Quantificational and modal interveners in degree constructions *

Daniel Lassiter
Stanford University

Abstract Heim (2001) points out that the relational semantics for degrees predicts ambiguities in sentences with comparatives and quantifiers such as *every girl* that are not observed. She also notes that the same ambiguities do appear with strong modals such as *must* and *have to*, but not with weaker modals such as *should*, *ought*, and *want*. The problem is to explain why these classes of expressions would behave differently, given that they are all standardly treated as universal quantifiers.

I present several counter-examples to Heim's account of this data and then argue that the puzzle involving ∀-DPs is the same as the puzzle of weak islands in amount *wh*-expressions, and that it yields to the analysis of weak islands due to Szabolcsi & Zwarts 1993, who argue that degree expressions are restricted in their interaction with the semantic operation *meet*. This accounts for ∀-DPs but leaves to be explained the possibility of modal intervention with strong modals. I argue that the split between ∀-DPs and strong modals supports recent work proposing that modals are not quantifiers over worlds but scalar expressions. An independently motivated scalar semantics for strong modals generates the ambiguity in a way that is compatible with Szabolcsi & Zwarts's theory, and that the predicted truth-conditions are correct for both readings with strong modals. The corresponding account of mid-strength modals explains their lack of ambiguity as merely apparent, due to the fact that the truth-conditions of the two readings are virtually indistinguishable, and neither embodies the missing reading that the quantificational theory leads us to look for. These results support both the scalar semantics for modality and Szabolcsi & Zwarts's semantic approach to intervention constraints.

Keywords: Intervention effects, weak islands, quantification, modality, scalar semantics

1 Introduction: Three Puzzles

Often theoretical analyses which account well for the particular phenomena that they are designed to explain generate unexpected predictions when their target phenomena interact. Whatever happens, we have much to learn: either the predicted effects emerge, and we have an additional argument in favor of the theories under

---

consideration; or something unexpected happens, and we are forced to refine the theories in ways we might not have considered. This paper confronts the predictions of theories developed, more or less independently, for three well-studied problem domains: the semantics of adjectival comparatives, quantificational and modal intervention in *wh*-questions, and the semantics of modals. I'll start by describing three puzzles discussed by Heim (2001) which are generated by standard semantic assumptions in these domains. I'll then argue that we can resolve the three puzzles by combining a semantic theory of quantifier intervention effects in degree expressions (Szabolcsi & Zwarts 1993) with a new theory of modality which treats modals as scalar expressions rather than quantifiers over possible worlds (Lassiter 2011a).

Puzzle 1: A missing ambiguity. The standard analysis of adjectival comparatives and *wh*-questions treats gradable adjectives as expressions of type $\langle d, et \rangle$, and comparative clauses and *wh*-operators alike as scope-taking expressions. This semantics predicts the existence of scope ambiguities which don't arise in either case.

(1) Every girl is less tall than Sam is.
   a. ✓ "Sam is taller than all the girls." $(every > CC_i > t_i)$
   b. ✗ "Sam is taller than the shortest girl." $(CC_i > every > t_i)$

(2) How tall is every girl?
   a. ✓ "For each girl x, please tell me: how tall is x?" $(every > wh_i > t_i)$
   b. ✗ "How tall is the shortest girl?" $(wh_i > every > t_i)$

I'll call the (b)-readings here, with another operator taking scope between the degree operator and its trace, the *intervention readings*. (The claimed predictions will be derived formally in the next section.) The first puzzle is: given that the standard semantics predicts that intervention readings should be readily available in these sentences, why don't we observe them?

Puzzle 2: Universal quantifiers vs. strong modals. Maybe the solution is to revise our assumption that the comparative clause is a scope-taking operator? No, because — as Heim (2001) points out — the predicted readings do appear when the ∀-quantifier is replaced by strong deontic and bouletic modals (which are usually treated as ∀-quantifiers ranging over possible worlds).

(3) This draft is 20pp. long. The final version has to be less long than that.
   a. ✓ "The maximum length is less than 20pp." $(has\ to > CC_i > t_i)$
   b. ✓ "The minimum length is less than 20pp." $(CC_i > has\ to > t_i)$

(If you have trouble getting both readings, try varying the prosody: reading (3a) is favored with stress on *less* and/or *long*, while reading (3b) is favored with stress on *has to*.)

