w \rangle \text{DOX} \langle c'', w' \rangle$, then $c' = c''$.

To say that $\langle\langle A, t \rangle, w \rangle \text{DOX} \langle\langle B, t' \rangle, w' \rangle$ is to say that it is compatible with what A believes at time t in world w that she is in world w' , that she is person B, and that the time is time t' . This means that if an agent is confused about who s/he is, then s/he is confused about what world s/he is in.³

Then effectively, insofar as the beliefs of two agents, or of one agent across different times, do not pertain to self-identity—are what Ninan (2010) calls “boring centered propositions”—we can

³ Lewis' (1979) argument for the possibility of belief sets containing two centered worlds $\langle c, w \rangle$ and $\langle c', w \rangle$, where $c \neq c'$, derives from his story about two distinct gods who are propositionally omniscient though neither can say which god he is. If the gods are propositionally omniscient, then they know which propositions are true, which false, and hence know which world they're in. Then, Lewis concludes, their confusion must lie in which center they are, *where they are in the actual world they've each correctly identified*. John Perry (p.c.), pointed out to me that the scenario Lewis paints is highly implausible from the point of view of the theory of action: If one of the gods throws down thunderbolts while the other throws down manna, and if both are omniscient, then the theory of action suggests that each *knows that he throws the substance he actually throws*.

compare those beliefs by ignoring the centers in the belief set, considering only the possible worlds in that set, as in the original Hintikka system:

[A] centered proposition p is boring just in case for any world w , and inhabitants x, y of w , $\langle w; x \rangle \in p$ iff $\langle w; y \rangle \in p$. Since boring centered propositions do not distinguish between worldmates, they are essentially equivalent to possible worlds propositions. (Ninan 2010)

Since in what follows we won't be focusing on the *de se*, we can present belief states—doxastic perspectives—as simple sets of possible worlds.

2.2 Doxastic Centers and Discourse Centers

In my earlier work on indexicality (Roberts 2015), I argued that interlocutors track the set of entities whose point of view is relevant at a given time, the set of **discourse centers** at that point in the discourse, and that this is part of the information available in the context of utterance. I proposed modeling these discourse centers with a special type of **discourse referent**.

The notion of a discourse referent is used in linguistics to track the available antecedents for anaphora in discourse. The notion goes back to Karttunen (1976); Heim (1982) modeled a discourse referent as a body of information—her ‘file’ if you will—that is purportedly about a single individual, which individual may or may not actually exist, or may be the arbitrary instantiation of the domain of a quantificational operator. Technically, discourse referents are represented as a set of special variables, each bearing an index which serves as the address for the file on one of these purported entities. A discourse referent has a “life-span”: if it is introduced, say, by an indefinite under the scope of an operator or as the arbitrary instance of an element of the domain of a quantificational operator, it is only available as a member of the set of discourse referents during the interpretation of the restriction and scope of the operator—its lifespan is restricted to the scope of the operator. So the set of available discourse referents is constantly changing during the course of a discourse; and the set of available discourse referents is a part of the specification of the context of utterance at any given time, as in the following simplified characterization:

Context of utterance at a given point in discourse $D = \langle CS_D, DR_D, \odot_D \rangle$:

CS_D : the Context Set, the set of worlds in $\cap CG$, where CG is the interlocutors' Common Ground (Stalnaker 1979), the set of propositions (sets of sets of worlds) that they all purport to take to be true

DR_D : the set of familiar discourse referents at that point in discourse. (Heim 1982, much subsequent work)

DR_D includes a distinguished set T_D of those times discussed in D , following much work on tense and aspect in discourse (Partee 1984, building on Reichenbach 1947),.

\odot_D is the set of discourse centers $\odot_{i,j}$, as defined below.

A discourse referent is **familiar** in the local context of utterance just in case all the interlocutors in the conversation are aware of its existence in that context; it may have been mentioned in that

context, or have otherwise been the focus of attention of the interlocutors, or it's some entity whose existence is widely recognized (e.g., *the sun*), at least among the interlocutors (e.g. *Aunt Jane*). A strongly familiar discourse referent is one which was introduced into the local context via an explicitly uttered NP. In contrast, a weakly familiar discourse referent need only be entailed to exist in the interlocutors' shared local context. All familiar discourse referents are weakly familiar, but only some are strongly familiar. In Gricean fashion, we reserve the term *weakly familiar* for those that are merely weakly familiar.

The set of discourse centers in D is constrained as follows:

$\odot_D \subseteq \{ \langle d_i, t_j \rangle \mid d_i, t_j \in DR_D \text{ \& } d_i \text{ is a doxastic agent whose beliefs at } t_j \in T_D \text{ are currently under discussion in } D \}.$

\odot_D always includes:

- a distinguished center $\odot_{i,j}^*$, corresponding to the speaker d_i at the time of utterance t_j ,
- a $\odot_{k,j}^@$ corresponding to the addressee at utterance time,
- and one corresponding to their join, inclusive 'we': $\odot^{*+@}$, with their joint (purported) point of view reflected in the CS.

Other centers are regularly introduced in conjunction with the interlocutors' consideration of alternative doxastic states:

- under the scope of an attitude predicate *pred*, \odot_D is expanded to include \odot^{pred} (e.g., \odot^{know} , $\odot^{believe}$, \odot^{hope} , \odot^{claim} , etc.). \odot^{pred} will be the ordered pair of the agent of the attitude at the time it obtains (the event time of the embedding clause).
- adverbials like *according to a* (at time t) or *from a's point of view* (at time t) introduce $\odot^{POV} = \langle a, t \rangle$
- in FID, reflecting the point of view of some third party agent under discussion, we have \odot^{FID} , the agent whose attitude at a given time is being adopted in the narrative.

Note that, as with discourse referents generally, when a discourse center is introduced under the scope of an operator or a predicate (in the latter case, in the predicate's propositional argument), the lifespan of that discourse center is limited to that scope. More generally, the set of available discourse centers in a given discourse D , \odot_D , is dynamic, updated regularly over time as (a) speakers and addressees change role, and as (b) we leave the scope of a predicate or operator that introduces a locally relevant \odot . For example, \odot^{fear} only stays in \odot_D in the local context consisting of the content of the complement of *fear*. When an FID passage has concluded, \odot^{FID} is removed from \odot_D , etc. This dynamic bookkeeping mechanism has a predecessor in the way that Reichenbachian Reference Time is modeled in Partee (1984), where it changes over the course of discourse as the interlocutors move from talking about one eventuality to another, usually with different event times.

What is crucial here for the account developed below is that (a) the pragmatics models how we track the set of doxastic centers currently salient in the context of utterance, and (b) that this set is dynamically updated in the course of conversation, even in the course of a single utterance. For formal realizations of how this works, consider Kamp's (1981) Discourse Representation Theory (DRT, Kamp & Reyle 1993), Heim's (1982) discourse referent update in her File Change Semantics, and Partee's (1984) related realization of Reichenbach's notion of Reference Time

using discourse referents in DRT. Discourse referents in these accounts are elegant book-keeping tools, as they are in many subsequent accounts of dynamic context change (e.g. recently AnderBois et al. 2015, Martin 2016, *inter alia*). But the particular way in which these are modeled is not especially important here. I have just chosen the simplest exposition of how to do this in a theory of how context influences interpretation.

I will call the conventional content of an expression its **CHARACTER**. This notion differs from that of Kaplan (1979), in that in keeping with dynamic theories of context like those cited above, context may be updated in the course of interpretation. Hence, indexical presuppositions needn't be satisfied by the global context of utterance, as they always are on Kaplan's account, but may be merely locally satisfied (Roberts 2015), so that an embedded indexical may have an interpretation which it could not have in the global context of utterance (as in languages that have so-called shifting indexicals). **CHARACTER** may involve at least the following types of content:

CHARACTER consists of:

Presupposed content: that which constrains the contexts of utterance in which utterance of the content is felicitous

Proffered content: that which enters into the compositionally calculated truth conditional content of the utterance in which it occurs—what is asserted, asked or directed

With these preliminaries, we can move on to discuss the **CHARACTER** of EMAS.

3. Evidential pragmatics and semantics for epistemic modals

Recall the informal discussion of the meaning of EMA *must* in the introduction, and in particular the evidential and indexical constraints on the presupposed modal accessibility relation which determines the domain of the modal in a given context. With the tools discussed in the previous section, we can now give the formal **CHARACTER** of *must* as follows:

CHARACTER of English epistemic *must*:

Given an utterance $must_{i,j} p$ in context $D = \langle CS_D, DR_D, \odot_D \rangle$, with world and time of evaluation w, t_j :

Presupposed content:

Indexical anchor: there is a $\odot_{i,j} = \langle d_i, t_j \rangle \in \odot_D$.

Evidential Modal Base: there is a function f mapping centered worlds to sets of propositions (each a set of worlds) s.t. $f(\langle \odot_{i,j}, w \rangle) = E^{vid}$, where:

$\cap E^{vid} \neq \emptyset$,

' E^{vid} is consistent'

$\forall p \in E^{vid}, \odot_{i,j}$ in w takes p to be plausibly true, and ' E^{vid} is plausible to d_i at t_j '

$E^{vid} \cap \text{Dox}(\odot_{i,j})(w) \neq \emptyset$.

' E^{vid} is consistent with d_i 's beliefs at t_j '

Proffered content: $\lambda p_{\langle s, t \rangle} \lambda w . p \subseteq \cap f(\langle \odot_{i,j}, w \rangle)$

Note that for simplicity I take E^{vid} to be a simple set of worlds, rather than a set of centered worlds. As noted in §2.1 above, this is possible because we don't consider here any *de se* propositions.

The CHARACTER of *must* is indexical and doxastic, because it is anchored to a presupposed discourse center $\odot_{i,j}$, whose belief state in w , $\text{Dox}(\odot_{i,j})(w)$, constrains the presupposed evidential modal base f , and hence the body of evidence E^{vid} it provides to restrict the domain of the modal. f will be related to the centering agent d_i 's belief state at t_j in the world of evaluation w : The conditions on the presupposed Modal Base f tell us that for $\langle \odot_{i,j}, w \rangle$ the function yields a consistent set of propositions that $\odot_{i,j}$ takes to be plausibly true in w ,⁴ a set which is consistent with all the propositions in $\odot_{i,j}$'s belief state in w .⁵ So E^{vid} is a body of evidence consistent with the anchor(s)' beliefs. As we will see, it is this possible extension of the center's belief state to include merely plausibly true propositions which makes the EMA statement evidential rather than truly epistemic or even strictly doxastic: the proffered content tells us that the center takes the evidence available to her to support the truth of ϕ , but this does not entail either that ϕ itself is true or that the center is firmly committed to its truth.

The proffered content of *must* is very simple: The prejacent is true in all the worlds in the presuppositionally restricted domain. As noted in the introduction, I do not offer evidence here for this simple semantics; but it will suffice to show the utility of the presupposed content that is crucial here.

Here is a comparable semantics for epistemic *might*, differing, as usual, only in that it merely requires that the prejacent be consistent with the modal base, instead of being entailed by it:

CHARACTER of English epistemic *might*:

Given an utterance *might* _{i,j} p in context $D = \langle \text{CS}_D, \text{DR}_D, \odot_D \rangle$, with world and time of evaluation w, t_j :

Presupposed content:

Indexical anchor: there is a $\odot_{i,j} = \langle d_i, t_j \rangle \in \odot_D$.

Evidential Modal Base: a function f mapping centered worlds to sets of propositions (each a simple set of worlds) s.t. $f(\langle \odot_{i,j}, w \rangle) = E^{\text{vid}}$, where:

$\cap E^{\text{vid}} \neq \emptyset$ &

' E^{vid} is consistent'

$\forall p \in E^{\text{vid}}, \odot_{i,j}$ in w takes p to be plausibly true, and

' E^{vid} is plausible to d_i at t_j '

$E^{\text{vid}} \cap \text{Dox}(\odot_{i,j})(w) \neq \emptyset$.

' E^{vid} is consistent with d_i 's beliefs at t_j '

Proffered content: $\lambda p_{\langle s, t \rangle} \lambda w . \cap [p \cup f(\langle \odot_{i,j}, w \rangle)] \neq \emptyset$

⁴ As is usual, assuming that the relevant belief state itself is consistent is an idealization. Though it is likely that most people's belief states are unwittingly inconsistent, without this idealization we cannot explore the respect in which the truth conditions of modals reflect what we take to be entailed by or consistent with a given premise set.

