

# Tokens of Meaning: Papers in Honor of Lauri Karttunen

Cleo Condoravdi and Tracy Holloway King (eds.)

July 3, 2018

CENTER FOR THE STUDY  
OF LANGUAGE  
AND INFORMATION

---

# Contents

1	Talking about (quasi-)higher-order uncertainty	1
	DANIEL LASSITER	



---

# Talking about (quasi-)higher-order uncertainty

DANIEL LASSITER

## 1.1 Introduction

Theories of the lexical semantics of modal language live in the shadow of modal logic. Being a junior partner in this relationship has many advantages for us linguists: we get to make use of technical results that come from modal logic, and we benefit from the linguistic insights of smart philosophers, many of whom might not bother with natural language if they didn't think that modal expressions hold some philosophical interest. Still, the technical apparatus of modal logic was developed by people whose main academic interest was definitely not the lexical semantics of one corner of the vocabulary of a single Germanic dialect. For those of us who do (for whatever reason) care deeply about the vocabulary of English for its own sake, we should of course be cautious about borrowing *in toto* theories that were designed to elucidate philosophical issues—issues whose import does not depend on the quirks of any natural language. For each such case, it is important to examine the borrowed theory carefully in order to ensure that its empirical predictions make sense of the grammatical and inferential properties of the item in question.

Lauri Karttunen was perhaps the first linguist to make this point clearly and forcefully, in his “Possible and Must” (Karttunen, 1972). He did so by exploring an extension of Hintikka (1962)'s influential logic of knowledge and belief to English expressions of epistemic modality—*possible*, *may*, *perhaps*, *must*, and *have to* among others—and showing

that the empirical predictions of the extended theory failed to match up with the intuitive interpretations of the English expressions in various ways. Lauri was careful to emphasize that these findings only impugned the application of standard systems of modal logic to the lexical semantics of epistemic modals in English: “I will argue that the principles on which our use of *possible* and *must* in ordinary language appears to be based are different from any of the well known standard systems that have been studied by logicians” (p.3). The same systems could still be relevant to the lexical semantics of other items, and to the elucidation of philosophical issues that are independent of the semantics of English: “[T]his is an interesting fact about ordinary language, but it should not be misconstrued as evidence against other modal systems” (ibid.).

Lauri’s insights in “Possible and Must” have generated a significant literature, especially around his novel observation that the intuitive interpretation of *must* is weaker than that of the necessity operator in standard modal logics of knowledge. A few of the papers in this still-growing collection are Groenendijk and Stokhof 1975, Kratzer 1991, von Fintel and Gillies 2010, Lassiter 2016, Mandelkern 2016, Goodhue 2018. Thanks to “Possible and Must”, there is now a consensus among linguists that *must* differs in some important ways from the  $\Box$  operator of modal logic S5—though there is still no consensus about exactly what the difference is. S5 is the most obvious way to interpret linguists’ usual gloss of *must*, *might*, *perhaps*, *possible* etc. as  $\forall$ - or  $\exists$ -quantifiers over a “set of epistemically possible worlds” which contains the actual world. Under certain common (though not uncontroversial) assumptions, these quantifiers correspond to the  $\Box$  or  $\Diamond$  of a normal modal logic where accessibility is an equivalence relation (reflexive, transitive, and symmetric).

In this paper I want to talk about another property of epistemic modals in English that apparently differentiates them from the S5 modal operators: the fact that they can be nested in a non-trivial way. A key property of S5 is that iterated modals reduce to the innermost. For example,

$$\begin{aligned}\Box\Box\phi &\Leftrightarrow \Box\phi \\ \Box\Diamond\phi &\Leftrightarrow \Diamond\phi \\ \Box\Diamond\Box\phi &\Leftrightarrow \Box\phi,\end{aligned}$$

and so on. In S5, piling modal operators on top of an already modalized formula has no semantic effect. This is due to the fact that accessibility in S5 is an equivalence relation. The worlds accessible from  $w$  are exactly the same as the worlds accessible from any world accessible from

$w$ . If all of the worlds you can access in one step from  $w$  verify  $\phi$ , then all of the worlds you can access in two steps will also verify  $\phi$ . So,  $\Box\phi$  holds at  $w$  if and only if  $\Box\Box\phi$  does.

However, as Moss (2015, 2018) discusses in detail, this property does not seem to characterize English accurately. In examples like those in (1), there are two epistemic expressions, one of which takes scope immediately above the other. Yet both appear to make a separate contribution to the meaning of the sentence, in violation of the S5 principle that the epistemic expression that takes wider scope should be semantically vacuous.

- (1) a. This book is an odyssey, a journey up through the mists of time from the remote past. It . . . explores what must perhaps be the most fundamental of all questions — who we are. (Dunbar, 2004, p. 7)
- b. Alice is a likely hire, and Bob might be a likely hire. (Moss, 2015)
- c. “At this point, it’s probably unlikely that we’ll trade Doc,” Ricciardi said Tuesday. (*San Jose Mercury News*, 7/21/2009)

This paper is an interim report on a project with two main components. Taking a probabilistic perspective on epistemic modality, I attempt to connect the interpretation of nested epistemics with theories of uncertainty that have been developed in philosophy and psychology, but without a focus on natural language. I then report on a study that explores the distribution of examples following the pattern in (1) in a large corpus of *New York Times* articles. I suggest that the examples of nested epistemics found in this corpus vary in the kind of uncertainty involved: in most of them, one or both of the modals is not epistemic (talking about information) but rather objective (‘worldly’, ‘stochastic’: talking about chance as a genuine feature of an indeterministic world). If this is right, then many or most naturalistic examples that follow the pattern in (1) do not technically constitute counter-examples to the S5 reduction property, because one or both of the modal operators is not being interpreted epistemically. While there are a few examples where true higher-order uncertainty may be implicated, I will suggest that the lower modal in these examples may be interpreted by ‘objectivizing’ the relevant probability measure, treating it as a contingent feature of the world rather than as a contextually fixed parameter of evaluation.

If the tentative generalizations floated here about the interpretation of nested epistemics hold up, then we have a situation that is interestingly different from the cases that Lauri discussed in his groundbreaking 1972 paper. There, Lauri showed that certain intuitions derived from

the study of epistemic logic were misleading and should be discarded when doing modal semantics for English. Here, we have a case where what appears to be a straightforward counter-example to the S5 reduction property—cases where such modals are nested but both make a semantic contribution—may not constitute a counter-example after all. Instead, they point to an interesting ambiguity in the kind of uncertainty that “epistemic” expressions trade in.

## 1.2 A compositional puzzle

When dealing with examples like (1), these are some key questions that we need to resolve:

- A) What do these sentences mean?
- B) How can we define a compositional interpretation procedure for the meanings in (A)?
- C) How do the meanings of uncertainty expressions relate to what we know about the psychological representation of uncertainty?