(4) How long does the paper have to be?

✓ "What is the minimum required length?" $(wh_i > have\ to > t_i)$

To generate this reading, *have to* must be able to scope below the *wh*-operator here, unlike in (2b).[1] The second puzzle, then, is: why can *have to* and other strong modals take scope where a $\forall$-DP cannot?

Puzzle 3: Strong vs. mid-strength modals.   Not all of the deontic and bouletic modals that are usually modeled as $\forall$-quantifiers over worlds behave this way. For instance, the intervention reading does not arise with *should* as it does with *have to*.

(5) This draft is 20pp. long. The final version should be less long than that.

     a. ✓ "The *maximum* preferred length is less than that." $(should > CC_i > t_i)$

     b. ✗ "The *minimum* preferred length is less than that." $(CC_i > should > t_i)$

Must, require and *need* pattern with *have to*, while *ought, want,* and *be supposed to* pattern with *should*. The third puzzle is: what is the relevant difference between these two groups of modal expressions that could explain the limited availability of intervention readings with the latter?

2   The relational semantics and quantificational interveners

On the relational semantics (von Stechow 1984; Rullmann 1995; Heim 2001, a.o.), the denotation of a gradable adjective which takes individual arguments is a function relating individuals to degrees. In general, an adjective A, applied to an individual $x \in \mathbf{D}_e$ and a degree $d \in \mathbf{D}_d$, returns True if and only if $\mu_A(x)$, the measure of x along the scale associated with A, is at least d. For instance, $tall(x)(d)$ is true iff x's height is at least d. We assume, for simplicity, that degrees are always arranged onto scales totally ordered by $\leq$.

(6) $[\![tall]\!] = \lambda x_e.\lambda d_d.\mu_{height}(x) \geq d$

Crucially, degrees are downward monotone in this semantics: anyone who is d-tall is also d'-tall for all heights $d' < d$.

    Theorists who adopt the relational semantics typically treat comparative clauses as scope-taking expressions. The idea is that the comparative operators *more/-er*, *less*, and *as ... as* compare the maxima of two sets of degrees, checking that the appropriate relation holds between them.

---

1 On the quantificational semantics, the reading of (4) analogous to (2a) should mean in effect: "Please tell me, for each deontically accessible world w, how long the paper is in w." I'm not sure if (4) has this reading. Maybe it is ruled out as bizarre, or perhaps this is the reading on which an appropriate answer might be "Between 8 and 14 pages."
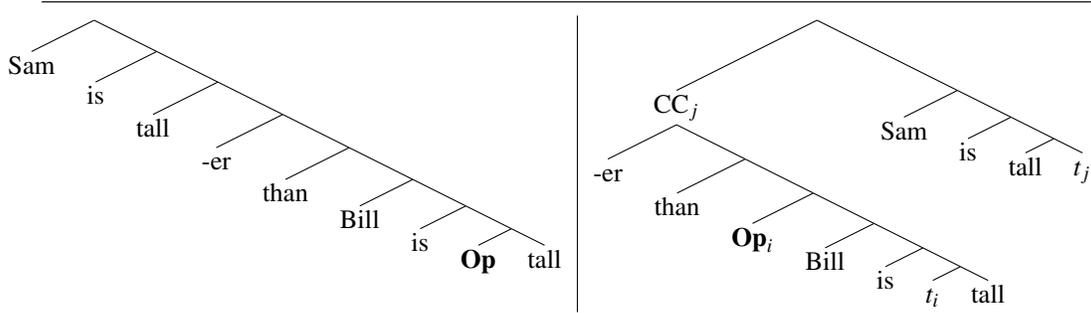
Table 1   (Slightly simplified) SS and LF of *Sam is taller than Bill is*.

