Probability and implicatures: A unified account of the scalar effects of disjunction under modals*

Paolo Santorio
University of Leeds

Jacopo Romoli
Ulster University

Abstract Sentences involving disjunction under epistemic modal adjectives — such as *possible*, *likely*, and *certain* — give rise to the inference that the disjuncts are epistemically possible. Inferences of this sort are often classified and treated differently, depending on the force of the embedding modal. Those triggered by possibility modals are singled out as 'free choice inferences' (Kratzer & Shimoyama 2002, Klinedinst 2007, Fox 2007, Chierchia 2013, a.o.), while those triggered by stronger modals are accounted for in a different way (Sauerland 2004, Fox 2007, Crnič, Chemla & Fox 2015 a.o.). In this paper, we pursue two goals. First, we develop and defend a degree semantics for epistemic modal adjectives, building on much recent work on the topic (Yalcin 2010, Lassiter 2011, 2014, Moss 2015a, Swanson 2015, a.o.). Second, we show that this semantics, in combination with the assumption that scalar implicatures can arise in embedded position (Fox 2007, Chierchia, Fox & Spector 2012, a.o.), can predict all the inferences triggered by disjunction under modals, including free choice ones, via a uniform mechanism. We conclude by outlining how the proposal can be extended to epistemic modal items in other syntactic categories, and to modals of different flavor.

1 Introduction

1.1 The inferences of disjunction under epistemic modals: An overview

Disjunctions in the scope of epistemic modal expressions give rise to scalar inferences to the effect that each disjunct is epistemically possible. For example, all the sentences in (1a)–(1c) are naturally heard as suggesting that the sentences in (2a) and (2b) are true.

(1)   a.   It's possible that we will hire Mary or Sue.
      b.   It's likely that we will hire Mary or Sue.
      c.   It's certain that we will hire Mary or Sue.

(2)   a.   It's possible that we will hire Mary.
      b.   It's possible that we will hire Sue.

Part of the evidence for these effects is that all of (1a)–(1c) are infelicitous in a context where one of (2a) or (2b) is false.

(3)   I talked to the committee, it's impossible that we will hire Mary but … #It's possible/likely/certain that we will hire Mary or Sue.

These observations generalize to epistemic modal expressions in different syntactic categories, including adverbs like *possibly, probably, certainly* and auxiliaries like *might, should* and *must*. They also arise with some modal expressions with non-epistemic flavor — in particular, with deontic modal adjectives like *allowed* and *required*, and with the corresponding auxiliaries *may* and *have to*.[1] Some examples are below in (4)–(7).

(4)   It might rain or snow tomorrow.
      a.   ⤳ *It might rain tomorrow*

---

1 Disjunction embedded under quantificational DPs like *all students* and *some students* also appears to give rise to similar inferences (Klinedinst 2007, Fox 2007 among others):

(i)   All students/Some students took Syntax or Logic this year.
      a.   ⤳ *some student took Syntax*
      b.   ⤳ *some student took Logic*

We focus on the case of modals in this paper and we leave to further work the task of investigating extensions of our account to cases like those in (i). See also footnote 8 below, where we point to arguments from Crnič, Chemla & Fox 2015, which call into question whether we should treat at least the case of *all* in the same way.

> b. ⤳ *It might snow tomorrow*

(5) John must have taken Syntax or Logic last year.

> a. ⤳ *John might have taken Syntax last year*
> b. ⤳ *John might have taken Logic last year*

(6) It is allowed to smoke or drink in this room.

> a. ⤳ *It's allowed to smoke in this room*
> b. ⤳ *It's allowed to drink in this room*

(7) John has to take Syntax or Logic this year.

> a. ⤳ *John may take Syntax this year*
> b. ⤳ *John may take Logic this year*

All these effects are unexpected from the viewpoint of classical modal semantics, which descends from modal logic and treats natural language modals as quantifiers over worlds. This analysis, in combination with the hypothesis that *or* simply corresponds to Boolean disjunction, predicts that the following inference patterns are invalid:[2]

$$\Box(p \vee q) \quad \nvDash \quad \Diamond p \wedge \Diamond q$$
$$\Diamond(p \vee q) \quad \nvDash \quad \Diamond p \wedge \Diamond q$$

There is broad (though not universal) agreement that a Boolean analysis of disjunction and classical modal semantics are essentially correct, and that the scalar effects we described should be captured as implicatures that arise on top of literal meaning (Kratzer & Shimoyama 2002, Simons 2005, Fox 2007, Klinedinst 2007, Chemla 2008, van Rooij 2010, Franke 2011, Alonso Ovalle 2005, Chierchia 2013, Crnič, Chemla & Fox 2015).[3] The main argument for this approach is that the problematic effects disappear in downward entailing contexts — a signature of scalar implicatures in general. For example, consider (8):

(8) It's not possible that we will hire Mary or Sue.

If the inferences in (2a) and (2b) arose also under negation, (8) would have the meaning in (9a), represented schematically in (9b):

---

2 The same point holds for *likely*; see §2 for details.

3 For some semantic accounts of free choice inferences, see Higginbotham 1991, Zimmerman 2000, Geurts 2005, Barker 2010, Starr 2016, among others. In this paper, we lack the space to set up a comparison between our analysis and these accounts.

(9)   a.   It's not true that: it's possible that we will hire Mary or Sue and it's possible that we will hire Mary and it's possible that we will hire Sue.

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

Notice that (9a) is compatible with it being possible that (say) we hire Mary. But this is clearly incorrect. (8) says that it's certain both that we will not hire Mary and that we will not hire Sue. This shows that the inferences we're interested in don't arise under negation.

To sum up: the scalar effects introduced above can be captured via a simple generalization:

> **Possibility Implicatures of Modals (PIM)**
>
> Sentences of the form MOD$[p \vee q]$ (where 'MOD' stands for any modal that is not downward entailing in its prejacent position) give rise to the inferences that each of the disjuncts is possible.
>
> $$\text{MOD}[p \vee q] \rightsquigarrow \Diamond p, \Diamond q$$

Let us notice right away that PIM applies to modal expressions of various flavors and syntactic categories. Throughout the core part of the paper, we focus on epistemic modal adjectives. This choice is dictated by a practical concern: our proposal makes crucial use of a degree semantics, and the most developed degree-based analyses of modal terms concern just items of this sort. We are going to sketch how the proposal may be extended to other items in §8.

1.2   The project: A uniform account

Strikingly, the great majority of analyses in the literature do not predict PIM via a uniform mechanism.[4] Instead, the scalar effects covered by PIM

---

4 Two notable exceptions to this are Chemla 2008 and Klinedinst 2007. Our proposal is in particular very similar in spirit to that of Klinedinst 2007, although we do not make the assumptions that modals quantify over pluralities of worlds. See §9.2 below for a brief comparison with Klinedinst's (2007) approach. We leave a detailed comparison with Chemla's (2008) proposal for future research. Notice also that Chierchia (2013) calls 'free choice inferences' both distributive and free choice inferences, but derives them in very different ways.

are generally classified and treated differently. The effects triggered by non-existential modals are sometimes called 'distributive inferences'—the intuition being that they 'distribute' the two propositions corresponding to each disjunct across the worlds in the quantificational domain. These inferences can be handled straightforwardly by some standard accounts of scalar implicature (e.g. Sauerland 2004 and Fox 2007). We go through the details in §2, but the basic algorithm consists simply in conjoining the basic meaning of the sentence with the negation of some of its stronger alternatives. This strengthened meaning entails the relevant possibility claims. Schematically:

$$\text{MODAL}(p \vee q)$$
$$\Downarrow$$
$$\text{MODAL}(p \vee q) \wedge \neg\text{MODAL}(p) \wedge \neg\text{MODAL}(q)$$
$$\Downarrow$$
$$\Diamond(p) \wedge \Diamond(q)$$

The possibility inferences triggered by possibility modals, on the other hand, are called 'free choice inferences' and are usually derived via a more complex route. One increasingly popular account (though by no means the only live option) involves exploiting a recursive mechanism for computing implicatures (see Kratzer & Shimoyama 2002, Fox 2007, Chierchia 2013 among others).

In this paper, we show that, by adopting a degree semantics for epistemic modal expressions, we can derive all the effects covered by PIM, in a uniform way, via the simple schematic process outlined above. The idea builds on recent work on the semantics of epistemic modals. Recent literature on modality has defended the claim that a degree semantics for modals might be empirically and theoretically superior to a classical quantificational semantics (Swanson 2006 and Lassiter 2011; see also Swanson 2015, Yalcin 2007, 2010, Lassiter 2014, Moss 2015a for related work). On one popular implementation, epistemic modals work as measures of probability, mapping propositions to a value on a probability scale. For concreteness, in this paper we adopt a probabilistic semantics for modals, though in principle other implementations of a degree semantics are possible. As mentioned, we focus on epistemic modal adjectives like *possible, likely* and *certain* as our main case study.

The rest of the paper is organized as follows. In §2, we set up some background on the semantics of modals and scalar implicature. In §3, we show how this account can straightforwardly derive distributive inferences, but not free choice ones. After a brief overview of the proposal in §4, in §5 we give

some background on degree semantics for adjectives in general, while in §6 we discuss in detail the case for a degree semantics for epistemic adjectives. In §7, we put forward our unified account of all possibility inferences and in §8 we show how it can be extended beyond epistemic modals. In §9, we discuss other remaining issues and make some brief comparisons with alternative proposals.

2 General background

2.1 The classical semantics of modals and probability operators

The contemporary benchmark for the semantics of modality is set by Kratzer's analysis (1981, 1991, 2012; see also Portner 2009). In outline, Kratzer treats modals as quantifiers over a contextually restricted domain of worlds. This domain of quantification is determined via two contextually provided sets of propositions (more precisely: functions from worlds to sets of propositions), which Kratzer calls *modal base* and *ordering source*. The modal base is used to single out the domain of quantification of the modal: worlds in the domain are all and only those that satisfy all the propositions in the modal base. The ordering source is (simplifying somewhat) used to induce an ordering on the worlds in this domain, singling out a set of 'best' worlds, along some relevant dimension.[5] What properties of worlds matter for the ordering depends on the flavor of modality involved. In particular, Kratzer takes epistemic modals to use an ordering that ranks worlds on the basis of how much they conform to stereotypical assumptions.

---

5 One standard way to extract an ordering $\preceq$ from a set of propositions S is straightforward (as pointed out by Lewis 1981). We say that, for any two worlds w_1, w_2, $w_1 \preceq w_2$ iff, for any $p \in S$, if $p(w_2) = 1$, then $p(w_1) = 1$. The set of 'best' worlds is the set of worlds such that there is no world that is better than them; i.e., $\text{BEST}_{\preceq} = \{w : \neg\exists w' : w' \prec w\}$. For simplicity, here we are assuming that deontic modals satisfy the so-called limit assumption (Lewis 1973, Stalnaker 1984). The limit assumption says that we can single out a set of 'best' worlds starting from any world in the modal base. Formally (and following the formulation in Kaufmann & Kaufmann 2015):

> A pair of a modal base f and an ordering source g satisfies the Limit Assumption iff for all possible worlds w, for all $v \in \bigcap f(w)$ there is a $u \in \text{BEST}_{g(w)}(f(w))$ such that $u \preceq_{g(w)} v$

Where $\text{BEST}_{g(w)}(f(w))$ is the subset of $\bigcap f(w)$ such that for all its elements there isn't a strictly $g(w)$-better world in $\bigcap f(w)$.

Using 'f' and 'g' to pick out respectively the modal base and the ordering source, an analysis of *possible* and *certain*, under this approach, is in (10) and (11).[6]

(10) $[\![$possible $\phi]\!]^{w,f,g} = 1$ iff $\exists w' \in \text{BEST}_{g(w)}(f(w))[[\![\phi]\!]^{w,f,g}(w')]$

(11) $[\![$certain $\phi]\!]^{w,f,g} = 1$ iff $\forall w' \in \text{BEST}_{g(w)}(f(w))\ [[\![\phi]\!]^{w,f,g}(w')]$

As Kratzer emphasizes, her view is well-equipped to handle graded modal expressions, i.e. modal expressions with intermediate force between possibility and necessity. *likely*, in particular, is treated as follows: *likely* ϕ is true iff, for every non-ϕ-world in the modal base, there is a ϕ-world in the modal base that is more highly ranked. Formally:

(12) $[\![$likely $\phi]\!]^{w,f,g} = 1$ iff
$\forall w' \in [\![\neg\phi]\!]^{w,f,g} \cap f(w)\colon \exists w'' \in [\![\phi]\!]^{w,f,g} \cap f(w)\colon w'' \prec_{g(w)} w'$

The analysis naturally generalizes to modal expressions in other syntactic categories, as well as expressions of different strength, modulo changes in quantificational force. Some examples of epistemically modalized expressions covered by Kratzer's analysis are modal auxiliaries like *might*, *should*, and *must* and adverbs like *possibly*, *probably*, and *certainly*.

2.2   Scalar implicatures

Before discussing distributive and free choice inferences, it is helpful to introduce scalar implicatures in general.

Grice (1975) first singled out implicatures as a category of interpretive effects going beyond the basic literal meaning of a sentence. Scalar implicatures are a subkind of implicatures. Roughly, scalar implicatures involve the denial of a logically stronger alternative to a sentence. Here is a classical example:

(13)   a.   We will hire Mary or Sue.
        b.   ⤳ *We won't hire both Mary and Sue*

Grice treated all implicatures as purely pragmatic in nature. On his view, implicatures are produced via general principles of reasoning, which take as

---

6 For simplicity, we treat adjectives as propositional operators. Nothing hinges on this assumption though. See also Swanson 2015 for discussion.

input the final result of the compositional computation of semantic value. Much recent work has questioned this account. On the alternative view, scalar implicatures are generated by semantic processes, i.e. processes that are implemented compositionally. The debate is far from being settled and we will not enter this discussion here. As we point out below, however, we need a view that allows scalar implicatures to arise from embedded positions. More precisely, we need a theory that allows for 'intermediate' scalar implicatures, in the sense of Sauerland (2012, 2014). A natural way to do this is to adopt a semantic view of scalar implicatures (Fox 2007, Chierchia, Fox & Spector 2012, Chierchia 2013, a.o.), though any theory that allows for intermediate scalar implicatures would work for our purposes.

Chierchia, Fox & Spector (2012) suggest that scalar implicatures are generated by a silent operator, which they represent as 'EXH' (for 'exhaustification'). EXH is roughly akin in meaning to natural language *only*. Like *only*, EXH combines with a sentence ϕ and it returns the proposition resulting from conjoining the meaning of ϕ with the negation of some of the alternatives of ϕ, $ALT(\phi)$ — what we call *excludable alternatives*. Schematically:

(14) $[\![\text{EXH } \phi]\!](w) = [\![\phi]\!](w) \wedge \ \forall \psi \in \text{EXCL}(\phi, ALT(\phi))[\neg[\![\psi]\!](w)]$

We take an alternative to be excludable just in case (a) negating it doesn't contradict the literal meaning of the sentence asserted, and (b) negating it doesn't force us to accept other alternatives on the list (Sauerland 2004, Fox 2007; see also Gazdar 1979). Here is the formal definition.[7]

(15) $\text{EXCL}(\phi, X) = \{\psi \in X : [\![\phi]\!] \not\subseteq [\![\psi]\!] \wedge$
$\neg \exists \chi [\chi \in X \wedge ([\![\phi]\!] \wedge \neg[\![\psi]\!]) \subseteq [\![\chi]\!]]\}$

The idea behind the functioning of EXH is this: we try to strengthen the sentence as much as possible, while at the same time avoiding both contradictions and arbitrary choices. For an example, consider the sentence in (13a), which has the alternatives in (16).

