

Extreme and Non-Extreme Deontic Modals*

Paul Portner and Aynat Rubinstein

July 14, 2014

1 Introduction

When we argue that our priorities militate in favor of a certain action, we can do so in a way that allows for competing alternatives, or in a way which does not. Suppose that A and B agree that no policy can both provide insurance for uninsured citizens and allow people to make their own choices:

- (1) A: It is important that our uninsured citizens get insurance.
B: It's also important that people make their own choices.
A: So how do we balance these things?
- (2) A: It is crucial that our uninsured citizens get insurance.
B: And it's crucial that we allow people to make their own choices.
A: So we're stuck.
- (3) A: Our uninsured citizens should get insurance.
B: People should also make their own choices.
A: So how do we balance these things?
- (4) A: Our uninsured citizens must get insurance.
B: And yet people must make their own choices.
A: What can we do?!

We refer to operators like *important*, *crucial*, *should* and *must* as DEONTIC NECESSITY MODALS.¹ Our goal in this paper is to contribute to the understanding of the differences

*We thank Graham Katz and Elena Herburger for discussions which have shaped our understanding of the topic and Jessica Rett, Nate Charlow, and Stephen Finlay for comments on previous versions. Any errors are of course our responsibility, not theirs. We gratefully acknowledge the support of National Science Foundation grant BCS-1053038 to both authors, and support from the Mandel Scholion Interdisciplinary Research Center in the Humanities and Jewish Studies to the second author.

¹This use of the term 'necessity' is common in linguistics but is more inclusive than is usual in philosophy. It reflects a theoretical assumption that the operators with weaker force (*should*, *important*) have the same

between the two types of deontic necessity statement that they produce. What is it about *must* and *crucial* that leads to an impasse, while *should* and *important* do not?

Our paper will focus on the semantic and pragmatic analysis of the modal elements seen in (1)-(4). We situate our work with respect to two broader issues in the semantics of modality. The first concerns the nature of differences in strength among necessity modals. In (5), we see strong necessity modals paired with their weak necessity counterparts:

- (5) (a) You must call Barbara. (*entails* You should . . .)
- (b) It is crucial to talk to Barbara. (*entails* It is important . . .)
- (c) It is certain that Barbara will win. (*entails* It is likely . . .)

- Question 1: Why are there **two strength levels** of modal necessity, and how do they differ?

The recent literature on this topic is diverse, and includes important work by, among others, von Stechow and Iatridou (2008); Finlay (2009, 2010); Kolodny and MacFarlane (2010); Rubinstein (2012).

The second broader issue is based on the observation that many modal operators have scalar semantics, and are gradable in the same way that concrete, non-modal adjectives are. The gradability properties of a number of modal operators are exemplified in (6). We see that *should*, *important*, and *likely* are gradable, similarly to a concrete adjective such as *big*.

- (6) a. You should call Barbara more than (you should call) Alice.² (cf. *bigger*)
- b. It is very important to talk to Barbara. (cf. *very big*)
- c. It is just as likely that Barbara will win as it is that Alice will. (cf. *as big*)

- Question 2: How can the **gradability properties** of modal operators be compositionally modeled within a general framework of gradability?

The issue of the gradability of modal expressions has received attention only fairly recently, by Portner (2009); Yalcin (2010); Lassiter (2010, 2011), and Klecha (2012), among others.

quantificational core as clearer examples of necessity modals like *must*. In a similar way, linguists often assume that attitude verbs such as *believe* and *want* and upper-range probability operators such as *likely* and *probable* have a necessity semantics (Hintikka, 1961; Kratzer, 1981, 1991).

The modals in question are not strictly limited to deontic interpretations, but rather have the full range of PRIORITY readings (deontic, teleological, and buletic; see Portner 2009). These subtypes frequently are similar by linguistic criteria, and the use of a particular modal in context is often indeterminate as to which subtype it is. For example, B's statement in (4) can be seen as relating to a moral principle (deontic) or a goal of policy (teleological); or the context might not specify either of these, in which case it is a priority modal with no specific subtype.

²Note that (6a) has an irrelevant reading where *more* compares frequency. We are focusing on the reading where it compares degrees of importance or priority. One naturally occurring example which arguably has this reading is the following: *A fillet more than a whole fish should have that beautiful fresh smell* (<http://www.smh.com.au/lifestyle/homestyle/blogs/tried-and-tasted/how-to-cook-the-perfect-piece-of-fish-20111020-1m9kn.html>). Accessed from GloWbE. All corpus data cited in this paper was produced by the BYU corpus project (Davies, 2004, 2008, 2013).

In this paper, we contribute a novel linguistic characterization of the difference between weak and strong deontic necessity modals. Among non-modal adjectives, the literature has identified a class known as the EXTREME ADJECTIVES; many relative adjectives have extreme counterparts, for example *big* has the extreme counterpart *huge* and *smart* has *brilliant* (Bolinger, 1967; Paradis, 2001; Rett, 2008; Morzycki, 2012). We will show that *must*, *crucial*, *certain* are also extreme elements, contrasting with their non-extreme counterparts *should*, *important*, *likely*. Since pairs of extreme vs. non-extreme adjectives are so common, it is to be expected that gradable modal operators would come in extreme and non-extreme versions as well. The existing literature on extreme adjectives then points us to a formal analysis of their gradability properties.³

The analysis here will focus on deontic and other priority modals, giving less attention to epistemic examples like (5c). Although we discuss epistemic modals in reviewing the relevant previous literature, our formal analysis will only apply to the priority types. We limit our attention in this way not only because deontic modality is the focus of the present volume, but also because priority modals show a logical profile distinct from epistemic and ability/opportunity modals.^{4,5}

Our paper is organized as follows: Section 2 reviews previous work on modal semantics, focusing in particular on the motivation for developing a theory with links to the scale-based semantics of adjectives, and then pushes this perspective further by incorporating the class of extreme expressions. We begin in 2.1 by reviewing the standard premise semantics classification and theory of modality within formal semantics due to Kratzer. We then provide in Section 2.2 some background on semantic classes of adjectives, discussing how different types of gradable adjectives can be characterized in terms of a scale-based semantics. Based on this background, we review and assess work which looks at modals from a similar perspective, such as Lassiter 2011 and Klecha 2012. A prominent idea in this literature is that many modal operators require a scale-based semantics, and that subclasses of modal operators should be identified with subclasses of gradable adjectives. In section 2.3 we use a variety of tests to establish a parallel between strong and weak necessity modals and the EXTREME and NON-EXTREME classes of adjectives exemplified by *huge* vs. *big*. We show that all of the strong necessity modals are in fact extreme words and point out that this perspective reconciles some of the theoretical tension in previous work on gradable modals.

In Section 3, we develop a detailed semantic and pragmatic analysis of extreme and non-extreme deontic and other priority modals. In this section, we propose that the scales of these modals are constructed from the same types of modal parameters which are familiar from the standard Kratzerian premise semantics for modality, that is conversational backgrounds

³Lassiter (2011) notes the similarity between certain strong necessity modals (in particular, deontics) and what he calls “high degree adjectives” like *huge*. He does not, however, note their unique grammatical properties or connect them to the literature on extreme adjectives.

⁴For example, while priority modals show the Union Property pattern, epistemic ones do not (Halpern, 2003; Yalcin, 2010; Lassiter, 2011): $\phi \geq \psi$, $\phi \geq \chi$ entail $\phi \geq (\psi \vee \chi)$. The fact that epistemics don’t show the Union Property is an important puzzle for the standard Kratzerian (1981; 1991) analysis.

⁵Dynamic modals do not show the same possibility/necessity distinction as other types. This can be seen in the fact that *can* is stronger than a modal logic existential \diamond (*can*(ϕ) does not entail *can*($\phi \wedge \psi$) \vee *can*($\phi \wedge \neg\psi$), cf. Kenny 1976, Portner 2009, sec. 4.4) and the fact that it has no obvious necessity counterpart (*will* is the best candidate).

which function as modal bases and ordering sources; this foundation in premise semantics allows us to draw connections to recent work on weak and strong necessity (von Stechow and Iatridou, 2008; Rubinsteyn, 2012) and to understand better the pragmatic difference between weak and strong – which is to say, non-extreme and extreme – deontic necessity statements. We summarize our results and discuss the prospects for future research in Section 4.

2 Semantic classes of modals

2.1 Standard classification of modals

Modal words have long, and correctly, been classified according to two major dimensions of meaning: their FORCE and their JUDGMENT TYPE (or simply TYPE, often informally called “flavor”). Modal logic traditionally distinguishes two levels of force, possibility and necessity, whereas linguists have noted further levels, such as weak necessity and slight possibility. Kratzer (1981, 1991) famously gives an analysis of several of these levels within a quantificational framework.

Turning to the (judgment) types, there is a wider range. Some well-known types include epistemic modality, deontic modality and related priority types, ability and circumstantial modality, to name a few. But this is not a closed class, since each of the named types can occur in an indefinite number of subvarieties; for example, among the deontics we can distinguish God’s law, national and international law, professional ethics, tribal custom, and classroom etiquette. Within deontic modality, we have different subtypes for each attitude holder (Alice’s desires, Bob’s desires, etc.), and possibly also different kinds of desire (things we want vs. things we hope for). Kratzer (1981, 1991, 2012) denies that these distinctions call for an analysis of modal auxiliaries in terms of lexical ambiguity; instead, she treats the modals as dependent on parameters of context which serve to narrow down the set of relevant possible worlds over which the modal quantifies in different contexts.