Compositionality is a crucial principle that should guide our investigation. That is, it is desirable to have a single theory of the meaning potential that *probably* contributes in examples (2a) and (2b)—not just methodologically, but also psychologically because such an analysis helps to explain the productive use and interpretation of modal language.

- (2) a. The time is now near at hand, which must probably determine whether Americans are to be free men or slaves . . . (G. Washington)<sup>1</sup>
- b. This time will probably determine whether Americans are to be free men or slaves.

I will assume that unembedded *probable*, *probably*, and *likely* are interpreted as indicating that the proposition that it embeds has high probability (Yalcin, 2005, 2007, 2010, Swanson, 2006, 2015, Lassiter, 2010, 2015, 2017). Roughly:

- (3) probably  $\phi \Leftrightarrow P(\phi) > .5$

A probabilistic interpretation is also appropriate for complex expressions built from *probable* and *likely* such as *improbable*, *very likely*, *probability*, and *more likely than not*: see Lassiter 2015, 2017 for compositional details.

---

<sup>1</sup>Example borrowed from Moss (2015). The quote is from a speech delivered on August 26, 1776, before the Battle of Long Island.

In addition, I will assume—rather more controversially—that certain other epistemic expressions including *possible*, *might*, *must* and *certain* are also interpreted directly in terms of probabilities. *Possible* and *might* mean roughly probability greater than 0, and *certain* and *must* mean probability 1 or something close to it. For a detailed analysis of these expressions and their relationship to *probable* and *likely* see Lassiter 2015, 2016, 2017. If you do not like the latter set of assumptions, feel free to think of it as a simplification for expository purposes, holding place for a more elaborate theory which captures the probabilistic entailments of these expressions in a more complicated way.

The question, then, is how to embed the interpretation of *probably* etc. in a way that correctly predicts the interpretation of Washington’s utterance in (2a). In principle, there is a straightforward way to make this happen: we can treat probabilities as being dependent on both the context and the world of evaluation. For example, we might suppose that example (1c) has roughly the truth-conditions sketched in (4).

- (4) *It’s unlikely that we’ll trade Doc* is true in context  $c$ , relative to world  $w$ , iff  $P_{c,w}$ , the probability measure associated with the beliefs of the speaker in  $c$ , assigns a value less than .5 to the proposition that “we” will trade Doc.

This kind of meaning would embed smoothly:

- (5) *It’s probably unlikely that we’ll trade Doc* is true in context  $c'$ , relative to world  $w'$ , iff  $P_{c',w'}$  assigns a value greater than .5 to the proposition generated by abstracting over the world variable in the interpretation of (4).

More compactly:

$$\llbracket 1c \rrbracket^{c',w'} = 1 \text{ iff } P_{c',w'}(\{w \mid P_{c',w}(\mathbf{we\ trade\ Doc}) < .5\}) > .5$$

On this approach, the probability measure associated with *unlikely* is world-dependent. So, we can abstract over the world variable and associate *unlikely to trade Doc* with an ordinary proposition: the set of worlds where the probabilistic beliefs of (e.g.) the speaker assign low credence to a trade.<sup>2</sup>

There are several reasons to be skeptical of this interpretation of (1c). Most importantly, it does not quite capture the right meaning.

---

<sup>2</sup>Of course, we might want to relax the assumption that the probability measure is uniquely associated with the speaker, reverting to, say, some kind of group belief. There are decent arguments for thinking of epistemic modals in this way, but it does not help here: the speaker of (1c) is equally unlikely to resolve his uncertainty about whether a trade is unlikely by polling the opinions of others around him, except inasmuch as they are informative about relevant worldly matters.



According to (5), *It's probably unlikely that we'll trade Doc* expresses the speaker's uncertainty about his own beliefs. In the context of the example—reporting a sports manager's deliberations about staffing—this is not what the speaker of (1c) seems to intend. If it were, the right way to resolve his uncertainty about Doc's prospects on the team would be soul-searching: to try to get clear about his own beliefs so that he could be sure whether he thought trading Doc was truly unlikely. However, the example does not convey the kind of epistemic self-doubt that interpretation (5) implies. The manager is more likely to resolve his uncertainty by keeping tabs on Doc's performance and other such worldly matters. In other words, example (1c) does not seem to express a failure of introspection, but rather uncertainty about how future events will unfold and so influence an eventual decision. This is the first hint of an idea that I will explore in more detail below: the interpretation of nested modal *unlikely* is not about anyone's credence, but about certain features of the world that are considered indeterminate and chancy.

Concerns about how to capture agreement and disagreement provide a second reason to question whether probabilistic expressions really make reference to anything like a context-and-world-dependent measure  $P_{c,w}$ . (See, for example, Knobe and Yalcin 2014 and references therein.) If two people argue about whether it is likely to rain, it does not seem right to suppose either that they are arguing about what one of them believes, or that they are talking about their own beliefs and thus really not disagreeing. This is an old puzzle that has generated a significant literature, and I do not want to add to it here—except to point out that many thinkers have suggested that disagreement phenomena to problematize the assumption that the world and context of evaluation jointly determine an information state relevant to evaluation. If it does not, then we may not be justified in assuming that epistemic sentences are always determinately true or false in a given context. In other words, the content of epistemic sentences may not always be associated with a classical proposition. The resolution of this contentious debate will bear heavily on the question of whether the simple-minded account of embedded epistemics sketched above is tenable.

Yalcin (2007) provides a now-famous argument against associating matrix epistemic expressions with a world-bound interpretation (along with several further arguments that I will not review here). If the interpretation described above were right, there should be a proposition associated with sentences like (6):

(6) It rained yesterday but it's unlikely that it rained yesterday.

Yalcin dubs these “epistemic contradictions”. They contrast with merely Moore’s-paradoxical sentences (e.g., *It rained yesterday but I don’t believe that it did*) in that the latter embed sensibly in contexts where epistemic contradictions remain odd.

- (7) a. ✓ Suppose that it rained yesterday but I don’t believe that it did.  
 b. # Suppose that it rained yesterday but it’s unlikely that it did.

(7a) makes perfect sense: we are asked to consider a situation where I fail to believe something true. In contrast, (7b) remains bizarre. The world-bound interpretation of epistemic expressions sketched above predicts a reasonable meaning for it, though: we are asked to suppose that we are in a world  $w$  which has the characteristics that (a) it rained, and (b) the speaker (group, etc.) assigns low credence to the proposition that it rained. So, (7b) should have a sensible interpretation akin to “Suppose that it rained yesterday but I (we, etc.) doubt that it did”.