---

7 This algorithm is called 'innocent exclusion' in Fox 2007. (15) is not the final version of innocent exclusion used by Fox, but it is enough for our purposes. For discussion see Fox 2007.

(16)
$$\left\{\begin{array}{ll}\text{We will hire Mary or Sue} & \mathbf{M} \vee \mathbf{S} \\ \text{We will hire Mary} & \mathbf{M} \\ \text{We will hire Sue} & \mathbf{S} \\ \text{We will hire Mary and Sue} & \mathbf{M} \wedge \mathbf{S}\end{array}\right\}$$

Of these alternatives, only the conjunctive one ($\mathbf{M} \wedge \mathbf{S}$) is excludable. As for the other two: if we excluded both, we would get a contradictory meaning (since, combining the implicatures with the assertion, we would have that one between Mary and Sue will be hired, but not Mary, and not Sue); if we excluded one, we would have to arbitrarily select one of the two disjuncts as the true one (e.g., if we ruled *We will hire Mary* as false, given the content of (13a) we would have to conclude that *We will hire Sue* is true). Hence only the conjunctive alternative is ruled out, and the strengthened meaning of the sentence is:

(17)   We will hire Mary or Sue and it's not true that we will hire both.

The kind of strengthening induced by EXH is generally called *exhaustification*, and sentences that have gained a stronger meaning in this way are called *exhaustified*.

So far, we haven't said what sentences enter the set of alternatives used to compute exhaustified meanings. This is a controversial issue in the literature, and one that cross-cuts the line between pragmatic accounts and semantic ones. This issue doesn't matter much for our purposes. For concreteness, we assume the complexity-based account given by Katzir 2007 and Fox & Katzir 2011. In outline, on this account the alternatives to *S* are those sentences that are no more complex than *S*, and that can be obtained from *S* by replacing *S*'s constituents with its subconstituents or relevant items from the lexicon.

3   Deriving distributive inferences (but not free choice ones)

3.1   Distributive inferences with *certain* and *likely*

We can now go back to sentences like (18a) and (18b) and show how we can derive the distributive inferences in (19a) and (19b) as straightforward scalar implicatures.

(18)   a.   It is certain that we will hire Mary or Sue.
         b.   It is likely that we will hire Mary or Sue.

(19) a. ⤳ *It is possible that we will hire Mary*
     b. ⤳ *It is possible that we will hire Sue*

Consider first (18a); its alternatives are in (20). All of the alternatives in (20) (aside from the one corresponding to the assertion) are excludable: they are all stronger than the basic meaning of (18a), and there is no alternative such that negating it compels us to accept another alternative.

(20) $\left\{ \begin{array}{ll} \text{It is certain that we will hire Mary or Sue} & \Box(\mathbf{M} \vee \mathbf{S}) \\ \text{It is certain that we will hire Mary} & \Box(\mathbf{M}) \\ \text{It is certain that we will hire Sue} & \Box(\mathbf{S}) \\ \text{It is certain that we will hire Mary or Sue} & \Box(\mathbf{M} \wedge \mathbf{S}) \end{array} \right\}$

By conjoining the basic meaning with the negated excludable alternatives, we obtain the distributive inferences that we are looking for, since (21) entails (22). (Here and in what follows, to avoid clutter we ignore the conjunctive alternative, which is superfluous since its negation is entailed by the negation of the other two alternatives.)[8]

(21) $\Box(\mathbf{M} \vee \mathbf{S}) \wedge \neg\Box(\mathbf{M}) \wedge \neg\Box(\mathbf{S})$

(22) $\Box(\mathbf{M} \vee \mathbf{S}) \wedge \Diamond(\mathbf{M}) \wedge \Diamond(\mathbf{S})$

This can easily extend to the case of *likely* in (18b), the alternatives of which are in (23). To our knowledge, the scalar inferences triggered by *likely* have not been discussed in the literature. But, as the reader can check, by adopting

---

8 As mentioned above, this account of distributive inferences is generally applied also to the parallel inferences of disjunction embedded in the scope of a universal quantifier like (i).

(i) Every student took Syntax or Logic

   a. ⤳ *some student took Syntax*
   b. ⤳ *some student took Logic*

Again, the inferences in (ia) and (ib) can be obtained by negating the alternative corresponding to each disjunct embedded under the universal quantifier: *every student took Syntax* and *every student took Logic.*

   Crnič, Chemla & Fox 2015, however, show that this derivation is problematic. In particular, they show that a sentence like (i) can be interpreted as giving rise to the distributive inferences even in contexts in which one of the two alternatives mentioned above is actually true (e.g., a context in which every student took Syntax). They then propose a different way of deriving distributive inferences, involving two exhaustifications and a stipulation about alternatives. Crucially, as they notice, the problem they identify does not extend to the case of modals.

the meaning in (12) and conjoining the sentence with negated excludable alternatives, as in (24), we predict the distributive effect (since (24) entails (25)).[9]

(23)
$$\left\{ \begin{array}{ll} \text{It is likely that we will hire Mary or Sue} & \triangle(\mathbf{M} \vee \mathbf{S}) \\ \text{It is likely that we will hire Mary} & \triangle(\mathbf{M}) \\ \text{It is likely that we will hire Sue} & \triangle(\mathbf{S}) \\ \text{It is likely that we will hire Mary and Sue} & \triangle(\mathbf{M} \wedge \mathbf{S}) \end{array} \right\}$$

(24) $\triangle(\mathbf{M} \vee \mathbf{S}) \wedge \neg\triangle(\mathbf{M}) \wedge \neg\triangle(\mathbf{S})$

(25) $\triangle(\mathbf{M} \vee \mathbf{S}) \wedge \Diamond(\mathbf{M}) \wedge \Diamond(\mathbf{S})$

In sum: we can derive the distributive inferences triggered by *certain* and *likely* in a straightforward way by adopting a standard account of scalar implicature.[10]

Before moving on, let us notice a potential problem concerning the effects triggered by *likely*. The account predicts that, whenever *likely (φ ∨ ψ)* triggers possibility inferences, it will also trigger the inference that *likely φ* and *likely ψ* are false. In fact, the derivation of the former inferences essentially relies on the latter. This may be problematic. Several reviewers have pointed out to us that they can distinguish an intermediate reading of *likely (φ ∨ ψ)* that triggers the relevant possibility inferences, but not the inferences that *likely φ* and *likely ψ* are false. Here is an example.[11] Suppose that I have removed some cards from a standard 52 card deck, and I ask you to draw a card at random. I add:

(26)   It is likely that you will draw a spade or a number.

(26) clearly triggers the inference in (27):

---

9 We are using here △ to represent the meaning of *likely* in (12).

10 Notice that a scalar implicature account of these inferences can also naturally account for why they tend to be absent in downward entailing contexts, given that this is a general property of scalar implicatures. One way to capture this property is assuming that the distribution of EXH obeys an economy condition along the lines of (i): that is, it is not added if it leads to an overall weaker meaning (this condition is a version of the Strongest Meaning Hypothesis; see Chierchia, Fox & Spector 2012 among others for discussion).

(i)   **Do not weaken!**: Do not insert EXH in *S* if the overall resulting meaning is weaker than ⟦S⟧.

11 This example is a refinement of one suggested by an anonymous referee.

(27)   a.   ⤳ *It is possible that you will draw a spade*
       b.   ⤳ *It is possible that you will draw a number*

While the judgment is subtle, it does not seem to trigger those in (28):

(28)   a.   ⤳ *It is not likely that you will draw a spade*
       b.   ⤳ *It is not likely that you will draw a number*

If this judgment is correct, this is an obvious problem for either standard theories of *likely*, or standard theories of distributive implicatures. One of the advantages we are going to claim for our proposal is that it improves on this prediction.[12]

3.2   The trouble with *possible*

Given that the account of distributive inferences sketched above covers both *certain* and *likely*, one might expect that it also covers the free choice inferences generated under *possible*. But this is not so if we stick to the classical semantics for *possible*.

Consider the counterpart of (18a) and (18b) involving *possible*. This sentence triggers the usual inferences in (29a) and (29b).

(29)   It is possible that we will hire Mary or Sue.

---

12 A referee and the editor point out an additional issue, distinct from the availability of what we've called 'intermediate reading'. They suggest that a sentence corresponding to the strengthened meaning sounds quasi-contradictory. Compare (i) to (ii), which involves *certain* and is unproblematic.

(i)   ?It's likely that we will hire Mary or Sue and it's not likely that we will hire Mary and it's not likely that we will hire Sue.

(ii)   It's certain that we will hire Mary or Sue and it's not certain that we will hire Mary and it's not certain that we will hire Sue.

We think the source of this contrast lies in the fact that *likely*, but not *certain*, is a neg-raising predicate (see Horn 1978, Gajewski 2007, Romoli 2013). This means that (i) is typically interpreted as in (iii). While (iii) is consistent, it is intuitively not easy to think of a situation that would make it true and this could be the reason why the sentence sounds slightly deviant.

(iii) ?It's likely that we will hire Mary or Sue and it's likely that we will not hire Mary and it's likely that we will not hire Sue.

      a.   ⤳ *It is possible that we will hire Mary*
      b.   ⤳ *It is possible that we will hire Sue*

Suppose we try to derive the free choice effect as a kind of implicature in the way we did above. Our alternatives are:

(30)
$$\left\{\begin{array}{ll} \text{It is possible that we will hire Mary or Sue} & \Diamond(\mathbf{M} \vee \mathbf{S}) \\ \text{It is possible that we will hire Mary} & \Diamond(\mathbf{M}) \\ \text{It is possible that we will hire Sue} & \Diamond(\mathbf{S}) \\ \text{It is possible that we will hire Mary and Sue} & \Diamond(\mathbf{M} \wedge \mathbf{S}) \end{array}\right\}$$

Here we get stuck. Differently from previous cases, the first two alternatives are *not* excludable, hence we cannot strengthen the assertion with their denial. If we deny $\Diamond(\mathbf{M})$, given the content of the assertion (i.e., $\Diamond(\mathbf{M} \vee \mathbf{S})$), $\Diamond(\mathbf{S})$ must be true. *Vice versa* if we deny $\Diamond(\mathbf{S})$. So the derivation of (29a) and (29b) is blocked.[13]

4 Overview of the proposal

One might wonder whether there is something wrong with our assumptions. Throughout §3, we presented a general kind of scalar reasoning that allowed us to derive distributivity implicatures under *certain* and *likely*. Why should that reasoning fail just for *possible*?

Why, indeed? In the next sections, we show that the simple scalar reasoning that we use to derive distributive inferences also gets us free choice, provided that we tweak the meaning of *possible*. On a standard, Kratzer-style semantics, *possible* is an existential quantifier. Among other things, this means that *possible* works as a scalar endpoint: *possible* is the weakest quantifier over epistemically possible worlds. As a result, when we deny *possible* ϕ we obtain a very strong claim, i.e. the claim that there are no ϕ-worlds. We suggest that we remedy the problem by introducing more structure into the semantics of *possible*. In particular, a degree-based treatment makes a difference to the alternatives generated by *possible*-sentences.

The next sections develop the proposal in detail, but here we give an overview. Following standard theories of gradable adjectives, we assume that *likely*, *certain* and *possible* work as measure functions. In particular, they all

---

13 To be sure, we still derive the implicature that it might not be that we hire both Mary and Sue, via the conjunctive alternative. But this is not enough to get the free choice effect: from $\neg\Diamond(\mathbf{M} \wedge \mathbf{S})$ it doesn't follow that $\Diamond(\mathbf{M})$ or $\Diamond(\mathbf{S})$.

map propositions to sets of degrees of probability. For example, this is the schematic meaning for *likely* ϕ:

(31) $[\![$likely $\phi]\!] = \lambda d. \, Pr([\![\phi]\!]) \geq d$

In addition, we are going to assume that gradable adjectives in the positive form combine with a covert morpheme, dubbed 'POS'. Roughly, POS sets a standard on the scale that the adjective operates on.

    For the case of *likely*, this standard is merely a degree of probability that is salient in the context. Hence *It is likely that* ϕ is true just in case the probability of ϕ is greater than the contextual standard for likelihood (which we denote as 's_{likely}').

(32) $[\![$It is POS likely that $\phi]\!] = 1$ iff $Pr([\![\phi]\!]) > s_{likely}$

For the case of *possible*, we assume that the standard is the minimum of the scale. As a result, *It is possible that* ϕ is true just in case there is some non-zero degree of probability such that the probability of ϕ is greater than that.[14]

(33) $[\![$It is POS possible that $\phi]\!] = 1$ iff $\exists d > 0$ such that $Pr([\![\phi]\!]) \geq d$

Our key observation is that we can derive possibility implicatures for all relevant cases by giving the exhaustivity operator intermediate scope between POS and the adjective. I.e., adopting a degree semantics for modal adjectives, the following configuration yields free choice:

(34)   POS [EXH [MODAL $(p \vee q)$]]

For a schematic example, consider:

(35) $[\![$POS [EXH [possible $(p \vee q)$]]$]\!] = 1$ iff $\exists d$ such that $d > 0$
                              and $Pr(p \vee q) \geq d$ and $Pr(p) < d$ and $Pr(q) < d$

    Informally, here is how we derive free choice from (35): if the probability of a disjunction is at least some degree d, and the probability of each disjunct is lower than d, then each disjunct must have positive probability. Given the

---

14 For some early predecessors of this semantics for *possible*, see Yalcin 2005, Lassiter 2010, as well as the close account in Swanson 2006.

meaning we're assuming for *possible*, this is equivalent to the claim that each disjunct is possible.[15]

The next three sections develop and defend the proposal in detail.

5 Background on degree semantics

5.1 Degrees, gradability, and scalarity

Let us start by clarifying what we mean by 'degree semantics.' We take a degree semantics for a category of lexical items I to be a semantics that represents the meanings of expressions in I by appealing to *degrees*, i.e. abstract entities that form a dense linear ordering, where the order is determined by a dimension (e.g., height, weight, length, etc). Following the custom, we call the orderings of degrees 'scales.'

On each of its two ends, a degree scale can be *closed* (when there is a first or a last degree on the scale) or *open* (when neither of the two previous conditions is met); see Kennedy & McNally 2005 for discussion. Different lexical items exploit different kinds of scales. In diagram form:

TOTALLY OPEN	○————————————○
LOWER CLOSED	●————————————○
UPPER CLOSED	○————————————●
TOTALLY CLOSED	●————————————●
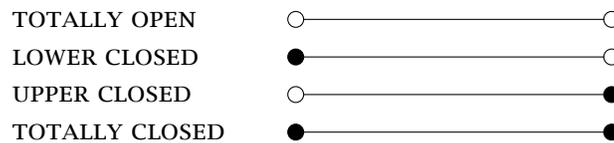
Items that have a degree semantics are generally *gradable*, i.e., they can be combined with a range of modifiers. Semantically, modifiers manipulate a degree threshold on the scale used by the adjective. The standard examples of gradable items are gradable adjectives, like *tall*, *open*, and *full*.

(36) Mary is very tall.

(37) The door is slightly open.

(38) The jar is almost full.

---

15 Note that this readily extends to disjunctions with more than two disjuncts, provided the assumption that each disjunct is an alternative to the whole disjunction.

(i) $[\![\text{POS } [\text{EXH } [\text{possible } (p \lor q \lor r)]]]\!] = 1$ iff $\exists d$ such that $d > 0$
and $Pr(p \lor q \lor r) \geq d$ and $Pr(p) < d$ and $Pr(q) < d$ and $Pr(r) < d$

The scale type associated with an adjective correlates with the availability or lack of availability of certain modifiers. For example, proportional modifiers like *partially*, *half*, or *60%* may only combine with closed scale adjectives, as (39) and (40) show.

(39)   #John is partially/half/60% tall.