⁵ The MB f could be characterized more generally: For *any* given $\langle \odot, w \rangle$, it maps that center to a set of propositions s.t. it contains all the propositions \odot believes in w plus a set that are merely taken by \odot in w to be plausibly true. But here, since the modal is indexical to $\odot_{i,j}$, the only value that will matter to the proffered content is that for the indexical anchor and the world of evaluation.

Here we find the same presupposed constraints on the evidential modal base E^{vid} , but the usual existential semantics of the proffered content. Again, the reader is invited to explore the proposal with their own preferred proffered content.

Even though the evidence that the doxastic center takes to bear on the truth of the prejacent may go beyond what is strictly believed, it has many of the properties of belief: One cannot consistently both believe p and take *not- p* to plausibly be true. So, for example, if we take the relations denoted by *presume* and *suppose* to involve the explicit assumption of plausible truth, then the following are explained:

(10) #It isn't raining. But I presume it is.

(11) #It's raining. But I suppose it isn't.

These have much the flavor of Moore's paradox, a point which bears on Yalcin's (2007) paradox to be discussed in §4.2 below.⁶ In other words, we have derived a meaning that gives rise to a parallel to Moore's paradox, as we see in (10) and (11), but without belief in (or, hence, knowledge of) the prejacent *per se*.

The fact that epistemic modals are perspectival, plus independently motivated assumptions about context, explains and even predicts a wide range of attested properties of *must*, several of which we will explore in the next section. Here, let me only note that the above CHARACTER of the EMAS *must* and *might* does not require that the evidence which supports the truth of the prejacent be **indirect**, unlike the EMA semantics proposed by von Fintel & Gillies (2010), or that the evidence be in any way inferior in quality to the agent's beliefs. It is consistent with the speaker believing the prejacent. It merely says that the presupposed evidence E^{vid} which entails p may go *beyond* what she believes, since $\cap E^{vid} \subseteq \text{Dox}(\odot_{i,j})(w)$, i.e. the set of propositions E^{vid} given by the Modal Base f needn't equal but only be consistent with the anchoring agent's beliefs at the relevant time. This naturally gives rise to a Gricean Quantity implicature:

So long as p is relevant to the Question Under Discussion (QUD) (Roberts 2012), an interlocutor *should* proffer it if she truthfully can and is cooperatively committed to resolving the QUD. Then assuming that the speaker *is* cooperative, saying *must p* implicates that she is not in a position to assert p , i.e. that she does not firmly believe it. Most often, if one doesn't yet accept some proposition p as true, yet has explicitly considered whether p (as is necessary in order to take it to be plausibly true), that would be because the evidence one had for the truth of p wasn't of sufficiently high quality to foster conviction. It might be merely circumstantial; or it might involve hearsay, requiring one to have confidence in the source of the report; or it might be based on reasoning to the best explanation, leaving open unforeseen factors. Any of these might explain why one merely takes p to be plausibly true, instead of properly believing it. And this is the source of the sense that *must p* is "weaker" than p alone. Since this is a conversational implicature, one would expect that insofar as what it is to be an epistemic modal is to be based on plausible evidence, we should find the same implicature across languages, as von Fintel & Gillies (2010) claim.

⁶ Note that the case is somewhat different in the imperative, as we see by replacing the second sentences in (10) and (11) with their imperative counterparts: *But suppose it is /it isn't (raining)*. We return to the counterfactuality of imperative *suppose* in §4.2.

In some languages, evidentials are specialized to target specific reasons for merely supposing: e.g. in Quechua, we have both reportative evidentials (indicating that *p* is supposed on the basis of hearsay) and inferential evidentials. EMAs like *must* and *may* are more general, less specific. But the corresponding modal base reflects what it is to be evidential in the general case, with evidence being “[t]he available body of facts or information indicating whether a belief or proposition is true or valid” (Oxford Living Dictionaries: English (<https://en.oxforddictionaries.com/definition/evidence>), or in the specialized legal lexicon: “something legally submitted to a tribunal to ascertain the truth of a matter” (Merriam-Webster on-line: <https://www.merriam-webster.com/dictionary/evidence>). Hence we expect similar behavior across languages, including the display of “weakness” relative to straightforward assertion of the prejacent, not only with EMAs, but with evidential particles. So far as I know, this is the case.

I also note in passing that the constraints captured by the indexical and doxastic perspective presuppositions triggered by *must* and *might* are not yet sufficient to capture how context constrains the modal base that determines the domain of one of these EMAs. Another crucial factor is the QUD, as argued explicitly by Moss (2015) for the interpretation of epistemic probability operators like *probably*. See also related work by Windschitl & Wells (1998) and Yalcin (2010).

Now we turn to a more detailed exploration of some of the predictions of this semantics for the English EMAs *must* and *might*, and how it shines light on several puzzles discussed elsewhere in the literature.

4. Application to some puzzles about the meanings of EMAs

4.1. Weak necessity

A well-known apparent weakness in epistemic modality is reflected in the pattern illustrated by von Fintel & Gillies’ (2010) examples (12) and (13):

- (12) [Seeing the pouring rain]
 - a. It’s raining.
 - b. ??It must be raining.
- (13) [Seeing wet rain gear and knowing rain is the only possible cause]
 - a. It’s raining.
 - b. It must be raining.

In such minimal pairs, asserting *must p* is consistently infelicitous when one would be in a position to simply assert *p* instead. Some have argued that this would be unexpected if we take *must* to have the force of simple necessity, since with unrestricted modality $\Box p$ entails *p* (see the useful overview in Portner 2009, Chapter 4). von Fintel & Gillies (2010) provide arguments that

English epistemic modals like *must* and *may* are strong but evidential, with the evidence in question indirect, this feature of their character explaining the apparent weakness.

Here is their semantics:

Def'n 4: **Kernels and bases:** K is a kernel for B_K , B_K is determined by the kernel K , only if:

- i. K is a set of propositions (if $P \in K$ then $P \subseteq W$).
- ii. $B_K = \cap K$

Def'n 5: **Strong *must* + evidentiality.** Fix a c -relevant kernel K :

- i. $[[must \phi]]^{c,w}$ is defined only if K does not directly settle $[[\phi]]^c$
- ii. If defined, $[[must \phi]]^{c,w} = 1$ iff $B_K \subseteq [[\phi]]^c$

Def'n 5 presupposes that K doesn't directly settle ϕ , and proffers that B_K entails it.

"We...see no choice but to stipulate the evidential component of *must* in its lexical semantics, and we have to leave as unsolved the mystery of why this seems to be happening with every epistemic necessity modal that we have come across." vF&G (2010:368)

Comparing von Fintel & Gillies' semantics with that offered in §3, we see that though both proposals take the EMAs to be evidential, theirs differs from the present proposal in several respects. The crucial difference is that the characterization of the EMAs in §3 takes it to be related to belief, so that the modal base of the auxiliaries is doxastic; while von Fintel & Gillies do not appeal to belief, let alone the beliefs of any particular agent. This is at the core of the following important features of my proposal, which differentiate it from theirs:

With respect to the indexical anchor:

- (a) The CHARACTER of *must* proposed in §3 carries the presupposition of an indexical anchor, a particular (if possibly arbitrary) discourse center whose doxastic state is under discussion. As with any anaphoric presupposition, felicitous use requires that this presupposition can be readily resolved in the context of utterance. In the usual case, only one or two of the contextually available discourse centers will be candidates to anchor the EMA.
- (b) Rather than a "cloud of admissible contexts" (von Fintel & Gillies 2008), I assume that in felicitous use of *must* the modal base will be understood to be a body of evidence available to the presupposed doxastic anchor, one consistent with their belief state.

With respect to the Modal Base (cf. their Kernel):

- (c) The premises given by the presupposed Modal Base must be consistent with the anchoring agent's beliefs. Further:
- (d) The conditions on the Modal Base permit and implicate that the evidential ground E^{vid} supporting the prejacent p goes beyond the anchoring agent's beliefs. Hence:
- (e) This CHARACTER conversationally implicates that the evidential ground is consistent with but qualitatively weaker than that of the agent's firm beliefs, as sketched in §3.

None of these are features of von Fintel & Gillies' Kernel.

This last difference, (e), I take to satisfy the desideratum they note in their quote above, that the evidential signal should be derived as a non-detachable, predictable conversational implicature. Like von Fintel & Gillies, I do not weaken *must* through use of an ordering source (Kratzer 1991), or by making the prejacent a test on the context (Veltman 1985). But the proposed semantics doesn't involve any notion of indirectness, unlike Def'n 5.i above. And it isn't strong in their sense, entailing the truth of the prejacent.

This is consistent with carefully developed experimental evidence due to Lassiter (2016), Del Pinel & Waldon (2018), and Goodhue (2018) which argues that *must* is relatively weak, contra the predictions of von Fintel & Gillies.

Insofar as what it is to be an epistemic modal, across languages, is to be indexical in the sense defined here, with an evidential modal base anchored to a contextually given doxastic center, then we expect that the conversationally derived weakness will be a general feature of epistemic modals across languages.

4.2. Variable anchoring

There has been a good deal of discussion and debate about whose epistemic state can be appealed to in the semantics of epistemic modals, e.g. in Hacking (1967), DeRose (1991), Egan et al. (2005), MacFarlane (2005), Stephenson (2007), and von Fintel & Gillies (2007a). Everyone notes that there is at least a default tendency to understand the speaker to be the relevant agent whose epistemic state is at-issue. DeRose (1991:5) proposes a *speaker inclusion constraint*:

speaker inclusion constraint: the relevant community [must] include the speaker.
Hence “whenever S truly utters *a might be F*, S does not know that *a* is not *F*.”

But as discussed by von Fintel & Gillies (2007a), this requirement is much too strong. Here is a sample of the types of anchors attested for epistemic modal auxiliaries:

- a) speaker at Utterance Time (©*):
 - (14) John might be the thief.
 - (15) This suggests that Angela must be in Austin right now.

The speaker's point of view is always the default, for all perspectival expressions. Especially out of context, this is the only reasonable way to anchor an EMA.

- b) Group containing speaker at some actual time ≠ Utterance Time:
 - (16) Given what we knew at the time, John might have been the thief. [von Fintel & Gillies]

Here the adverbial explicitly shifts the doxastic point of view to that of the speaker and others in the anaphorically retrieved denotation of *we* at some time *prior* to the utterance time.

But there are cases where the grounds appealed to appear to be solely those of the addressee:

c) Addressee:

- (17) Where might you have put the keys? [von Fintel & Gillies]

As with evidential particles in many languages, this amounts to Interrogative Flip (Speas & Tenny 2003)—wherein the anchoring perspective naturally becomes that of the addressee when seeking information they might offer.

d) Arbitrary group containing addressee at some hypothetical time:

- (18) [Military trainer:] Before you walk into an area where there are lots of high trees, if there might be snipers hiding in the branches use your flamethrowers to clear away the foliage. [von Fintel & Iatridou 2003]

Note the imperative mood and indexical *your*: (18) constitutes generic instructions issued to the addressees. The *if*-clause constitutes a precondition for carrying out the instructions—they are applicable in circumstances in which the available evidence leaves open the possibility that there are snipers (Roberts 2018). Then the only reasonable anchor for *might* would be the trainee(s) in some actual (future) situation, who have to decide whether the correct conditions obtain.

e) Explicit third person orientation:

- (19) As far as Bill knows, John might be the thief. [von Fintel & Gillies]
(20) From John's point of view, it must be raining.
(21) This suggested to George that the polar ice cap might be melting.
(22) John thinks it must be raining.

(19), (20) and (21) illustrate several ways to explicitly shift the intended point of view. The center's time may be the present (19) (20), or some past time (21) (the past suggestion-time). With attitude predicates like *think* (22), shifting to the agent is the default (Stephenson 2007). But this is not a necessary shift. For example, in (23) and (24) the anchor probably includes the speaker (with or without the addressee), or perhaps a group that includes both the speaker and John, though the latter hasn't yet drawn the relevant conclusion from the information to which he's privy:

- (23) John won't acknowledge/hasn't yet conceded/hasn't realized that it must be raining.
(24) Has John realized that it must be raining?