This framework has been the basis for studying many different kinds of modals in different languages, and has proven fruitful in understanding their linguistic and logical properties as well as crosslinguistic variation. However, a number of important puzzles and problems for the quantificational approach have accumulated over the years. One of these is one of the main issues addressed in this paper: the fact that modal expressions are often gradable. Kratzer’s treatment of levels of modal force provided a partial solution to this problem, but it fell short of a full, compositional analysis of natural language modal expressions. For example, the theory does not compose the meaning of *more likely than* using a plausible derivation of the comparative construction more generally. The grammar of gradable elements has been extensively studied in the literature on adjectives, but the theories of modality and gradability have not been put together until very recently (Portner, 2009; Yalcin, 2010; Lassiter, 2010, 2011; Klecha, 2012; Katz et al., 2012).

2.2 Scalar classification of modals

It is natural to think of gradable predicates as subclassifying objects into groups by measuring them along a lexically encoded dimension: for example, *long* classifies objects according to how they measure in length. There are competing views about how to formally implement gradability, some of which assume DEGREES as part of the semantic ontology, and others which do not. In this paper we adopt a degree-based approach to gradability, following Kennedy and McNally (2005); Kennedy (2007) and others. We do so for concreteness, and under the assumption that this choice does not affect our main claims about the distinction between extreme and non-extreme predicates, and modals in particular. As an entry point to the extensive literature on adjective meaning and gradability, including the debate about the semantic status of scales and degrees and the cognitive basis of these notions, we refer the reader to recent overview articles on these topics (Beck, 2011; Demonte, 2011).

Some of the most common reflexes of gradability crosslinguistically, exemplified in (7), are the ability of a predicate to occur in comparative constructions, to be modified by degree modifiers like *almost*, *very* or *slightly*, and to appear in degree *how*-questions.

- (7) a. The green vase is prettier than the brown vase.
- b. The glass is almost full.
- c. How long was the *Beagle*?

Languages may also allow modification of certain gradable predicates by measure phrases. *Long* in English allows this type of modification, whether in its comparative form (8a), or in its morphologically unmarked form (8b), known as the *positive form*.

- (8) a. The *Mayflower* was (about) 5 meters longer than the *Beagle*.
- b. The *Beagle* was 27 meters long.

Gradability has been most closely studied with respect to adjectives, and indeed most of the modal expressions we focus on in this paper belong to this category. It is well known, however, that other grammatical categories admit of gradability as well (Bolinger, 1972; Doetjes, 1997; Kennedy and McNally, 2005; Kennedy and Levin, 2008; Sassoon, 2007, 2013). The verbal elements *should* and *ought to* attest the cross categorial nature of gradability in the modal domain (see (6a) above and data below).

It is also noteworthy that not all adjectives pass the tests for gradability, or at least that some pass them more easily than others (Sapir 1944; Burnett 2014). Good candidates for non-gradability are adjectives like *true* and *pregnant* and privative adjectives like *fake*:

- (9) a. ?Your sister is more pregnant than this woman
- b. ?It is more true that crows are black than it is that apples are red.

While these examples are not ungrammatical, they require contextual support to be interpreted. As is often assumed, we take it that coercion (a semantically or pragmatically induced shift in meaning) is involved (see, for example, Kennedy 2007; Burnett 2011, 2014), although in this domain we do not find explicit theories of the coercion process on a par with those discussed in the aspectual domain (Moens and Steedman, 1988; de Swart, 1998).

The literature is somewhat ambivalent about the gradability status of one particular class of adjective, particularly important here, the extreme adjectives.

- (10) a. ?Godzilla is more gigantic than Mothra. (Morzycki, 2012, (10))
 b. ?The Tesla is more excellent than the Volt. (cf. Paradis 2001)

Both Morzycki and Paradis rate extreme adjectives as questionable in the comparative, while Rett (2008) assumes that they are grammatical. (As we'll see below, Morzycki discusses the fact that they are more acceptable in some contexts than others.)

Among the predicates we focus on here, we observe that the strong necessity modals (*must*, *crucial*, *certain*) are somewhat degraded in the comparative, unlike their weak counterparts:

- (11) a. You ought/?have to call Zoe more than (you ought/?have to call) Barbara.
 You should/?must call Zoe more than (you should/?must call) Barbara.⁶
 b. It is more important/?crucial to talk to Zoe than it is to meet her face to face.
 c. It is more likely/?certain that the sun will set tonight than it is that will rise tomorrow.

Quantitative data bears out our feeling that strong necessity modals are less natural than weak ones in the comparative: *important* and *likely* occur in the comparative much more frequently than *crucial* and *certain*, and this pattern is comparable to that observed with non-modal extreme/non-extreme pairs.⁷

2.2.1 Background on semantic classes of adjectives

Following Kennedy and McNally (2005, 349), we assume degrees as abstract representations of measurement; a scale S_Δ is a pair $\langle D, R \rangle$ consisting of a set of degrees D that are organized by a total order R along a dimension Δ . Gradable predicates relate the members of their domain (objects or, in the case of modals, propositions) to degrees on their lexically associated scale. The denotation of the predicate includes a measure function, μ , whose role is to measure objects on the relevant dimension. A gradable adjective like *long* thus denotes a relation between objects x that have a length and degrees of length d .

- (12) a. $\mu_{length} = [\lambda x : x \text{ has a length } . \text{ the degree } d \text{ such that } x \text{ measures } d \text{ in length}]$
 b. $[[\textit{long}]] = [\lambda x \lambda d . \mu_{length}(x) = d]$ (informally = $\lambda x \lambda d . x \text{ is } d\text{-long}$)

Comparative morphemes and degree modifiers place restrictions on the degree returned by the measure function, typically by relating it to another degree (Kennedy, 2007, 5). These morphemes are essentially the glue that derives appropriate truth conditions for sentences

⁶We note that some speakers find the examples with modal auxiliaries worse on the whole than those with semi-modals. The relative pattern of acceptability judgments reported in the text is stable across these types, however.

⁷The percent of occurrence in the comparative and mutual information scores between target items and the comparative in the British National Corpus (Davies, 2004) are as follows: *important*=6.92% (PMI=5.062), *crucial*=.98% (2.251), *likely*=16.38% (6.306), *certain*=.88% (2.087). As examples of non-modal adjectives: *big*=14.13% (6.093), *gigantic*=.25% (.283), *pretty*=3.37% (4.026), *gorgeous*=0.50% (1.282).

in which gradable predicates occur. In (13), assuming a simple-minded syntax, we see how the denotation for a phrasal comparative morpheme *-er/more* works:⁸

- (13) a. $\llbracket \text{-er/more} \rrbracket = \lambda G \lambda y \lambda x. \exists d \exists d' [d > d' \wedge G(x)(d) \wedge G(y)(d')]$
 b. The *Mayflower* was longer than the *Beagle*.
 c. $\llbracket (13b) \rrbracket = \llbracket [Mayflower \llbracket \text{-er long} \rrbracket \text{ than } Beagle] \rrbracket = \text{True iff}$
 $\exists d \exists d' [d > d' \wedge \mu_{length}(\mathbf{Mayflower}) = d \wedge \mu_{length}(\mathbf{Beagle}) = d']$

In order to account for gradable adjectives in their simplest, positive form, the grammar needs to provide some semantic glue, parallel to *-er/more ... than*, which can relate the measured degree to another degree not explicitly mentioned, the STANDARD. For some adjectives (namely, relative adjectives), the standard is contextually determined, while for others (absolute adjectives), the standard is lexically determined by the properties of the scale (Unger, 1975; Rotstein and Winter, 2004; Kennedy and McNally, 2005). A null morpheme POS (for “positive”) introduces the standard and expresses that the measured degree is at least as high (Cresswell, 1977; Bartsch and Vennemann, 1973; Bierwisch, 1989; von Stechow, 1984; Kennedy, 1997).

- (14) $\llbracket \text{POS} \rrbracket = \lambda G \lambda x. \exists d [d \geq std(G) \wedge G(x)(d)]$

Example (15a) could be judged true in the context of a conversation about old exploration ships, but it would likely be judged false in a discussion about the evolution of non-military ships in the past two centuries. Thus, *long* has a contextual standard in this case; $std(\llbracket long \rrbracket)$ is given by context. In contrast, (15b) has a lexical standard, and is true if the ship was completely full; hence $std(\llbracket full \rrbracket)$ is the maximum degree of fullness.⁹

- (15) a. The *Beagle* was long.
 b. The *Beagle* was full.

- (16) $\llbracket (15a) \rrbracket = \llbracket [Beagle \llbracket \text{POS long} \rrbracket] \rrbracket = \text{True iff}$
 $\exists d [d \geq std(\llbracket long \rrbracket) \wedge \mu_{length}(\mathbf{Beagle}) = d]$

The theory of scales and standards provides the basic toolset—dimensions, degrees, orderings, standards of comparison—for capturing the meaning of gradable predicates. Focusing specifically on adjectives, Rotstein and Winter (2004) and Kennedy and McNally (2005) argued that a linguistically relevant property of scales is whether they are OPEN or CLOSED, that is, whether or not they contain maximal and minimal elements. The full typology of structures ranges from totally open to totally closed scales and the two partly closed scales in between.