(40)   The jar is partially/half/60% full.

For a comprehensive survey of the relationship between modifiers and scale structure, we refer the reader to Kennedy & McNally 2005 and Kennedy 2007.

5.2   Semantics of gradable adjectives: Degrees and POS

We assume a semantics for gradable adjectives in the style of Cresswell 1976, von Stechow 1984, Heim 1985, 2000 (among others). On this semantics, gradable adjectives denote functions from degrees and individuals to truth values (type $\langle d, et \rangle$).[16] As an example, here is the lexical entry for *tall*.[17]

(41) $[\![tall]\!] = \lambda d.\, \lambda x.\, [\text{HEIGHT}(x) \geq d]$

Informally, *tall* maps an individual and a degree to true just in case the individual's degree of height is equal to or greater than that degree.

Following the literature, we assume that the degree argument of the adjective is provided by a separate morpheme, which appears in a syntactic position labeled 'Deg(ree)P(hrase).' When the adjective appears in the positive form, we assume the presence of a covert morpheme 'POS' (for 'positive form'), which relates the degree argument of the adjective to a salient standard of comparison.

Following Heim 2000, we assume that all degree phrases, including POS, are of type $\langle dt, t \rangle$ and that syntactically they are generated as a sister of the adjective. This produces a type mismatch with adjectives (which are of type $\langle d, et \rangle$). As a result, DegP moves leaving a trace of type d and combines with

---

16 These functions are monotonic with respect to their degree argument, in the sense that, given a choice of *e*-type argument, if they yield truth when applied to a degree, they also yield truth when applied to all degrees that are lower on the scale. Formally (see also Heim 2000):

(i) $f_{\langle d, et \rangle}$ is monotone iff $\forall x \forall d \forall d'$ (if $f(d)(x) = 1$ and $d' < d$, then $f(d')(x) = 1$).

17 To remove clutter, we ignore index parameters wherever they're not needed.

the rest of the sentence after lambda abstraction.[18] For illustration: the LF of the simple sentence in (42a) after movement is in (42b).

(42)   a.   Mary is tall.
          b. $[_{\text{DEGP}}$ POS$]$ $[$ λd_1 $[_{\text{DP}}$ Mary is $[d_1$ tall $]]]$

As Heim points out, DegP-movement has to be severely constrained. In particular, it cannot outscope quantificational DPs and negation (Kennedy 1997, Heim 2000, Beck 2012, Romero 2015 a.o.). To see this, notice that, if POS were allowed to scope over the quantifier, as in (44), we would predict that (43) has the reading in (45) (where $\mathbf{s}_{tall}$ represents the contextual standard for tallness).

(43)   No student is tall.

(44) $[_{\text{DEGP}}$ POS$]$ λd_1 $[[_{DP}$ No student$]$ is $[_{\text{AP}}$ d_1 tall$]]$

(45) $\exists d[d > \mathbf{s}_{tall} \wedge \neg\exists x[\mathbf{student}(x) \wedge \mathbf{tall}(d)(x)]]$

(45) says that there is a degree of height d above the standard of tallness such that no student is d-tall. This requires merely that we can find a degree of height (say, 9 feet) such that no student is tall to that degree. Obviously (43) does not have this reading.

     From now on, we are going to assume that all syntactic constraints applying to gradable adjectives in general will carry over to the case of epistemic modal adjectives. We refer the reader to Heim's paper for extended discussion of these issues.

     Let us say more about POS. Following Kennedy & McNally 2005, Kennedy 2007 (see also Bochnak 2015), we assume that the denotation of POS involves a contextually provided function $\mathbf{R}$, which restricts the possible denotations of d depending on the clause that POS takes as one of its arguments.

(46) $[\![\text{POS}]\!] = \lambda G_{\langle d,t \rangle}. \exists d[\mathbf{R}(G)(d) \wedge G(d)]$

As an example, the truth conditions for a sentence like *Mary is tall* are in (47), '$\mathbf{s}_{tall}$' indicates the standard of comparison for *tall* in the context.

---

18   For an alternative perspective on degree phrases that doesn't use movement, see Alrenga & Kennedy 2014. We think our account could be made compatible with this alternative treatment, though we don't develop this variant here.

(47) $\exists d[\mathbf{R}(\lambda d'[\mathbf{tall}(d')(m)])(d) \wedge \lambda d'[\mathbf{tall}(d')(m)](d) =$
$\exists d[d > \mathbf{s}_{tall} \wedge \mathbf{tall}(d)(m)] =$
$\exists d[d > \mathbf{s}_{tall} \wedge \text{HEIGHT}(m) \geq d]$

The standard of comparison is fixed in different ways for different adjectives. Adjectives that exploit a totally open scale, like *tall*, are evaluated relative to a contextually given threshold (and are hence dubbed 'relative standard adjectives'). Adjectives that exploit closed scales are evaluated relative to the endpoints, and are hence dubbed 'minimum standard' or 'maximum standards.' Overall, the correlations between scale structure and adjective type are mapped in the following diagram:

SCALE TYPE	ADJECTIVE TYPE	EXAMPLE
open	relative standard	*tall*
lower closed	minimum standard	*dirty*
upper closed	maximum standard	*safe*
totally closed	minimum standard maximum standard min or max standard	*open* *closed* *transparent*

Notice that, for the case of adjectives exploiting totally closed scales, some adjectives are ambiguous between a minimum and a maximum standard reading, while some others invariably pick out one of the endpoints of the scale.

Discussing the correlation between scale structure and adjective type would take us too far from our main target. So we are simply going to assume that **R** takes a degree and an adjective and restricts the possible denotations of d on the adjective scale in the following ways: **R** applied to a degree d and a $\langle d, t \rangle$ property involving *possible* makes sure that d is higher than the bottom of the scale ($d > 0$), in the case of *certain*, d is restricted to be the maximum of the scale ($d = 1$), while with *likely* d is restricted to be larger than a contextually salient degree ($d > \mathbf{s}_{likely}$).[19]

---

19 Something has to be said about how POS distinguishes between adjectives that use the same scale, like *possible*, *certain* and *likely*. On the account we are assuming, these adjectives turn out to have the very same lexical entry; hence the difference in truth conditions they produce

Presumably, these facts should be explained, but providing this explanation is orthogonal to issues specific to modal adjectives.[20]

6 Degree semantics for epistemic modal adjectives

This section lays out our analysis of the epistemic modal adjectives *likely, certain* and *possible*. We start from a degree-based probabilistic semantics for *likely* (inspired by Yalcin 2010 and Lassiter 2011, 2014) and extend it, with some changes, to *certain* and *possible*. The idea of using a degree semantics for these modals is controversial, so we will spend substantial space defending our assumptions.

6.1 *Likely*

6.1.1 The arguments for a degree semantics

Existing arguments for a degree semantics for epistemic modal adjectives fall into two categories.[21] The first concerns the logical properties of these adjectives; the second concerns their ability to combine with scalar modifiers.

The item that has received greatest attention recently is *likely*.[22] Kratzer's semantics for *likely* in (12) predicts the validity of inference patterns that are obviously incorrect. For example, it predicts the validity of the following:

(48) ϕ is as likely as $\neg\phi \vDash \phi$ is as likely as ψ

---

has to be explained via differences in the value of the **R** parameter. Crucially, this is not a problem arising specifically with modal adjectives. Rather, it extends to all adjectives that exploit the same scale (e.g. *hot* and *warm*) and is part of the general problem of correlating scale structures and adjective types.

20 So far as we know, the only attempt at providing a systematic account is due to Kennedy 2007, who (building also on Kennedy & McNally 2005) hypothesizes that the correlations we described are generated by a principle of 'Interpretive Economy', which roughly dictates using the endpoints of a scale whenever they are available. We should note that even this account cannot explain in full the range of variability observed for closed scale adjectives.

21 For relevant work, see Swanson 2006, 2011, 2015, Yalcin 2007, 2010, 2012, Lassiter 2010, 2011, 2014, Moss 2015b,a. See also Cariani, Santorio & Wellwood 2016 for an up-to-date summary of the arguments for a probabilistic semantics for *likely*. For concreteness, we are going to adopt as our benchmark the approach in Yalcin 2010.

22 Arguments that the logical features of *likely* demand a probabilistic semantics are put forward by Yalcin (2010) as well as Lassiter (2011), though the formal work that undergirds them precedes them (see in particular Halpern 1997, 2003).

This is incorrect, since it entails that any proposition that is as likely as its negation is at least as likely as any other proposition (and hence, that it has probability 1). Conversely, a probabilistic semantics invalidates the problematic pattern.[23]

Arguments from compositional interactions focus on the fact that *likely* may combine with degree modifiers and may appear in comparatives.

(49)   It is very likely that Mary will hand in her paper in time.

(50)   It is more likely that Mary gets an A than that Sue gets an A.

In addition, *likely* also appears in combination with certain proportional modifiers, namely percentage modifiers of the form *n%*, as in (51):

(51)   It is 20/50/80/100% likely that it will snow.

The presence of adjectival modifiers is the hallmark of degree semantics. In particular, the fact that *likely* can combine with proportional modifiers as in (51) is evidence that the relevant scale is closed on both ends.

In sum, there appears to be a happy convergence between the two strands of evidence for a semantics for *likely* that exploits degrees of probability. A probability function maps each proposition in a given space to a real number lying in the closed interval [0,1]. Hence a probability function can be characterized as a measure function mapping propositions to a closed scale, exactly as the data about modifiers suggests. Thus the logical and the compositional properties of *likely* seem to be predicted, in one stroke, by a degree semantics based on probability.

We should note that, even if we accept a degree analysis, the claim that *likely* invokes a probabilistic scale is not uncontroversial. In particular, Klecha (2014) has argued that *likely* is gradable but exploits an open scale. Given that Klecha's semantics still retains some relevant logical properties[24], the choice between Klecha's semantics and a probabilistic one is irrelevant for us. So we leave an evaluation of Klecha's proposal to other work. For concreteness

---

23 Holliday & Icard (2013) point out that there are semantics that are based on a qualitative ordering of worlds (different from Kratzer's) that manage to capture the logical *desiderata* laid out by Yalcin. But these qualitative semantics suffer from other logical problems (as pointed out by Lassiter 2014; see also the discussion in Cariani, Santorio & Wellwood 2016).

24 In particular, that semantics still vindicates the following principle:

$$p \text{ is more likely than } \bot \vDash (p \lor q) \text{ is more likely than } q \text{ (given that } q \nvDash p)$$

(and because we think it is eventually empirically superior), in this paper we opt for a fully probabilistic semantics.

6.1.2 Semantics for *likely*

We assume that the denotation of *likely* is analogous to the denotation of standard gradable adjectives, modulo a different argument type (propositions rather than individuals), and a switch to a probability scale.[25] Following the semantics for *probably* in Yalcin 2010, we assume that semantic values are relativized to functions e from worlds to probability spaces (besides being relativized to worlds, as is customary.) Probability spaces are pairs $\langle E, Pr \rangle$ of a set of possible worlds E and a probability measure Pr. Officially, Pr takes as input subsets of E, but for simplicity we will assume that Pr directly maps each world in E to a numerical value in the interval [0,1] in accordance with the standard constraints on probability distributions.[26] Below is our semantics for *likely* (type $\langle d, \langle st, t \rangle \rangle$).

(52) $\quad [\![\text{likely } d \ \phi]\!]^{w,e} = 1$ iff $Pr_{e(w)}(\{w' : [\![\phi]\!]^{w',e} = 1\}) \geq d$

($Pr_{e(w)}$ is the function Pr in $e(w)$.) We will use the simplified entry in (53).

(53) $\quad [\![\text{likely}]\!] = \lambda d. \ \lambda p. \ [Pr(p) \geq d]$

Differently from existing analyses of *likely*, we assume that the meaning of *likely* does not involve the specification of a threshold on the scale. (This would make it difficult to account for the compositional interactions of *likely*.) Rather, we assume that, on a par with other gradable adjectives, the standard of comparison used by *likely* is fixed by a DegP item, and in particular by

---

25 A semantics that appeals to a notion of probability needs to explain how that notion is interpreted. (For an overview of the options, see Hájek 2012). For current purposes, we simply assume that this is a notion of subjective probability capturing the credences of the speaker. This assumption is simplistic and runs into a series of problems (including the arguments from retractions and disagreements put forward by MacFarlane 2011, and the puzzles generated by so-called epistemic contradictions discussed by Yalcin 2007). But the point is orthogonal to our main concerns in this paper.

26 Specifically, the constraints we need to impose are:

    (a) $Pr(E) = 1$

    (b) If p and q are disjoint sets of worlds, $Pr(p \cup q) = Pr(p) + Pr(q)$.

For simplicity, we also assume with Yalcin that the space of all possible worlds W is finite.

POS for the positive form. So the LF of (say) (54) is in (55), with POS moving out for the usual type-mismatch reasons.

(54)   It is likely that we will hire Mary.

(55)   [$_{\text{DEGP}}$ POS] [λd_1 [it is [d_1 likely] [that we will hire Mary]]]

As explained, we assume that the value $\mathbf{s}_{likely}$ is simply a degree of probability that stands out in the context.[27]

   The truth conditions resulting from these assumptions appear adequate: a sentence like (54) is true iff the probability that we will hire Mary is higher than the contextual threshold for *likely*.

(56) $[\![(54)]\!] = 1$ iff $\exists d[d > \mathbf{s}_{likely} \wedge Pr([\![\text{we will hire Mary}]\!]) \geq d]$

In the next subsection, we show how this analysis of *likely* can straightfor-wardly be extended to *certain*.

6.2   *Certain*

It is uncontroversial that *certain* is gradable: it may appear in comparative constructions, as in (57), and compose with degree modifiers, as in (58).

(57)   It is more certain that we will hire Mary than that we will hire Sue.

(58)   It is pretty certain that we will hire Mary.

We suggest that *certain* and *likely* work on scales that are overlapping, if not identical. Here are three pieces of evidence for this claim. First, *It is certain that p* works as a congruent answer to a question asking about the probability of p.[28]

(59)   a.   A: How likely is it that we will hire Mary?
       b.   B: It is certain that we will hire Mary.

Second, *certain* entails *likely*.

(60)   It is certain that we will hire Mary $\vDash$ It is likely that we will hire Mary.

---

27 This choice departs from several existing semantics for *likely* (e.g., Yalcin's official semantics in his 2010), which set this threshold at .5. But there is evidence that the threshold can be lower or higher than this, depending on context, as discussed by Yalcin (2010) himself.

28 We learned of examples of this kind from Klecha 2014. (Klecha uses a variant of (59) to argue that *likely* and *must* are not scalemates.)

Third, modified instances of *likely* may entail modified instances of *certain*. For example, *extremely likely* seems to entail *almost certain*, as witnessed by the awkwardness of (62).

(61)   It is extremely likely that we will hire Mary ⊨ It is almost certain that we will hire Mary.

(62)   ?? It is extremely likely, but not almost certain, that we will hire Mary.

These data motivate a degree semantics for *certain* that uses a scale that overlaps with that of *likely* — hence, a probabilistic scale.[29]

Extending our probabilistic semantics to *certain* is easy. The entry (in (63), simplified version in (64)), is analogous to that of *likely*. The only difference is that the meaning of *certain* dictates that the maximum degree on the scale 'stands out,' in Kennedy's sense.

(63) $[\![$ certain $d\ \phi]\!]^{w,e} = 1$ iff $Pr_{e(w)}(\{w' : [\![\phi]\!]^{w',e} = 1\}) \geq d$

(where $Pr_{e(w)}$ is the function Pr in $e(w)$)

(64) $[\![$ certain $]\!] = \lambda d.\ \lambda p.\ [Pr(p) \geq d]$

The resulting semantics yields adequate truth conditions for sentences involving *certain*: the probability of the prejacent of *certain* is 1. That is, (65) is true iff the probability that we will hire Mary is 1 (a very similar analysis is defended in Lassiter 2010, 2016 among others).