Dowell (2011) notes that it's certainly possible to shift to an agent other than the agent of the embedding attitude when we use explicit shifting adverbials, as in this variation on one of her examples:

- (25) Leiter believes that, for all Blofield and No.2 know, Bond might be in Zurich.

(25) may be true even though Leiter knows that Bond is not in Zurich. And Dowell gives other convincing cases without explicit adverbial shifting (see her pp.23ff). But the most

convincing case I know involves the Mastermind scenario used extensively by von Fintel & Gillies (2008:83-84), the variation below suggested by Kai von Fintel (p.c.). This scenario consists of (constructed) dialogues between players of the game Mastermind. The game involves a pair of players, one of whom, the codemaker, has full access to an array of colored pegs; the other player, the codebreaker, tries to guess the colors and pattern within a limited period. Play usually involves a series of partial guesses by the codebreaker about the number of pegs of particular colors and their locations. At each turn, the knowledgeable codemaker can give positive feedback about what the codebreaker gets right. For this variation, imagine that the master codemaker Mordecai is giving a tutorial about how to be an effective codemaker. The class has been working through a mock game, with Pascal acting the role of codebreaker. Each class member, playing at codemaker, is supposed to track the information Pascal has to that point. Mordecai says:

- (26) You're tracking Pascal's evidence. At this point in play, you should know that there might be three reds, though in fact there are only two.

Here, the only coherent reading is one where *might* is anchored to Pascal rather than to either the speaker Mordecai or the agent of *know*, the addressee. If you find the example felicitous, as I do, then this is evidence that anchoring to an embedding agent is optional.

f) Modal subordination (Roberts 1989, to appear):

- (27) Local meteorologists believe that a hurricane is in the offing.
It might hit the coast, or it might veer out to sea.
I myself doubt this forecast.

In (27), the pronoun *it* in the second sentence requires an antecedent, the only plausible one being the indefinite hurricane that takes narrow scope under *believe*. Also, *might* requires a doxastic anchor. Taking *might* to be modally subordinate to the content of the complement of *believe*, so that the first clause of the second sentence means 'if a hurricane is in the offing, it might hit the coast', it is then natural to take the agent whose beliefs reportedly include the subordinating content—and hence reportedly believes in the existence of the forthcoming hurricane—to be the anchor for *might*. That is the set of local meteorologists. This is made clear here by the speaker's subsequent denial of that content. Hence, subordination to a doxastic context facilitates anchoring to the agent whose beliefs include the subordinating content.

g) Third person in Free Indirect Discourse (FID):

- (28) John pondered his situation. Where was Clarissa now? She might be in New Orleans with Sidney. But she might be in Chicago.

Here, as typical in FID, there is no explicit suggestion to the addressee to shift point of view from that of the speaker to that of the reported agent, John (Sharvit 2008, Eckardt 2014). Assume that the novel takes place in the 19th century, but we are reading it in the 21st. Then the use of *now* in querying Clarissa's location, despite past tense, suggests the shift.

- h) Multiple bodies of evidence, one more “objective” than the other:
 (29) Given the results of the DNA tests, John might be the thief. But if we take the eyewitness seriously, John can’t have been the thief. [von Fintel & Gillies]

The speaker in (29) explicitly suggests first one body of evidence, the DNA tests, and then another, the eyewitness’ account, which make contradictory predictions about whether John might have been the thief. This is a case where the speaker (and probably the addressee, as well) knows the DNA evidence, and knows that it can be interpreted as implying that John might have done it. The *if*-clause in the last sentence suggests that the speaker and addressee haven’t yet decided whether to accept the eyewitness’ evidence. The evidential semantics of EMAS argues for the speaker(+addressee) anchor ©*+@ for *can’t*: If the eyewitness saw the theft, that person would have the best available evidence and the speaker could simply report *But according to the eyewitness, John was the thief*. So the EMA is used because from the interlocutors’ point of view the evidence is hearsay. This example shows how context not only suggests the identity of the intended discourse center anchor, but in some cases may suggest the particular body of evidence supposed by that agent. Here two bodies of evidence available to the same anchor ©*+@—both plausible but neither absolutely determining the guilt of John—would lead to different conclusions.

von Fintel & Gillies (2007a) also suggest that EMAS might be anchored to non-human evidential sources (logs, charts, etc.), and that in such cases this leads to a more “objective” interpretation where the speaker’s opinion doesn’t count:

- (30) The hulk might be in these waters. [von Fintel & Gillies 2007a, after Hacking 1967]

But whatever the evidential sources for (30), if the speaker knows of the evidence they provide, then s/he can serve as anchoring center. If there’s an implication that the body of evidence is “objective” or “consensual” or generic, then the modal statement may be taken to have more force—after all, evidentiality is about the speaker’s judgment of the quality of her evidence for certain beliefs, and higher quality evidence makes the prejacent seem more likely to be true. Such examples seem particularly prone to the inclusive ‘we’ interpretation, where the Common Ground contains such “objective” evidence; and I suspect the common consensus about the value of the evidence strengthens the sense of objectivity, as well. But (30) is still evidential in the sense defined here—suppositional rather than confidently assumed to be true; and the non-human provenance of the evidence doesn’t mean that the anchor itself is non-agentive; it may be the epistemic agent who has that information.

Let me reiterate the central hypothesis about discourse centers: For an individual to be a discourse center, it does not suffice that that individual be an agent who is familiar and salient to the interlocutors and is capable of having a doxastic point of view. Hence, e.g., generally we don’t readily get exocentric readings of EMAS under attitudes, unlike what we find in Predicates of Personal Taste. We see the distinction in these examples from Stephenson (2007) (her (23) and (24)):

- (31) Mary: How’s that new brand of cat food you bought?
 Sam: I think it’s tasty, because the cat has eaten a lot of it. [Kai von Fintel, p.c.]

(32) [Context: The dog is eating its food with enthusiasm that he usually saves for table scraps.]

Mary: Wow, the dog really likes the dog food you're feeding him.

Sam: # I think it might be table scraps.

Evidently, for Predicates of Personal Taste like *tasty*, the anchor only need be some agent whose taste is at-issue, as is arguably the case in (31). But for EMAs, the anchoring individual's doxastic state itself must be relevant; that arguably isn't the case for the dog in (32), so that the dog-anchored interpretation of *might* is unavailable, infelicitous in this context. Compare this with (26), where Pascal's evidential state *is* arguably at-issue, and the exocentric reading is more readily available.

Lest one think that the perspective of a dog isn't sufficiently relevant to be reflected in a discourse center, undercutting the value of (32), consider the following extension of the Mastermind scenario:

(33) [During a break in a game of Mastermind, the codemaker Mordecai is talking to a friend, Bart, about the game. Mordecai knows that there are only two reds. Talking about the codebreaker Pascal, he tells Bart:]

Pascal is playing pretty well, but his memory isn't the best. He remembers the evidence for most of the colors, but has forgotten that there's one green. #Therefore, there still might be three reds.

Here Pascal is quite salient, but he still cannot felicitously serve as anchor for *might* unless we add *according to him* to the last sentence. Pascal isn't an interlocutor in the conversation, the clause in which the EMA occurs is unembedded, and there is no evidence of FID, so there is no discourse center corresponding to Pascal at this juncture in the discourse. Unlike in (26), Mordecai needn't know how well Bart is tracking the state of the board, so has no evidence to anchor *might* to Bart's perspective. Even if Mordecai were willing to mislead Bart about what was on the board, (a) an assertion about Mordecai's own state wouldn't be relevant here (where they're talking about Pascal's) and (b) Bart knows that Mordecai knows *exactly* how many reds there are, so that it would be misleading at best to assert the proposition we get by anchoring *might* to Mordecai. Hence, the example is infelicitous because *might* remains unanchored. So just as in Barlew's (9), where Peyton Manning's location in Denver is salient but that doesn't suffice to anchor *come* felicitously to Manning, here Pascal's perspective is salient, but that doesn't suffice to felicitously anchor *might* to Pascal.

Hence, the restriction of doxastic anchors to the set of discourse centers predicts the pattern of interpretation observed above, typical of anchors for indexical expressions generally (Roberts 2015, Barlew 2015). With EMAs, not only does the anchor vary, in this restricted way, but it gives clues about the body of evidence which is intended to restrict the modal's domain.

Summarizing, we see that for EMA domain restriction:

- in root declaratives, evidence to which the speaker or inclusive 'we' is privy is the default (14) – (16)

- in root interrogatives, evidence to which the addressee is privy is the default, i.e. interrogative flip (17)
- in imperative conditional protases, as in (18), evidence to which the addressee is hypothetically privy is the default, as this information is useful for determining the applicability of conditional advice or directions
- in attitude complements (22) – (24), evidence to which the agent of the attitude is privy is the default
- with explicit adverbials like those in (19), (20) and (21), we see a shift similar to that in attitude complements
- in modal subordination (27) the evidentiality is anchored to the relevant agent in the doxastic subordinating context
- in FID (28) the default is the evidence of the doxastic agent whose perspective is adopted by the author.

In most of these contexts, where the indicated anchoring is a mere default, it can be pragmatically over-ridden by other available discourse centers (McDowell 2011). The resolution of the anchoring presupposition of an EMA is anaphoric presupposition satisfaction (Heim 1983), hence subject to the usual constraints on anaphora resolution: familiarity and salience of the intended discourse referent antecedent (here, a discourse center), pragmatic plausibility, and coherence of the resulting resolution. A useful contrast is between the anchoring of EMAs in attitude complements and those governed by explicit shifting adverbials like *from Abe's point of view*. With the adverbials, the shift appears to be obligatory, conventional, whereas in the attitude complements it is only the pragmatically governed default. In the latter it is just that the agent of the attitude tends to be the most salient center in that context, so that anaphora resolution naturally points to that center, all other things being equal.

4.3 **Epistemic modals in interaction with the attitudes**

In the recent literature, proofs involving semantic interactions between EMAs and attitudes like knowledge and belief have been used to argue for Relativism—the post-truthconditional assessment of content relative to a Judge—and for a kind of Expressivism about EMAs—a failure of the clauses in which they occur to have their usual truth conditional content. Here I will consider two prominent accounts along these lines, due to Egan et al. (2007) and Yalcin (2007), and argue that on the present account we need neither Relativism nor Expressivism to predict the kinds of interpretations attested.

4.3.1 **Granger's puzzling proof**

Determining the intended anchor for a given modal auxiliary not only makes a difference to truth conditions, but, as we might expect, also to the patterns of inference associated with the use of epistemic modal auxiliaries. Given the typically privileged status of an anchoring agent's evidence, this is especially apparent in examples where the anchor shifts across different uses of the same modal in a single argument. Moreover, the pragmatic basis of the determination of the anchor for an epistemic modal auxiliary differs from the way that the anchor for an attitude

predicate like *know* is determined compositionally. When all these factors interact, the result can lead to some complexity in analyzing the resulting argument.

We see this in the following example from Egan et al. (2005), which is the basis of their central argument for relativism about epistemic modals:

- (49) [Context: Professor Granger is in the South Pacific and knows it. She heard her friend Myles, in Boston, on the radio speculating about where the missing Granger might be. Myles knew that she'd originally planned to go to Prague, but not where she actually went. She puzzles about the following statements, each apparently true:]
- (a) When he says, "She might be in Prague" Myles says that I might be in Prague.
 - (b) When he says, "She might be in Prague" Myles speaks truly iff neither he nor any of his mates know that I'm not in Prague.
 - (c) Neither Myles nor any of his mates know that I'm not in Prague.
 - (d) If Myles speaks truly when he says that I might be in Prague, then I might be in Prague.
 - (e) I know I'm not in Prague.
 - (f) It's not the case that I know I'm not in Prague if I might be in Prague.

Supposedly, in (49) steps (a) – (e) lead to a contradiction in (f), but that seems counterintuitive.

The puzzle comes from looking at surface form only, giving rise to the following apparently sound argument (taking *I* in (49) to consistently refer to Granger):

Premises: the evidently true (a) – (f)

- (b)+(c) |= Myles speaks truly when he says "She might be in Prague" = (i)
- (i)+(a) |= Myles speaks truly when he says Granger might be in Prague = (ii)
- (ii)+(d) |= Granger might be in Prague = (iii)
- (f)+(iii) |= It's not the case that Granger knows Granger is not in Prague. = (iv)
- (e)+(iv) |= Granger knows Granger is not in Prague and it's not the case that Granger knows Granger is not in Prague.