⁸The phrasal analysis of the comparative is coupled with a straightforward semantic interpretation (Heim, 1985), but it is unclear whether it is to be preferred over a more elaborate clausal analysis involving ellipsis (Lechner, 2004). See Bhatt and Takahashi (2011) and Beck (2011) for an overview of the debate.

The particular formalizations given in (12b), (13a), and (14) are not as simple as they could be, but are given in these ways to make them compatible with the presentation of Morzycki’s ideas below.

⁹Normal issues of granularity mean that, in different context, there can be variation in what counts as reaching the maximum or minimum on the scale. For example, a glass may count as full even if it could hold slightly more liquid. These uses are known as “loose” or “imprecise” (Kennedy and McNally, 2005).

- (17) a. $\langle D_{()} , R \rangle$ (Totally open scale)
 b. $\langle D_{\downarrow} , R \rangle$ (Lower closed scale)
 c. $\langle D_{\uparrow} , R \rangle$ (Upper closed scale)
 d. $\langle D_{\square} , R \rangle$ (Totally closed scale)

Patterns of modification and entailment have been used to argue that all four types of scales are represented among gradable adjectives.¹⁰ For example, compatibility with proportional modifiers (*half, mostly, most of the way*) is used as a diagnostic of totally closed scales.

- (18) Totally closed scale
 a. The glass is half/mostly full.
 b. Her eyes were half/most of the way closed.
 (Kennedy and McNally, 2005, 352)

Adjectives that do not accept proportional modifiers can be further subclassified using end-point modifiers. Some of these are said to diagnose the existence of a maximum degree (e.g., *100%, completely, fully*), while others have been used to diagnose the existence of a minimal degree (e.g., *barely, slightly*). Assuming in addition that antonym pairs have the same set of degrees but reverse orderings of these degrees, we can identify adjectives with all four scale types in (17):¹¹

- (19) a. Totally open scale
 i. Her brother is completely ??tall/??short.
 ii. Her brother is barely/slightly ??tall/??short
 b. Lower closed scale
 i. The pipe is fully ??bent/straight.
 ii. The pipe is barely bent/??straight.
 c. Upper closed scale
 i. The treatment is completely safe/??dangerous.
 ii. The treatment is slightly ??safe/dangerous.
 [(i) examples from Kennedy and McNally 2005, 355]

With this finer-grained theory of scale structure, we can say a bit more about the interpretation of gradable adjectives in the positive form. Kennedy (2007) argues that, when an adjective has a lexical standard, this lexical standard is always identified with an endpoint of a closed scale. Thus, *tall*, with its open scale, must use a contextual standard. In contrast, *full*, with its totally closed scale, takes a lexical standard given as the (upper) endpoint of its scale.¹²

¹⁰In this paper, we present only a subset of the data concerning modification patterns. See Rotstein and Winter (2004); Kennedy and McNally (2005) for additional tests.

¹¹On the details of the reversal of ordering relations and antonymy, see Sassoon (2007, 2010); Rett (2008).

¹²Whether an adjective with a totally closed scale has a maximum or minimum standard seems to be systematically related to its meaning in complex ways (Kennedy and McNally, 2005; Kennedy, 2007). *Loaded (with)* contrasts with *full* in having a minimum standard, although their meanings are very similar.

2.2.2 The scalar approach to modals

Many modals—both adjectives and auxiliaries—are grammatical in constructions that diagnose gradability (Portner, 2009; Yalcin, 2010; Lassiter, 2011; Klecha, 2012). Gradable interpretations affect the full range of modality (judgment) types mentioned in section 2.1: epistemic, priority, and as shown in (20), also ability.

- (20) After global warming takes hold, cockroaches will be more able to survive than people.

These observations are not new. As noted above, the idea that possibility and necessity are gradable notions has been part and parcel of formal semantic theories of modality since at least Kratzer 1981. Over the years a debate has developed over how to formally analyze gradability in the modal domain, and in recent years two main approaches have been developed. One approach has aimed to marry Kratzer-style modal semantics with the independently evolving scale-based theory that was introduced in the previous section. The other has argued for the rejection of the traditional quantificational approach in favor of an analysis of modality based on the mathematical formulation of probability. In the remainder of this section, we will briefly review some of the work along these lines.

Villalta (2000, 2006, 2008) makes one of the first attempts to articulate a fully scalar account of modals building on the Kratzerian semantics. As part of another ambitious project, namely explaining the selection of indicative and subjunctive complements by Spanish attitude verbs, she proposes that subjunctive-selecting predicates, including many of interest here (such as *important*, *necessary*, and *probable*) are similar to gradable adjectives in taking a degree argument. The degrees they reference are defined as sets of propositions that form equivalence classes according to the Kratzerian ordering of ‘better possibility’, which is used to compare propositions according to how desirable, necessary, or likely they are.¹³

Villalta took an important step in the direction of providing a fully scalar and compositional semantics of gradable modals, but she did not examine the properties of modal scales in detail. Moreover, her account is closely tied to Kratzer’s (1981) definition of comparative possibility, and as such it inherits problems that have been pointed out for that approach concerning the interaction of priorities and likelihood (e.g., Professor Procrastinate and the Miners Paradox; see Lassiter 2011 for a summary). Other scholars have worked to refine the system to account for this interaction, either by exploring how the interaction of priorities and likelihood might affect the content of ordering sources (Katz et al. 2012; von Stechow 2012; Charlow 2013), by positing alternative definitions of comparative possibility (Kratzer, 2012), or by adding a contextual parameter that may affect the ordering of worlds depending on an information state (Cariani et al., 2013).

A second type of response has been to abandon the pre-orders of Kratzer’s semantics (i.e., ordering sources and the comparative possibility relation they induce) in favor of a weak ordering of propositions based on probability and utility. Swanson (2006) and Yalcin (2007, 2010) motivate this approach for explicit expressions of probability like *probable*, *probably*, *likely*, while Goble (1996); Finlay (2009, 2010); Lassiter (2010, 2011) apply it more generally to a wider range of epistemic and deontic expressions. Lassiter in particular highlights the

¹³Portner (2009) sketches an approach to probability expressions based on the same method of constructing degrees.

natural connection between the probabilistic semantics he adopts for gradable epistemic modals and the scalar theory of non-modal gradable adjectives, assuming that all epistemic modals, adjectives and auxiliaries alike, are interpreted on a single scale, the totally closed scale of ordinary probability, but differ in their associated standards. As pointed out by Klecha (2012), however, this assumption runs counter to the principle of scalar semantics that standards are determined by the adjective’s scale structure and other properties inherent to an item’s meaning. Adjectives that are interpreted on totally closed scales, in particular, have been argued not to take relative standards (Kennedy, 2007).

For the interpretation of deontic and buletic modals, Lassiter (2011) proposes a scale that measures the expected utility of a proposition. According to Lassiter, both weak necessity *should/ought* and strong necessity *must*, as priority modals, are interpreted relative to a “very high” contextual standard in the positive form. (The lexical entry also states other conditions on the probability of the proposition and its alternatives.) Given the differences between this treatment for priority-type modality and the treatment of epistemic modality outlined above, a lexical ambiguity seems to be postulated for items with both senses like *must*.

Klecha (2012) analyzes gradable modality in a way which borrows from both of the above approaches. Focusing on epistemics, he denies that *possible* is a scalar adjective to begin with (see also Herburger and Rubinstein 2014), and argues that *likely* and *certain* are interpreted on different scales, despite the entailment relation between them. The crucial observation is that the modification patterns of these modals cannot be explained with reference to a single scale. *Likely* contrasts with *certain* in that it cannot occur with modifiers that diagnose upper closed or totally closed scales (**completely likely*, **half likely*).¹⁴ Moreover, he argues that *certain* cannot be interpreted on the totally closed probability scale given that it accepts high-value proportional modifiers but not low-value ones (e.g., *It is 95%/*30% certain*).¹⁵ Because Klecha does not analyze the adjectives based on a single scale, the entailment from *certain* to *likely* is not automatic, and must be stipulated in the meanings.

The debate between Lassiter and Klecha leaves us somewhat uncertain about the scales and standards for epistemic adjectives, as neither approach seems to give a full explanation of their modification patterns. In the next section, we will argue that another parameter of classification can help (in combination with the properties related to scales and standards which we surveyed in this section) to elucidate the properties of gradable modals.

2.3 Extreme modals

The scalar approach to modal semantics has both increased our understanding of the properties of modal expressions and inspired new ideas for their semantic analysis. However, there are differences between classes of gradable adjectives not captured by the idealized perspective reviewed so far. One such difference is exemplified by the pair *pretty/gorgeous*,

¹⁴See Portner (2009) for similar arguments concerning *probable*.