Yalcin (2007)’s solution relies on a different parameterization of the semantic evaluation function, where epistemic modals are associated with an information state that is not determined by the world of evaluation. For present purposes, we can treat this as a probability measure that is directly supplied as a parameter of evaluation. In effect, Yalcin replaces the schema (8) that we assumed above with (9).

- (8)  $\llbracket \phi \text{ is likely} \rrbracket^{c,w} = 1$  iff  $P_{c,w}(\llbracket \phi \rrbracket^{c,w}) > .5$   
 (9)  $\llbracket \phi \text{ is likely} \rrbracket^{c,w,P} = 1$  iff  $P(\llbracket \phi \rrbracket^{c,w,P}) > .5$

Since worlds and probability measures are independent parameters of evaluation, they can now be shifted independently by embedding operators. Yalcin’s explanation for the oddity of (7b) makes crucial use of this fact. The embedding verb *suppose* not only operates on the world parameter, but also on the probability measure. In effect<sup>3</sup>:

- (10)  $\llbracket x \text{ supposes } \phi \rrbracket^{c,w,P} = 1$  iff  $\forall w' \in S(x, w): \llbracket \phi \rrbracket^{c,w',P(\cdot | S(x,w))}$ ,

where  $S(x, w)$  is the set of worlds compatible with what  $x$  is supposing at world  $w$ , and  $P(\cdot | S(x, w))$  is the probability-measure parameter restricted to the set of worlds compatible with what  $x$  is supposing at  $w$ . Crucially, we do *not* interpret the embedded occurrence of  $\phi$  relative to a probability measure associated with the universally-quantified world

---

<sup>3</sup>This presentation is somewhat modified and simplified from Yalcin’s official system, but it should be enough to make the crucial point that epistemic contradictions can be used to motivate a non-world-bound interpretation of subjective probability expressions.

variable (like so:  $\forall w' \in S(x, w) : \llbracket \phi \rrbracket^{c, w', P_{c, w}}$ ). Instead the  $P$ -parameter is modified so as to reflect the global changes in information required by supposing that it's raining.

The payoff of these modifications is that it predicts the contrast in (7) between embedded Moore's-paradoxical sentences and epistemic contradictions. To fulfill the order in (7a), you have to make sure that (11) is satisfied:

$$(11) \quad \forall w' \in S(\mathbf{you}, w) : \llbracket \text{rain yesterday} \rrbracket^{c, w', P'} \text{ and} \\ \llbracket \text{I don't believe (rain yesterday)} \rrbracket^{c, w', P'},$$

where  $P' = P(\cdot \mid S(\mathbf{you}, w))$ . This is readily satisfiable. On the other hand, (7b) requires you to suppose something impossible:

$$(12) \quad \forall w' \in S(\mathbf{you}, w) : \llbracket \text{rain yesterday} \rrbracket^{c, w', P'} \text{ and} \\ \llbracket (\text{rain yesterday}) \text{ is unlikely} \rrbracket^{c, w', P'} \\ = \forall w' \in S(\mathbf{you}, w) : \llbracket \text{rain yesterday} \rrbracket^{c, w', P'} \text{ and} \\ P'(\llbracket \text{rain yesterday} \rrbracket^{c, w', P'}) < .5$$

Recall that  $P'$  is  $P$  conditioned on the set of worlds compatible with what is being supposed. So, to fulfill the condition in (12) one would have to find a set of supposition-worlds with the property that it rained in each, but the probability of rain conditional on these same supposition-worlds is low. It is not possible to fulfill this command because the probability of rain, conditional on a set of worlds in which it is raining, can only be 1. This accounts for the oddity of (7b), in contrast to the unremarkable (7a).

The key property of Yalcin's system that makes this solution possible is the presence of independently shiftable parameters providing a world of evaluation and a relevant informational background (probability measure). Now, this account has been challenged on both empirical and theoretical grounds: see Dorr and Hawthorne 2013, Moss 2015, Stojnić 2017, Ninan 2018, Mandelkern 2018. Rather than considering whether there is some way out, I would like to consider what Yalcin's conclusion implies for the analysis of nested epistemics like *probably unlikely*. Here it is:

$$(13) \quad \llbracket \text{probably (unlikely } \phi) \rrbracket^{c, w, P} = 1 \text{ iff } P(\{w' \mid \llbracket \text{unlikely } \phi \rrbracket^{c, w', P}\}) \\ > .5 \\ = 1 \text{ iff } P(\{w' \mid P(\{w'' \mid \llbracket \phi \rrbracket^{c, w'', P}\}) < .5\}) > .5$$

Note that nothing in the interpretation of  $\phi$  depends on the variable  $w'$  that is extracted over in the interpretation of the outer modal *probably*. As a result the set of worlds  $w'$  that verify the condition  $P(\{w'' \mid \llbracket \phi \rrbracket^{c, w'', P}\}) < .5$  is either all of them (if this condition holds) or none of them (if it fails). If the condition holds and all worlds have this property,

the outer condition has probability 1 and the sentence is true. If it fails and no worlds have this property, the outer condition has probability 0 and the sentence is false. In effect, *It's probably unlikely that we'll trade Doc* is true if and only if *It's unlikely that we'll trade Doc* is.

In general, the prediction of Yalcin's domain semantics is that the interpretation of nested probability operators should be controlled by the meaning of the innermost operator. This is very close to the S5 reduction property discussed above.<sup>4</sup> As we observed above, this prediction does not seem to be correct for nested epistemics: both modals make separate contributions to the meaning of *It's probably unlikely that we'll trade Doc*.

Accepting Yalcin's analysis puts us in a bind, then: getting the right interpretation of epistemics embedded under attitude verbs like *suppose* seems to require a semantics that makes the wrong prediction when epistemics are embedded under other epistemics. Moss (2015, 2018) responds to this tension by proposing a radically revised semantic theory built around sets of probability measures, with deep revisions to the type-theoretic properties of the system, the interpretations of connectives, and a number of other adjustments. While Moss' system is attractive in many ways and has significant philosophical motivation, I want to explore whether it is possible to account for the interpretation of nested epistemics in a way that is compatible with Yalcin's explanation of epistemic contradictions, without major revisions to other aspects of the semantic theory.

In what follows I will explore the idea that we can deal with nested epistemics by importing a distinction among several different kinds of probability. These distinctions have considerable independent philosophical, psychological, and linguistic motivation. Independence of the probability measure from the world of evaluation makes sense for one prominent kind of probability, subjective credence. However, epistemic language can also make reference to mind-independent features of the

---

<sup>4</sup>Indeed, Yalcin's system has an intimate connection to S5: "informational consequence" as he defines it is equivalent to necessary consequence in S5, i.e., classical consequence when a  $\Box$  is prefixed to every sentence. See Schulz 2010 for a detailed discussion of this point.