Contradiction.

But as we have seen in earlier sections, epistemic modal auxiliaries are indexical, anchored to an available doxastic agent. This means that in assessing truth conditions and propositions expressed, we have to take into account not just the surface form of the premises in the argument, but the way the modals are anchored. Ignoring (f) for the moment, it is clear from the story that the following anchors are understood, where subscript *M* indicates that Myles is the doxastic anchor, *G* that it is Granger, and *I* is always Granger:

- (a) When he says, "She might_M be in Prague" Myles says that I might_M be in Prague.
- (b) When he says, "She might_M be in Prague" Myles speaks truly iff neither he nor any of his mates know that I'm not in Prague.
- (c) Neither Myles nor any of his mates know that I'm not in Prague.
- (d) If Myles speaks truly when he says that I might_M be in Prague, then I might_M be in Prague.

(e) I know I'm not in Prague.

(f) It's not the case that I know I'm not in Prague if I might be in Prague.

In (a) and (b) as uttered in the story, *might* is clearly anchored to Myles, because Granger is reporting what Miles said when he was talking about the missing Granger and answering questions about what *he* knew about where she might be. The only reasonable way to understand *might* in (d) is with the anchoring indicated: In the *if*-clause of (d), *might* has to be anchored to Miles because, as we saw in the abbreviated logical form of the argument above, this is the only way for (d) to combine with (ii) to allow us to deduce (iii) via Modus Ponens; i.e., it is crucial to have the same anchoring in (ii) as in the *if*-clause of (d). And the plausibility of (d) itself depends on it reflecting the natural assumption that if one speaks truly in claiming that *p*, then *p* is true. Hence, the proposition in the consequent must have the same logical form as that in the complement of *says* in the antecedent. And thus, in the logical form of (iii), so deduced, *might* is consistently anchored to Miles, as it was in (ii) and (d).

(e) involves no epistemic modal, but it has the verb *know*, itself epistemic. Given its lexical semantics, *know* has to take its subject's denotation, Granger, as the anchoring agent for the reported belief state: It's true in all Granger's belief-worlds that she's not in Prague.

Then what does (f) mean? For the puzzle to arise, Egan et al. again need this to be a fairly obvious truth, so they need a logical form like the following:

(f') if I might_A be in Prague [then] it's not the case that I know I'm not in Prague

if-clause:

I might_A be in Prague:

$\Diamond_A P(G)$

main clause:

scope of *it's not the case that*:

I know_A I'm not in Prague

$\Box_A \neg[P(G)]$

then:

it's not the case that I know_A I'm not in Prague $\neg(\Box_A \neg[P(G)]) =$

$\Diamond_A P(G)$

This logical form requires that the same agent A that anchors the doxastic state appealed to in the lexical content of *know* also anchors *might*, so that the content of the two clauses is essentially equivalent. But in the story, the speaker Granger is the denotation of the subject of *know*. Hence, to get the plausible logical form, we have to take Granger to anchor *must* in the *if*-clause: A = Granger. Then (f') means 'if for all Granger knows Granger might_G be in Prague, then it's not the case that Granger knows she's not in Prague'.

But plugging this back into (49) will not yield the contradiction. In the schematic derivation, that contradiction was entailed by (f) plus (iii) under Modus Ponens. MP is only applicable if (iii) has the same logical form as the *if*-clause of (f). But we have seen that in (iii) *must* is anchored by Miles, while in (f), it is anchored by Granger.

Instead, in the story we have something like the following logical form for (f):

(f'') if I might_M be in Prague [then] it's not the case that I know_G I'm not in Prague

$I_G \text{ might}_M \text{ be in Prague:}$

it's not the case that $I_G \text{ know}_G I'm \text{ not in Prague}$ $\neg(\Box_G \neg[P(G)]) = \Diamond_M P(G)$
 $\neq \Diamond_G P(G)$

'if for all Miles knows Granger might be in Prague, then it's not the case that Granger knows that Granger is not in Prague'

But (f'') simply isn't true. Just because Miles is in the dark, that doesn't mean that Granger doesn't know where she is. And the force of the argument depended on the plausibility of each of the premises, including (f).

So, on the only reasonable interpretation of (f), (f'), there's no contradiction with (e).

The crucial premise here, then, was (f). And it was important for its plausibility that the epistemic agent of the modality implicit in *know* was the same as the anchor of *must*. But though both *know* and *must* are doxastic, they differ in how they may be anchored: *must* is variably anchored, the intended anchoring contextually resolved, while the agent of *know* is given lexically and must always be its subject.

Egan et al., of course, are well aware of these issues. But their Relativist account takes a different tack than the present one. Relativism assumes that the proposition expressed does not take into account the kind of contextual anchoring assumed here. Instead, in evaluating the truth of the proposition, one considers the point of view of some salient judge. So the judge is, as it were, post hoc to the compositional semantics, whereas here, indexical anchoring is part of the determination of the proposition expressed. One might say that Relativist semantics for EMAs is like an inverse of the presuppositional anchoring account proposed here: there, the judge stands aside from the proffered content of the expression, and here the anchor does as well. But in Relativist accounts the judge comes after the fact, after the propositional content has been determined; while on the present account the anchor is presupposed, hence part of the usual process of felicity determination, and it indirectly plays a role in the resulting truth conditions—the proposition expressed. On the present, indexical approach, there is no need for Relativism to explain the attested meanings.

The account I propose here is a Contextualist account of the EMAs. I note that Egan et al. (2005) argue against a Contextualist account of *might* and *must* partly on the grounds that their context sensitivity is different from that of purely anaphoric elements like *local* or *enemy*, which are much more promiscuous in how their implicit arguments are anchored. That is, their (50) would be felicitous without prior discussion or relevance of any particular locale:

(50) Many local bars are full of Browns fans.

Professor Granger (still in the South Pacific), hearing this spoken over Skype by her friend Jason in Cleveland, will not hesitate to say "that's right". The fact that the relevant bars aren't local to Granger doesn't interfere with her willingness to agree with (50), unlike how the fact that she knew that she wasn't in Prague interfered with her willingness to agree with Myles' claim that she might be in Prague.

But there is independent evidence that *local* is *not* indexical in the sense defined here. Put another way, *local* is locative and anaphoric, but not perspectival, unlike the EMAs. Consider Partee's (1989) examples:

- (51) After the game, every sports fan stopped at a local bar.
- (52) After the game, every sports fan stopped at a bar here for a drink.

In (51) the locale may be that of the speaker, the game, or the (widely dispersed) sports fans, only the first of them corresponding with a discourse center under the rules for center introduction discussed above. This contrasts with the behavior of *here* in (52), which can only be anchored to ©*. Hence, *local* is not indexical in the sense defined here. See Barlew (2014, 2017) for extensive discussion of the differences between anaphoric locatives like *local* and indexicals like *here*. Cross-linguistically, the difference between the EMAs and the merely anaphoric elements is real, but the reason is that the EMAs, like other indexicals—pure, impure, and shifted—require a discourse center as antecedent anchor, while *local* and its ilk do not.

Hence, we do not expect the EMAs to pattern with *local* in this respect.

4.3.2 **Yalcin's (2007) puzzle**

Yalcin (2007) observes that pairs of examples like the following appear to pose problems for the usual semantics of epistemic modality, for example Stephenson's (2007) account assuming a *de se* perspective for the modal:

- (53) #It is raining and it might not be raining.
- (54) #It is raining and possibly it is not raining.

These are “epistemic contradictions”, said to be defective in a way intuitively parallel to Moore's paradox:

- (55) It is raining and I do not know that it is raining.
- (56) It is not raining and for all I know, it is raining.

Yalcin points out that the paradox is usually treated as involving a “pragmatic conflict”: “in making an assertion in a normal discourse context, one usually represents oneself as knowing what one says”, so each of these examples represents the speaker as “both knowing something and also as knowing that one does not know it”. So if *might* ϕ means ‘for all I know ϕ is true’, we get the same pragmatic conflict in (53) and (54).

Yalcin then introduces (57) and (58), where in the latter (53) is embedded under *suppose*.

- (57) Suppose it is raining but you don't believe it is.
- (58) #Suppose it is raining but it might not be.

The facts about (58) and its negative counterpart, in schematic form:

Suppose (ϕ & $\Diamond\neg\phi$)

Suppose ($\neg\phi$ & $\Diamond\phi$)

He points out that we cannot just “piggyback” an explanation of the embedded examples on that for the Moore’s paradox examples, because the Moore-type examples (55) and (56) are “perfectly acceptable in the embedded contexts just described”:

(59) Suppose it is raining and I do not know that it is raining

(60) Suppose it is not raining and for all I know, it is raining

Yalcin offers parallel examples involving *might* in the antecedents of conditionals, where, once again, there is a marked contrast with examples involving *know* instead of *might*:⁷

(61) If it is raining and I don’t know that it is raining, then...

(62) # If it is raining and it might not be raining, then...

Yalcin describes the situation thus (p.987ff):

Moore-paradoxical sentences serve to describe totally clear possibilities, possibilities we can readily imagine obtaining. The same apparently does not apply to epistemic contradictions. These sentences do not seem to describe coherent possibilities, as witness the fact that an invitation to suppose such a conjunction strikes us as unintelligible. . . . Like Moore-paradoxical sentences, epistemic contradictions are not assertable; but unlike Moore-paradoxical sentences, they are also not *supposable*, not *entertainable as true*. . . . Evidently there is no coherent way to entertain the thought that it is not raining and it might be raining.

That suggests that we should drop the supposition that the two conjuncts actually are compatible. If we take it instead that $\neg\phi$ is truth-conditionally *incompatible* with $\Diamond\phi$, then we will have a ready explanation for our inability to entertain their conjunction. If there simply is no possible situation with respect to which ($\neg\phi$ & $\Diamond\phi$) is true, then that explains why it is so hard to envisage such a situation. The conjunction is just semantically a contradiction. But although this line of explanation covers our intuitions about epistemic contradictions in embedded contexts, it comes at an unacceptably high price. If $\neg\phi$ and $\Diamond\phi$ are contradictory, then the truth of one entails the negation of the other. On ordinary classical assumptions, this means that $\Diamond\phi$ entails the negation of $\neg\phi$ — that is, it means $\Diamond\phi$ entails ϕ . But that result is totally absurd. It would imply that the epistemic possibility operator \Diamond is a factive operator, something it very clearly is not.

...

So it appears we face a dilemma.

- $\neg\phi$ and $\Diamond\phi$ should be modelled as having incompatible truth conditions, in order to explain why it is not coherent to entertain or embed their conjunction; but

⁷ Rich Thomason (p.c. to Bill Harper, brought to my attention by Nate Charlow) noticed around 1975 that “conditionally entertaining that P is not conditionally entertaining that KP” (Rich’s characterization, p.c.), offering as an example *If my wife were cheating on me, I wouldn’t know it*, subsequently much cited in the literature.

- $\neg\phi$ and $\Diamond\phi$ should be modelled as having compatible truth conditions, in order to block the entailment from $\Diamond\phi$ to ϕ .

Yalcin argues that such examples cannot be addressed within the standard relational semantics of the type developed by Kratzer, and instead offers an alternative, “domain semantics” account based on acceptance. But I will argue that on reasonable assumptions about the meanings of the expressions involved, these examples are not problematic for the indexical Kratzerian account proposed here, and in fact are exactly what it would lead us to predict. In other words, Yalcin is ignoring another possible resolution to the problem.