¹⁵Stephen Finlay (p.c.) suggests that *95% certain* has to be idiomatic and non-compositional. Klecha (2012) argues that a compositional analysis is possible given an analysis of the adjective that relates it to an appropriate non-probabilistic scale.

both members of which presumably are related to a basic scale of beauty. In this section, we argue that weak/strong modals like *important/crucial* and *should/must* (holding constant their judgment type) differ in the same way as *pretty/gorgeous*. Like *gorgeous*, *must* and *crucial* are EXTREME EXPRESSIONS (Bolinger, 1967; Paradis, 2001; Rett, 2008; Morzycki, 2012), describing arguments which have the property denoted by their non-extreme counterparts (*pretty*, *should*, *important*) to a degree higher than would normally be conveyed by the non-extreme words. Extreme adjectives have a number of properties which show that they have scale-based meanings which are intimately related to the semantics of non-extreme expressions, but they are also distinct both semantically and pragmatically from ordinary gradable adjectives.

Section 2.3.1 reviews the relatively small literature on extreme adjectives, outlines some of their important properties, and summarizes one prominent analysis of them in the literature. Then in Section 2.3.2, we use these properties to show that strong necessity modals should be classified as extreme expressions, while weak necessity modals are non-extreme. These points set the stage for our own effort to provide a semantic analysis for deontic and other priority necessity modals in Section 3.

2.3.1 Extreme adjectives

In an often-cited early discussion, Paradis (2001) states that extreme adjectives involve scales in their meaning (diagnosed by the fact that they come in antonym pairs), but that they only occur with a subset of modifiers associated with gradability. Specifically, they are used with what she calls “totality modifiers” (*absolutely*, *totally*, *utterly*), but not with approximators (*almost*). According to Paradis, some but not all people accept them with the comparative and superlative.

- (21) a. The salsa is absolutely terrible.
 b. *The salsa is almost terrible.
 c. ?The salsa is more terrible than the guacamole.

To explain this pattern Paradis says that extreme adjectives “represent the ultimate point of a scale”, even though they can be coerced into meanings which apply to elements not at the maximum point of the relevant scale.

Paradis’ intuitions about the semantics of extreme adjectives are appealing, but it does not seem correct to say simply that they describe their arguments as being at the endpoint of their scale. For example, (21a) does not mean that the salsa is of the highest possible (or even the highest actually realized) degree of badness. There certainly are more terrible things, and quite possibly more terrible condiments. Later work on extreme adjectives has therefore pursued the intuition that extreme adjectives utilize subscales related to the more normal scales of their non-extreme counterparts.

Rett (2008) assumes that extreme adjectives utilize totally closed scales (both upper and lower bounded) which are subscales of their non-extreme counterparts. *Terrible* would utilize a scale which begins at some threshold t and reaches to the top of the scale. The idea that extreme adjectives’ scales are totally closed is motivated, it seems, by two factors: first, that there must be a lower bound defined by t , and second that there must be an upper

bound to explain the appearance of maximality modifiers like *totally*. However, against this vision is the fact that modifiers which normally go with lower bounded scales, like *barely*, are not acceptable (**barely terrible*). Moreover, as we have seen, maximality operators are not interpreted as expected if they diagnose closed scales for extreme adjectives. Just as (21a) does not mean that the salsa is maximally terrible, (22) does not mean that Janie is at the highest degree of beauty, as it would if *totally* placed the adjective’s argument at the top of the top portion of the beauty scale; it is quite possible that even within the context (a fashion show) there are individuals recognized to be even more gorgeous. Rather, it means something like that she is without question in the range of gorgeous individuals.¹⁶

(22) Janie is totally gorgeous.

Given these points about *barely* and *totally/absolutely*, it seems that we do not have clear evidence concerning the scale type for extreme adjectives. Morzycki (2012) presents an analysis of extreme adjectives which, though he does not emphasize the fact, allows for extreme adjectives to be either bounded or unbounded at both top and bottom. His approach is that scalar adjectives are interpreted with respect to a contextually salient set of degrees D_C on their scales; non-extreme adjectives relate to values within D_C , while extreme adjectives relate to values above all of the values in D_C .^{17,18} The standards follow suit: the standard for a non-extreme adjective is in D_C , while that for an extreme adjective is above D_C .

One of Morzycki’s major contributions was a discussion of a range of properties which differentiate extreme from non-extreme adjectives and which also can serve as diagnostics for whether a term is an extreme expression. We will look at five such properties in detail:

1. Extreme adjectives occur with extreme degree expressions, or EXTREME MODIFIERS, like *flat-out*, *positively* and *downright*, while non-extreme adjectives (even those with maximum standards) are less natural.¹⁹

(23) a. ??flat-out good
 b. flat-out excellent
 c. ??flat-out straight

¹⁶As noted by Stephen Finlay (p.c.), the example also has a reading ‘Janie is gorgeous in every respect’. This shows that *totally gorgeous* can grade an object on a scale other than the scale of *gorgeous*. Many adjectives are like this, and we note that *totally pretty* has a similar reading.

¹⁷“Salience” could be interpreted in different ways, and Morzycki glosses it in this context as “degrees . . . that we regard, for the purposes of the discussion, as reasonable candidates for values we might want to consider” (Morzycki, 2012, 582-583). The issue of how to properly characterize the pragmatics of “salient” vs. “non-salient” degrees is difficult and closely connected to the analysis of exclamationatives (see for example Zanuttini and Portner 2003; Miró 2006; Rett 2008).

¹⁸Whether the set of extreme degrees ($D - D_C$) is open and unbounded, open and bounded, or closed at the top simply depends on the scale S . As for the lower portion of the extreme subscale, whether it is lower closed or open (but bounded) depends on how the boundary between the extreme and non-extreme subscales is specified. If the set of extreme degrees is open on the top, we would expect a maximizer like *totally* not to have its core scalar use; this may be what’s going on with (22). If it is open on the bottom, we do not expect an adjective like *barely* to be felicitous in its normal use, and this may explain the relative oddness of *?barely gorgeous*.

¹⁹Morzycki points out that Cruse (1986) can be seen as having first identified this property.

Though it's not common to find a maximum standard adjective with an extreme counterpart, in such cases, only the extreme adjective is natural with these modifiers:

- (24) a. ??downright full
b. downright brimming

Though there is a clear contrast, between *good* and *full* on the one hand, and *excellent* and *brimming*, on the other, it seems that there is a difference in acceptability with extreme modifiers among the non-extreme adjectives between the relative standard *good* and the maximum standard *full*. The latter seems better with an extreme modifier than the former, in particular if it is not expected that the container is full (“We’ve had bad luck fishing cod lately, but today our nets were downright full!”).²⁰

2. The near inverse of the previous property is that extreme adjectives strongly resist modification with *very*. They contrast in this respect with relative adjectives such as *good* and absolute ones like *straight*.

- (25) a. very good
b. ??very excellent
c. very straight

- (26) a. very full
b. ??very brimming

Note that Kennedy and McNally (2005, 371) argue that coercion is involved when *very* occurs with a minimum or maximum standard adjective; the key point here, though, is that these nevertheless differ from extreme adjectives.

3. Extreme adjectives are often degraded the comparative (as in (21c), repeated below as (27a), but they do occur and improve with *even*, (27b).

- (27) a. ?The salsa is more terrible than the guacamole.
b. The salsa is even more terrible than the guacamole.

4. Though they resist the comparative, extreme adjectives fully accept the equative construction.

- (28) The salsa is as terrible as the guacamole.

5. Extreme adjectives are unacceptable when placed in comparatives with their non-extreme counterparts.

²⁰Even *good* can be used as an extreme expression, but more contextual support is needed. Morzycki labels adjectives which can become extreme expressions in context CONTEXTUAL EXTREME ADJECTIVES, a class which might include all gradable adjectives if we grant an indefinite amount of contextual support. Our discussion here focuses on the LEXICAL EXTREME ADJECTIVES, those which behave as extreme in any context in which they are used.

- (29) a. *The salsa is more terrible than the guacamole is bad.
 b. *The salsa is worse than the guacamole is terrible.
- (30) a. *The bucket is more brimming than the cup is full.
 b. *The cup is fuller than the bucket is brimming.

As Morzycki notes, the data in (29)-(30) contrast with other examples where two adjectives on related scales are compared:

- (31) The field is longer than it is wide.

Morzycki (2012) lists several other properties which distinguish extreme and non-extreme adjectives, but does not provide a formal analysis of them. Briefly, these properties are: (i) Extreme adjectives can readily be intensified by prosodic means (Cruse 1986); (ii) in discourse, an extreme adjective is perceived as disagreeing (not agreeing) with an understatement; and (iii) extreme adjectives are readily used in hyperbole. We will not discuss these properties further.

Recall that within Morzycki’s analysis, non-extreme adjectives associate their argument with a contextually salient degree, while extreme adjectives associate their argument with a degree above the contextually salient portion of the scale. He assigns *big* and *gigantic* the following meanings, where C refers to the contextually salient set of degrees (= D_C above):

- (32) a. $\llbracket big_C \rrbracket = \lambda x \lambda d . d \in C \wedge x$ is d -big
 b. $\llbracket gigantic_C \rrbracket = \lambda x \lambda d . d > max(C) \wedge x$ is d -big

This proposal allows Morzycki to provide explanations for the patterns involving extreme and non-extreme adjectives observed above.

1. Extreme modifiers consider the interpretation of the adjective relative to a new set of degrees C^+ which is an extension of the salient set C , and they presuppose that the contextual standard is within the extreme portion of C^+ , namely $C^+ - C$.