There is an extra wrinkle when we expand the account to include probability operators. If the outer modal is a downward-monotone probability operator that excludes probability 1—such as *unlikely* or *less than 30% probability*—it maps a true inner condition to false because it has probability 1, and a false inner condition to true because it has probability 0. In effect, downward monotone nesting probability operators function as negation. Sentences where the outermost operator is non-monotone—like *exactly 50% chance* or *between 80 and 90% likely*—are predicted to be trivially false since they exclude both probability 0 and probability 1, which are the only values available in a vacuous abstraction.

world—facts about propensity and about the distribution of properties in a class—that are conceptually determined by objective features of the world of evaluation. On the account that I will pursue, epistemics that are nested under other epistemics generally take on one of these worldly interpretations. This gives us a surprisingly simple theoretical account of these initially vexing constructions: nested epistemics are generally associated with ordinary propositions and embed just like non-modal sentences. In addition, it allows us to maintain Yalcin’s account of epistemic contradictions while also leaving room for non-trivializing interpretations of nested epistemics.

To begin, I will discuss how the distinction between various kinds of probability has been made in philosophy and psychology, and then use this discussion to inform the classification of some examples drawn from a large corpus of news text.

### 1.3 Two kinds of probability

Philosophers of probability have distinguished between subjective and objective probability (e.g., Hacking 1975, Lewis 1980, Gillies 2000). Subjective probability, also known as “credence” or “degree of belief”, is a property of individuals’ psychological states. Pure subjective uncertainty might involve questions about unique events that have a determinate answer, but for which we lack certain key information. For example, I have no idea whether Plato was taller than Aristotle, but I do know that there is a determinate answer to this question. My uncertainty is of the purely subjective type.

Objective probability, also known as “chance” or “propensity”, is a non-psychological property of the world. A truly stochastic event—say, the time that it takes for a certain radioactive particle to decay—is intrinsically chancy, independent of what anyone believes about it, or indeed whether any agents exist. Along similar lines, the particle might be said to have a certain propensity to decay within a given period of time. While the extension of this kind of reasoning to non-quantum events is metaphysically controversial, it is natural to think of certain kinds of macroscopic events as non-deterministic and intrinsically random as well. A flip of a fair coin has a 0.5 chance of coming up heads, independent of what anyone thinks about it; this is a fact about the world. If I draw a ball at random from a well-mixed urn with 3 red and 7 green balls in it, there is a 0.3 chance of getting a red ball. The roll of a die, a deal of a card from a well-shuffled pack, and the result of a future sporting event whose outcome depends on many chance factors might all be thought to have a similar character. Naturally, we can also have

mixed uncertainty: for instance, I might be uncertain about whether a given coin was fair, so that I am subjectively uncertain about the value of a certain objective chance (Lewis, 1980).

One way to gain an intuitive feel for the distinction is to imagine yourself being in possession of all immediately causally relevant information about an event. For example, suppose that you were in immediate possession of all relevant information about the factors influencing an event's outcome. The only thing that you do not have direct observational evidence of is the actual outcome. Would you still have any uncertainty about the outcome? If not, you are dealing with a case of purely subjective probability. Otherwise, the situation includes at least some degree of inherent stochasticity, i.e., objective chance. Of course, the problem with this heuristic is that it relies on the intuitive notion of "immediately causally relevant" information. If the event in question involves a coin flip or weather conditions, a layperson and a physicist might resolve this expression in radically different ways. But, for most of us, it is natural to suppose that there would be a certain amount of inherent stochasticity in coin flips, draws from an urn, the outcomes of sporting events, atmospheric events, and the like.

This latter point brings out one of the reasons why the extent of objective chance is controversial in philosophy: does it exist beyond the quantum level? I want to sidestep such metaphysical questions by taking a psychological perspective on meaning, like much work in semantics and reasoning that investigates the implicit metaphysical assumptions of human language and cognition without addressing questions of the ultimate accuracy of these assumptions (cf. Bach 1986). We are not worried here about whether coin flips really involve objective chance, but whether people typically believe that it does, and whether such a belief influences the way that they communicate about uncertainty. So, it makes sense to turn to the psychological literature for clues about whether and to what extent people are sensitive to such a distinction in reasoning, decision-making, and communication.

Psychologists have distinguished two kinds of uncertainty that mirror the philosophical discussion. "Singular external" (Kahneman and Tversky, 1982) or "aleatory" (Ülkümen et al., 2015) uncertainty involves events that are considered to be intrinsically random and essentially unpredictable. This type of uncertainty is intuitively the same as objective chance, as long as we are willing to extend the latter concept well beyond the quantum level. These authors argue that our knowledge of external uncertainties is generally due to reasoning about specific cases, e.g., about the physical characteristics of a certain coin, or about the likely trajectory of an arrow fired from a bow pointed in a par-

ticular direction, under specific atmospheric conditions. Corresponding to subjective probability is what psychologists have called “internal” (Kahneman and Tversky, 1982) or “epistemic” (Ülkümen et al., 2015) uncertainty.<sup>5</sup> This is the kind of uncertainty that we have about in-principle knowable features of the world, such as the length of the Nile, or Plato and Aristotle’s relative heights.

In any case, given the proliferation of terminology in this domain, it will be useful to declare the labels that I will use at this point.

- **Objective probability**, a world-bound property of events involving inherent randomness
  - abbreviation: “**O**-probability” or just “**O**”
  - cf.: chance, propensity, singular external uncertainty, aleatory uncertainty, stochasticity, randomness, indeterminism
- **Subjective probability**, relating to an agent’s information about an in-principle knowable event
  - abbreviation: “**S**-probability” or just “**S**”
  - cf.: credence, degree of belief, internal uncertainty, epistemic uncertainty, (lack of) information

(The best labels would perhaps be “chance” and “credence”, which are widely used in the philosophical literature on probability, but these will be less useful as abbreviations later on since they begin with the same letter.)

There is considerable evidence for the linguistic and psychological reality of this distinction. On the psychological side, we have, for example, decision-making experiments which suggest that the distinction influences choice behavior. People prefer to place bets on events whose objective chance is known, rather than on those which involve pure subjective uncertainty or a mixture of subjective uncertainty and objective chance. For instance, Ellsberg (1961) famously observed that people generally prefer to bet on a binary event whose objective probability is known to be 50% rather than an event about which they have no information about objective probabilities, so that it could be anywhere in the  $[0, 1]$  range—even if they are indifferent between the two possible outcomes in both cases.