For contrasts like that between (57) and (58), the key to explaining the puzzle illustrated lies in explicating the meaning of *suppose*, which in its imperative form instructs the addressee to entertain a hypothetical doxastic state in which the complement is true. To do so for (58), the addressee would have to entertain a state in which the conjoined complement is true, i.e. one in which both the proposition that it’s raining and the proposition that it might not be raining are true. But what is the anchor for *might* in the complement? Recall that under the scope of an attitude predicate, we regularly introduce a new discourse center, the agent of the attitude at the time it is held, which is the default anchor for any epistemic modal in the complement. So under *suppose* the salient doxastic center is $\textcircled{C}^{\text{suppose}}$, the hypothetical doxastic agent whose belief state entails the truth of the complement at the time the supposition is realized. Note that in imperative (58), though the addressee is the understood agent of *suppose*, the doxastic perspective of $\textcircled{C}^{\text{suppose}}$ and that of the actual addressee $\textcircled{C}^{\text{@}}$ are in principle distinct—we’re usually asked to suppose something we don’t already believe; and because the complement of *suppose* tends to be counterfactual, the perspective entertained by $\textcircled{C}^{\text{suppose}}$ is possibly even incompatible with the addressee’s actual belief state. So though they have the same agent, the addressee, $\textcircled{C}^{\text{suppose}}$ and $\textcircled{C}^{\text{@}}$ are distinct doxastic centers. The embedded *might* in (58) is anchored to $\textcircled{C}^{\text{suppose}}$, whose entertained doxastic state then has to entail both the proposition that it is raining and the proposition that it might (from her own point of view) not be raining, i.e. that it’s consistent with all she knows that it isn’t raining. But this, even more clearly than in Moore’s original examples, is an inconsistent doxastic state: one both containing only worlds in which it’s raining *and* containing some worlds in which it’s not raining. Since the addressee cannot rationally entertain such a state, these directions are anomalous, and cannot be realized. This is very different from (57), which merely requires that in the revised doxastic state the addressee’s counterfactual counterpart is confused.

That’s the intuitive account I will offer. Now the details about the proposed analysis of *suppose*. (If you wish, you can skip to the end of this brief technical sub-section.)

A formal analysis of *suppose*

Here is a relatively informal sketch of the conventional content of English *suppose*:

CHARACTER of *suppose*:

INFORMAL FIRST PASS

In context D , *suppose* takes a proposition p and an agent a and yields a proposition that is true at a given world w and time t just in case at t in w , a doxastically entertains p .

We might say that to doxastically entertain a proposition p is to contemplate what would be the case were p true—the way the world would be in that case. But as that paraphrase suggests, this is essentially a counterfactual attitude toward the complement proposition p .⁸ Uttering *I suppose that p* implicates a failure to confidently accept the truth of p . Someone can felicitously issue the directive (57) when the Common Ground entails that it's not raining. Then the attitude involves entertaining this counterfactual proposition.

What is it to entertain a proposition, as opposed to believing it, doubting it, etc.? Common usage,⁹ semantic intuition, and the fact that *suppose p* is so often followed by *then...* argue that it involves putting oneself in the hypothetical doxastic state wherein one takes the world to be as characterized by p , perhaps so that one can then consider what follows. To do this one doesn't entertain the proposition p alone, in isolation. Instead, one is interested in what it implies about the world *against the background of one's other beliefs about the way the world is*. Of course, this may be a counterfactual attitude, since the entertained proposition may be inconsistent with some of one's actual beliefs. And one cannot reasonably simultaneously entertain inconsistent propositions, for *ex falso quodlibet*: from inconsistency nothing interesting follows, because everything follows. Thus, crucially, rationally entertaining a proposition (or a set of propositions, via their intersection) presupposes that it is a non-empty set of worlds. Accordingly, to entertain p against the backdrop of one's belief state, first p itself must be consistent, and then one must revise those beliefs to yield a state that's consistent with p but in other respects as much as possible like the way things actually are so far as one knows.¹⁰

Capturing this similarity requirement, as usual in counterfactual modality, requires a similarity metric over possible worlds. I adopt a very simple notion of similarity, adapted from Heim (1992), and intended, as is hers, mainly as a place-holder for whatever notion is most appropriate in the general case. One is invited to substitute one's favorite similarity metric. SIM takes a proposition p (a set of worlds) and a world w , and yields the most w -like of the p worlds:

$$\text{SIM}(p)(w) =_{\text{def}} \{w' \mid w' \in p \ \& \ \forall w'' \in p: w' \text{ resembles } w \text{ at least as much as } w'' \text{ does}\}$$

Using SIM, we derive a belief state (a set of worlds) which is based on the doxastic state of an agent, revised only as much as required to make p true, though p may be counterfactual from the point of view of that doxastic state. We do this using Heim's (1992) **REVDOX**:

$$\begin{aligned} \text{REVDOX} =_{\text{def}} \lambda p \lambda \langle \langle x, t \rangle, w \rangle \lambda w'. w' \in (p) \ \& \\ \exists w'' \in \text{DOX}(\langle \langle x, t \rangle, w \rangle): w' \in \text{Sim}(\cap \{X: \text{Dox}(\langle \langle x, t \rangle, w \rangle) \subseteq X \ \& \ X \cap p \neq \emptyset\}) (w'') \end{aligned}$$

RevDOX takes a proposition p and a centered world $\langle \langle x, t \rangle, w \rangle$ to yield the set of worlds w' in which p is true that are most like those worlds that characterize $\langle x, t \rangle$'s belief state in w .¹¹ As its

⁸ So-called counterfactual conditionals in English themselves are only possibly counterfactual, as recognized at least since Karttunen & Peters (1979). This will be the case with *suppose* as well.

⁹ Merriam-Webster (<http://www.merriam-webster.com/dictionary/entertain>) offers this definition for *entertain* (definition 3a): “to keep, hold, or maintain in the mind”.

¹⁰ As above, this implicitly presupposes that those beliefs themselves are consistent. I would maintain that our use of *believe* itself presupposes that convenient fiction, as evidenced by what we take to follow from the truth of a belief report. Thus, the problem of inconsistent beliefs is one for epistemology, not semantics.

¹¹ Whether there always is a unique such set is, again, a question that also goes beyond current considerations.

name is intended to suggest, this is itself the revision of a doxastic state (the perspective of the center $\langle x, t \rangle$ in w) to entail p , which may be itself counterfactual in that doxastic state. REVDOX tells us to construct the relevant proposition as follows:

- Take all the propositions that are both entailed by the center's doxastic state in that centered world and consistent with p . Call this the p -consistent set for $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$:

$$\{X: \text{Dox}(\langle \langle x, t \rangle, w \rangle) \subseteq X \ \& \ X \cap p \neq \emptyset\}.$$
- Consider the worlds where all of the propositions in the p -consistent set are true. Call this the p -revised belief set for $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$:

$$\cap \{X: \text{Dox}(\langle \langle x, t \rangle, w \rangle) \subseteq X \ \& \ X \cap p \neq \emptyset\}$$

Note that if p is consistent with all the worlds in $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$, this set is just a subset of $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$ itself. But if p is false in the the actual belief set $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$, then there will be a null-intersection between $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$ and the p -revised belief set.
- Among the p -revised belief set, consider only those that are most like one or another of the worlds w'' where the center's beliefs are all true, i.e. those in $\text{Dox}(\langle \langle x, t \rangle, w \rangle)$:

$$\text{Sim}(\cap \{X: \text{Dox}(\langle \langle x, t \rangle, w \rangle) \subseteq X \ \& \ X \cap p \neq \emptyset\}) (w'')$$
- Then the worlds in $\text{RevDox}(p)(\langle \odot \rangle, w)$ are those among those most-actual_belief_set-like worlds in the p -revised $\langle \odot, w \rangle$ -belief-set in which p is true.

Then the formal CHARACTER of *suppose* is:

CHARACTER of *suppose*: type $\langle \langle s, t \rangle, \langle e, t \rangle \rangle$ ¹²
 In context D :

$$\lambda p_{\langle s, t \rangle} \lambda x_e \lambda w \lambda t. \ ||\text{ENTERTAIN}[\langle \langle x, t \rangle, w \rangle, \text{REVDOX}(p)(\langle \langle x, t \rangle, w \rangle)]||^{w, t} = 1$$

This takes a proposition p (type $\langle s, t \rangle$) and an agent a (type e) to yield a proposition that is true in a given w , t just in case at t in w , a doxastically entertains the REVDOX-revision of p . I leave the notion of entertaining a doxastic state undefined, relying on the intuitive characterization above.

In the examples of interest here, *suppose* is used in the imperative grammatical mood. I assume without argument that imperatives have realization conditions, rather than truth conditions—circumstances in which they would count as being realized because the corresponding declarative with the addressee as subject would be true. Further, the realization conditions of an imperative require that the time of realization be at or after the time of issuance of the directive involving the imperative.¹³ So roughly, an imperative utterance denotes a special type of property, type $\langle s, \langle e, t \rangle \rangle$, in English always anchored to the addressee:

An **imperative** $S!$ with LF $[s \ ! \ VP]$ addressed to a at time t^* in context D is realized in world w just in case: $\exists t \leq t^*: a \in |VP|^D(w)(t)$

The realization of an imperative can only take place at some time t which is immediate or in the future with respect to the utterance time t^* , and (in English) this realization involves the

¹² In actual implementation, the agent is an individual concept, type $\langle s, e \rangle$. See Roberts (2015). But I simplify here for ease of exposition.

¹³ See Roberts (2015b, 2018) for details, but these assumptions are broadly in keeping with other accounts of imperatives, e.g. Kaufmann 2010.

imperative property holding of the addressee a at t in the world of evaluation w . We can say that the use of the imperative presupposes that there is an addressee, who serves to anchor the realization conditions.

Then imperative *suppose*, in keeping with its general CHARACTER given above, has the following realization conditions:

CHARACTER of imperative *suppose*:

Assume that $[_S ! \text{suppose } p]$ is addressed to a at time t^* in world w in context D :

Presupposed:

$$\langle a, t^* \rangle = \odot^@ \in \odot_D$$

‘there is an addressee’

Proffered:

(type $\langle \langle s, t \rangle, \langle s, \langle e, t \rangle \rangle \rangle$)

$$\lambda p_{\langle s, t \rangle} \lambda w \lambda x_e . x = a \ \& \ \exists t \leq t^* \ \& \ \text{ENTERTAIN}[\langle \langle x, t \rangle, w \rangle, \text{REVDox}(p)(\langle \langle x, t \rangle, w^* \rangle)]$$

Imperative *suppose*, like imperatives generally, presupposes the addressee at utterance time, which will serve as the agent of the realization conditions. It takes a complement proposition p to yield a property, type $\langle s, \langle e, t \rangle \rangle$, as the imperative’s proffered content. This is a special property, one that can only be realized by the anchoring agent $\odot^@$ ($x = a$), and anchored to $\odot^@$ ’s doxastic state at realization time ($\text{REVDox}(p)(\langle \langle x, t \rangle, w^* \rangle)$). The directive to suppose p will be realized at some future time t just in case at t the addressee entertains the revision of her actual belief state at t to make p true.

Note that the CHARACTER of *suppose* does not presuppose that the complement p is *not* part of the addressee’s actual belief state. But generally we assume that a directive is not felicitous if the addressee is already realizing the intended end-state. So issuance of a directive to *suppose* p implicates that so far as the speaker knows, the addressee doesn’t already believe that p and/or hasn’t already entertained it in the current circumstances.

There is another, pragmatic consequence of proffering *suppose*: As usual with attitude predicates, considering an attitude leads to the introduction of \odot^{suppose} to the set of discourse centers \odot_D , a center consisting of the agent of the attitude at the time and world in which it holds. But since the projected realization time t needn’t equal the speech time t^* , the addressee discourse center at speech time $\langle \langle x, t^* \rangle, w \rangle \neq \langle \langle x, t \rangle, w \rangle$, the discourse center for \odot^{suppose} . That is, the agent who entertains the revised belief state is not the actual addressee, $\odot^@$: they may be entertaining distinct, possibly even incompatible belief states.

In (58), \odot^{suppose} both entertains a belief state in which it’s raining and anchors *might*, hence taking it to be consistent with all she entertains that it might not be raining. In other words, the addressee is directed to entertain an inconsistent belief state, and this explains the infelicity of the example: it is unrealizable, a pragmatic anomaly, because the belief state to be entertained is anomalous.

But what about (57)? Why is it felicitous? Note that the anchor of *know* in the complement is conventionally given to be its subject, here the denotation of *you*. But who is that individual in this hypothetical context? It is yet another center, \odot^{REVDox} , the counterfactual knower *in the revised hypothetical state being entertained*, and hence the counterpart in that state of $\odot^@$. Thus,

\odot^{REVDOX} is the ordered pair of the counterfactual counterpart of the actual addressee (in the arbitrary counterfactual (REVDOX) world w' in which what \odot^{suppose} supposes is true) at supposition time t . But there is no anomaly in \odot^{suppose} entertaining a doxastic state in which p is true though her counterpart in that state doesn't know that p is true. Thus, (57) is felicitous.

Supporting this account of the felicity of (57), compare (57) and (58) with (63):¹⁴

(63) Suppose it's raining and for all you know it might not be.