- (33) $\llbracket downright_C \rrbracket = \lambda a \lambda x : std(a_{C^+}) \in (C^+ - C) . \exists d[d \geq std(a_{C^+}) \wedge a_{C^+}(x)(d)]$

In the positive construction, the contextual standard of an extreme adjective will have to be within $C^+ - C$, since if it were in C , it would have no effect; thus, extreme adjectives naturally satisfy the presupposition of the extreme modifier. In contrast, in the positive construction the standard for a non-extreme adjective will be in C , and so for a non-extreme adjective to satisfy the presupposition of an extreme modifier, the standard relative to C must be different from the standard relative to C^+ . In other words, the shift of context produced by the extreme modifier must cause a shift of standard for the adjective. Morzycki’s idea is that this happens precisely in the cases where a non-extreme adjective is coerced into compatibility with an extreme modifier. For example, when *positively full* is acceptable, it implies that the argument is full relative to a standard of fullness which is beyond the range of degrees of fullness previously thought to be relevant in the conversation (“We didn’t think it could be literally full, but it turns out it was!”).

2. Morzycki (2012, 101) analyzes *very+ADJ* as asserting that the individual it takes as argument is at the top of the contextually salient set of degrees of the scale of *ADJ*:

$$(34) \quad \llbracket \textit{very}_C \rrbracket = \lambda a \lambda x . \exists d [d \geq \textit{std}(a_C) \wedge \textit{small}(\textit{max}(C) - d) \wedge a_C(x)(d)]$$

The infelicity of *??very gigantic* is thus explained on this analysis through the contradictory requirements it places on the measure of the object described; specifically, it asserts that the argument is big to a degree which is both below $\textit{max}(C)$ (via *very*) and above $\textit{max}(C)$ (via *gigantic*).

3. Morzycki’s explanation for why extreme adjectives often resist the comparative is complex, but the basic idea is that it is not natural to compare degrees which are not in the salient range. Ideally, we would have a theory of salience of degrees to provide a more specific reason why this would be so,²¹ but intuitively it makes sense that if the salsa is terrible, it is so overwhelmingly bad that it might be difficult or pointless to decide whether it is better or worse than the (also terrible) guacamole. After all, if it’s terrible, you know all you need to know: that you’re not going to eat it. But this difficulty or pointlessness can be overcome pragmatically, as in the context of a Mexican cooking class where an instructor’s detailed feedback on the student’s failed salsa and guacamole would be relevant. Moreover, it seems that *even* (specifically the presupposition it triggers) supports the kind of context in which extreme degrees can be compared; see Morzycki (2012, 4.3) for further discussion.

4. In contrast to comparatives, Morzycki argues that the semantics for equatives does not involve comparison. For example, (28) states that there is a degree of badness beyond the salient set C which is the degree of the salsa, and which is the degree of the guacamole. Seen this way, the equative only asserts the existence of a single degree, and this is why it does not show the pragmatic incongruity sometimes seen with comparatives.

Morzycki’s explanation for the comparative and equative facts may not be entirely convincing, for two reasons: it’s not clear why comparing extreme degrees in the comparative should be less felicitous than referring to them in the equative, and the formulation of the equative semantics without comparison is non-standard, and so might be seen as ad hoc. Our intuitions about the comparatives and equatives are slightly different from his. It seems to us that, in a simple context, any distinctions above C along the relevant dimension are collapsed, and so there is only a single extreme degree. In such contexts, a sentence comparing two extreme degrees would be logically false, unless they are equated. Context could, however, articulate the portion of the scale above C , making multiple extreme degrees available, and in such a context, a comparative could be informative. Further work will be needed to determine the best analysis of the behavior of extreme adjectives in comparatives and equatives.

5. In Morzycki’s analysis, comparisons involving an extreme and a non-extreme adjective, like (29), have trivial truth conditions. (29a) states that salsa’s degree of badness is

²¹Our analysis will actually be framed in terms of a more explicit explanation of what “salience” amounts to for modal adjectives.

beyond *C*, and moreover beyond the guacamole’s degree of badness, which is within *C*. This must be the case. (29b) states that the salsa’s degree of badness, which is within *C*, is above the guacamole’s degree of badness, which is beyond *C*. This cannot be the case.

2.3.2 Strong necessity modals are extreme expressions

Strong necessity modals, but not weak ones, have the properties associated with extreme adjectives:

1. Strong necessity modals readily take extreme modifiers:²²

- (35)
- a. Susan positively/flat-out/downright must/*should call her mother.
 - b. It is positively/flat-out/downright crucial/*important for Mary to call her mother.
 - c. It is positively/flat-out/downright certain/*likely that Mary will call her mother.

2. Strong necessity modals are less natural with *very*:²³

- (36)
- a. Susan very much *must/should call her mother.
 - b. It is very *crucial/important for Mary to call her mother.
 - c. It is very *certain/likely that Mary will call her mother.

3. Strong necessity modals resist the comparative, and in the comparative, they are better with *even*:²⁴

- (37)
- a. Susan ?must/should call her mother more than she ?must/should call her father.
 - b. It is more ?crucial/important for Mary to call her mother than her father.

²²In the corpora we have examined, extreme modifiers are extremely rare with the modal elements on which we have focused our discussion. In the BNC, there is one occurrence of *positively must* and no other combinations of interest. In COCA the raw counts and percentages of *flat-out/positively/downright* with the modal elements are: *must*=6 (.003%), *should*=3 (.001%), *certain*=3 (.027%), *likely*, *crucial*, and *important*=0. Overall, the quantitative trends conform to our hypothesis, but the numbers are too small to count as strong evidence by themselves.

²³Corpus patterns in the BNC support this difference between adjective classes. Percentage occurrence with *very* and association measures are as follows: *crucial*=.21% (PMI=.801), *important*=5.57% (5.556), *certain*=.08% (-.552), *likely*=1.146% (3.28); non-modal adjectives for comparison: *gorgeous*=.50% (1.282), *pretty*=3.37% (4.026).

We used *very much* in place of *very* alone with the auxiliaries (36a). The intuition data in this case confirm our predictions, but the combinations are very rare. Neither occurred in the BNC, while the raw counts in COCA were *very much must*=0 and *very much should*=1 (<.001%). The single example is interesting, though, and fits our analysis: *You should know – actually, for complete, suspenseful enjoyment of the film, you very much should not know, but the word is out, so we’re obliged to tell you – that Heavenly Creatures is based on a notorious murder case.* (‘A Heavenly Trip Toward Hell’, *Time*, 1994).

²⁴Corpus data related to the modals in (37) was provided in footnote 7 above.

- c. It is more ?certain/likely that Mary will call her mother than her father.
- (38)
- a. Susan must call her mother even more than she must call her father.²⁵
 - b. It is even more crucial for Mary to call her mother than her father.
 - c. It is even more certain that Mary will call her mother than her father.

In evaluating the example with *must*, it is important to keep in mind that the relevant reading is one where degrees of obligation are being compared; the example has a grammatical reading where frequencies of calling are being compared.

Notice that *more crucial* and *more certain* are not entirely unacceptable in (37), but this was the case with the non-modal extreme adjectives like *more terrible* in (21) as well. When they are usable, they seem synonymous with the versions containing *even* in (38).

4. Both strong and weak necessity modals are acceptable in the equative:

- (39)
- a. Susan must call her mother just as much as she must call her father.
Susan should call her mother just as much as she should call her father.
 - b. It is as crucial/important for Mary to call her mother as it is for her to call her father.
 - c. It is as certain/likely that Mary will call her mother as it is that she will call her father.

5. Strong and weak necessity modals cannot be placed in the comparative with one another:

- (40)
- a. *Susan must call her mother more than she should call her father.
*Susan should call her mother more than she must call her father.
 - b. *It is more crucial for Mary to call her mother than it is important for her to call her father.
*It is more important for Mary to call her mother than it is crucial for her to call her father.
 - c. *It is more certain that Mary will call her mother than it is likely that she will call her father.
*It is more likely that Mary will call her mother than it is certain that she will call her father.

We have shown that the strong necessity modals pattern with extreme adjectives, and conclude that they should be included within this class. The weak necessity modals, in contrast, should be treated as non-extreme expressions.

²⁵This sentence is somewhat degraded in comparison to the examples with *crucial*, *certain*, but considerably better than the corresponding (36a) without *even*.

2.4 Summary and plan

The observation that strong necessity modals are like extreme adjectives is important because it points the way towards a better analysis of the strong/weak necessity distinction and, more broadly, the gradability of modal expressions. Before we turn to a formal analysis of the strong and weak priority modals *must/should* and *crucial/important*, we think it's useful to think about the intuitive basis for assigning a scale structure to deontic expressions divided into salient/non-extreme vs. non-salient/extreme portions. What would the degrees be, and how would they be ordered? And what would be the difference between those “within the salient range” and those “beyond the salient range”?