While this example involves a mixture of subjective and objective uncertainty, it is possible to separate the two cleanly by considering how people’s betting behavior is affected by whether the chancy event has already occurred. Brun and Teigen (1990) show that, when given

---

<sup>5</sup>The label “epistemic” is unfortunate, since it is really subjective beliefs that are relevant, regardless of whether the beliefs in question constitute knowledge.

the choice between placing a bet on a an event with known chances before or after it has occurred, people prefer to bet in advance. This is surprising: we would expect a strict Bayesian to be indifferent, since the subjective probability of winning the bet remains the same if there is no prospect of gaining information about the outcome once it has occurred. This result suggests that there is a preference for betting under pure *O*-uncertainty rather than pure *S*-uncertainty, and so, *a fortiori*, that people make this distinction psychologically. Relatedly, Rothbart and Snyder (1970) found that experimental participants had greater confidence in their predictions about the roll of a die if the die had not yet been rolled. Their participants were also willing to bet more money on the outcome of the roll before it had happened.

The latter findings are consonant with Kahneman and Tversky (1982)'s claim that people believe in a sharp asymmetry between past and future: past events are in-principle knowable, but future events—unless they are impossible or inevitable—involve external uncertainty. Coming from a different perspective, Lewis (1980) likewise makes a past/future asymmetry central to his theory of the relationship between subjective probability and objective chance. Note that the temporal dimension is also important in resolving what is “in-principle knowable”. As a rule, past events are in-principle knowable, and uncertainty about them is of the purely subjective type. Future events may or may not be chancy: for instance, there is an objective chance of rain tomorrow, but no chance that  $3 + 3$  will fail to equal 6 tomorrow. However, many interesting future events are chancy, and involve a mixture of subjective and objective uncertainty.

Turning to linguistic traces of the subjective/objective distinction, Ülkümen et al. (2015) addressed this question with a corpus study of uncertainty expressions in *New York Times* articles published in 2008–2009. They extracted all instances of a predetermined list of phrases from these articles, including “sure”, “confident”, “certain”, “chance”, “likely”, “likelihood”, and “probability”, and asked annotators to rate them on a number of dimensions including perspective, control over the event's outcome, source of uncertainty, and locus of uncertainty (internal or external). A key finding was that uncertainty expressions vary in whether they are typically used to discuss subjective or objective probabilities. The most interesting result was that there was a clear preference for *likely*, *chance*, *likelihood*, and *probability* when the locus of uncertainty was external (objective), and for *sure*, *confident*, *certain* when the locus of uncertainty was internal (subjective). While the result was far from categorical, it was highly reliable.

This aspect of Ülkümen et al. (2015)'s study reveals two interesting



facts. First, the subjective/objective distinction is immediately relevant for communicating about uncertainty, and not just indirectly via its effects on reasoning and decision-making. Second, the association between particular items and particular types of uncertainty is not categorical. While it is not clear how it should be explained semantically, it would be a mistake to hard-code this association into the semantics of the various expressions. For example, while Ülkümen et al. (2015) found that *likely* is much more frequently used to talk about objective chances, it is easy to find examples where the only sensible reading involves pure subjective uncertainty.

- (14) a. An accident report into a light aircraft’s crash landing at East Midlands Airport earlier this year has revealed the most likely cause was an incorrectly fitted screw in the landing gear. (web)
- b. [T]he residential-scale reservoirs . . . were likely used around 900 B.C. It’s likely that the systems were lined with a thick, clay “plaster” that allowed the areas to hold the water instead of seeping away. (web)

Presumably the cause of the plane crash under discussion is a fixed fact about the world: investigators simply have subjective uncertainty about what this cause was, and an incorrectly fitted screw is the best hypothesis. Similarly, viewed from the present day, the way that certain reservoirs were constructed and used thousands of years ago is clearly not a matter of objective chance. Subjective uncertainty is all we can muster.

Contrariwise, it may also be that *certain* can express pure objective chance in examples such as “It is certain that everyone will eventually die”. (But not, of course, in the frame “*x* is certain that . . .”.) However, it is more difficult to show that this is strictly objective since someone who is in a position to assert the ubiquity of death will presumably also be in a state of maximal subjective confidence about the truth of the same proposition.

#### 1.4 The distributional interpretation

A third kind of probability that seems to be linguistically relevant is what I will call “distributional” probability, involving information about frequencies or proportions in a reference class. (Abbreviation: **D**-probability.) This corresponds to what Kahneman and Tversky (1982) call “Distributional external” uncertainty, and is related to some versions of the frequentist interpretation of probability. *D*-probability is the kind that is relevant when we imagine ourselves constructing an

ad hoc group of real-world individuals or events, and considering the distribution of properties within this group. It can be operationalized as the proportion of individuals in the group that satisfies a certain predicate, analogous to the proportional reading of the quantifiers *few* and *many* (Partee, 1988). Equivalently, *D*-probability can be thought of as the chance that a randomly sampled member of the group would satisfy the predicate. The following web examples illustrate.

- (15) a. The eight anti-seizure medications in the study came in 37 colors and four shapes. Overall, a change in color saw about a 20 percent additional likelihood that a prescription would lapse.
- b. At the University of Nebraska College of Law [in 2006] ... [t]he probability that a Hispanic resident with the credentials of the average black admittee would be admitted was 43 percent.

In these and many other examples of *D*-readings, there is an indefinite or a bare plural within the scope of a probability expression. The indefinite is apparently not interpreted as an existential quantifier. For example, neither scopal order in (16) captures the interpretation of (15a), on either an objective or a subjective interpretation of the probability measure *prob*. Letting *b* be the implied baseline from which “additional” deviation is being measured in this example:

- (16) a.  $\exists x[\text{prescription}(x) \wedge \text{prob}(\text{lapse}(x)) = b + 0.20]$   
 b.  $\text{prob}(\exists x[\text{prescription}(x) \wedge \text{lapse}(x)]) = b + 0.20$

The first reading in (16) deals with uncertainties involving some particular prescription; the second is about the probability that *any* prescription would lapse. Neither corresponds to the natural interpretation of (15a). Rather, this example is about the probability that a randomly selected member of a certain real-world class (those prescriptions, among those on which the study gathered data, for which the color of the pills had changed) had a certain property (not being renewed by the patient). Similarly, the claim being made in (15b) is that, among the class of Hispanic applicants who had the overall credentials of an average black admittee, 43% were admitted. This is obviously different from the claim that, for some Hispanic applicant *x* with certain credentials, the probability that *x* would be admitted was or is 43% (subjectively or objectively). It is also different from a claim about the probability that a Hispanic applicant (any Hispanic applicant) with these credentials would be admitted. The chance that there would be *at least one* such applicant admitted was presumably much higher prospectively,

and retrospectively it is either 1 or 0. (This holds both objectively, and subjectively given that all relevant admissions data were available.)

*D*-readings are not restricted to sentences with indefinites, or to *likely*, *probable*, and their derivatives. For example, (17) is an example with generic *you* and the noun *chance*.