(65)   It is certain that we will hire Mary.

(66) $[\![(65)]\!] = 1$ iff $\exists d[d = 1 \wedge Pr([\![$ we will hire Mary $]\!]) \geq d]$

6.3 *Possible*

This section lays out our analysis of *possible*. From a formal point of view, this analysis is unproblematic and closely connected to our analysis of *likely*. But the analysis is empirically controversial; in particular, Klecha 2014 has argued extensively that *possible* is not gradable and that hence it doesn't have a degree semantics. Here we take up the empirical challenge. We are going to agree with Klecha that the evidence presented so far in the literature is

---

29 This idea is not problem-free. In particular, Klecha points out some disanalogies between *likely* and *certain* in comparative constructions and in certain downward entailing environment. We don't have the space to address these worries here.

unsatisfactory, but we are going to produce new data in favor of gradability from both English and Italian. We will conclude that a gradable analysis is not problem-free, but it is our best bet.

6.3.1 Skepticism about gradability

We start from Lassiter's (2010, 2011, 2016) defense of the idea that *possible* is gradable. Lassiter treats *possible* as a minimum standard adjective exploiting a closed scale. This places *possible* in a category of items that includes adjectives like *acquainted*, *protected*, and *documented* (Kennedy & McNally 2005); though, interestingly, Lassiter doesn't draw a direct close comparison with adjectives in this group. The data Lassiter uses to support this claim includes the following:

(67)　　It is slightly possible that the Jets will win.　　　　(Lassiter 2010)

(68)　　I felt that if it was 80-90 percent possible that [the cancer] hadn't spread, I didn't want the hysterectomy.　　(Lassiter 2011; web data)

(69)　　In fact, it is more possible that tomorrow is the zombie apocalypse than people magically floating away into the clouds.　　(Lassiter 2016; web data)

Lassiter's empirical claims have been criticized from a number of directions. A first, flat-footed challenge simply concerns sentences like (67)–(69). A number of informants judge that these sentences are not fully felicitous. In addition, Klecha (2014) provides both corpus and experimental evidence against Lassiter's gradability claim. The experimental evidence is particularly significant for us: Klecha tests the acceptability of *possible* in combination with a four modifiers — *very*, *pretty*, *too*, and the comparative form (*more than*). The finding is that modification negatively affects acceptability and it crucially does so for *possible* more than it does with *likely*.

We agree that these are substantial challenges. But we think that an investigation of other modifiers provides evidence that *possible* is gradable after all.

6.3.2 Evidence for the gradability of *possible*

To our knowledge, the debate about *possible* has focused on a relatively restricted set of modifiers: *very*, *slightly*, comparatives like *more than*, and

proportional modifiers like *60%*. But this choice seems arbitrary. We have independent knowledge that some of these modifiers do not combine with closed scale adjectives. Conversely, some modifiers that combine specifically with closed scale adjectives have so far been ignored.

Kennedy & McNally (2005) discuss a class of minimum standard adjectives that exploit a closed scale. This class includes the adjectives *aware*, *able*, *acquainted*, *documented*, *understood*, *publicized*. Kennedy and McNally notice that the default intensifier for these adjectives is *well*. Conversely, most of them may not combine with *slightly* and *very*, as the following sentences show.

(70)   ??John is slightly/very able to do his homework.

(71)   ??Sara is slightly/very acquainted with this theory.

(72)   ??The election is slightly/very documented/publicized.

On a probabilistic analysis, *possible* turns out to be exactly a minimum standard adjective operating on a closed scale. As a result, we expect it to pattern with (70)–(72). Hence it's not surprising that *slightly possible* and *very possible* are awkward. (Let us hasten to point out that the adjectives appearing in (70)–(72) *do* combine with proportional modifiers, hence if we had a full analogy we would expect, say, *60% possible* to be felicitous. More on this shortly.)

Conversely, we expect that *possible* combines with the kind of intensifiers that are acceptable with the adjectives in (70)–(72). In the next paragraphs, we suggest that this prediction is in part borne out, if we look at crosslinguistic data. In particular, we claim that: (i) *possible* combines with modifiers that are clearly scalar in meaning, and in fact select specifically for closed scales; (ii) the truth-conditional effects of *possible*-modification involve shifting a threshold along a scale that interacts with the scale used by *likely*. The emerging pattern of modification is still spotty, but on balance supports a degree analysis for *possible*.

The items we focus on are the Italian modifier *ampiamente* and the English modifiers *well* and *very well*. We argue that, on the most natural analysis, these modifiers work by shifting a degree parameter introduced by *possible*.

Italian *ampiamente*   Our first example is the Italian intensifier *ampiamente* (translatable as 'amply'). *Ampiamente possibile* is grammatical and intuitively conveys that the relevant event is both possible and also somewhat likely.

(73) E' ampiamente possibile che Samanta   verra'   alla   festa.
      is amply   possible that Samantha come.FUT to.the party
      'it is amply possible that Samantha will come to the party'

There is clear evidence that *ampiamente* is a genuine scalar modifier.[30] In particular, the licensing of *ampiamente* is sensitive to scale structure: *ampiamente* is infelicitous when combined with adjectives that operate on an open scale (like *alto/tall*), adjectives that operate on a lower-closed, upper-open scale (like *bagnato/wet*), and adjectives that operate on a lower-open, upper-closed scale (like *asciutto/dry*).

(74) ??Samanta   è ampiamente alta
        Samantha is amply   tall

(75) ??Il   tavolo è ampiamente bagnato
        The table   is amply   wet

(76) ??Il   bucato è ampiamente asciutto
        The laundry is amply   dry

By contrast, *ampiamente* is felicitous with adjectives (and, in some cases, locutions) that exploit fully closed scales. Below are some examples that directly translate the items in Kennedy and McNally's list.

(77) L'esistenza   di una corrispondenza fra   Churchill e
      The.existence of a   correspondence between Churchill and
      Mussolini è ampiamente documentata.
      Mussolini is amply   documented

---

30 The evidence that we present here sets *ampiamente* aside from catch-all modifiers like *quite* or the italian suffix *-issimo*, which combine with virtually all scalar items:

(i)   Sarah is quite tall.

(ii)   The counter is quite wet.

(iii) The laundry is quite dry.

(See the discussion in chapter 5 of Klecha 2014, as well as Beltrama & Bochnak 2015). As several researchers have pointed out, the fact that these modifiers are entirely insensitive to scale structure suggests a different analysis, on which they do not operate on a scale at all. For example, Klecha 2014 treats *quite* as a quantifier over contexts; *quite* works as an intensifier because it quantifies over a number of possible input values to the 'standard' parameter of POS.

(78) Questo fenomeno   astronomico è ampiamente compreso.
      This   phenomenon astronomical is amply   understood

(79) La mostra   è ampiamente pubblicizzata.
      The exhibition is amply   advertised

(80) Le autorità   sono ampiamente a conoscenza dei   fatti.
      The authorities are amply   acquainted   with.the facts

This distributional pattern virtually coincides with the pattern of the English intensifier *well*, studied by Kennedy & McNally 2005. The fact that its distribution correlates with a certain kind of scale structure is a typical hallmark of degree modifiers.

In addition, there is truth-conditional evidence that modification with *ampiamente* has the effect of shifting a threshold on a degree scale that is at least overlapping with that used by *probabile* (which translates *likely*). In particular, the use of *ampiamente possibile* suggests that the degree of likelihood of a proposition is not low. This is showed by the contrast between (81) and (82).

(81) È   possibile, ma molto improbabile, che i   Bulls arrivino ai
      It.is possible, but very   unlikely,   that the Bulls get   to.the
      playoffs
      playoffs.

(82) ??È   ampiamente possibile, ma molto improbabile, che i   Bulls
      It.is amply   possible, but very   unlikely,   that the Bulls
      arrivino ai   playoffs
      get   to.the playoff

This contrast survives when *possibile* appears in combination with modifiers different from *ampiamente*. For example, the following case shows that *del tutto possibile* (which roughly translates *entirely possible*) is compatible with low probability of the prejacent.

> *Scenario*: Giovanni is fond of buying lottery tickets, despite the fact that his chances of winning the lottery are very low. Maria is irritated by this behavior and comments: "What an idiot! It's obvious that he will never win".

In this scenario, (83) is a fine rebuttal to Maria, but (84) sounds contradictory.

(83)   È del tutto possibile che Giovanni vinca, anche se   è molto
        is entirely   possible that Giovanni win,   even   though is very
        improbabile.
        unlikely
        'It is entirely possible that Giovanni wins, even though it is very
        unlikely'

(84)   #È ampiamente possibile che Giovanni vinca, anche se   è molto
        is amply   possible that Giovanni win,   even   though is very
        improbabile.
        unlikely
        'It is amply possible that Giovanni wins, even though it is very unlikely'

Given these data, the natural suggestion is that *ampiamente* has a scalar
meaning, and that, on a par with modifiers like *very*, it works by shifting
upwards a threshold on a degree scale.

Back to English: *well, scarcely* and *impossible*   Can we find a counterpart
of Italian *ampiamente* in English? The natural candidate, especially given the
pattern individuated by Kennedy and McNally, is *well* itself. And indeed sev-
eral of our informants (in particular, native speakers of British English from
some areas of Northern England and Northern Ireland) find modifications of
possible with *well* or *very well*, as in (85), grammatical:

(85)   It is (very) well possible that it will rain.

Moreover, *well* and *very well* are very commonly used to modify epistemic
might:

(86)   It might well be that it rains.

(87)   Mary might very well bring her girlfriend to the party.

There are two other suggestive pieces of evidence that English *possible* has
a scalar semantics. The first is that *possible* combines, at least for some
speakers, with another modifier that selects for fully closed scales, namely
scarcely.

(88)   It is scarcely possible that Rubio will win at this point in the race.

The licensing pattern of *scarcely* follows closely that of *well*, and of *ampiamente* in Italian. On the one hand, *scarcely* is infelicitous with all adjectives that don't employ a closed scale.

(89) ??Samantha is scarcely tall.

(90) ??The table is scarcely wet/dry.

On the other, *scarcely* seem to be licensed with all the adjectives in the Kennedy and McNally closed scale list.

(91)   Homer's life is scarcely documented.

(92)   Earthquakes are still scarcely understood.

Therefore *scarcely* appears to be a genuine scalar minimizer which combines with *possible*.[31]

Second, we observe that *possible* has a gradable antonym, which can be formed by *in-* prefixation, namely *impossible*. *Impossible* picks out a maximum standard on a scale, as the following data show:

(93)   It is almost impossible that Rubio will win the nomination.

(94)   It's completely impossible that we will hire a semanticist this year.

On the plausible assumption that *in-* prefixation triggers polarity reversal, this suggests that *possible* also has a scalar meaning.

Spotty gradability   In summary: *possible* may combine with genuine scalar modifiers and this modification has the truth-conditional effects we expect on a scalar semantics. At the same time, *possible* is somewhat awkward with a number of other modifiers, including comparatives and proportional modifiers. So there seems to be gradability, although in an uncharacteristically spotty way.

We think that, in this predicament, the best solution is adopting a genuine degree semantics. A semantics that is not based on degrees has two major disadvantages. First, it must treat scalar modifiers like *well* and *ampiamente* as ambiguous. These modifiers would have a scalar meaning, which would be the one used in compounds with genuinely gradable adjectives like *documented/documentato*, and a nonscalar one, which would be the one used in

---

31 Although it should be noticed that *scarcely* can apply also to verbs, like in 'John scarcely ran.'

combination with *possible*. This is obviously a stipulation. Second, it's just unclear how a non-degree semantics can account for the truth-conditional effects of modification by *well* and *ampiamente*. In particular, it's unclear how to account for the fact that *well possible φ* and *ampiamente possibile φ* entail that φ is somewhat likely in any straightforward way.

Conversely, we think that a theory that treats *possible* as using degrees will have the resources to explain the spotty pattern we observe. Here we tentatively follow Lassiter 2016 in hypothesizing that there is a constraint dictating a preference for relative standard adjectives over minimum standard adjectives, whenever the two can be used to express the same meaning. On this view, *more possible* is blocked by the competition with *more likely*; similarly *40/50/60% possible* would be blocked by *40/50/60% likely*.

Finally, let us flag another route that deserves investigation. We may revise some of our background assumptions about scale structure. In particular, it might be that the degree scale exploited by *possible* has a different structure from the one we assume — for example, it might not be a total ordering. While this would be incompatible with a fully probabilistic semantics, it would be in principle compatible with our account of the scalar inferences triggered by *possible*. This said, for current purposes we set this hypothesis aside and proceed with a probabilistic analysis.

6.3.3 Semantics for *possible*

We use the semantics in (95) (simplified version in (96)). In this analysis, *possible*, exactly like *likely* and *certain*, has type $\langle d, \langle st, t \rangle \rangle$. As usual, the only difference with respect to *likely* and *certain* is that *possible* makes salient the lowest degree on the probability scale. Composition with POS works in the usual way.

(95) $[\![\text{possible } d \ \phi]\!]^{w,e} = 1$ iff $Pr_{e(w)}(\{w' : [\![\phi]\!]^{w',e} = 1\}) \geq d$

(where $Pr_{e(w)}$ is the function Pr in $e(w)$)

(96) $[\![\text{possible}]\!] = \lambda d. \lambda p. [Pr(p) \geq d]$

As a result, a sentence like (97) has the truth-conditions in (98): it is true just in case the probability of us hiring Mary is non-zero.

(97)   It is possible that we will hire Mary.

(98) $\exists d [d > 0 \wedge Pr([\![\text{we will hire Mary}]\!]) \geq d]$

6.4   Summary

We have argued that *certain, likely* and *possible* are gradable adjectives. We model their semantics on the semantics of, respectively, maximum, relative, and minimum standard adjectives. They all operate on a closed scale, which is a probability scale. We also assume that the positive form of *likely, certain*, and *possible* involves a covert morpheme POS, which sets the standard of comparison using either information contained in lexical entries or (when the latter is not available) contextual information.

7   A unified account of possibility implicatures

In this section, we first show how our account replicates the results of the standard analysis, straightforwardly predicting the distributive inferences triggered by *likely* and *certain*. We then show how, unlike the standard analysis, our account naturally generalizes to the free choice inferences triggered by *possible*.

7.1   Possibility implicatures with *likely* and *certain*

To start, consider a sentence with *likely* whose prejacent includes a disjunction like (99). Suppose that we exhaustify it as in (99), via an exhaustivity operator taking scope over the whole sentence.

(99)   It's likely that we will hire Mary or Sue.

(100)   EXH[POS [λd_1 [it is [d_1 likely] [that we will hire Mary or Sue]]]]

Assuming for concreteness that the standard of *likely* in the context is set to .5, the alternatives to (100) are in (101).

(101)
$$
\left\{
\begin{array}{ll}
\text{It is POS likely \ldots Mary or Sue} & \exists d[d > .5 \wedge Pr(\mathbf{M} \vee \mathbf{S}) \geq d] \\
\text{It is POS likely \ldots Mary} & \exists d[d > .5 \wedge Pr(\mathbf{M}) \geq d] \\
\text{It is POS likely \ldots Sue} & \exists d[d > .5 \wedge Pr(\mathbf{S}) \geq d] \\
\text{It is POS likely \ldots Mary and Sue} & \exists d[d > .5 \wedge Pr(\mathbf{M} \wedge \mathbf{S}) \geq d]
\end{array}
\right\}
$$

It is easy to see that all the alternatives are excludable (e.g., the negation of *it's likely that we will hire Mary* does not entail that it's likely that we will hire Sue). By adding their negation to the sentence, we get (102).[32]

(102) It is likely that we will hire Mary or Sue and it is not likely that we will hire Mary and it is not likely that we will hire Sue.