As to the first question, we propose that degrees of necessity correspond to the weight of the requirements which favor a particular proposition. For example, ‘you don’t murder’ has a higher degree of necessity than ‘you do the dishes’, because the requirements favoring it are weightier. And as to the second question, the idea is that the salient degrees are those assigned to propositions which are worth discussing, and they are worth discussing because they are based on requirements the interlocutors might or might not follow. Non-salient degrees, in contrast, are based on settled or unquestioned requirements. For example: On an elementary school assignment students have a hard requirement that the student write his name on the paper. With no name, the assignment will not be graded at all, an outcome these good students would never contemplate, and so the degree of necessity of ‘I write my name’ is “beyond the salient range”. Assuming it is graded, though, points will be assigned according to a rubric with elements like “Illustrate your report with a hand-drawn, colored picture: 5 points”. If the student fails to color the picture, he might receive only four points based on this criterion. For the students, the alternatives ‘I draw but do not color the picture’ and ‘I draw and color the picture’ are worth discussing, and hence the degrees of necessity of the alternatives are salient in the context. In this context, writing your name has an extreme degree of necessity, while coloring your picture has a non-extreme degree.

In the next section, we work towards a formal analysis of necessity modals which incorporates the intuition that the scale of necessity is divided into a salient, or non-extreme, part and a non-salient, or extreme, part. This approach not only provides an explanation for the empirical differences between strong and weak modals enumerated in this section. It also promises a pragmatic grounding for the difference in the notion of salience. On our approach, *must* and *should*, and *crucial* and *important*, are all semantically necessity modals; they differ in that the strong ones paint an alternative as necessary and reject further discussion, while the weak ones portray it as being necessary but still up for discussion.

In terms of issues of scales and standards, our approach can be seen as pursuing a synthesis of the earlier ideas of Lassiter and Klecha.²⁶ In considering how to represent the relationships between semantically related modal adjectives, these authors represent two reasonable strategies. Lassiter follows the strategy of putting the pair *likely* and *certain* on the same scale, and thus makes the differences between them fairly minimal (literally, a matter of degree). However, as we have seen, the differences between extreme and non-extreme adjectives amount to more than that. Klecha follows a different strategy, where *likely*

²⁶It should be kept in mind, though, that their debate focused on epistemic modals, and so we must extrapolate a bit to attribute them contrasting positions concerning deontics.

and *certain* are associated with different scales, stipulating a relation between those scales. However, because the relation between the scales is not motivated by any broader picture of adjective semantics, this leaves an explanatory gap. Our approach puts the extreme and non-extreme pairs on subscales of the same basic scale. Hence, as in Lassiter’s analysis, the members of the pairs differ in degree. However, they are on pragmatically different parts of this scale, and so there is a sense in which Klecha is right as well in saying they use different scales. Crucially, the difference between the extreme and non-extreme parts of modal scales is based on a motivated pragmatic distinction and is parallel to what is observed with many non-modal adjectives.

3 Analysis

The intuition behind theories of extreme adjectives is that the argument somehow goes beyond the normal or expected range of its scale, and as we have discussed, this intuition is apt for extreme modals; for example, just as downright huge things are big to a degree beyond the expected, the positively crucial actions are so important as not to be in need of debate. We aim to capture this intuition by defining a general scale-type labeled the SCALE OF NECESSITY, where degrees are modeled as sets of propositions which are necessary relative to subsets of an ordering source. This scale of necessity will be used as the basis for the analysis of both weak and strong necessity modals. The scale of necessity may be compared to a general scale of size, which would provide the core of the analysis of both *big* and *huge*.

Our formulation builds on existing work on strong and weak necessity within a premise semantics framework. Specifically, von Stechow and Iatridou (2008) proposed an important analysis of *must* and *should* which enriched the palette of conversational backgrounds from the two (modal base and ordering source) familiar from Kratzer (1981), to a set consisting of a modal base and at least two ordering sources, one primary and one secondary. With a strong necessity modal, only the primary ordering source is relevant, and it determines the modal’s domain of quantification in the way familiar from Kratzer’s work. With a weak necessity modal, the secondary ordering source operates on the set of worlds which are relevant according to the modal base and primary ordering source, to further restrict the domain of quantification. Specifically, the semantics first identifies the set of worlds compatible with the modal base $m(w)$ which are best-ranked according to the primary ordering source, $o_1(w)$, and then picks out the best-ranked worlds among those according to the secondary ordering source $o_2(w)$, as in (42). We assume, following von Stechow and Iatridou (2008), that there is always a set of “best” worlds among those ordered by an ordering source.²⁷

$$(41) \quad \begin{aligned} \text{a. } & u \leq_O v \text{ iff } \{p \in O : v \in O\} \subseteq \{p \in O : u \in O\} \\ \text{b. } & \text{Best}(M, O) = \{v \in M : \neg \exists u [u <_O v]\} \end{aligned}$$

$$(42) \quad \text{Best}(M, O_1, O_2) = \text{Best}(\text{Best}(M, O_1), O_2)$$

In this way, they account for the fact that strong necessity modals logically entail their weak counterparts, and provide some degree of intuition about what the difference between them

²⁷See Katz et al. (2012) for an alternative way of combining multiple ordering sources. The definitions cite the premise sets given by the modal base and ordering source(s) at the world of evaluation, e.g. $M = m(w)$.

amounts to.

Although they advance our semantic understanding, von Stechow and Iatridou do not provide much explanation of the crucial difference between primary and secondary ordering sources (see their footnote 11). Building on their work, Rubinstein (2012) proposes that this difference is to be understood in terms of pragmatics, where the difference between the two ordering sources tracks the distinction between what is treated as non-negotiable versus what is treated as negotiable. The primary ordering source, in these terms, contains priorities that the speaker presents as ones she is not willing to give up, while the secondary ordering source contains priorities that might give way in the face of other factors.²⁸ We can illustrate Rubinstein’s analysis by considering the examples at the beginning of the paper:

- (43) a. Our uninsured citizens should get insurance.
 b. Our uninsured citizens must get insurance.

The sentence in (43a) is appropriate in a context in which the speaker is contributing to a negotiation about health care reform, while (43b) is appropriate when the speaker wants to rule out further debate. Combining these ideas with the semantics from von Stechow and Iatridou (2008), the prediction is that *must*(ϕ) says that ϕ follows from non-negotiable requirements, while *should*(ϕ) says that ϕ follows taking into account requirements about which the speaker allows negotiation.

Our analysis connects the semantic and pragmatic differences between primary and secondary ordering sources to the conceptualization of scale structure developed in work on gradable and extreme adjectives. Specifically, we propose that the extreme portion of the scale (the non-salient portion) is defined only in terms of the priorities in the primary ordering source. This makes sense of the intuition that in the extreme portion of the scale, distinctions between degrees are not salient. They are not salient because any such distinctions would require treating one non-negotiable requirement as more important than another. The non-extreme (salient) portion of the scale, in contrast, is defined by taking into account both the primary and secondary ordering sources, and given Rubinstein’s understanding of the secondary ordering source, this captures the intuition that distinctions within this portion of the scale are salient. They are salient because they are based on priorities which the speaker treats as comparatively important, but not set in stone.

Formally speaking, the non-extreme part of the scale is derived by considering subsets of the secondary ordering source $o_2(w)$. In what follows f is a contextually provided function from sets of premises o to sets of non-null subsets of o , where the members of $f(o)$ are totally ordered by \subseteq :

$$(44) \quad S_{c,w} \text{ is only defined if } f(o_2(w)) \text{ is defined, for every world } w. \text{ When defined,}$$

$$S_{c,w} = \langle D_w, \leq \rangle: \quad D_w = \{ \{ p : \text{Best}(m(w), o_1(w), o') \subseteq p \} : o' \in f(o_2(w)) \}$$

$$\leq = \{ \langle d_1, d_2 \rangle : d_1, d_2 \in D \wedge d_1 \supseteq d_2 \}$$

According to (44), a degree in D is a set of propositions, specifically the set which are best

²⁸Stephen Finlay (p.c.) makes the interesting suggestion that the priorities in o_2 might be secondary not because they are seen as unsettled, but rather because it is at issue whether they are achievable. This suggestion might allow us to make a connection between this analysis and the information sensitivity of deontic modals (Kolodny and MacFarlane 2010; Charlow 2013; Cariani et al. 2013; Cariani to appear).

according to the modal base, primary ordering source, and a particular subset of the secondary ordering source. Degrees are ordered by the superset relation, and because $f(o_2(w))$ is totally ordered by assumption, the set of degrees is also totally ordered.

We construct an extension of the scale $S_{c,w}$ to include extreme degrees in a similar way by considering subsets of the primary (non-negotiable) ordering source $o_1(w)$. The complete scale which includes the extreme extension is $S_{c,w}^+$:

$$(45) \quad S_{c,w}^+ \text{ is only defined if } f(o_1(w)) \text{ and } S_{c,w} \text{ are defined, for every world } w. \text{ When defined,}$$

$$S_{c,w}^+ = \langle D_w^+, \leq^+ \rangle: \quad D_w^+ = D_w \cup \{ \{ p : Best(m(w), o') \subseteq p \} : o' \in f(o_1(w)) \}$$

$$\leq^+ = \{ \langle d_1, d_2 \rangle : d_1, d_2 \in D^+ \wedge d_1 \supseteq d_2 \}$$

Notice that $S_{c,w}^+$ is defined using the two-place version of *Best*, since the secondary ordering source is not relevant. The extreme part of the scale $S_{c,w}^+$ is $\langle (D_w^+ - D_w), \leq^+ \cap (D_w^+ - D_w)^2 \rangle$.

(46a) defines a measure function μ_Σ which will allow us to give a scalar semantics for necessity modals. The lexical entries in (46b-c) capture the two levels of strength observed with necessity modals: weak ones, specifically *should* and *important*, are measures of propositions of non-extreme degree, whereas strong ones, namely *must* and *crucial*, are restricted to extreme values.