With the explicit epistemic shifter *for all you know*, (63) contrasts sharply with (58), is more like (57). Here the *you* that controls \odot^{know} is the hypothetical counterpart of the addressee, and so is not the same center as the \odot^{suppose} that anchors *might* in (58).

Other forms of the puzzle:

Pretty much the same thing is going on in the conditional forms of Yalcin's puzzle, (61) revised here with *you* instead of *I* to draw out the parallels with (63):

(61') If it is raining and you don't know that it is raining, then...

(62) # If it is raining and it might not be raining, then...

Take the antecedent of a conditional to give instructions for construction of a state which the interlocutors are to entertain, so as to consider what follows. Then in the scope of the conditional, we have a new discourse center, \odot^{if} , like $\odot^{*+@}$ except that from its perspective the antecedent of the conditional is true. So in (61') and (62), both conjuncts of the antecedent must be true in the hypothetical information state of \odot^{if} , relative to which the consequent will be assessed in determining the truth of the whole conditional. It cannot be true in a single consistent information state both that it's raining *and* that it might not be raining, so these 'instructions' for building such a state are unrealizable, hence infelicitous in (62). But in (61'), as in (57), *you* denotes the addressee's counterpart in the entertained state and is the agent of the attitude of knowing, again introducing a distinct discourse center, \odot^{know} , for the addressee's counterpart in the entertained hypothetical scenario (or that of the arbitrary agent, in case we take *you* to be generic). There is nothing inconsistent in the contrast between the rain and the counterpart's ignorance in that scenario. We see the same shift to acceptability from (62) to (64) with the explicit shifter *for all you know*, in which the hypothetical counterpart of the addressee \odot^{know} is the agent of *know*, and hence anchors *might*.

(64) If it is raining but for all you know it might not be raining, then...

With different anchor discourse centers, the two clauses yield a state that one can consistently entertain.

¹⁴ I would prefer replacing *and* in (55) and (63) with *but*, as more consistent with the contrast; but I don't take this as essential, and in any case it wouldn't change their truth conditional content.

Now consider what happens when we reverse the order of the two conjuncts of the *suppose* complement in (58), or the *if*-clause in (62):¹⁵

- (65) # Suppose that it might not be raining and it's raining. [cf. (58), same judgment]
(66) # If it might not be raining and it's raining, then. . . [cf. (62), same judgment]

Here the results are just as unacceptable. One might be tempted initially just to say that it's the same problem: these are the same two conjuncts in each case as in (58) and (62), so the same pragmatic inconsistency arises. But it's a bit trickier than that. Consider your judgments for this pair of examples from Dorr & Hawthorne (2013), where one order of the antecedent conjuncts is felicitous, the other not:

- (67) #If Jack and Jill fail and they might pass, then their preparation is to blame.
(68) If Jack and Jill might pass and they fail, then their preparation is to blame.

and the comparable pair with *suppose*:¹⁶

- (69) #Suppose that Jack and Jill fail and they might pass.
(70) Suppose that Jack and Jill might pass and they fail.

Though (67) and (69) are like (62) and (58) in being unacceptable, presumably for the same kind of reason, changing the order of the conjuncts in (68)/(70) results in an acceptable conditional or supposition. Why should that be?

The explanation for the felicity of (68) and (70) lies in the role of aspect and aktionsarten in the dynamics of temporal update, as worked out in Hinrichs (1986), Partee (1984), and Dowty (1986) using Reichenbach's (1947) notion of **Reference Time**. Reference Time may be thought of as the contextually given interval in which we take the event denoted by a clause to occur. Partee shows how by systematically updating Reference Time throughout the course of a discourse we can account for a variety of effects of context on the determination of the intended event times for clauses in the discourse. In a sequence of clauses, Reference Time is updated after each clause is interpreted; crucially for our purposes, the way the Reference Time is updated is a function of whether the clause just completed displays telic or atelic aktionsarten. In an atelic clause, the corresponding eventuality is a state or on-going process which displays a sort of homogeneity (Dowty 1979): nothing happens, or nothing changes. But in a telic, we have an event in which there is a change of state: something occurs. Accordingly, telics typically move along the action, so that the Reference Time for interpretation of the following clause is moved forward to an interval immediately after the change of state. But atelics do not move the Reference Time forward; a sequence of atelics will typically be understood to describe one and the same interval. We see the difference in discourses like the following:

¹⁵ Nate Charlow and Phil Kremer (p.c.) helped me to understand the importance of the reversed order examples, for which I am grateful.

¹⁶ Again, I think that both (68) and (70) are a bit better with *but* instead of the second *and*, but they're also acceptable as-is, so that doesn't undercut the puzzle.

- (71) That afternoon it was snowing. The wind was blowing and the temperature was below freezing. John was snoring in his armchair in front of the fire.
- (72) Stefan was startled awake by a loud noise in the next room. He jumped up, afraid. He called the front desk and demanded an explanation. The clerk apologized.

(71) consists of a sequence of atelic clauses (progressives corresponding to on-going processes), and it is understood as describing a single period of time, when it was simultaneously snowing, windy, and below freezing, while inside John was snoring. Nothing changes, and the Reference Time (the interval corresponding to *that afternoon*) stays the same from clause to clause. But in (72) we have a sequence of telic clauses, and accordingly we understand them as describing events which occur one after the other. Partee's account uses this difference and a dynamically updated Reference Time as a bookkeeping device to track how the action moves forward.

Now consider (68). Here we're asked to entertain a state in which initially it's possible that Jack and Jill will pass. But *then* they fail. Though the possibility described by the first conjunct is atelic, the property denoted by *fail* is telic—it describes a state in which something occurs, a change of state, so that the *then* in my paraphrase just above is the result of the Reference Time being moved slightly forward by the regular mechanics of Reference Time update. We're presented with a state in which something occurs, a change of state denoted by the second conjunct in which (assuming both conjuncts pertain to one and the same test) the state of possibility denoted by the first conjunct is terminated. In other words, the mere possibility in the first conjunct is compatible with the occurrence of the event denoted by the second, resulting in a new state in which the possibility no longer obtains. But this kind of sequential interpretation is precluded in (67): we have to entertain a situation in which Jack and Jill fail, and *then* they might pass, the Reference Time having moved forward after telic *fail* to yield the 'then' in the resulting sense. Again assuming that we take the failure/passage to pertain to one and the same test given on a single occasion, such a situation isn't possible. Once one has failed a given test, it's no longer possible that one might pass that test. So the scenario described is impossible, and the resulting utterance is anomalous. The same account explains the felicity patterns in the parallel *suppose* examples in (69) and (70).

The reason order of conjuncts doesn't matter in the original 'rain' pairs—(58) and (65), (62) and (66)—is that in these examples both conjuncts are atelic in aktionsarten—rain is just a process, and a possibility is just a state, and in neither does anything occur to result in a change of state. Atelic aktionsart does not move the Reference Time forward, so that in the absence of any adverbial indication to the contrary, one takes a sequence of two atelic clauses to have the same Reference Time, and hence to overlap in event time, leading to the pragmatic contradiction in these cases.

Dorr & Hawthorne (2013) consider a version of Yalcin's puzzle with disjunction. We see the same pattern in the disjunctions in (73) – (74) as discussed for the *suppose* and *if*-clause versions above, and the same kind of contrast between Dorr & Hawthorne's (75), with the inverted order of the two conjuncts in the first disjunct of (74), and my (76) with telic conjuncts:

- (73) Either it's raining and I don't know it, or it's a good day for a picnic
- (74) # Either it's raining and it might not be, or it's a good day for a picnic

- (75) # Either it might not be raining and is raining, or it's a good day for a picnic
 (76) Either Jack and Jill might pass but they fail, or they'll just squeak by

In (75) we have two atelic eventualities in the first disjunct, with no implication that they pertain to different times/situations. Since one situation cannot be consistently characterized by both these atelic properties, the first disjunction is anomalous, and hence the disjunction is infelicitous. But (76) is better because the possibility of passing can be understood to be the background situation in which the telic eventuality of failing occurs, a change of state which then terminates the possibility.

Hence, order in the conjuncts only makes a difference when the example also involves a change in Reference Time between the conjuncts. What the examples involving telic aktionsarten tell us, then, is that the original puzzle depended partly for its plausibility on the use of atelic aktionsarten in the two conjuncts in the relevant clauses. The moral I would draw from this is that we must be very careful in using natural language examples not to treat them analytically as if they were simple logical formulae, say $\neg\phi$ and $\Diamond\phi$. Natural language clauses are much more complex than logical formulae, with CHARACTERS involving presupposed as well as proffered content, and hence potentially multiple contextual dependencies to take into account. This is what Yalcin ignores in posing his dilemma: EMAS carry presuppositions which differentiate them from attitude predicates and offer a solution to his problem.

Mandelkern (2019a) calls examples like (53) “Wittgenstein sentences”, attributing them to Wittgenstein (2001 [1953], II.x.109). He considers this variation with disjunction (77):

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

This case is interesting because, of course, the speaker is reasonable to assume that either he won't win or he won't lose (on the assumption that winning and losing are the only two possibilities). And the corresponding Moore example is acceptable:

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

An explanation for the anomaly of (77) is simple in the present account of *might*. On the reading of (77) that Mandelkern has in mind, neither disjunction can be true in much the same way as a Moore sentence cannot be true as asserted. Each disjunction dynamically takes its first conjunct as part of the modal base for *might* in the second conjunct. So for the second conjunct of the first disjunction to be true, the prejacent *I will win* must be consistent with the assumed belief state that includes the first conjunct ‘I won't win’. But of course, this is not consistent. The second disjunct similarly implies an inconsistent belief state. Hence the anomaly. Again, this is a pragmatic anomaly, involving in each disjunction the inclusion of the first conjunct to the modal base of *might* in the second.

But note that there is a perfectly reasonable interpretation of this string where the second conjunct of each disjunct is read as an optimistic/pessimistic aside, so not taking the first conjunct as part of the modal base of *might*. (Linguists: try putting a L*H on *might* along with a LH% to force this reading.) I find that perfectly felicitous.

On the analysis just given, Yalcin’s (2007) “dilemma” fails to arise. As usual, we take $\neg\phi$ and $\Diamond\phi$ to have compatible truth conditions: both may be true in a single possible world. However, as in Moore’s original examples, it is not possible for one and the same doxastic agent to entertain a state in which it is simultaneously true both that $\neg\phi$ and that it might be the case that ϕ —the result would be an inconsistent doxastic state, or in the case of *if*, an inconsistent premise set, because entertaining $\neg\phi$ results in a quasi-belief state in which ϕ is not true in any of the possible worlds in the state. The difference between the cases with *might* and those with *know* is that the latter introduces a counterpart of the agent entertaining the truth of the conjuncts, and that counterpart’s doxastic state needn’t be compatible with the agent’s. Hence, the apparent puzzle only arises if we ignore the complexities of doxastic anchoring under iterated attitudes. By giving them their due, we have accounted for the puzzle without resorting to the kind of non-standard modal semantics that Yalcin adopts. The account is simple and independently motivated by what we have already said about the EMAs themselves, and by the semantics of *suppose*, with a bit of Reference Time resolution (in a well-accepted account) thrown in. And *pace* Yalcin (2007,2010) and Stalnaker (2014:139ff), an account of this phenomenon requires no expressivist treatment of modality or special posterior context.

In this discussion we have focused on cases involving embedding under *suppose*, the predicate used in Yalcin’s original examples. But the problem he pointed to arises with other attitude predicates. Consider the examples in (79) – (84):

- (79) #George believes that it’s raining and it might not be.
- (80) #/✓ George believes that it’s raining(,) though it might not be.
- (81) #Stefan doubts that it’s raining and it might not be.
- (82) Marcia expects that it’ll be raining when we get there.
- (83) Marcia expects that it might not be raining when we get there.
- (84) #Marcia expects that it’ll be raining when we get there and it might not be.

In (79), as we would expect from the discussion above, we see that since George cannot entertain both the propositions denoted by the complement’s conjuncts, he cannot consistently believe them. The string in (80) yields the same results if we take the *though* clause to modify the complement *it’s raining*, but is fine if we understand it to modify the root sentence, hence reflect the belief state of the speaker rather than George. In (81), *doubts* yields the same infelicity: it would be inconsistent to both doubt that it’s raining and doubt that it might not be. (82) and (83) show that both conjuncts of (84) are felicitous as complements of *expects*, while their conjunction in (84) is not. Again, this is what we predict if we take the default anchor of *might* in the complement of *expect* to be the agent of the attitude. To confirm all these intuitions, of course, we would need to work out a detailed lexical semantics for each of these predicates, something I will not do here. But the general approach seems promising.