Given our semantics for *likely*, (102) is equivalent to the schematic (103), which we can simplify as in (104).

(103) $\exists d[d > .5 \wedge Pr(\mathbf{M} \vee \mathbf{S}) \geq d] \wedge \neg \exists d[d > .5 \wedge Pr(\mathbf{M}) \geq d] \wedge$
$$\neg \exists d[d > .5 \wedge Pr(\mathbf{S}) \geq d]$$

(104) $Pr(\mathbf{M} \vee \mathbf{S}) > .5 \wedge Pr(\mathbf{M}) \leq .5 \wedge Pr(\mathbf{S}) \leq .5$

From (104) we can infer that the probability that we will hire Mary is non-zero and the probability that we will hire Sue is non-zero.

(105) It is likely that we will hire Mary or Sue.
    a. ⤳ *The probability that we will hire Mary is non-zero*
    b. ⤳ *The probability that we will hire Sue is non-zero*

But now, recall out semantics for *possible*: *possible* ϕ is true iff there is a degree of probability d such that the probability of ϕ is greater than d. This means that (105a) and (105b) entail, respectively:

(106) a. *It's possible that we will hire Mary*
    b. *It's possible that we will hire Sue*

Hence we predict that *likely* ($\phi \vee \psi$) generates standard possibility implicatures.

The derivation that we just illustrated can be replicated straightforwardly when we replace *likely* with other epistemic modal expressions. First, it can be replicated with *certain*. By following analogous steps, we derive (107), from which it follows again that the probability of each disjunct has to be non-zero, and therefore that both disjuncts are possible.

(107) $Pr(\mathbf{M} \vee \mathbf{S}) = 1 \wedge Pr(\mathbf{M}) < 1 \wedge Pr(\mathbf{S}) < 1$

---

32 We are omitting the negation of the conjunctive alternative which is entailed by the negation of the other two alternatives.

Second, it can be replicated when *likely* appears in combination with various modifiers, as in the following cases.[33]

(108)   It is very likely that we will hire Mary or Sue
            It is 60% likely that we will hire Mary or Sue
            It is 8% likely that we will hire Mary or Sue
            It is .00000001% likely that we will hire Mary or Sue
            . . .

In all these cases, the only difference is the probability threshold that is involved in the meaning of the relevant modal adjective. But this doesn't affect the computation of implicatures; hence, in all these cases, we are able to derive probabilistic inferences, exactly as we did for *likely*. We leave it to the reader to work out the details.

7.2   Giving EXH intermediate scope

So far, we have showed how a degree semantics can mimic the predictions of a Kratzer-style approach. Let us now start showing how we get divergent predictions. Recall the judgment we elicited at the end of §3: (26) seems to trigger the inferences in (27), but not those in (28).

(26)   It is likely that you will draw a spade or a number.

(27)   a.   ⤳ *It is possible that you will draw a spade*
            b.   ⤳ *It is possible that you will draw a number*

(28)   a.   ⤳ *It is not likely that you will draw a spade*
            b.   ⤳ *It is not likely that you will draw a number*

Yet a Kratzer-style semantics, together with the view of implicature we're adopting throughout the paper, is unable to predict this. Conversely, we can predict this pattern in the framework that we have developed. The key maneuver is letting EXH take intermediate scope between POS and *likely*.

Suppose that, rather than parsing (99) as (100), we parse it as follows, letting EXH take scope below the λ-abstractor over degrees:

(109)   POS [λd_1 [EXH [it is [d_1 likely] [that we will hire Mary or Sue]]]]

---

33 Here we make the (very natural) assumption that the semantic contribution of the modifiers is to shift the threshold appearing in the entry of *likely*.

Let us sketch the derivation. First, the alternatives that EXH uses are in (110) (ignoring 'it is' here). Notice that, since POS is not in the scope of EXH, the alternatives involve a free variable d.[34]

(110)
$$\left\{ \begin{array}{ll} d \text{ likely} \dots \text{Mary or Sue} & Pr(\mathbf{M} \vee \mathbf{S}) \geq d \\ d \text{ likely} \dots \text{Mary} & Pr(\mathbf{M}) \geq d \\ d \text{ likely} \dots \text{Sue} & Pr(\mathbf{S}) \geq d \\ d \text{ likely} \dots \text{Mary and Sue} & Pr(\mathbf{M} \wedge \mathbf{S}) \geq d \end{array} \right\}$$

Once more, all alternatives are excludable. The outcome of the computation is (111), where d is a free variable at this stage of the derivation. Assuming again that the relevant contextual standard is .5, the resulting truth conditions are in (112).

(111) $Pr(\mathbf{M} \vee \mathbf{S}) \geq d \wedge Pr(\mathbf{M}) < d \wedge Pr(\mathbf{S}) < d$

(112) $\exists d[d > .5 \wedge Pr(\mathbf{M} \vee \mathbf{S}) \geq d \wedge Pr(\mathbf{M}) < d \wedge Pr(\mathbf{S}) < d]$

Notice that the truth conditions in (112) are crucially weaker than the ones we got by letting EXH take wide scope (in (104)). The latter truth conditions ruled out that the probabilities of the individual disjuncts could be higher than the contextual threshold. The truth conditions in (112), conversely, only

---

34 We are assuming that the value of the free variable d is fixed by the context at the relevant stage of the computation and that, once fixed, it remains constant across alternatives. Notice that the assumption that EXH takes scope over clauses involving free variables is independently needed to account for some cases of embedded implicature. For example, consider (i), on the reading in which it is equivalent to (ii).

(i)   a.   Every student did some of her homework.
       b.   Every student did some of her homework and no student did all of her homework.

To account for the implicatures of (109), we need to assume that EXH takes scope below the DP *every student* (see Chemla & Spector 2011 for discussion of analogous examples).

(ii)   Every student [λx_1 [EXH [x_1 did some of x_1's homework]]]

Let us also consider a potential objection. One might worry about considering more alternatives for the sentence in (105) — schematically, all the ones meaning that $Pr(p) > d'$, for all degrees d' such that $d' > d$. This is not a problem, however. Even if we allowed these alternatives, we would conclude that the probability of p is greater than or equal to d but not equal to any d' greater than d: that is, we would conclude that the probability of p is exactly d (where d is above the standard for **likely**). See Fox & Hackl 2006 for discussion of a similar situation with the (absence of) implicatures in sentences like *John weighs more that 120 pounds*.

require (i) that the disjunction **M** ∨ **S** has a degree of probability d higher than the contextual threshold for *likely*, and (ii) that the probability of both individual disjuncts is lower than d. They leave open that the probability of the individual disjuncts are intermediate between d and the contextual standard, as in the following diagram:

$$\begin{array}{ccc} & \textbf{S} \ \ \textbf{M} & \textbf{M} \vee \textbf{S} \\ \vdash\!\!\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!\!\vdash\!\!\!\vdash\!\!\!-\!\!\!-\!\!\!-\!\!\!\vdash\!\!\!-\!\!\!-\!\!\!-\!\!\!\dashv \\ \mathbf{s}_{likely} & & d \end{array}$$

Hence a degree semantics for *likely* manages to predict that *likely (φ ∨ ψ)* triggers possibility inferences without also triggering the inferences that *likely φ* and *likely ψ* are false. The key assumption that we used to get this result is that POS can outscope EXH. This same assumption will do substantial work in the derivation of the possibility inferences of *possible*.[35]

7.3 Predicting free choice under *possible*

Let us start by showing that, by letting EXH take wide scope with respect to POS, we run into the same problems as standard modal semantics. Consider (113), and assume that it is parsed as in (114).

(113)   It is possible that we will hire Mary or Sue.

(114)   EXH[POS [λd_1 [it is [d_1 possible] [that we will hire Mary or Sue]]]]

Schematically, the alternatives to (114) are as in (115).

(115)
$$\left\{ \begin{array}{ll} \text{POS } \lambda d \ [d \text{ possible } \ldots \text{ Mary or Sue]} & \exists d[d > 0 \wedge Pr(\mathbf{M} \vee \mathbf{S}) \geq d] \\ \text{POS } \lambda d \ [d \text{ possible } \ldots \text{ Mary]} & \exists d[d > 0 \wedge Pr(\mathbf{M}) \geq d] \\ \text{POS } \lambda d \ [d \text{ possible } \ldots \text{ Sue]} & \exists d[d > 0 \wedge Pr(\mathbf{S}) \geq d] \\ \text{POS } \lambda d \ [d \text{ possible } \ldots \text{ Mary and Sue]} & \exists d[d > 0 \wedge Pr(\mathbf{M} \wedge \mathbf{S}) \geq d] \end{array} \right\}$$

---

35 Notice that this assumption is in line with the syntactic constraints in §5. We noticed that, while there are constraints on the movement of items in DegP position (which include POS), those constraints are limited to certain categories of items, like DPs and negation. Moreover, even this relatively uncontroversial assumption will become unnecessary on slightly different accounts of local implicatures. For example, a theory in the style of Chierchia 2004, on which exhaustified meanings are computed in parallel with basic meanings and carried along with them in the compositional computation, we don't need any syntactic assumptions at all.

It is easy to see that the alternatives involving the two disjuncts are not excludable. If we negate (say) *it's* POS *possible that we'll hire Mary*, we would get that the probability of us hiring Mary is zero, which in turn entails that *it's* POS *possible that we'll hire Sue* is true. As a result, no possibility inferences are computed if we parse the sentence as in (114).[36]

Now suppose that (113) is parsed as in (116), with POS taking scope over EXH. The alternatives in this case are in (117).

(116)   POS [λd_1 [EXH [it is [d_1 possible] [that we will hire Mary or Sue]]]]

(117)
$$
\left\{
\begin{array}{ll}
\text{it is } d \text{ possible } \ldots \text{Mary or Sue} & Pr(\mathbf{M} \vee \mathbf{S}) \geq d] \\
\text{it is } d \text{ possible } \ldots \text{Mary} & Pr(\mathbf{M}) \geq d] \\
\text{it is } d \text{ possible } \ldots \text{Sue} & Pr(\mathbf{S}) \geq d] \\
\text{it is } d \text{ possible } \ldots \text{Mary and Sue} & Pr(\mathbf{M} \wedge \mathbf{S}) \geq d]
\end{array}
\right\}
$$

Crucially, on the new meaning all the alternatives are excludable. To see this, notice that the negation of $Pr(\mathbf{M}) \geq d$, i.e. $Pr(\mathbf{M}) < d$, together with the content of the assertion, i.e. $Pr(\mathbf{M} \vee \mathbf{S}) \geq d$, does *not* entail that $Pr(\mathbf{S}) \geq d$. So we can proceed with exhaustification. Schematically, this is the strengthened meaning. (Again, we omit the negation of the conjunctive alternative, since it is entailed by the negation of the other two.)

(118) $Pr(\mathbf{M} \vee \mathbf{S}) \geq d \wedge Pr(\mathbf{M}) < d \wedge Pr(\mathbf{S}) < d$

(118) is then fed to POS; the resulting meaning is in (119).

(119) $[\![$ POS [λd_1 [EXH [it is [d_1 possible] [that we will hire Mary or Sue]]]] $]\!]$ =
$\exists d[d > 0 \wedge Pr(\mathbf{M} \vee \mathbf{S}) \geq d \wedge Pr(\mathbf{M}) < d \wedge Pr(\mathbf{S}) < d]$

(119) says that the probability of the disjunction is greater than d, and that the probability of each of its disjuncts is less than or equal to d. It follows that the probability of each disjunct must be greater than zero. In other words, (119) entails (120):

(120) $Pr(\mathbf{M}) > 0 \ \wedge \ Pr(\mathbf{S}) > 0$

Given our semantics for possibility claims, the relevant possibility inferences follow from (120):

(121)   a.   ⤳ *It is possible that we will hire Mary*

---

36 Though the conjunctive alternative is innocently excludable, so (113), on the parse in (114), does give rise to a run-of-the-mill scalar implicature that we will not hire both Mary and Sue.

    b.   ⤳ *It is possible that we will hire Sue*

Hence we predict free choice, using exactly the same mechanics we used to predict other possibility inferences.

8   Extending the account

8.1   Extension to epistemic modals in other syntactic categories

Given the degree semantics adopted above, we can account for all the possibility inferences generated by epistemic modal adjectives (recapitulated below) via a uniform mechanism.

(122)   a.   It's possible that we will hire Mary or Sue.
           b.   It's likely that we will hire Mary or Sue.
           c.   It's certain that we will hire Mary or Sue.

(123)   a.   ⤳ *It's possible that we will hire Mary*
           b.   ⤳ *It's possible that we will hire Sue*

We have illustrated our analysis using the epistemic modal adjectives *possible, likely* and *certain*. But, in principle, what we say can be generalized to modals in other syntactic categories — including modal auxiliaries like *might, should* and *must*. For instance, we may adopt a semantics for *might* that is exactly analogous to that of *possible*, with a degree variable, which is then existentially quantified over.

(124) $[\![\text{might } d\ \phi]\!]^{w,e} = 1$ iff $Pr_{e(w)}(\{w' : [\![\phi]\!]^{w',e} = 1\}) \geq d$

    Whether we can run the same account of possibility implicatures will depend on the compositional details. In particular, it will depend on whether we can compositionally separate the degree element from the position in which existential closure would happen, and insert an exhaustivity operator in between. In part, this question depends on whether modal auxiliaries should be treated as gradable. Pursuing this question in full is beyond the scope of this paper. Let us just notice again that, at first sight, epistemic modal auxiliaries are perfectly felicitous when modified by *well* and *very well*.

(125)   It might very well be that we will hire Sue.

8.2 Extension to deontic modals

So far, we have focused on free choice inferences arising from epistemic modals. But possibility inferences are generated also by other modals, in particular deontic modals — in fact, the problem of free choice was first pointed out just in connection with the latter (von Wright 1968, Kamp 1973). Below are some examples (see also (6) and (7) in §1).

(126) You are allowed to take Syntax or Logic.

    a. ⤳ *You are allowed to take Syntax*

    b. ⤳ *You are allowed to take Logic*

(127) You are required to take Syntax or Logic.

    a. ⤳ *You are allowed to take Syntax*

    b. ⤳ *You are allowed to take Logic*

The argument that we're running in this paper is, in part, an argument from generality: our account deserves consideration because it is able to predict the full array of data in a unified way. So it's important that we manage to extend the account to deontic modals as well. Can we do this?

From a technical point of view, there is no difficulty. We can simply treat *allowed*, *required*, and the like as degree expressions, mapping propositions to a degree scale that has the same formal properties as a probability scale. This will allow us to derive probability inferences via exhaustification exactly as in the case of epistemic modals. At the same time, even granting that deontic modals have a degree semantics, it's unclear that the scale that they exploit has the right formal properties.

To illustrate the problem, take a naïve idea: deontic modal expressions map propositions to a scale of permissibility (set aside, for the moment, exactly what this notion of permissibility amounts to). At an intuitive level, it doesn't seem that the degree of permissibility of a disjunction like (126) is greater than or equal to the degree of permissibility of one of its disjuncts, like (126a) and (126b).[37] In fact, it's not even clear what could be meant by saying that a disjunction is 'more permissible' than its disjuncts. Yet

---

37 Thanks to Paul Portner (p.c.) for suggesting to us this way of putting the point.

this property is exactly what we need to derive possibility inferences via exhaustification.[38]

In this section, we discuss the prospects for formulating a degree semantics of the right kind for deontic modals. Given that our main focus is elsewhere, our discussion will have two significant limits. First, we won't be able to present empirical evidence in support of a degree semantics for these modals. For evidence of this sort (including the observation that the deontic modal auxiliary *should* can appear in the comparative) see, among others, Portner & Rubinstein 2016. Second, since there is no ready-made semantics we can use, our proposal will be just an outline. Nevertheless, what we say should be enough to show that there are good prospects for generalizing our account to the deontic case.