- (46) a. **Measure function:** For any proposition p , world w , and necessity scale Σ , $\mu_\Sigma(p, w)$ is the highest degree d in D_w^+ such that $p \in d$.
- b. Weak necessity modal: $\llbracket N \rrbracket^c = [\lambda p \lambda d \lambda w . d \in D_w \wedge d = \mu_\Sigma(p, w)]$
- c. Strong necessity modal: $\llbracket N^+ \rrbracket^c = [\lambda p \lambda d \lambda w . d \in (D_w^+ - D_w) \wedge d = \mu_\Sigma(p, w)]$

We can see how the sketched system works with a concrete example. Suppose that the speaker is supportive of three goals: eating ramen (R), seeing Casablanca (C), and meeting up with their friend who's in town for the weekend (F). $f(o_2(w))$ reflects how strongly held these goals are, and can be construed as a series of sets of goals which result from successively dropping the weaker ones. Assume that $f(o_2(w)) = \{ \{ F \}, \{ C, F \}, \{ R, C, F \} \}$, reflecting that F is the most important and R the least, and also assuming (for the time being) that the three goals are compatible. These goals require going to different parts of town: F requires going to Georgetown; C requires going to Silver Spring; and R requires going to Adams Morgan. In this context, 'We go to Georgetown' has the maximal non-extreme degree on the scale of necessity. 'We go to Silver Spring' has an intermediate degree, and 'We go to Adams Morgan' has the lowest degree. *It is more important that we go to Silver Spring than Adams Morgan* is therefore true.

We assume that unmodified necessity modals combine with their argument via the covert degree specifier POS in (14) (Kennedy, 2007). In parallel to how POS interacts with ordinary relative adjectives, when it combines with a necessity modal, it will require the prejacent to exceed a contextually determined standard. Plausibly, the contextual standard is set at the degree of 'We go to Georgetown', and in such a context *It is important that we go to Georgetown* is true but *It is important that we go to Silver Spring/Adams Morgan* are false.

This analysis makes some good predictions: Since extreme degrees are constructed as subsets of non-extreme degrees in (44)-(45), our analysis derives the entailment relations

between strong and weak necessity modals (5). It derives the meaning proposed for *should* and *must* by von Fintel and Iatridou (2008) and Rubinstein (2012) as special cases: the contextually specified degree for *should* is derived from a non-null subset of o_2 , while that specified for *must* is derived using no premises from o_2 . In addition, Morzycki’s (2012) explanations of the properties of extreme adjectives and modifiers directly carry over to the present account. For example, *very* requires a degree in the non-extreme range and *positively* requires one in the extreme range.

The way of constructing the scales in (44)–(45) has a serious logical flaw, however. It does not allow that incompatible propositions can have the same degree of necessity. Change the scenario so that eating ramen and seeing *Casablanca* are incompatible yet equally strong goals. As a result, going to Adams Morgan (to have ramen) and going to Silver Spring (to see *Casablanca*) are equally important, but we cannot do both. In this kind of scenario, $f(o_2(w))$ is the set of premise sets $\{\{F\}, \{R, C, F\}\}$, from which the premise set $\{R, C, F\}$ generates two disjoint maxima in the ordering, worlds where we eat ramen (in Adams Morgan) and worlds where we see *Casablanca* (in Silver Spring). (In both sets of worlds, we also go to Georgetown to see the friend.) Neither destination is a necessity with respect to $\{R, C, F\}$, so neither has any degree of necessity. This is incorrect.

We would like to be able to use both of the maxima defined by $\{R, C, F\}$ to associate propositions with the same degree of necessity. In other words, because eating ramen entails going to Adams Morgan, and seeing *Casablanca* entails going to Silver Spring, $\{R, C, F\}$ should determine the same degree for going to Adams Morgan and going to Silver Spring. One approach would be to have f select not just subsets of the ordering source, but sets of consistent subsets of the ordering source, for example $f(o_2(w)) = \{\{\{F\}\}, \{\{R, F\}, \{C, F\}\}\}$. Here, the second member encodes the fact that $\{R, F\}$ and $\{C, F\}$ are equally important collections of goals. For each set of sets of premises (e.g., $\{\{R, F\}, \{C, F\}\}$), we would then collect the necessities with respect to each member ($\{R, F\}$ and $\{C, F\}$) to define a degree of necessity. Going to Adams Morgan and going to Silver Spring would thus have the same degree of necessity (and this degree would be lower on the scale than the degree of going to Georgetown).

As an alternative to the above, we get a similar effect by allowing inconsistent premise sets to be selected by f , and defining degrees by looking at all of the maxima in the world-ordering defined by each premise set. In particular, the following definition might work, where $GP(A, B, C)(p)$ indicates that p is a “good possibility” (in the sense of Kratzer 1991) with respect to A , B , and C :

$$(47) \quad S_{c,w} \text{ is only defined if } f(o_2(w)) \text{ is defined, for every world } w. \text{ When defined,}$$

$$S_{c,w} = \langle D_w, \leq \rangle: \quad D_w = \{\{p : GP(m(w), o_1(w), o')(p) : o' \in f(o_2(w))\}$$

$$\leq = \{\langle d_1, d_2 \rangle : d_1, d_2 \in D \wedge d_1 \supseteq d_2\}$$

Assuming that $f(o_2) = \{\{F\}, \{R, C, F\}\}$ (encoding that R and C are equally important), this definition assigns the highest degree of necessity to going to Georgetown, and equal but lower degrees to going to Adams Morgan and Silver Spring. This definition has the surprising property that the scale of necessity is defined not in terms of standard necessity (entailed by the union of the maxima), but rather in terms of the weaker notion of good possibility (entailed by at least one maximum).

The revised definition allows us to account for examples with incompatible goals like the following.

- (48) a. It is as important to preserve the wetlands as it is to build the new housing (which would drain the wetlands).
b. We should preserve the wetlands just as much as we should build the new housing.

As Morzycki (2012) pointed out, equatives are also natural with extreme expressions, and there seems to be little problem in interpreting cases where they describe incompatible propositions as having the same degree:

- (49) a. It is as crucial to preserve the wetlands as it is to build the new housing.
b. We must preserve the wetlands just as much as we must build the new housing.

These sentences are clearly acceptable, and describe situations which are true dilemmas. They could be analyzed using the same strategy as (47) (or the related alternative mentioned in the text).

Although extreme and non-extreme modals are both perfectly acceptable with equatives, Cruse (1986); Paradis (2001); Bolinger (1967); Morzycki (2012) note that extreme expressions are less natural with comparatives. This suggests that the extreme part of the scale has only a single degree (though it can be coerced into multiple degrees).

- (50) a. ?It is more crucial to preserve the wetlands than it is to build the new shopping center.
b. ?We must preserve the wetlands more than we must build the new shopping center.

We can model this idea by assuming that $f(o_1(w))$ as a default is a singleton; this stipulation makes sense given the assumption that $o_1(w)$ contains only non-negotiable propositions, and so in many context there would be no useful way for f to distinguish among them. However, under coercion we can make distinctions among these propositions, in which case f can select multiple subsets of $o_1(w)$, giving rise to multiple extreme degrees.

Morzycki (2012) also noted that it is infelicitous to compare or equate extreme and non-extreme adjectives, as in (40). This fact follows in the same way as Morzycki proposes for non-modal adjectives: since the degree of the prejacent of the extreme adjective is guaranteed to be higher than that of the non-extreme one, such sentences are trivially true or false in any context in which they are interpretable.

Our analysis handles positive forms as well. For a non-extreme modal, the standard is set at the degree of some relevant member of $f(o_2(w))$, and in the right context, (51a) will be true. With extreme modals, in a non-coerced context there is a single extreme degree, and the standard is that degree. Supposing we follow the scale-definition using good possibility, (51b) is true if preserving the wetlands is a good possibility with respect to the sole member of $f(o_1(w))$.

- (51) a. It is important to preserve the wetlands.
b. It is crucial to preserve the wetlands.

Our account predicts that (51b) can be true even if a proposition inconsistent with preserving the wetlands also has the extreme degree. This is clearly possible, for example when it's uttered after (49), but not normal. More typically, (51b) would imply that we preserve the wetlands in all accessible worlds, and so build the new housing in none. We take this to be a scalar implicature, though it would also be possible to build it in as a component of the meaning of POS. Either way, this follows from the assumption that, in the normal case, the extreme degree is based on the entire $o_1(w)$, and that $o_1(w)$ is consistent. In this case, our semantics is equivalent to the standard Kratzerian analysis of *must*.

We have seen how our analysis accounts for some of the subtle differences between strong and weak deontic modals in a way which parallels the treatment of other extreme/non-extreme expressions. Specifically, we have mentioned the distinction between *very* and extreme modifiers like *positively* and discussed contrasts in the acceptability and interpretation of equatives, comparatives, and the positive form. We have not returned to the final three properties listed in Section 2.3, and doing so must await another occasion.