4.3.3 (Dis)agreement

Stalnaker (2014:143,fn.12) points out a problem with Yalcin's (2007) theory. Yalcin accounts for shifting epistemic anchoring via an indexed information state. He claims that this state is obligatorily shifted under attitudes to the information state of the agent of the attitude. Stalnaker takes issue with this obligatory shift, giving the following example:

(85) Alice: Jones believes that it might rain, but I disagree—I think we can count on fine weather all day.

If the information state relative to which *might* in the complement of *believes* is interpreted is that of Jones, then by asserting the first conjunct Alice is committed to the truth of *it might rain* in Jones' information state. But (85) strikes us as consistent, so we must conclude that Alice's disagreement is not over the proposition expressed by *it might rain* so-anchored, but over what possibilities are live options. This, of course, is a particularly sophisticated exemplar of the problem of faultless agreement, one of the central arguments for Relativist approaches to the semantics of EMAs (Laserson 2005; Egan et al. 2005; MacFarlane 2005, 2011; Egan 2007; Stephenson 2007, etc.).

This example points up a respect in which the debate over faultless disagreement as it has been framed to this point begs a deeper question: What does it mean to agree or disagree, either in response to what someone has said or with what someone believes?

I argue elsewhere (Roberts 2017) that (dis)agreement isn't about the truth of the proposition expressed by the speaker in a speech act or in an attitude complement, but about a comparison of features of the belief states of the relevant parties. What the complement of *believes* in (85) tells us is that it is compatible with Jones' belief state that it will rain. And this is the sense in which Alice disagrees: she fails to have a belief state that admits of the (future) possibility of rain. On this view, agreement, wherein two agents share a belief, is not about *what an agent says* so much as about the possibility that it points to. Thus in (85) Alice disagrees with Jones in that while his belief state admits of a certain possibility, hers does not. Agreement and disagreement gauge belief states as to whether they are congruent in the relevant way. Moreover, this cannot be reduced in the general case to shared assessments of the simple truth of a proposition—'what is said'. In fact, agreeing or disagreeing with some possibility under consideration has a much richer potential than just assessing truth—we can compare the probability, likelihood, plausibility, etc. of a given proposition across the two belief states: *Jones believes that there's a very good chance it will rain, but I disagree: I don't think it's that likely*. But space precludes developing this argument here.

What is relevant here is that there is reason to think that agreement and disagreement are about whether two agents share the same relevant attitude. They are *not* about the truth conditions of an utterance used to express that opinion. Therefore, agreement and disagreement do not constitute convincing arguments for a particular semantics for EMA statements.

5. Conclusions and prospects

The present proposal amounts to a partial answer to question (i) posed in the introduction:

- i. What constraints, if any, are there on the character of the premise set for an EMA?
- ii. What is the nature of the relationship between premises and conclusion that is required for truth of the EMA statement?

This leaves question (ii) relatively open. Nonetheless, teasing out the parameters of variation appealed to by the modal base presuppositions of EMAs helps to make sense of the complexity of the data, and permits one to abstract from it in determining the appropriate proffered content—the semantics of the EMA, if you will. If you don't like the simple proffered content in the CHARACTERS I proposed, experiment with something different. But I think you will find that taking into consideration the presuppositions proposed herein will simplify the task of accounting for the full range of attested interpretations, and that in fact this context-sensitivity must be controlled for in order to get consistently correct predictions for meaning-in-context.

We started from the usual Kratzerian assumption that an EMA presupposes a contextually-given accessibility relation that yields the set of worlds over which the modal operator ranges; this relation is given by a modal base (function from worlds to sets of propositions), which is presupposed in using the EMA. This paper argues for two principal theses about constraints on the modal base for an EMA like English *must* and *might*:

1. INDEXICALITY: The interpretation of an EMA is presuppositionally anchored to a discourse center, an agent-time pair $\langle a, t \rangle$ such that a 's beliefs at t are reflected in the EMA's modal base. As evidence for this thesis we observed that though the speaker is the default anchor of an EMA, this is readily shiftable under attitudes, in questions, and in extended supposition contexts, modal subordination and FID. These contexts are just those in which we expect an alternative doxastic discourse center to be available and relevant, as independently attested for deictic motion verbs and for indexical pronouns in shifting-indexical languages.
2. EVIDENTIALITY: Following von Stechow & Gillies (2010), I argued that epistemic modals are evidential. But the current proposal differs from theirs in significant ways, largely because evidentiality here is taken to be doxastic (and hence “weak”), rather than epistemic (“strong”). The resulting weakness is consistent with recent experimental evidence about the way that native speakers understand EMAs. Further, I do not need to require that the evidence be “indirect”. Rather, we can capture the intuitive indirectness straightforwardly as a non-detachable Gricean implicature.

The first thesis makes EMAs indexical in the sense of Roberts (2015)—presuppositionally anchored to the salient doxastic perspective of an agent. Portner (2009, §2.3.5) and Moltmann (2012) both tentatively suggest that EMAs may be shiftable indexicals. But these informal proposals lack detail, and fail to capture important differences between EMAs and shiftable indexicals. Both these types of expressions are anchored by a doxastic center, and both are perspectival—pertaining to centered worlds. But Roberts (2015) argues that in any given language which has them, specific shiftable indexicals carry selectional restrictions limiting which discourse centers can anchor them to a lexically specified subset of the set of possible

discourse centers. E.g. Amharic (Leslau 1995, Schlenker 2003) shiftable 1st person –*ññ* can only be anchored to ©*, ©@ or ©^{say}, the last center corresponding to the agent of a verb of saying. But in Slave (Rice 1986, Anand & Nevins 2004), the 1st person shiftable indexical can be anchored to the agents of ‘tell’, ‘say’, ‘think’, or ‘want’. The wide variety of constraints in other languages with shiftable indexicals give evidence of yet other inventories of possible anchors (see Roberts 2015 for summary; Deal 2017 for a taxonomic comparison). In contrast, EMAs shift anchoring more promiscuously. They are more like the deictic motion verb *come* (Barlew 2015), in that they may be anchored to any contextually salient and RELEVANT perspective, as captured in the present account by the set of available discourse centers.

There is a third thesis which I argue for in Roberts (2017):

- a) AGREEING, ASSESSING, AND THE QUD: It is widely attested that some EMA utterances of the form *modal p* can provoke a felicitous response which assesses or agrees with not the entire modal statement, but the prejacent *p* alone; while other responses assess or agree with the truth of the entire EMA. The discourse status of the modality *per se* in such cases, as attested in felicitous response patterns and agreement, is a function of the question under discussion, the QUD. Epistemic modality is quite often backgrounded. However, unlike the presupposed anchoring to a discourse center, the tendency to be backgrounded is not *conventional*, but is instead a function of what’s typically at-issue—the kinds of questions we tend to pose. Another way of putting this is that **we respond to an assertion *qua* answer**.

The three theses above are related. For example, which agent the modal is understood to be anchored to is also partly a function of the QUD, as argued by Moss (2015). And Kneer’s (2015, in preparation) work suggests that anchoring to an evidential perspective and weakening to a doxastic-like semantics, though *prima facie* independent, may be related. He offers experimental evidence that “belief attributions do manifest a perspective effect, though knowledge attributions do not.” Here’s one way we might think of this: Knowledge is about facts, and it is assumed that knowledge correctly reflects the actual world, a correlation that is independent of the agent. But belief is an opinion about a possibility, arrived at from a certain perspective which may skew the image of the world. And so belief is inherently and essentially perspectival, in the sense defined here. Or at least, this is how we humans talk about knowledge and belief, as reflected in the semantics and pragmatics of the corresponding predicates.

One of the virtues of the proposed account, from a linguistic point of view, is that all of the tools used in the analysis are independently motivated, so that nothing is *ad hoc*. On the basis of these simple assumptions, we built an account with the following virtues:

- (a) It correctly predicts the contextually limited range of candidates for the anchoring agent of such a modal, as attested in the literature.
- (b) It thereby constrains what body of evidence is understood to be relevant (that of the anchor).
- (c) It counters purported arguments for modal relativism due to Egan, Hawthorne & Weatherson (2005).
- (d) It sheds light on Yalcin’s (2007) version of Moore’s paradox for embedded epistemic modals.

Accounts which address the kinds of puzzles we've considered in §4 above by taking EMAs to be expressive, or to be interpreted relative to an Assessor or speech act operators, all fail to be able to account for the full range of puzzles considered there. Especially, most of them have difficulties with the fact that EMAs work quite as well in embedded contexts, without many of the features that motivate special semantics in root clauses.

It may be useful here to compare the present proposal with the approach proposed in Mandelkern (2019c). Mandelkern's "guiding observation" (p.3) is that "a claim of $\lceil \text{Might } p \rceil$ is a proposal to make p compatible with the common ground, and to make this fact itself common ground." This is a nice, clear intuition about what is typically asserted in uttering an indicative whose main clause contains an EMA in a context where the QUD doesn't pertain to the epistemic state of some agent other than the speaker. That is, in fact, what the present proposal predicts about such utterances, when understood against the backdrop of Roberts (2018) on speech acts and their effects on context. But as is the case with many other proposals in the literature, Mandelkern's work focuses on too narrow a range of examples involving EMAs. And he takes this intuition to indicate what the *proffered content* of the utterance should be. But as Roberts (2018) argues, though the conventional content of an indicative clause lends itself most straightforwardly to service as an assertion (in pretty much the sense of Stalnaker 1978), that does not mean that this function is part of that conventional content, and in fact, if we are to construct a fully adequate account of the wide range of interpretations and uses even of just the indicative clauses containing EMAs, then the conventional content Mandelkern proposes would prove too inflexible. We have seen this above both in §4.2 (Variable anchoring) and especially in §4.3 in the analysis of the contribution of EMAs embedded under attitudes. In most of the examples in those sections, the semantics of the EMA is, as Willer (2015) calls it, *descriptive*, conveying something about the anchoring agent's attitude toward the prejacent. The problem is even more striking when we consider EMA uses in interrogatives like (17) or (24), or in imperatives like (57) – (60) above. So Mandelkern's observation is too limited to serve as adequate grounds for the proffered content of clauses containing an EMA.

The present essay presents the latest phase of an investigation of the interpretation of utterances containing modal elements that I began in Roberts (1989). Throughout, I have explored whether it might be possible to maintain that the relevant sentences have a simple, classical truth conditional content if we take into account the role of context in contributing to the broad variation in attested interpretations, including some that pose *prima facie* problems for the classical account. As before, the key to this approach is a careful consideration of the conventional but non-proffered content of the expressions involved, here their presupposed content. I have not directly argued for the standard, quantificational account. Instead, I offer evidence that most of the purported problems with such an account disappear if we understand the presuppositions triggered by the use of an EMA. Insofar as the classical account is simpler than others proposed in the literature, and easier to integrate elegantly into the overall compositional semantics for natural languages in the truth conditional tradition, this would appear to be a point in its favor.