8.2.1 Deontic modals as measures of expected utility?

One natural attempt at a degree semantics for deontic modals exploits a scale of *expected utility*.[39] Roughly, expected utility is a measure of how valuable a certain proposition is for an agent, given what else they know about the world. And in fact, an expected utility semantics has been proposed for deontic modals (see Goble 1996 and Lassiter 2011; see also Aloni 2005 for related work on imperatives). But this kind of semantics, at least without substantial tweaks, is a nonstarter for our purposes. Let us explain why.

The expected utility (henceforth, EU) of a proposition for an agent is defined on the basis of their credence function (represented, as usual, as an assignment of probabilities to propositions) and their utility function. Roughly, you can think of utility as a measure of how valuable for the agent a certain state of the world is. Since we won't be using EU in our positive proposal for deontic modals, we are not going to define explicitly EU here; rather, we refer the reader to any of the classical discussions or overviews.[40] Any mainstream way of defining EU will do for our purposes. What matters

---

38 In general, this property is a kind of additivity property. Probability measures satisfy a finite additivity property, but even a weaker additivity property will do; for discussion, see Holliday & Icard 2013.

39 In a sense, any semantics for deontic modals that exploits a notion of comparative closeness (starting from Lewis's semantics in Lewis 1973) counts as a degree semantics. But here we intend to pick out degree semantics that have a kind of quantitative structure, going beyond orderings.

40 For a classical account of expected utility theory, see Jeffrey 1983; for a more recent and very popular version of decision theory see Joyce 1999. See Briggs 2015 for a recent overview.

to us is that (as we're going to point out) EU accounts yield predictions that are problematic for a theory of the scalar features of modals.

Let us sketch two EU-based semantics for deontic modals. We are going to use *required* as our sample item. Since this example just serves the purposes of illustrating the problem for EU theories, we won't be concerned with the distinction between weak and strong necessity modals.[41]

In its general form, an EU semantics for *required* analyzes *required* ϕ as saying that the EU of ϕ is higher than a certain threshold value. Formally (and introducing a utility function parameter u besides familiar ones):

(128) $[\![\text{required } \phi]\!]^{w,e,u} = 1$ iff $EU_{e(w),u(w)}(\{w' : [\![\phi]\!]^{w',e,u} = 1\}) > t_{ALT}$

t_{ALT} is a 'threshold' value that is set in relation to a set of alternatives specifying mutually exclusive and jointly exhaustive courses of action. Different analyses diverge on what alternatives should be used. One simple option is that the only relevant alternative is the negation of the prejacent (as in the first account in Goble 1996; a similar idea, though not in the context of an EU semantics, is used by Jackson 1985, Jackson & Pargetter 1986). On this account, the schematic analysis in (128) becomes:

(129) $[\![\text{required } \phi]\!]^{w,e,u} = 1$ iff $EU_{e(w),u(w)}(\{w' : [\![\phi]\!]^{w',e,u} = 1\}) >$
$\qquad\qquad\qquad\qquad\qquad\qquad EU_{e(w),u(w)}(\{w' : [\![\neg\phi]\!]^{w',e,u} = 1\})$

I.e., *required* ϕ is true iff the expected utility of ϕ is greater than the expected utility of $\neg\phi$. A second option is to take the relevant alternatives to be a contextually supplied set of mutually exclusive and jointly exhaustive propositions, which specify alternatives courses of action (as in the second proposal in Goble 1996, and in the vicinity of the proposal in Lassiter 2011). On this proposal, (128) is specified as:

---

41 As a reminder: it is widely acknowledged (see e.g. von Fintel & Iatridou 2008) that some deontic modals are stronger than others. For example, *required* and *have to* (which are so-called strong necessity modals) are stronger than *ought* (a so-called weak necessity modal). The following contrast illustrates the difference.

(i)   You ought to wash the dishes. In fact you're required to do so/you have to.

(ii)   #You're required to/you have to wash the dishes. In fact, you ought to do so.

(130) $[\![$required $\phi]\!]^{w,e,u} = 1$ iff $\{w : [\![\phi]\!]^{w,e} = 1\} = \bigcup BEST_{EU_{e(w),u(w)}}(ALT)$
(where $BEST_{EU_{e(w),u(w)}}(ALT)$ is the set of members of ALT that maximize expected utility, relative to $e(w)$ and $u(w)$)

Informally: *required ϕ* is true iff ϕ coincides with the union of the alternatives that maximize expected utility. The choice between different versions of (128) doesn't matter for our purposes, so we don't pursue the issue further.

Expected utility accounts of deontic modality (henceforth, "EU accounts") have much to recommend them. But they also have drawbacks, some of which have been pointed out in existing literature (see, e.g., Cariani 2016a,b). For reasons of space, here we limit ourselves to pointing out two problems that relate closely to the scalar properties of modals.

The first problem is very simple and very general. The problem is that EU accounts yield systematically wrong predictions when deontic modals appear embedded in DE environment. (131) entails the claims in (131a) and (131b):[42]

(131) Sally is not required to discuss her paper with John or Mary.

   a. ⤳ *Sally is not required to discuss her paper with John*
   b. ⤳ *Sally is not required to discuss her paper with Mary*

To better see the empirical point, notice that (132) sounds inconsistent.

(132) #Sally is not required to discuss her paper with John or Mary. But she is required to discuss it with Mary.

These facts are fully expected on classical analyses for *required*. These analyses are upward monotonic in the prejacent position, and hence validate the entailment *required $\phi \vDash$ required (ϕ or ψ)*. Since downward entailing contexts reverse the direction of entailment, these analyses predict that (131) entails (131a) and (131b). But all versions of (128) are nonmonotonic, and hence fail to validate the relevant entailment. As a result, EU accounts miss the entailment in (131), and predict that the discourse in (132) is consistent.

The second problem is less general — since it presupposes a Sauerland/Fox-style account of distributivity inferences — but still substantial. We saw that,

---

42 See also von Fintel 2012 for discussion and see Cariani 2013 for an alleged counterexample. While we feel the pull of Cariani's example, we hasten to point out that isolated counterexamples to the entailment we discuss can be easily accommodate via embedded implicatures. The key point is that EU accounts predicts that entailments like that in (131) should fail systematically, since the semantics of *required* is nonmonotonic. We think that the overall data clearly speaks against this prediction.

on EU accounts, *required φ* is not in general stronger than *required* (*φ* or *ψ*). This is sometimes claimed to be an advantage of EU accounts (e.g., by Lassiter 2011).[43] But it creates problems for the computation of implicatures, since it blocks the derivation of the distributive inferences in (127a) and (127b).

(127)   You are required to take Syntax or Logic .

    a.   ⤳ *You are allowed to take Syntax*
    b.   ⤳ *You are allowed to take Logic*

Recall: the inferences in (127) are derived by (i) assuming that (133a) and (133b) (below) express stronger propositions than (127), and (ii) assuming that implicatures are computed via exhaustification of stronger alternatives that are excludable.

(133)   a.   You are required to take Syntax.
      b.   You are required to take Logic.

On EU accounts, the computation stops because step (i) is blocked. On the one hand, this seems a drawback of EU accounts. On the other, it makes these accounts look like nonstarters for our purposes. If it can't get us even standard distributivity inferences, *a fortiori* this analysis won't get us free choice.

8.2.2   A probabilistic semantics for deontic modals

We suggest a minimal variant of classical ordering semantics for deontic modals. (See Kratzer 1986, 2012, as well as Lewis 1973.) Recall a schematic Kratzer-style analysis of deontic *required*:

(134) $[\![\text{required } \phi]\!]^{w,f,g} = 1$ iff $\forall w' \in \text{BEST}_{g(w)}(f(w))$, ϕ is true at w'

We suggest that we move to a degree semantics simply by replacing the quantificational element of (134) with a mapping to a probability scale. According to (134), *required φ* is true iff the set of best worlds entails *φ*. On our

---

43 One of the reasons is that it blocks problematic-sounding inferences like the one from (i)-a to (i)-b, which are validated by quantificational semantics. (This is known in classical literature as 'Ross's puzzle,' after Ross 1944.)

(i)   a.   You are required to post this letter.
    b.   You are required to post this letter or burn it.

proposal, *required* ϕ is true iff the probability of ϕ, conditional on the world being one of the best worlds, is 1.[44] Here is a schematic entry (for simplicity, we incorporate the compositional contribution of POS directly in the truth conditions).

(135) $[\![$POS required $\phi]\!]^{w,f,g} = 1$ iff
$$Pr(w : [\![\phi]\!]^{w,f,g} = 1 | \{w' : w' \in \text{BEST}_{g(w)}(f(w))\}) = 1$$

For current purposes, we leave open how the notion of probability in (135) should be interpreted. Plausibly, in some cases (broadly, subjective uses of deontic modals) it will represent the credences of the speaker or a group of agents, while in others (broadly, objective uses) it will capture an objective notion of probability (be it chance, evidential probability in the sense of Williamson 2000, or something else).[45]

As usual, *allowed* can be defined as the dual of *required*: *allowed* ϕ is true just in case the probability of ϕ, conditional on the proposition including all best worlds in the modal base, is non-zero .

(136) $[\![$POS allowed $\phi]\!]^{w,f,g} = 1$ iff
$$Pr(w : [\![\phi]\!]^{w,f,g} = 1 | \{w' : w' \in \text{BEST}_{g(w)}(f(w))\}) > 0$$

It's easy to check that this analysis renders modals monotonic on their prejacent, hence the two problems we pointed out in 8.2.1 are sidestepped.

8.2.3   Predicting possibility inferences

Let us show that the analysis, like its counterpart for epistemic modals, predicts all the possibility inferences we observe via a unified mechanism. Let's focus first on *required*, which we take to have the entry:

(137) $[\![$required$]\!]^{w,f,g} = \lambda d. \lambda p. Pr(p | \{w' : w' \in \text{BEST}_{g(w)}(f(w))\}) \geq d$

---

44 We are assuming a standard way of defining conditional probability as a ratio:
$$Pr(q|p) = \frac{Pr(p \wedge q)}{Pr(p)}$$

45 Let us point out that our basic maneuver — replacing a quantifier with a probability function — is compatible with a more radical departure from Kratzer. Our semantics still singles out a set of 'best' worlds via an ordering source. But one might also single out best worlds in a different way — e.g., as worlds via expected utility. Our main reason for sticking to a more conservative setup here is that the latter already yields all that we need.

Assuming composition of POS in the usual way, (127) is predicted to have the truth-conditions in (138):

(127)   You are POS required to take Syntax or Logic.

(138) $[\![(127)]\!] = 1$ iff $Pr(\,\mathbf{S} \vee \mathbf{L}\,|\,$ BEST $) = 1$

As usual, we derive distributive inferences via exhaustification. We assume that, on its strengthened reading, (127) is parsed as including an occurrence of EXH (which we assume to be below POS, though the issue is of no consequence in this case).

(139)   POS [λd_1 [EXH[[d_1 required] [that you take Syntax or Logic]]]]

Given standard assumption about alternatives, the truth conditions of (140) are:

(140) $[\![(139)]\!] = 1$ iff $Pr(\,\mathbf{S} \vee \mathbf{L}\,|\,$ BEST $) = 1 \wedge$
$$Pr(\,\mathbf{S}\,|\,\text{BEST}\,) < 1 \wedge Pr(\,\mathbf{L}\,|\,\text{BEST}\,) < 1$$

As above, (140), in turn, entails (141).

(141) $Pr(\,\mathbf{S} \vee \mathbf{L}\,|\,$ BEST $) = 1 \wedge \exists d[d > 0 \wedge Pr(\,\mathbf{S}\,|\,$ BEST $) > d] \wedge$
$$\exists d[d > 0 \wedge Pr(\,\mathbf{L}\,|\,\text{BEST}\,) > d]$$

(141), given our meaning for *allowed*, yields the relevant possibility inferences.

(127)   You are required to take Syntax or Logic.

    a.   ⤳ *You are allowed to take Syntax*
    b.   ⤳ *You are allowed to take Logic*

The same account holds for possibility inferences with *allowed*. In analogy with what we have done so far, we assume the following entry:

(142) $[\![\text{allowed}]\!]^{w,f,g} = \lambda d.\,\lambda p.\,Pr(p \mid \{w': w' \in \text{BEST}_{g(w)}(f(w))\}) \geq d$

The basic LF of (126) (assuming that the modal takes wide scope with respect to the subject) is in (143):

(126)   You are allowed to take Syntax or Logic

(143)   POS [λd_1 [[d_1 allowed] [you take Syntax or Logic]]]

Given our assumptions about POS and the entry in (142), the schematic truth conditions of (126) are:

(144) $[\![(126)]\!] = 1$ iff $\exists d[d > 0 \wedge Pr(\mathbf{S} \vee \mathbf{L} \mid \text{BEST}) \geq d]$

Free choice is derived, exactly as it happens for epistemic modals, by using an exhaustivity operator that takes scope above the modal, but below POS. Hence, on the exhaustified reading, (126) has the LF:

(145) POS $[\lambda d_1 [$ EXH $[[d_1$ allowed] [you take Syntax or Logic]]]]

Given standard assumption about alternatives, the truth conditions of (145) are:

(146) $[\![(145)]\!] = \exists d \, [d > 0 \wedge Pr(\mathbf{S} \vee \mathbf{L} \mid \text{BEST}) \geq d \wedge$
$Pr(\mathbf{S} \mid \text{BEST}) < d \wedge Pr(\mathbf{L} \mid \text{BEST}) < d]$

The truth conditions in (146) entail, among other things:

(147) $\exists d[Pr(\mathbf{S} \mid \text{BEST}) > d] \wedge \exists d[Pr(\mathbf{L} \mid \text{BEST}) > d]$

which gives us the free choice inferences, repeated below in (148).

(148) You are allowed to take Syntax or Logic.
    a. ⤳ *You are allowed to take Syntax*
    b. ⤳ *You are allowed to take Logic*

8.3 Other modals

Epistemic and deontic modals are the paradigm examples of modals giving rise to possibility inferences. But these inferences have been discussed also in connection with modals of other flavors. One interesting case is that of ability modals. Ability modals give rise to possibility inferences in some cases, but block them in others. For a case of the latter kind, consider (149) (from Nouwen 2017).[46]

(149) ??My son can walk or talk.

---

46 See also van Tiel 2011 for experimental evidence suggesting that ability modals are not associated to free choice inferences like deontic and epistemic modals are.

(149) is pretty odd with the intended free choice interpretation that my son can walk and my son can talk. On the other hand, there are clear cases where ability modals clearly give rise to possibility inferences, like (150) or that in (151), from Geurts 2010 and Klinedinst 2007, respectively.

(150)   Betty can balance a fishing rod on her chin or her nose.

    a.   ⤳ *Betty can balance a fishing rod on her chin*
    b.   ⤳ *Betty can balance a fishing rod on her nose*

(151)   Jenny can outsmart a doctor or a lawyer.

    a.   ⤳ *Jenny can outsmart a doctor*
    b.   ⤳ *Jenny can outsmart a lawyer*

Here we don't have space to discuss these modals, in particular with respect to the contrast between (149), on the one hand, and (150) and (151), on the other (see Nouwen 2017 for a proposal). But let us point out that our account could in principle be extended to ability modals. The trick is, again, to decompose the meaning of ability modals in a measure function and a kind of quantifier over degrees, and insert an exhaustivity operator between the two — i.e., to treat ability modals that give rise to free choice as exploiting the usual configuration:

(152)   POS [EXH [MODAL $(p \lor q)$]]

Again, this can be done by replacing all-or-nothing logical notions like consistency and entailment with probabilistic notions. For a toy example,[47] suppose that we treat *can* as a possibility modal operating over an appropriately restricted set of possibilities. Much like the semantics for epistemic modals above, *can φ* will say that the probability of its prejacent is non-zero . The resulting truth conditions for (153) are in (154):

(153)   Jenny can outsmart a doctor.