4 Discussion

We have given strong evidence that strong and weak necessity modals with priority readings should be classified as extreme and non-extreme expressions, respectively. Moreover, we have argued that this observation, when combined with other ideas from the scalar approach to modals, points to a better understanding of the semantic and pragmatic differences between the two types of modal necessity. At an empirical level, our analysis accounts for some of the subtle differences between strong and weak deontic modals in a way that parallels the treatment of other extreme/non-extreme expressions. Specifically, we have mentioned the distinction between *very* and extreme modifiers like *positively* and discussed contrasts in the acceptability and interpretation of equatives, comparatives, and the positive form.

We have focused specifically on deontic and other priority readings of necessity operators. In future work, we hope to extend these ideas to other judgment types, in particular epistemics. Once an analysis of epistemics is in hand, it will be important to compare the kind of approach we have advocated here with the probabilistic semantics for gradable epistemic modals developed by Yalcin (2010) and Lassiter (2011). We also hope to extend our ideas to non-priority root modals such as *can*, *able*, and adjectives formed with *-able*. These are not necessity operators, but are nonetheless gradable. The distinct logical properties of these two groups of non-priority modals will allow us to further refine the formal analysis of gradable modality.

References

- Bartsch, R. and Vennemann, T. (1973). *Semantic structures: A study in the relation between syntax and semantics*. Athenäum, Frankfurt.
- Beck, S. (2011). Comparison constructions. In Heusinger, K. v., Maienborn, C., and Port-

- ner, P., editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 2, pages 1341–1390. de Gruyter, Berlin.
- Bhatt, R. and Takahashi, S. (2011). Reduced and unreduced phrasal comparatives. *Natural Language and Linguistic Theory*, 29:581–620.
- Bierwisch, M. (1989). The semantics of gradation. In Bierwisch, M. and Lang, E., editors, *Dimensional Adjectives*, pages 71–262. Springer, Berlin.
- Bolinger, D. (1972). *Degree words*. Mouton, The Hague.
- Bolinger, D. L. (1967). Adjective comparison: a semantic scale. *Journal of English Linguistics*.
- Burnett, H. (2011). The puzzle(s) of absolute adjectives: On vagueness, comparison and the pragmatic origin of scale structure. In Paperno, D., editor, *UCLA Working Papers in Linguistics: Special Issue on Semantics and Mathematical Linguistics*. University of California, Los Angeles.
- Burnett, H. (2014). A delineation solution to the puzzles of absolute adjectives. *Linguistics and Philosophy*, 37(1):1–39.
- Cariani, F. (to appear). Deontic modals and probabilities: One theory to rule them all? In Charlow, N. and Chrisman, M., editors, This volume. Oxford University Press.
- Cariani, F., Kaufmann, M., and Kaufmann, S. (2013). Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy*, 37:225–259.
- Charlow, N. (2013). What we know and what to do. *Synthese*, 190:2291–2323.
- Cresswell, M. J. (1977). The semantics of degree. In Partee, B. H., editor, *Montague grammar*, pages 261–292. Academic Press, New York.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press, Cambridge.
- Davies, M. (2004). BYU-BNC. (based on the British National Corpus from Oxford University Press). Available online at <http://corpus.byu.edu/bnc/>.
- Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990-present. Available online at <http://corpus.byu.edu/coca/>.
- Davies, M. (2013). Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries. Available online at <http://corpus.byu.edu/bnc/>.
- de Swart, H. (1998). Aspect shift and coercion. *Natural Language and Linguistic Theory*, 16(347-85).
- Demonte, V. (2011). Adjectives. In Heusinger, K. v., Maienborn, C., and Portner, P., editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 2, pages 1314–1340. de Gruyter, Berlin.

- Doetjes, J. S. (1997). *Quantifiers and Selection: On the distribution of quantifying expressions in French, Dutch and English*. Holland Academic Graphics, Leiden. PhD thesis.
- Finlay, S. (2009). Oughts and ends. *Philosophical Studies*, 143:315–340.
- Finlay, S. (2010). What ought probably means, and why you can’t detach it. *Synthese*, 177:67–89.
- Goble, L. (1996). Utilitarian deontic logic. *Philosophical Studies*, 82(3):317–357.
- Halpern, J. (2003). *Reasoning about Uncertainty*. MIT Press, Cambridge, Mass.
- Heim, I. (1985). Notes on comparatives and related matters. Ms., MIT.
- Herburger, E. and Rubinstein, A. (2014). Is ‘more possible’ more possible in German? To appear in *The Proceedings of SALT 24*.
- Hintikka, J. (1961). Modality and quantification. *Theoria*, 27:110–28.
- Katz, G., Portner, P., and Rubinstein, A. (2012). Ordering combination for modal semantics. In Chereches, A., editor, *The Proceedings of SALT 22*, pages 488–507. CLC Publications.
- Kennedy, C. (1997). *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. PhD thesis, University of California Santa Cruz.
- Kennedy, C. (2007). Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30(1):1–45.
- Kennedy, C. and Levin, B. (2008). Measure of change: The adjectival core of degree achievements. In McNally, L. and Kennedy, C., editors, *Adjectives and Adverbs: Syntax, Semantics and Discourse*, pages 156–182. Oxford University Press.
- Kennedy, C. and McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, 81(2):345–381.
- Kenny, A. (1976). Human abilities and dynamic modalities. In Manninen, J. and Tuomela, R., editors, *Essays on Explanation and Understanding: Studies in the Foundations of Humanities and Social Sciences*, pages 209–232. D. Reidel, Dordrecht.
- Klecha, P. (2012). Positive and conditional semantics for gradable modals. In Guevara, A. A., Chernilovskaya, A., and Nouwen, R., editors, *Proceedings of Sinn and Bedeutung 16*, volume 2, pages 363–376. MITWPL.
- Kolodny, N. and MacFarlane, J. (2010). Ifs and oughts. *The Journal of Philosophy*, 108(3):115–143.
- Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H.-J. and Rieser, H., editors, *Words, Worlds, and Contexts*, pages 38–74. de Gruyter, Berlin.

- Kratzer, A. (1991). Modality. In von Stechow, A. and Wunderlich, D., editors, *Semantik/Semantics: An International Handbook of Contemporary Research*, pages 639–650. de Gruyter, Berlin.
- Kratzer, A. (2012). *Modal and Conditionals: New and Revised perspectives*. Oxford University Press.
- Lassiter, D. (2010). Gradable epistemic modals, probability, and scale structure. In Li, N. and Lutz, D., editors, *The Proceedings of SALT 20*, pages 197–215. CLC Publications, Ithaca, NY.
- Lassiter, D. (2011). *Measurement and Modality: The Scalar Basis of Modal Semantics*. PhD thesis, New York University.
- Lechner, W. (2004). *Ellipsis in Comparatives*. Mouton de Gruyter, Berlin.
- Miró, E. C. (2006). *Wh-exclamatives in Catalan*. PhD thesis, University of Barcelona.
- Moens, M. and Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14:15–28.
- Morzycki, M. (2012). Adjectival extremeness: degree modification and contextually restricted scales. *Natural Language and Linguistic Theory*, 30(2):567–609.
- Paradis, C. (2001). Adjectives and boundedness. *Cognitive Linguistics*, 12(1):47–65.
- Portner, P. (2009). *Modality*. Oxford University Press, Oxford.
- Rett, J. (2008). *Degree Modification in Natural Language*. PhD thesis, Rutgers University.
- Rotstein, C. and Winter, Y. (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics*, 12(3):259–288.
- Rubinstein, A. (2012). *Roots of Modality*. PhD thesis, University of Massachusetts, Amherst.
- Sapir, E. (1944). Grading. a study in semantics. *Philosophy of Science*, 11(2):93–116.
- Sassoon, G. W. (2007). *Vagueness, Gradability and Typicality: A Comprehensive Semantic Analysis*. PhD thesis, Tel Aviv University.
- Sassoon, G. W. (2010). Measurement theory in linguistics. *Synthese*, 174:151–180.
- Sassoon, G. W. (2013). *Vagueness, gradability and typicality: The interpretation of adjectives and nouns*. Brill, Leiden.
- Swanson, E. (2006). *Interactions with context*. PhD thesis, MIT.
- Unger, P. (1975). *Ignorance*. Clarendon Press, Oxford.
- Villalta, E. (2000). Spanish subjunctive clauses require ordered alternatives. In Jackson, B. and Matthews, T., editors, *The Proceedings of semantics and Linguistic Theory X*, pages 239–256. CLC Publications, Ithaca NY.

- Villalta, E. (2006). *Context dependence in the interpretation of questions and subjunctives*. PhD thesis, Tübingen.
- Villalta, E. (2008). Mood and gradability: an investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy*, 31(4):467–522.
- von Fintel, K. (2012). The best we can (expect to) get? Challenges to the classical semantics for deontic modals. Ms, MIT. Paper presented at the Central APA, February 17, 2012.
- von Fintel, K. and Iatridou, S. (2008). How to say ought in foreign: The composition of weak necessity modals. In Guéron, J. and Lecarme, J., editors, *Time and Modality*, pages 115–141. Springer, Berlin.
- von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics*, 3:1–77.
- Yalcin, S. (2007). Epistemic modals. *Mind*, 116(464):983–1026.
- Yalcin, S. (2010). Probability operators. *Philosophy Compass*, 5(11):916–37.
- Zanuttini, R. and Portner, P. (2003). Exclamative clauses: At the syntax-semantics interface. *Language*, 79(1):39–81.