- (17) In fact, in 2000, which was one of the last years that I was mayor, it was 191 for crime in the country. For example, in Boston, there was a 59 percent greater chance you'd be the victim of a crime than in New York City. (Rudy Giuliani, Republican Presidential Primary debate, September 5, 2012)

(18) contains two instances, one with a bare plural and *chance*, and one with a gerundive phrase and *likelihood*.

- (18) There was a 67 percent greater chance that black patients would not receive treatment compared with white patients; and being single, divorced, separated, or widowed as compared with being married or having a domestic partner also increased the likelihood of not getting definitive treatment.

It would be interesting to develop a compositional interpretation for distributional readings of probability expressions—building, perhaps, on theories that treat indefinites as denoting free variables or choice functions, and on theories of genericity and their interaction with indefinites, plurals, etc. It might even turn out that apparent *D*-readings could be reduced to a generic interpretation of *S*- or *O*-readings. Since indefinites and bare plurals play an important role in the theory of genericity, this might help to explain the frequent occurrence of indefinites with an apparent free-variable interpretation in examples of *D*-readings.<sup>6</sup> However, I do not have a clear sense of how such a reduction would work in the details.

Absent some theoretical reduction along the lines just mentioned, examples like (15) appear to motivate a third type of uncertainty in the interpretation of English probability expressions, corresponding to frequentist probability, and which has not been studied previously from a linguistic perspective. *D*-probability is, like *O*-probability, a non-psychological feature of the world—as real as rain, sleet, or snow once the relevant class of individuals has been specified. Since objective facts about the world are sufficient to determine the proportion of individuals in a certain class that satisfy a certain property, it is natural to expect the interpretation of distributional probability expressions to be indexed to the world of evaluation.

---

<sup>6</sup>Many thanks to a helpful reviewer for pressing me to consider this possibility.

## 1.5 An informal corpus study of nested epistemics

Our main goal here is to understand the interpretation of examples involving nested epistemics, where one epistemic expression occurs in the immediate scope of another. I pointed out that a compositional interpretation of these examples would be very simple if we had a way to “objectivize” probabilities, tying them to the world of evaluation rather than some parameter of evaluation or other non-world-bound feature of interpretation. Now we have two such ways—objective probability and distributional probability—and some evidence that at least some “epistemic” expressions are genuinely ambiguous in whether they express subjective, objective, or distributional probability.

If we wish to explain embedded occurrences of epistemics in the most conservative possible way—with epistemic modals taking a propositional argument—this would suggest the following strong hypothesis: *Embedded epistemics always have a world-bound interpretation*. That is, we should find examples of modals indicating *O*-, *D*-, and *S*-uncertainty in matrix position, but those embedded under another epistemic expression should always have an *O*- or *D*-interpretation.

As a result, we do not expect singly embedded examples to show the full 3x3 typology of ambiguities that is in principle available. The  $X > S$  reading (for any  $X \in \{S, O, D\}$ ) is ruled out if the relevant subjective probability measure is provided by a parameter of evaluation that is independent of the world of evaluation, as we are assuming. This assumption is borrowed from previous work, especially Yalcin 2007, and could of course be discarded by relaxing this assumption or by giving a more complex theory of embeddings (as in Moss 2015, 2018).

On the other hand, we might find that *S*-readings are not totally unembeddable:  $S > S$  readings might occur when an agent is genuinely uncertain about her own beliefs, due to psychological limitations (e.g., time pressure rendering it difficult to reason through a problem, so that one’s own implicit beliefs have not yet emerged). This is in fact what I expect to find: when embedded *S*-readings occur, they should be in rare cases where it is pragmatically sensible for someone to express (un)certainty about her own subjective (un)certainty.

In this section I will evaluate these hypotheses in a rough-and-ready way against examples of nested epistemics. Except where noted explicitly, the examples are drawn from the NYT-Eng section of the Gigaword corpus, comprising 1.4 million words of *New York Times* newswire. The NYT examples were selected by retrieving all sentences in which two or more of the following items occurred within the same clause (as delimited by the punctuation marks [!,:;?]).

probable, probably, improbable, likely, certain, sure, definitely, possible, possibly, impossible, must, might, may, maybe, perhaps.

To keep the study to a manageable size I did not include many relevant items such as *confident, clear, clearly, dubious, think, believe, know, doubt, chance, likelihood, probability, possibility, confidence*.

This search technique returned a large number of hits, most of them irrelevant: for example, examples of *must* with a deontic interpretation, *certain* in expressions such as *a certain individual*, and examples in which the two modals were not in a scopal relation (e.g. *It is very probable or even certain that . . .*). However, within this collection I was able to find a substantial number of examples which do appear to be genuine nested epistemics.

I selected for discussion here examples that had a relatively clear intended meaning, and which were relevant to a theoretically interesting point. All interpretations are my own. A controlled evaluation using trained, hypothesis-naïve annotators would of course be preferable, but this will have to wait for future work.

### 1.5.1 Subjective uncertainty about casually relevant chance factors ( $S > O$ )

It is relatively easy to find examples in which an author expresses subjective uncertainty about chance factors. For example, (19) appears to have this character.

- (19) Nuclear war is as likely, and perhaps more likely, in our time than at any point during the Cold War because superpower constraints are much weaker . . . (perhaps > more likely)

The message is that complex, interacting geopolitical forces have conspired such that the objective probability of nuclear war is (definitely) at least as high as at any point during the cold war—and that the information available to the author renders it possible that this objective probability is even higher. So, *perhaps more likely* has a  $S > O$  reading.

Here is another, somewhat more complex example.

- (20) Despite starting 29th in today's Samsung/Radio Shack 500, Gordon had a spring in his step on Saturday at Texas Motor Speedway after posting the fourth-fastest lap during the final practice session. "We're probably the dark horse for sure," Gordon said. "Our cars have been really, really fast lately. . . . Maybe Sunday will be our day."

I believe that the intended scopal order here is *for sure > probably*. Assuming this, Gordon's message seems to be that he is subjectively

certain that objective features of the world (his team’s abilities, other teams’ performances) will be arranged so that his team probably performs above expectations on Sunday.

- (21) To build a large database of people who are interested in participating as links in the chain, Watts’ team has set up a Web site . . . Participants will receive e-mail messages asking them to help locate a target person by forwarding e-mail to a personal contact who might be likely to know the target.

As I understand this example, Watts’ team is asking participants to identify people who might (given their information about the individuals in question) be likely (given objective facts about that individual, such as their location and activities) to know the target.

The following is a relatively clear case, where subjective *definitely* embeds the epistemic expression *much more likely*. The latter is explicitly justified in terms of current conditions in the global economy—a radically indeterminate macrostructure if there ever was one.

- (22) “You can be guaranteed of rocketing commodity prices if and when the global economy starts working together,” said Bond. “We’ve definitely got conditions now where commodity price squeezes are much more likely than they were in the past.”