(154) $\exists d[d > 0 \land Pr(\llbracket \text{Jenny outsmarts a doctor} \rrbracket) \geq d]$

---

47 See, among many, Mandelkern, Schultheis & Boylan 2015 for arguments to the effect that semantics in this vein suffer from serious empirical problems.

9 Other issues

9.1 Wide scope disjunction and some other cases of free choice

Up to now, we have only considered instances of free choice where the modal operator takes wide scope over disjunction. But it is well-known that free choice inferences are generated also in cases where disjunction appears to take scope over modals. E.g., (155) suggests that rain is possible, and that so is snow (Legrand 1975, Zimmerman 2000, Geurts 2005, Simons 2005, Fox 2007 among others).

(155)   It's possible that it will rain or it's possible that it will snow.
    a.   ⤳ *It's possible that it will rain*
    b.   ⤳ *It's possible that it will snow*

The present proposal, like most other implicature-based accounts, doesn't address this puzzle.[48] But let us observe that it is compatible with independent accounts of wide scope free choice, and in particular with attempts at reducing wide scope free choice to narrow scope free choice via syntactic or semantic means. (See, e.g., the proposal in Simons 2005, on which in a sentence like (155) the modal undergoes covert movement and takes scope over disjunction; though see also Alonso Ovalle 2005, §4.13, for some substantial problems with Simons' proposal.)

Second, free choice obtains also with negated conjunction, as noticed by Fox 2007.

(156)   It's not certain that we will hire Mary and Sue.
    a.   ⤳ *It's possible that we will not hire Mary*
    b.   ⤳ *It's possible that we will not hire Sue*

In our framework, the inferences in (156a) and (156b) can be derived by using two exhaustivity operators — one below *and* one above POS, as in (157).

(157)   EXH[not [POS [EXH[certain [that we will hire Mary and Sue]]]]]

To see how this works, consider the lower EXH first. The relevant alternatives are:

---

48 Ciardelli, Groenendijk & Roelofsen (2009) is an exception to this. Some semantic accounts of free choice, e.g. Zimmerman 2000 or Geurts 2005, also do not have this problem.

$$(158) \quad \left\{ \begin{array}{ll} d \text{ certain that we will hire Mary or Sue} & Pr(\mathbf{M} \vee \mathbf{S}) \geq d \\ d \text{ certain that we will hire Mary} & Pr(\mathbf{M}) \geq d \\ d \text{ certain that we will hire Sue} & Pr(\mathbf{S}) \geq d \\ d \text{ certain that we will hire Mary and Sue} & Pr(\mathbf{M} \wedge \mathbf{S}) \geq d \end{array} \right\}$$

The prejacent is stronger than all alternatives, so EXH is vacuous here. It does, however, change the alternatives which are then fed to the second EXH (represented in (159)). In particular, the negated exhaustified disjunctive alternative is crucial once the second exhaustivity operator enters the derivation.

$$(159) \quad \left\{ \begin{array}{ll} \text{not[POS[EXH[}d\text{ certain} \ldots \text{Mary and Sue]]]} & \neg[Pr(\mathbf{M} \wedge \mathbf{S}) = 1] \\ \text{not[POS[EXH[}d\text{ certain} \ldots \text{Mary]]]} & \neg[Pr(\mathbf{M}) = 1 \wedge Pr(\mathbf{S}) < 1] \\ \text{not[POS[EXH[}d\text{ certain} \ldots \text{Sue]]]} & \neg[Pr(\mathbf{M}) = 1 \wedge Pr(\mathbf{S}) < 1] \\ \text{not[POS[EXH[}d\text{ certain} \ldots \text{Mary or Sue]]]} & \begin{array}{l} \neg[Pr(\mathbf{M} \vee \mathbf{S}) = 1 \\ \quad \wedge Pr(\mathbf{M}) < 1 \wedge Pr(\mathbf{S}) < 1] \end{array} \end{array} \right\}$$

The second EXH quantifies over the alternatives above and the result is in (160). That is, we conclude that the probability of hiring Mary and Sue is less than one, that hiring one or the other is 1, and, crucially, that the probability of neither disjunct is 1. As a result, we derive the relevant possibility inferences.

(160) $\neg[Pr(\mathbf{M} \wedge \mathbf{S}) = 1] \wedge Pr(\mathbf{M} \vee \mathbf{S}) = 1 \wedge Pr(\mathbf{M}) < 1 \wedge Pr(\mathbf{S}) < 1$

9.2 Brief comparison with alternative proposals

We don't have the space to compare in detail the present proposal to other accounts in the literature. But let us make some brief remarks about its relation to two successful scalar approaches, i.e. those in Fox (2007) and Klinedinst (2007).[49]

Let us start with Fox's account. Fox assumes a standard existential analysis of possibility modals. Free choice is derived by assuming that sentences like (161) are parsed as involving two exhaustivity operators, as in (162).

(161) It is possible that we will hire Mary or Sue.

(162) EXH[EXH[it is possible that we will hire Mary or Sue]]

---

[49] We leave a comparison with other scalar-implicature based accounts (e.g., Chemla 2008, Franke 2011, van Rooij 2010) and with semantic accounts (e.g., Zimmerman 2000, Geurts 2005, Barker 2010, Starr 2016) for future research.

The crucial element of the account is that the outermost EXH exploits alternatives that have been already exhaustified. In particular, the alternatives for the outermost EXH include the exhaustified disjuncts.

(163)
$$\left\{ \begin{array}{ll} \ldots & \\ \text{EXH[It is possible that we will hire Mary]} & \Diamond \mathbf{M} \wedge \neg \Diamond \mathbf{S} \\ \text{EXH[It is possible that we will hire Sue]} & \Diamond \mathbf{S} \wedge \neg \Diamond \mathbf{M} \\ \ldots & \end{array} \right\}$$

This in turn means that the second EXH winds up negating the exhaustified alternatives above, giving rise to free choice as in (164).

(164) $[\![\text{EXH[EXH[it is possible that we will hire Mary or Sue]]}]\!] =$
 $\Diamond (\mathbf{M} \vee \mathbf{S}) \wedge \neg [\Diamond \mathbf{M} \wedge \neg \Diamond \mathbf{S}] \wedge \neg [\Diamond \mathbf{S} \wedge \neg \Diamond \mathbf{M}] =$
 $\Diamond (\mathbf{M} \vee \mathbf{S}) \wedge \Diamond \mathbf{M} \wedge \Diamond \mathbf{S}$

Fox's account extends to all types of modals and to nominal quantifiers.[50]

The advantage that we claim for our account over Fox's is that we give a unified story for deriving all possibility inferences. This covers both distributivity inferences and free choice ones. By contrast, Fox has to resort to different mechanisms for the two cases. His account of distributive inferences is the one involving ordinary exhaustification, and sketched in §3; conversely, free choice requires double exhaustification. Of course, this is not enough to establish that our account is overall better. But it suggests that it deserves consideration as a competitor.

Let us move on to Klinedinst's (2007) proposal. This is probably the closest proposal to ours, at least in spirit. Klinedinst derives free choice inferences as a kind of embedded distributivity implicature, by making a key background assumption, i.e., that all modals (and, among them, possibility modals) are plural quantifiers over worlds containing a distributivity operator DIST.[51] As a result, (165) says (roughly) that there is a plurality of worlds such that in each of them we will hire Mary or Sue.

(165) It is possible that we will hire Mary or Sue.

---

50 Notice also that Fox's (2007) account is in principle compatible with ours. That is, it is easy to show that global recursive exhaustification would give rise to free choice in the same way as above.

51 For an introductory discussion of the semantics of plurals and distributivity operators, see Nouwen 2016, Winter & Scha 2015.

Against this background, Klinedinst's suggestion is that we embed an exhaustivity operator below the modal, but above DIST.

(166)   It is possible EXH [DIST [we will hire Mary or Sue]]

Crucially, DIST works as a universal quantifier. Since EXH takes scope over a universal quantifier, Klinedinst can derive free choice much like a standard distributive inference. The alternatives and the computation of the exhaustified clause are below.

(167)
$$\left\{ \begin{array}{ll} \text{DIST[we will hire Mary or Sue]} & \forall w'[\mathbf{M}_{w'} \vee \mathbf{S}_{w'}] \\ \text{DIST[we will hire Mary]} & \forall w'[\mathbf{M}_{w'}] \\ \text{DIST[we will hire Sue]} & \forall w'[\mathbf{S}_{w'}] \\ \text{DIST[we will hire Mary and Sue]} & \forall w'[\mathbf{M}_{w'} \wedge \mathbf{S}_{w'}] \end{array} \right\}$$

(168) $[\![\text{EXH[DIST[we will hire Mary or Sue]]}]\!] =$
$\forall w'[\mathbf{M}_{w'} \vee \mathbf{S}_{w'}] \wedge \neg \forall w'[\mathbf{M}_{w'}] \wedge \neg \forall w'[\mathbf{S}_{w'}] =$
$\forall w'[\mathbf{M}_{w'} \vee \mathbf{S}_{w'}] \wedge \exists w'[\mathbf{M}_{w'}] \wedge \exists w'[\mathbf{S}_{w'}] =$

Putting all this together we have (169), which entails the usual free choice inferences.

(169) $[\![\text{it is possible that [EXH[DIST[we will hire Mary or Sue]]]}]\!] =$
$\exists W[\forall w' \in W[\mathbf{M}_{w'} \vee \mathbf{S}_{w'}] \wedge \exists w' \in W[\mathbf{M}_{w'}] \wedge \exists w' \in W[\mathbf{S}_{w'}]]$

Like Fox's account, Klinedinst's (2007) extends to quantifiers over individuals. Hence Klinedinst can account for free choice-type inference triggered by plural quantifiers (as in (170)) with the very same tools.

(170)   Some students took three months or didn't finish at all.
    a.   ⤳ *Some students took three months*
    b.   ⤳ *Some students didn't finish at all*

A natural worry about Klinedinst's proposal is that it crucially relies on a plural semantics for modals, and moreover on the presence of a covert distributivity operator DIST in modal sentences. While there is some evidence for the former, the latter assumption comes with no independent motivation. In fact, one can easily raise overgeneration worries about DIST. A theory that assumes the presence of a distributivity operator predicts that various items should be able to take scope between the modal and DIST — i.e., exactly where

EXH is located on Klinedinst's theory. To take a concrete example, consider (171).

(171)   Nobody is allowed to enter.

Klinedinst predicts that (171) has, among others, a reading with the following schematic truth conditions:

(172) $\exists W[\neg\exists x[\forall w \in W[\mathbf{enter}_w(x)]]]$ $\exists >$ Nobody $>$ DIST

(172) says that there is a plurality of (deontically ideal) worlds such that no individual entered in all of them. I.e., (172) is false only if there is some individual who enters in all (deontically ideal) worlds; hence, on this reading it should be equivalent to the most prominent reading of *No one is required to enter*. Obviously (172) doesn't have a reading of this sort.[52]

Also in this case, this point alone is not enough to give our account a definitive advantage. For one thing, Klinedinst could appeal to scope constraints that are similar to the ones we postulate for POS. (Though the latter constraints are independently motivated in the literature, while to our knowledge no analogous arguments have been given about DIST.)

Let us also emphasize again that one advantage that, at present, both Fox's and Klinedinst's accounts have over our own account is that both naturally generalize to free choice-type inferences triggered by nominal quantifiers, like

---

52 Thanks to an anonymous referee for suggesting this way of putting the point.

Let us notice that Klinedinst could appeal to whatever constraints on scope disallow a similar reading with other cases involving distributive operators, like bare plurals. Similarly, (i), does not have a reading which would be true if there is a plurality of students such that no individual met all of the member of such plurality (but, say, each individual met one of the students).

(i)   Nobody met students.

It is unclear that the parallelism between plural definites and modals hold when we move to other contexts. For instance, a non-monotonic case like (ii) can be shown not to have a reading in which *exactly two* scopes in between existential quantification and DIST. But it is unclear that this reading is not there for its plural definite counterpart in (iii).

(ii)   Exactly two people are allowed to enter.

(iii)   Exactly two people talked to the students.

(170). Conversely, it remains to be seen how our account could be extended beyond modals.[53] This is one of the issues we have to put off to future work.

9.3 The optionality of free choice

It is an established point in the literature that all possibility inferences, including free choice ones, are optional effects. For example, consider:

(173)   Mary is allowed to take Syntax or Logic.

(173) has a reading that doesn't imply that Mary is allowed to take Syntax and that she is allowed to take Logic. Rather, it suggests that the speaker is simply ignorant about which disjunct is true.[54] This reading is sometimes called an 'ignorance reading,' and is brought out by continuations like that in (174).

(174)   Mary is allowed to take Syntax or Logic, but I don't remember which one.

Ignorance readings of this kind are available also for epistemic modals, though they are somewhat harder to obtain. To generate them, we need a context that makes salient a body of information that entails, but it not entailed by, the speaker's information. For an example, consider the following scenario.

> Department X has voted an offer to a candidate. The details are not public, but my friend, who is in Department X's search committee, has let slip information suggesting that the offer might have gone to a certain candidate. Unfortunately, I have forgotten whether this candidate is Mary or Sue.

Against this background, one says:

(175)   It's possible that Department X made an offer to Mary or Sue, but I don't remember which.

---

[53] Notice also that in principle our account is compatible with Klinedinst's account for the nominal quantifier cases, since the possibility of having a distributivity operator for nominal quantifiers is not in question.

[54] Ignorance inferences standardly arise from a standard Gricean maxim of Quantity not based on alternatives; see Sauerland 2004, Fox 2007 among others for discussion.

(175) is perfectly appropriate in this scenario. Moreover, notice that (thanks again to the conjunct *I don't remember which*) the usual possibility inferences are blocked. Of course, upon hearing (175) we can infer that it's compatible with the speaker's knowledge both that Mary is the recipient of the offer and that Sue is. But *possible* in (175) evidently doesn't merely describe the speaker's knowledge.

Here we want to point out that our account, on a par with other exhaustivity-based accounts, has no difficulties predicting the optionality of possibility inferences. We predict that a sentence like (176) has a number of available parses.

(176)   It's possible that Department X made an offer to Mary or Sue.

In 7.1, we have pointed out that (176) has two parses that give rise to (slightly) different possibility inferences:

(177)   a.   EXH [POS [likely [DX made an offer to Mary or Sue]]]
         b.   POS [EXH [likely [DX made an offer to Mary or Sue]]]

But we also predict the availability of a parse that does not involve EXH in any position:

(178)   POS [likely [DX made an offer to Mary or Sue]]

In this case, we predict no implicature at all.

Of course, we also need to say something about when the parse involving an exhaustivity operator is available. For concreteness, here we simply rely on the suggestion in Fox 2007. Fox's idea is that speakers adopt the exhaustified parses to avoid ignorance inferences. More precisely: Fox assumes that speakers will first attempt to parse a sentence without an exhaustivity operator. If this parse gives rise to ignorance inferences, the sentence is reparsed as involving one (or possibly more) occurrences of EXH.

10   Conclusion

Disjunctions in the scope of epistemic modals give rise to scalar inferences to the effect that each disjunct is epistemically possible. These inferences are usually called and treated differently. Those generated by non-possibility modals are labeled 'distributive inferences,' and derived as scalar implicatures. Those generated by possibility modals are called 'free choice infer-

ences' and are derived via more complex mechanisms. We have shown that, by adopting a degree semantics for modals, we can give a unified account of all these effects as standard scalar implicatures.