References:

- Anand, Pranav & Andrew Nevins (2004) Shifty operators in changing contexts. In K. Watanabe & R.B. Young (eds.) *Proceedings of SALT 14*. Ithaca, NY: CLC Publications.
- AnderBois, Scott, Adrian Brasoveanu & Robert Henderson (2015) At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32 (1):93–138. doi: 10.1093/jos/fft014.
- Barlew, Jefferson (2014) Coming toward a doxastic agent: A doxastic analysis of the motion verb *come*. *Proceedings of Sinn und Bedeutung 19*. Georg August University at Göttingen. Göttingen, Germany.
- Barlew, Jefferson (2017) *The Semantics and Pragmatics of Perspectival Expressions in English and Bulu: The Case of Deictic Motion Verbs*. Ph.D. thesis, The Ohio State University.
- Brennan, Virginia M. (1993) *Root and Epistemic Modal Auxiliary Verbs*. Ph.D. dissertation, University of Massachusetts at Amherst.
- Deal, Amy Rose (2017) Shifty asymmetries: universals and variation in shifty indexicality. Ms., UC Berkeley.
- Del Pinel, Guillermo & Brandon Waldon (2018) Experimenting with epistemic tensions: *Must* might be weak, *might* must be strong. Ms., University of Michigan and Stanford University.
- DeRose, Keith (1991) Epistemic possibilities. *The Philosophical Review* 100(4):581-605.
- Doron, Edit (1991) Point of view as a factor of content. In S. Moore & A.Z. Wyner (eds.) *Proceedings of Semantics and Linguistic Theory (SALT) 1*. Cornell University Working Papers in Linguistics.
- Dorr, Cian & John Hawthorne (2013) Embedding epistemic modals. *Mind* 122.488:868-913.
- Dowell, Janice (2011) A Flexibly Contextualist Account of Epistemic Modals. *Philosopher's Imprint* 2011, pp.1-25. https://www.academia.edu/1125714/A_Flexible_Contextualist_Account_of_Epistemic_Modals
- Dowty, David (1979) *Word Meaning and Montague Grammar*. Reidel, Dordrecht.
- Dowty, David R. (1986) The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and Philosophy* 9:37-61.
- Eckardt, Regine (2014) *The Semantics of Free Indirect Discourse*. Brill, Leiden, The Netherlands.
- Egan, Andy (2007) Epistemic Modals, Relativism, and Assertion. *Philosophical Studies* 133:1–22.
- Egan, Andy, John Hawthorne & Brian Weatherson (2005) Epistemic modals in context. In G. Preyer & G. Peter (eds.) *Contextualism in Philosophy: Knowledge, Meaning and Truth*. Oxford University Press, 131-170.
- Fillmore, C. J. (1975) *Santa Cruz lectures on Deixis*. Indiana University Linguistics Club, Bloomington, Indiana.
- von Fintel, Kai & Anthony S. Gillies (2007a) *Might* made right. In Andy Egan & Brian Weatherson (eds) *Epistemic modality*. Oxford University Press.
- von Fintel, Kai and Anthony S. Gillies (2007b) An opinionated guide to epistemic modality. In T.S. Gendler & J. Hawthorne (eds.) *Oxford studies in epistemology: volume 2*. Oxford: Oxford University Press, 32–62.
- von Fintel, Kai & Antony S. Gillies (2008) CIA leaks. *The Philosophical Review* 117 (1): 77–98.
- von Fintel, Kai & Anthony S. Gillies (2010) *Must*. . .stay. . .strong! *Natural Language Semantics* 18:351-383.

- von Fintel, Kai & Sabine Iatridou (2003) Epistemic containment. *Linguistic Inquiry* 34(2):173-98.
- von Fintel, Kai & Sabine Iatridou (2008) How to say *ought* in foreign: The composition of weak necessity modals. In J. Guéron & J. Lecarme (eds.) *Time and Modality*. Springer, 115-141.
- Ginzburg, Jonathan (1994) An Update Semantics for Dialogue. *Proceedings of the Tilburg International Workshop on Computational Semantics*.
- Ginzburg, Jonathan (2012) *The Interactive Stance: Meaning for Conversation*. Clarendon Press, Oxford.
- Glass, Leila (2013) Deriving indirectness and questioning entailment for epistemic *must*. Handout of a talk at the LSA, Boston, January 2013.
- Goodhue, Daniel (2018) *Must* ϕ is felicitous only if ϕ is not known. *Semantics and Pragmatics* 10, Article 14. <https://doi.org/10.3765/sp.10.14>
- Grice, H.P. (1957) Meaning. *The Philosophical Review* 66 (1957).
- Hacking, Ian (1967) Possibility. *Philosophical Review* 76(2):143–68.
- Hacquard, Valentine (2006) *Aspects of Modality*. Ph.D. dissertation. MIT.
- Hacquard, Valentine (2013) On the grammatical category of modality. In M. Aloni, M. Franke & F. Roelofsen (eds.) *Proceedings of the 19th Amsterdam Colloquium*.
- Heim, Irene (1982) *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Heim, Irene (1992) Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183-221.
- Hinrichs, Erhard (1986) Temporal Anaphora in Discourses of English, *Linguistics and Philosophy* 9:63--82.
- Hockett, Charles F. (1990) Bring, take, come, and go. *Journal of English Linguistics* 23:239-244.
- Kamp, Hans (1981) A Theory of Truth and Semantic Representation. In Jeroen Groenendijk, Theo M. V. Janssen and Martin Stokhof (eds.) *Formal Methods in the Study of Language, Vol. I*, Mathematische Centrum, Amsterdam. Reprinted in Groenendijk, Janssen and Stokhof (eds.), (1984) *Truth, Interpretation and Information*, Foris, Dordrecht.
- Kamp, Hans & Uwe Reyle (1993) *From Discourse to Logic*. Kluwer, Dordrecht.
- Kaplan, David (1977) Demonstratives, draft #2. Ms., UCLA Philosophy Department. Revised and published as "Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals," in Joseph Almog, John Perry & Howard Wettstein (eds.) *Themes from Kaplan*, Oxford University Press, 1989, pp.481-563.
- Karttunen, Lauri (1976) Discourse Referents. In J. McCawley (ed.) *Syntax and Semantics* 7, Academic Press, New York.
- Karttunen, Lauri & Stanley Peters (1979) Conventional implicatures. In C.-Y. Oh & D. Dineen (eds.) *Syntax and Semantic, Volume 11: Presupposition*. Academic Press, New York, 1-56.
- Klein, Wolfgang (1994) *Time in Language*. Routledge, NY.
- Kneer, Marcus (2015) *Perspective in Language*. Ph.D. dissertation, EHESS/ENS, Paris.
- Kneer, Marcus (in preparation) Epistemic modals are not assessment-sensitive. Ms., University of Zurich.
- Kratzer, Angelika (1981) The notional category of modality. In H. J. Eikmeyer and H. Rieser (eds) *Words, Worlds and Contexts*. de Gruyter, Berlin, pp.38-74.

- Kratzer, Angelika (1991) Modality. In A. von Stechow and D. Wunderlich (eds) *Semantics: An international handbook of contemporary research*. de Gruyter, Berlin, pp.639-650.
- Kratzer, Angelika (1998) Scope or pseudoscope? Are there wide-scope indefinites? In S. Rothstein (ed.) *Events and Grammar*. Kluwer, Dordrecht, 163-196.
- Lassiter, Dan (2016) *Must*, knowledge and indirectness. *Natural Language Semantics* 24:117–163.
- Leslau, Wolf (1995) *Reference Grammar of Amharic*. Harrassowitz, Wiesbaden.
- Lewis, David (1979) Attitudes *de dicto* and *de se*. *The Philosophical Review* 88.4:513-543.
- Lycan, William (2001) *Real conditionals*. Oxford: Oxford University Press..
- MacFarlane, John (2005) The assessment sensitivity of knowledge attributions. In T. Szabó Gendler & J. Hawthorne (eds.) *Oxford Studies in Epistemology 1*. Oxford University Press, pp.197-233.
- MacFarlane, John (2005) The assessment sensitivity of knowledge attributions. *Oxford Studies in Epistemology* 1:197-233. Reprinted in Ernest Sosa, Jaegwon Kim, Jeremy Fantl & Matthew McGrath (eds.) *Epistemology: An anthology*. Oxford: Blackwell.
- MacFarlane, John (2011) Epistemic modals are assessment-sensitive. In B. Weatherson & A. Egan (eds.) *Epistemic Modality*. Oxford University Press, 144-178.
- Mandelkern, Matthew (2019a) Bounded modality. *Philosophical Review* 128.1:1-61.
- Mandelkern, Matthew (2019b) What ‘must’ adds. *Linguistics and Philosophy* 42(3):225-266.
- Mandelkern, Matthew (2019c) How to do things with modals. *Mind & Language* early view, April 29, 2019.
- Martin, Scott (2013) *The Dynamics of Sense and Implicature*. Ph.D. dissertation, The Ohio State University.
- Martin, Scott (2016) Supplemental update. *Semantics and Pragmatics* 9, Article 5, 1-61.
- Moltmann, Friederike (2012) Two kinds of first-person-oriented content. *Synthese* 184:157-177.
- Moore, G. E. (1993) Moore's Paradox. In Baldwin, Thomas. *G. E. Moore: Selected Writings*. London: Routledge. pp. 207–212. ISBN 0-415-09853-X.
- Moss, Sarah (2015) On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8.5:1-81.
- Ninan, D. (2010) De se attitudes: Ascription and communication. *Philosophy Compass* 5(7):551-567.
- Oshima, D. Y. (2006a). Motion deixis, indexicality, and presupposition. In Gibson, M. & Howell, J. (eds.) *Proceedings of Semantics and Linguistic Theory (SALT) 16*. Cornell University Press, Ithaca, NY, 172-189.
- Oshima, D. Y. (2006b) *Perspectives in Reported Discourse*. PhD thesis, Stanford University.
- Palmer, F. R. (1990) *Modality and the English Modals*, 2nd edition. Longman.
- Partee, Barbara H. (1973) Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70:601-609.
- Portner, Paul (2009) *Modality*. Oxford Surveys in Semantics and Pragmatics. Oxford University Press.
- Reichenbach, Hans (1947) *Elements of Symbolic Logic*. University of California Press, Berkeley.
- Rice, Keren (1986) Some remarks on direct and indirect speech in Slave (Northern Athapaskan). In Florian Coulmas (ed.) *Direct and Indirect Speech*. Berlin: Mouton de Gruyter. pp. 161–178..

- Roberts, Craige (1989) Modal Subordination and Pronominal Anaphora in Discourse. *Linguistics and Philosophy* 12.6:683-721. Reprinted in Javier Gutierrez-Rexach (ed.) *Semantics: Critical concepts in linguistics*, Routledge, 2003.
- Roberts, Craige (1996/2012) Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In Jae Hak Yoon and Andreas Kathol (eds.) *Ohio State University Working Papers in Linguistics* Volume 49. Reprinted in the 1998 version with a new Afterword in *Semantics and Pragmatics*, Volume 5, 2012.
- Roberts, Craige (2015) *Indexicality: de se semantics and pragmatics*. Ms., OSU.
- Roberts, Craige (2015b) Plans and imperatives: A semantics and pragmatics for imperative mood. *Proceedings of the Amsterdam Colloquium 2015*. ILLC, University of Amsterdam.
- Roberts, Craige (2017) Agreeing and Assessing. Ms., NYU and OSU.
- Roberts, Craige (2018) Speech acts in discourse context. In Daniel Fogal, Daniel Harris & Matt Moss (eds.) *New Work on Speech Acts*. Oxford University Press.
- Roberts, Craige (to appear) Modal Subordination: *It would eat you first!* Submitted to Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.) *Companion to Semantics*. Wiley.
- Schlenker, Philippe (2003) A plea for monsters. *Linguistics and Philosophy* 26:29-120.
- Sharvit, Yael (2008) The puzzle of free indirect discourse. *Linguistics and Philosophy* 31:353-395.
- Stalnaker, Robert (1974) Pragmatic Presuppositions. In M. Munitz and Peter Unger (eds.), *Semantics and Philosophy*. NY: New York University Press, 197-214.
- Stalnaker, Robert (1978) Assertion. In *Syntax and Semantics* 9. Academic Press, NY.
- Stalnaker, Robert C. (2008) *Our Knowledge of the Internal World*. Oxford University Press. Chapter 3: "Locating Ourselves in the World".
- Stalnaker, Robert (2014) *Context*. Oxford University Press.
- Stephenson, Tamina (2007) A Parallel Account of Epistemic Modals and Predicates of Personal Taste'. In Puig-Waldmüller, E. (ed.) *Proceedings of Sinn und Bedeutung 11*. Barcelona: Universitat Pompeu Fabra, pp. 583–97.
- Veltman, Frank (1985) *Logics for conditionals*. PhD thesis, University of Amsterdam.
- Willer, Malte (2015) An update on epistemic modals. *Journal of Philosophical Logic* 44(6):835-849.
- Windschitl, Paul D. & Gary L. Wells (1998) The Alternative-Outcomes Effect. *Journal of Personality and Social Psychology*. 75.6:1411-1423.
- Wittgenstein, Ludwig (2001) [1953]. *Philosophical Investigations*. 3rd ed. Malden, MA: Blackwell.
- Yalcin, Seth (2007) Epistemic modals. *Mind* 116:983-1026.