(23) has a similar character:

- (23) The new federal law makes most legal immigrants ineligible for Supplemental Security Income . . . All those people might be likely applicants for Home Relief.

If I understand correctly, the intended interpretation of (23) is that, for each newly ineligible immigrant  $x$ , it is subjectively possible that  $x$  would be likely (in light of objective facts about the individual and the application process) to apply for Home Relief. (As an aside, (23) is a counter-example to von Fintel and Iatridou (2003)’s “epistemic containment” constraint, which forbids quantifiers from taking scope directly above an epistemic modal. See Swanson 2010 for additional counter-examples.)

### 1.5.2 Subjective uncertainty about distributional probabilities ( $S > D$ )

The only good candidate of a potential  $S > D$  example that I have located is (24):

- (24) We do have an enemy, and it is definitely possible that what you consider a “word from God” could in fact be demonic and be “a word from satan.” (web)

An  $S > O$  reading of this example is not very plausible, since it is presumably a determinate present fact whether a given communication is from God or from Satan (or indeed from neither). If so, only  $S > S$  and  $S > D$  would make sense here. My interpretation of this example goes as follows: the author is expressing maximal confidence in the claim that Satan has the ability to deceive people into thinking that his words come from God. In other words, the author is confident that—among the class of communications that his audience consider to be “words from God”—a non-zero proportion of these are in fact “words from Satan”. If so, the quantity that he is maximally confident of is a distributional probability.

### 1.5.3 $D > S?$

I have found no corpus examples that clearly indicate a  $D > S$  interpretation. Here is an invented example that has been suggested to me:

(25) He is a Republican; so, it’s unlikely that he is probably a pacifist.

To the extent that the example makes sense, it suggests distributional reasoning: if you sample an individual at random from the class of Republicans, it is unlikely that the individual sampled will be one who is probably a pacifist. We might imagine, after sampling an individual, being forced to infer from observables whether the individual is a pacifist. In any case, if example (25) is intelligible it may provide another argument that subjective readings of probability expressions can be embedded under other epistemic operators.

### 1.5.4 Multiple layers of objective uncertainty ( $O > O$ )

Most of the examples that follow could also be interpreted as  $S > O$ . The best  $O > O$  candidates seem to involve well-informed speculation about the future, which (as we noted above) people seem to treat as open and in-principle unknowable.

(26) NBA Commissioner Adam Silver said Wednesday night that it’s “probably unlikely” for Orlando to host the 2017 NBA All-Star Game because the city hosted the exhibition just several years ago, in 2012. . . . Silver answered that Orlando’s chances are slim “only because the game was here in 2012 and there are some other teams that haven’t had an All-Star Game in a long time . . . But having said that, this was a great experience down here in 2012, and I’m sure we won’t be waiting to go to 29 other teams before we come back here.” *(Orlando Sentinel, 2/25/15)*

Here Silver expresses uncertainty about the future value of a set of

variables—other teams’ levels of interest—which are known to be causally relevant to Orlando’s chances of getting the Game. (This choice is ultimately under Silver’s control.) It’s hard to be sure that the “probably” here is of the objective sort, but if it is, this is an  $O > O$  example.

- (27) Tom McGowan, the town planner, suggested that if critical sites needed protection that could not be provided in any other way, “a very focused proposal might be likely to receive favor.” He cited such vulnerable open areas as Routes 118 and 254, both entry roads into the village center.

We might think to treat (27) as  $D > S$ , but I think that this would be incorrect: McGowan is not discussing the prospects of a proposal randomly selected from a known class, but rather noting that, if a focused proposal is written, it *might* (depending on its characteristics) be *likely* (depending on the planning commission’s internal deliberations) to receive favor. If so, this can be read as an  $O > O$  example.

### 1.5.5 Genuine higher-order uncertainty ( $S > S$ )

Many theorists have wished to ban true higher-order uncertainty, of the  $S > S$  variety, for philosophical or computational reasons. It’s true that allowing higher-order uncertainty makes the semantics more complicated, and that it conflicts with certain theoretical positions about the nature of belief or introspection. Still, it may be a real thing. Here are some examples that could be analyzed as  $S > S$  readings.

- (28) “I believe in a ‘yes’ victory, but I’m definitely not certain,” said Prime Minister Goran Persson . . .

I do not see any real alternative but to understand this example as an expression of Persson’s higher-order certainty about the non-maximal value of his own subjective degree of belief. Here is another similar example:

- (29) This book is an odyssey, a journey up through the mists of time from the remote past. It explores what must perhaps be the most fundamental of all questions - who we are. (web)

As I understand it, the author is not at all sure about what the most fundamental of all questions is—but he is virtually certain that it *might* be “who we are”. So, this looks like another  $S > S$  reading.

Once we have opened this can of worms, further candidates start to appear.

- (30) Mr. Nolan must certainly be rubbing his conspiratorial hands with glee as he hears that the arrival of Primark has meant that



other stores are now looking to make a new home in Truro. (web)

This example would typically be interpreted as modal concord. This is not an unreasonable analysis, but note that examples (28) and (29) cannot be treated in this way since the modals are not matched in strength. Perhaps some constructions that could plausibly be used to express concord can also yield to a nested analysis. I am tempted to read (30) as expressing higher-order confidence, either about the author's own certainty or about the appropriate degree of certainty for conversational partners to adopt in light of the evidence available.

I do not know whether any of these examples will hold up in the end: perhaps they can all be explained away in some other fashion. But it may also be that authors sometimes wish to express (un)certainty about their own beliefs, about how their beliefs will stabilize upon reflection, or about what the appropriate degree of belief is for people with access to a given body of information.

## 1.6 Distributional considerations

There appear to be significant gaps in the distribution of nested modals in the Gigaword-NYT corpus. For example, none of the following occur in the same clause in the corpus.

- *definitely ... may/(im)probable*
- *impossible ... definitely/maybe*
- *perhaps ... definitely*
- *possibly ... improbable/probable*
- *improbable ... probably/certain/definitely/may/must/perhaps/possibly/sure*

In a corpus as large as this one, where many combinations are attested by numerous examples, these gaps are striking. Some of them may be explicable in terms of Ülkümen et al. (2015)'s observations about preferred interpretations of modals: as I mentioned above, they point out that *certain*, *sure*, *confident*, and expressions derived from them tend to deal with subjective uncertainty, while *likely*, *likelihood*, *chance*, *probability* and their derivatives prefer an external (objective or perhaps distributional) interpretation. If such preferences do exist, then we might be able to explain certain gaps by pointing to the semantic ill-formedness, or pragmatic infelicity, of the preferred readings of certain combinations.