References

Aloni, Maria. 2005. Utility and implicatures of imperatives. In *Proceedings of DIALOR'05*. Loria, Nancy, France.

Alonso Ovalle, Luis. 2005. *Disjunction in alternative semantics*. University of Massachusetts, Amherst dissertation.

Alrenga, Peter & Christopher Kennedy. 2014. No more shall we part: Quantifiers in English comparatives. *Natural Language Semantics* 22(1). 1–53. https://doi.org/10.1007/s11050-013-9099-4.

Barker, Chris. 2010. Free choice permission as resource-sensitive reasoning. *Semantics and Pragmatics* 3(10). 1–38. https://doi.org/10.3765/sp.3.10.

Beck, Sigrid. 2012. DegP scope revisited. *Natural Language Semantics* 20(3). 227–272. https://doi.org/10.1007/s11050-012-9081-6.

Beltrama, Andrea & Ray Bochnak. 2015. Intensification without degrees cross-linguistically. *Natural Language and Linguistic Theory* 33(3). 843–879. https://doi.org/10.1007/s11049-015-9294-8.

Bochnak, Ryan. 2015. The degree semantic parameter and cross-linguistic variation. *Semantics and Pragmatics* 8(6). 1–48. https://doi.org/10.3765/sp.8.6.

Briggs, Rachael. 2015. Normative theories of rational choice: Expected utility. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Winter 2015.

Cariani, Fabrizio. 2013. *Ought* and resolution semantics. *Noûs* 47(3). 534–558. https://doi.org/10.1111/j.1468-0068.2011.00839.x.

Cariani, Fabrizio. 2016a. Consequence and contrast in deontic semantics. forthcoming in *Journal of Philosophy*.

Cariani, Fabrizio. 2016b. Deontic modals and probabilities: One theory to rule them all? In Nate Charlow & Matthew Chrisman (eds.), *Deontic modality*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198717 928.003.0002.

Cariani, Fabrizio, Paolo Santorio & Alexis Wellwood. 2016. Foundational and empirical questions in degree semantics with modality as a case study. lectures notes for NASSLLI 2016, Northwestern University and University of Leeds.

Chemla, Emmanuel. 2008. Similarity: Towards a unified account of scalar implicatures, free choice permission and presupposition projection. Unpublished manuscript. http://www.emmanuel.chemla.free.fr/Material/Chemla-SIandPres.pdf.

Chemla, Emmanuel & Benjamin Spector. 2011. Experimental evidence for embedded scalar implicatures. *Journal of Semantics* 28(3). 359–400. https://doi.org/10.1093/jos/ffq023.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, 39–103. Oxford: Oxford University Press.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697977.001.0001.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: an international handbook of natural language meaning volume 3.* Berlin: Mouton de Gruyter.

Ciardelli, Ivano, Jeroen Groenendijk & Floris Roelofsen. 2009. Attention! *Might* in inquisitive semantics. In Satoshi Ito Ed Cormany & David Lutz (eds.), *Semantics and Linguistic Theory (SALT) 19.* https://doi.org/10.3765/salt.v19i0.2520.

Cresswell, Max J. 1976. The semantics of degree. In Barbara Partee (ed.), *Montague grammar*, 261–92. New York: Academic Press.

Crnič, Luka, Emmanuel Chemla & Danny Fox. 2015. Scalar implicatures of embedded disjunction. *Natural Language Semantics* 23(4). 271–305. https://doi.org/10.1007/s11050-015-9116-x.

von Fintel, Kai. 2012. The best we can (expect to) get? Challenges to the classic semantics for deontic modals. Paper for a session on Deontic Modals at the Central APA. http://web.mit.edu/fintel/fintel-2012-apa-ought.pdf.

von Fintel, Kai & Sabine Iatridou. 2008. How to say *ought* in Foreign: The composition of weak necessity modals. In Jacqueline Gueron & Jacqueline Lecarme (eds.), *Time and modality*, 115–141. Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-1-4020-8354-9_6.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Com-*

positional Semantics, 71–120. Palgrave. https://doi.org/10.1057/978023021
0752_4.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586. https://doi.org/10.1007/s10988
-006-9004-4.

Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19(1). 87–107. https://doi.org/10.1007/s1105
0-010-9065-3.

Franke, Michael. 2011. Quantity implicatures, exhaustive interpretation, and rational conversation. *Semantics and Pragmatics* 4(1). 1–82. https://doi.or
g/10.3765/sp.4.1.

Gajewski, Jon Robert. 2007. Neg-raising and polarity. *Linguistics and Philosophy* 30(3). 289–328. https://doi.org/10.1007/s10988-007-9020-z.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form*. New York: Academic Press.

Geurts, Bart. 2005. Entertaining alternatives: Disjunctions as modals. *Natural Language Semantics* 13(4). 383–410. https://doi.org/10.1007/s11050-005-
2052-4.

Geurts, Bart. 2010. *Quantity implicatures*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511975158.

Goble, Lou. 1996. Utilitarian deontic logic. *Philosophical Studies* 82(3). 317–357. https://doi.org/doi.org/10.1007/bf00355312.

Grice, Paul. 1975. Logic and conversation. In Donald Davidson & Gilbert Harman (eds.), *The logic of grammar*, 64–75. Encino, CA: Dickenson.

Hájek, Alan. 2012. Interpretations of probability. In Edward N. Zalta (ed.), *The stanford encyclopedia of philosophy*, Winter 2012. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2012/en
tries/probability-interpret/.

Halpern, Joseph Y. 1997. Defining relative likelihood in partially-ordered preferential structures. *Journal of Artificial Intelligence Research*. 1–24. https://doi.org/10.1613/jair.391.

Halpern, Joseph Y. 2003. *Reasoning about uncertainty*. Cambridge, MA: MIT Press.

Heim, Irene. 1985. Notes on comparatives and related matters. Unpublished manuscript. http://semanticsarchive.net/Archive/zcoZjYoM/Comparati
ves%2085.pdf.

Heim, Irene. 2000. Degree operators and scope. In Brendan Jackson & Tanya Matthews (eds.), *Semantics and Linguistic Theory (SALT) 10*, 40–64. Cornell

University, Ithaca, NY: CLC Publications. https://doi.org/doi.org/10.3765/salt.v10i0.3102.

Higginbotham, Jim. 1991. Either/or. In *North East Linguistic Society (NELS) 21*. University of Massachusetts, Amherst: GLSA.

Holliday, Wesley & Thomas Icard. 2013. Measure semantics and qualitative semantics for epistemic modals. In *Semantics and Linguistic Theory (SALT) 23*, 514–534. https://doi.org/doi.org/10.3765/salt.v23i0.2670.

Horn, Laurence. 1978. Remarks on neg-raising. In Peter Cole (ed.), *Syntax and semantics 9: Pragmatics*. Academic Press.

Jackson, Frank. 1985. On the semantics and logic of obligation. *Mind* 94(374). 177–195. https://doi.org/10.1093/mind/xciv.374.177.

Jackson, Frank & Robert Pargetter. 1986. Oughts, options, and actualism. *Philosophical Review* 95(2). 233–255. https://doi.org/doi.org/10.2307/2185591.

Jeffrey, Richard. 1983. *The logic of decision*. Chicago, IL: University of Chicago Press. 250–253.

Joyce, James. 1999. *The foundations of causal decision theory*. Cambridge: Cambridge University Press. 545–551. https://doi.org/10.1017/cbo9780511498497.

Kamp, Hans. 1973. Free choice permission. *Proceedings of the Aristotelian Society* 74. 57–74. https://doi.org/10.1163/9789004252882_008.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistic and Philosophy* 30(6). 669–690. https://doi.org/10.1007/s10988-008-9029-y.

Kaufmann, Magdalena & Stefan Kaufmann. 2015. Conditionals and modality. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, vol. 3, 237–270. John Wiley & Sons. https://doi.org/10.1002/9781118882139.ch8.

Kennedy, Christopher. 1997. *Projecting the adjective: The syntax and semantics of gradability and comparison*. University of California, Santa Cruz dissertation. 243. https://doi.org/10.4324/9780203055458.

Kennedy, Christopher. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and philosophy* 30(1). 1–45. https://doi.org/10.1007/s10988-006-9008-0.

Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381. https://doi.org/10.1353/lan.2005.0071.

Klecha, Peter. 2014. *Bridging the divide: Scalarity and modality*. University of Chicago dissertation.

Klinedinst, Nathan. 2007. *Plurality and possibility*. University of California, Los Angeles dissertation.

Kratzer, Angelika. 1981. The notional category of modality. In Hans J. Eikmeyer & Hannes Rieser (eds.), *Words, worlds, and contexts: New approaches to word semantics*. Berlin, Germany: de Gruyter. https://doi.org/10.1515/9783110842524-004.

Kratzer, Angelika. 1986. Conditionals. In *Chicago Linguistics Society: papers from the parasession on pragmatics and grammatical theory*, vol. 22, 1–15. University of Chicago, Chicago IL: Chicago Linguistic Society.

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650.

Kratzer, Angelika. 2012. *Modals and conditionals: New and revised perspectives*. Vol. 36. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199234684.001.0001.

Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In Yukio Otsu (ed.), *Proceedings of the Tokyo conference on psycholinguistics*, vol. 3, 1–25. Tokyo: Hituzi Syobo.

Lassiter, Daniel. 2010. Gradable epistemic modals, probability, and scale structure. In Nan Li & David Lutz (eds.), *Semantics and Linguistic Theory (SALT) 20*, 197–210. https://doi.org/10.3765/salt.v20i0.2557.

Lassiter, Daniel. 2011. *Measurement and modality: The scalar basis of modal semantics*. New York University dissertation.

Lassiter, Daniel. 2014. Epistemic comparison, models of uncertainty, and the disjunction puzzle. *Journal of Semantics*. https://doi.org/10.1093/jos/ffu008.

Lassiter, Daniel. 2016. Graded modality. In Lisa Matthewson, Cécile Meier, Hotze Rullmann & Ede Zimmerman (eds.), *Companion to semantics*. Blackwell.

Legrand, Jean Ehrenkranz. 1975. *Or and Any: The semantics of two logical operators*. University of Chicago dissertation.

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

Lewis, David. 1981. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic* 10(2). 217–234. https://doi.org/10.1017/cbo9780511625237.008.

MacFarlane, John. 2011. Epistemic modals are assessment-sensitive. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*. Oxford: Oxford Uni-

versity Press. https://doi.org/10.1093/acprof:oso/9780199591596.003.00 06.

Mandelkern, Matthew Aaron, Ginger Schultheis & David Boylan. 2015. I believe I can ϕ. In Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 256–265.

Moss, Sarah. 2015a. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8(5). 1–81. https://doi.org/10.3765/sp.8.5.

Moss, Sarah. 2015b. Probabilistic knowledge. book forthcoming for Oxford University Press, University of Michigan.

Nouwen, Rick. 2016. Plurality. In Maria Aloni & Paul Dekker (eds.), *Cambridge handbook of semantics*, 267–284. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139236157.010.

Nouwen, Rick. 2017. Free choice and distribution over disjunction: The case of free choice ability. manuscript, June 2017, University of Utrecht. http://ric knouwen.org/rwfn/wp-content/uploads/2014/01/ability2-nonanon.pdf.

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

Portner, Paul & Aynat Rubinstein. 2016. Extreme and non-extreme deontic modals. In Nate Charlow & Matthew Chrisman (eds.), *Deontic modality*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/978 0198717928.003.0010.

Romero, Maribel. 2015. The conservativity of *many*. In Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 20–29. Amsterdam.

Romoli, Jacopo. 2013. A scalar implicature-based approach to Neg-raising. *Linguistics and philosophy* 36(4). 291–353. https://doi.org/10.1007/s10988 -013-9136-2.

van Rooij, Robert. 2010. Conjunctive interpretation of disjunction. *Semantics and Pragmatics* 3(11). 1–28. https://doi.org/10.3765/sp.3.11.

Ross, Alf. 1944. Imperatives and logic. *Philosophy of Science* 11(1). 30–46. https://doi.org/doi.org/10.1086/286823.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367–391. https://doi.org/10.1023/b:ling.0000023378 .71748.db.

Sauerland, Uli. 2012. The computation of scalar implicatures: Pragmatic, lexical or grammatical? *Language and Linguistics Compass* 6(1). 36–49. https://doi.org/10.1002/lnc3.321.

Sauerland, Uli. 2014. Intermediate Scalar Implicatures. In Salvatore Pistoia Reda (ed.), *Pragmatics, Semantics and the Case of Scalar Implicatures*,

72–98. London: Palgrave Macmillan UK. https://doi.org/10.1057/97811373
33285_4.

Simons, Mandy. 2005. Semantics and pragmatics in the interpretation of 'or'. In Effi Georgala & Jonathan Howell (eds.), *Semantics and Linguistic Theory (SALT) 15*, 205–222. Cornell University, Ithaca, NY: CLC Publications. https://doi.org/10.3765/salt.v15i0.2929.

Stalnaker, Robert. 1984. *Inquiry*. Cambridge, MA: MIT Press. https://doi.org/1 0.2307/2185468.

Starr, William. 2016. Expressing permission. In Jacob Collard Mary Moroney Carol-Rose Little & Dan Burgdorf (eds.), *Semantics and Linguistic Theory (SALT) 26*. https://doi.org/10.3765/salt.v26i0.3832.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3(1). 1–77. https://doi.org/10.1093/jos/3.1-2.1.

Swanson, Eric. 2006. *Interactions with context*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Swanson, Eric. 2011. How not to theorize about the language of subjective uncertainty. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199591 596.003.0009.

Swanson, Eric. 2015. The application of constraint semantics to the language of subjective uncertainty. English. *Journal of Philosophical Logic*. 1–26. https://doi.org/10.1007/s10992-015-9367-5. http://dx.doi.org/10.1007/s 10992-015-9367-5.

van Tiel, Bob. 2011. Universal free choice. In Anna Chernilovskaya Ana Aguilar Guevara & Rick Nouwen (eds.), *Sinn und Bedeutung (SuB)*, vol. 16, 627–638.

Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford: Oxford University Press. https://doi.org/10.1093/019925656x.001.0001.

Winter, Yoad & Remko Scha. 2015. Plurals. In Shalom Lappin & Chris Fox (eds.), *The handbook of contemporary semantic theory*, 77–113. John Wiley & Sons. https://doi.org/10.1111/b.9780470670736.2015.00005.x.

von Wright, Georg. 1968. An essay in deontic logic and the general theory of action. *Acta Philosophica Fennica* 21. 110–133.

Yalcin, Seth. 2005. Epistemic modals. In *MIT Working Papers in Linguistics*, vol. 51. Cambridge, MA: MIT.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. https://doi.or g/10.1093/mind/fzm983.

Yalcin, Seth. 2010. Probability operators. *Philosophy Compass* 5(11). 916–937. https://doi.org/10.1111/j.1747-9991.2010.00360.x.

Yalcin, Seth. 2012. Bayesian expressivism. *Proceedings of the Aristotelian Society* 112(2, part 2). 123–160. https://doi.org/10.1111/j.1467-9264.2012.00329.x.

Zimmerman, Thomas Ede. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8. 255–290. https://doi.org/10.1023/A:1011255819284.

Paolo Santorio
School of Philosophy
University of Leeds
Woodhouse Lane
Leeds LS2 9JT, United Kingdom
paolosantorio@gmail.com

Jacopo Romoli
School of Communication
Ulster University
Shore Road
Newtownabbey, BT37oQB, United Kingdom
jacopo.romoli@gmail.com