## 1.7 Embedding epistemics simple-mindedly

One of the goals of this paper is to consider the prospects of a simple-minded analysis of nested epistemics: probability expressions of all stripes embed propositions (sets of possible worlds), and anything that can serve as their direct argument must pick out a set of possible worlds. If we start with Yalcin (2007)'s hypothesis that subjective probabilities come into the composition via a parameter that is independent of the world of evaluation, we expect that subjective probability expressions should not embed non-trivially, but objective and distributional probability expressions—being world-bound—should.

The examples tentatively analyzed as  $S > S$  in section (1.5.5) do not yet refute this position, but they do problematize it somewhat. If we wish to hold onto a simple theory of embeddings, we may have to “objectivize” subjective probabilities in some cases, treating them as features of the world about which we can be uncertain. There does not seem to be any technical barrier to doing so, though there might be philosophical reasons to try to find a way around it. Alternatively, we could frame a more complicated theory of embeddings, as Moss (2015, 2018) does. While this theory does many things well, it is definitely much more complex and radical in its semantic and philosophical implications than the simple-minded approach suggested here.

This brings us back to the vacuity issue from section 1. Do multiple epistemics reduce to the innermost in English, as they do in S5 and (in a slightly more complicated way) on Yalcin (2007)'s account? On face, the answer is a clear “no”: the nested examples that we have analyzed have complex, subtle interpretations that definitely are not equivalent to the same sentences with the outermost modal deleted. However, the English data are not relevant to this question if there is a subtle ambiguity in the interpretation of the modal expressions, as I argued that there is in many cases. A convincing test of the empirical claim would have to involve  $S > S$ ,  $O > O$ , or  $D > D$  interpretations.  $D > D$  readings do not seem to occur among the examples that I have found. (I am unsure what they would even mean). Various examples did occur that could plausibly be analyzed as  $S > S$  or  $O > O$ , but these could probably be re-interpreted in some other way by a dedicated proponent of the S5 reduction property.

This conclusion is somewhat unsatisfying, since we do not know for sure yet whether reduction holds. Reduction does have two nice features that might make it worth pursuing, though. First, it may allow us to maintain Yalcin (2007)'s account of epistemic contradictions in light of problem cases involving nested epistemics noted by Moss (2015,

2018) . Second, it means that we will have to continue doing what Lauri taught us many years ago: continue to take inspiration from logic and philosophy, but evaluate every idea inspired from these sources carefully against empirical data. The only major methodological development to report is that taking a probabilistic perspective on modal semantics allows to draw additional inspiration from the rich tradition of philosophical and psychological work on belief, reasoning and choice that is tied in to probabilistic modeling and statistics.

## References

- Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9(1):5–16.
- Brun, Wibecke and Karl Halvor Teigen. 1990. Prediction and postdiction preferences in guessing. *Journal of Behavioral Decision Making* 3(1):17–28.
- Dorr, Cian and John Hawthorne. 2013. Embedding epistemic modals. *Mind* 122(488):867–913.
- Dunbar, Robin. 2004. *The Human Story*. Faber & Faber.
- Ellsberg, Daniel. 1961. Risk, ambiguity, and the Savage axioms. *The quarterly journal of economics* pages 643–669.
- Gillies, Donald. 2000. *Philosophical theories of probability*. Psychology Press.
- Goodhue, Daniel. 2018. *Must p* is felicitous only if *p* is not known. To appear in *Semantics & Pragmatics*.
- Groenendijk, Jeroen and Martin Stokhof. 1975. Modality and conversational information. *Theoretical linguistics* 2(1-3):61–112.
- Hacking, Ian. 1975. *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press.
- Hintikka, Jaako. 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press.
- Kahneman, Daniel and Amos Tversky. 1982. Variants of uncertainty. *Cognition* 11(2):143–157.
- Karttunen, Lauri. 1972. Possible and must. In J. Kimball, ed., *Syntax and semantics*, vol. 1. Seminar Press.
- Knobe, Joshua and Seth Yalcin. 2014. Epistemic modals and context: Experimental data. *Semantics and Pragmatics* 7:1–21.
- Kratzer, Angelika. 1991. Modality. In A. von Stechow and D. Wunderlich, eds., *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, pages 639–650. Walter de Gruyter.
- Lassiter, Daniel. 2010. Gradable epistemic modals, probability, and scale structure. In N. Li and D. Lutz, eds., *Semantics & Linguistic Theory (SALT) 20*, pages 197–215. CLC Publications.

- Lassiter, Daniel. 2015. Epistemic comparison, models of uncertainty, and the disjunction puzzle. *Journal of Semantics* 32(4):649–684.
- Lassiter, Daniel. 2016. *Must*, knowledge, and (in)directness. *Natural Language Semantics* .
- Lassiter, Daniel. 2017. *Graded Modality*. Oxford University Press.
- Lewis, D. 1980. A subjectivist’s guide to objective chance. In R. Jeffrey, ed., *Studies in inductive logic and probability*, pages 263–293. University of California Press.
- Mandelkern, Matthew. 2016. A solution to Karttunen’s problem. To appear in *Proceedings of Sinn und Bedeutung 2016*.
- Mandelkern, Matthew. 2018. Bounded modality. To appear in *The Philosophical Review* .
- Moss, Sarah. 2015. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8:1–81.
- Moss, Sarah. 2018. Probabilistic knowledge .
- Ninan, Dilip. 2018. Relational semantics and domain semantics for epistemic modals. *Journal of Philosophical Logic* 47(1):1–16.
- Partee, Barbara. 1988. Many quantifiers. In *Proceedings of ESCOL*, vol. 5, pages 383–402.
- Rothbart, Myron and Mark Snyder. 1970. Confidence in the prediction and postdiction of an uncertain outcome. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* 2(1):38.
- Schulz, Moritz. 2010. Epistemic modals and informational consequence. *Synthese* 174(3):385–395.
- Stojnić, Una. 2017. One’s modus ponens: Modality, coherence and logic. *Philosophy and Phenomenological Research* 95(1):167–214.
- Swanson, Eric. 2006. Interactions With Context. Ph.D. thesis, MIT.
- Swanson, Eric. 2010. On scope relations between quantifiers and epistemic modals. *Journal of Semantics* .
- Swanson, Eric. 2015. The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic* pages 1–26.
- Ülkümen, Gülden, Craig R Fox, and Bertram F Malle. 2015. Two dimensions of subjective uncertainty: Clues from natural language. To appear in *Journal of Experimental Psychology: General*.
- von Fintel, Kai and Anthony Gillies. 2010. Must... stay... strong! *Natural Language Semantics* 18(4):351–383.
- von Fintel, Kai and Sabine Iatridou. 2003. Epistemic containment. *Linguistic Inquiry* 34(2):173–198.
- Yalcin, Seth. 2005. A puzzle about epistemic modals. *MIT Working Papers in Linguistics* 51:231–272.
- Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464):983–1026.
- Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11):916–937.