(7) $\mathbf{max} = \lambda D_{\langle d,t\rangle}.\iota d_d.\forall d'[D(d') \rightarrow d' \leq d]$

(8)   a. $[\![\textit{more}/\textit{-er}]\!] = \lambda D_{\langle d,t\rangle}.\lambda D'_{\langle d,t\rangle}.\mathbf{max}(D') > \mathbf{max}(D)$

    b. $[\![\textit{less}]\!] = \lambda D_{\langle d,t\rangle}.\lambda D'_{\langle d,t\rangle}.\mathbf{max}(D') < \mathbf{max}(D)$

    c. $[\![\textit{as}\ldots\textit{as}]\!] = \lambda D_{\langle d,t\rangle}.\lambda D'_{\langle d,t\rangle}.\mathbf{max}(D') \geq \mathbf{max}(D)$

We assume that there is movement of a silent degree operator **Op** inside the comparative clause (Chomsky 1977), corresponding semantically to λ-abstraction of the degree variable of the lower copy of the gradable adjective. The comparative clause must undergo QR in order to resolve type-mismatch (Heim 2001). The assumed surface structure and LF of this sentence are depicted in Table 1. The associated truth-conditions simplify to the assertion that Sam's height is greater than Bill's.

(9) $[\![\textit{Sam is taller than Bill is}]\!]$

    a. $= [\![\textit{-er}_j]\!]([\![\mathbf{Op}_i\ \textit{Bill is } t_i \textit{ tall}]\!])([\![\textit{Sam is tall } t_j]\!])$

    b. $= \mathbf{max}(\lambda d.\mu_{height}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{height}(\mathbf{Bill}) \geq d)$

    c. $= \mu_{height}(\mathbf{Sam}) > \mu_{height}(\mathbf{Bill})$

Less-comparatives and equatives have the same analysis, replacing *-er* with the appropriate operator from (8), with the semantic effect of changing the relation used to compare the two maxima.

The rather roundabout way of deriving the truth-conditions in (9c) becomes useful when there are other operators in the sentence with which the maximality operation can interact (e.g., von Stechow 1984; Rullmann 1995). These interactions are an important source of evidence for the relational analysis. Without reviewing these positive arguments here, let's proceed to the problem: the account predicts scope ambiguities in sentences with comparatives and universal quantifiers, but one of these readings is systematically missing. To see this gap we have to look at *less*-comparatives and comparatives modified by *exactly* because the two readings of
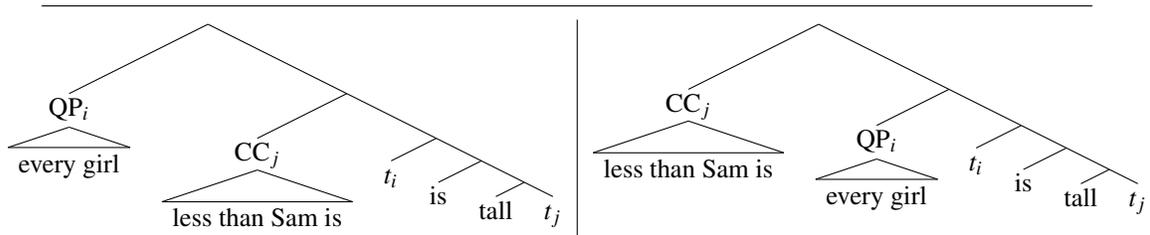
QP$_i$: every girl — CC$_j$: less than Sam is — t_i — is — tall — t_j

CC$_j$: less than Sam is — QP$_i$: every girl — t_i — is — tall — t_j

Table 2   Two predicted LFs of *Every girl is less tall than Sam is.*

unmodified *more/-er* comparatives are truth-conditionally equivalent (Heim 2001). But consider the two LFs in Table 2 and the associated truth-conditions in (10a) and (10b) respectively:[2]

(10) Every girl is less tall than Sam is.

    a. Reading 1: $QP_i > CC_j > t_i > t_j$

        $\forall x[\mathbf{girl}(x) \rightarrow \mathbf{max}(\lambda d.\mu_{height}(x) \geq d) < \mu_{height}(\mathbf{Sam})]$

        ✓ "Every girl is less tall than Sam, i.e., he is taller than the *tallest* girl."

    b. Reading 2: $CC_j > QP_i > t_i > t_j$

        $\mathbf{max}(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mu_{height}(x) \geq d]) < \mu_{height}(\mathbf{Sam})$

        ✗ "The greatest height d s.t. every girl is at least d-tall is less than Sam's height; i.e., he is taller than the *shortest* girl."

The intervention reading in (10b), with the QP taking scope below the comparative clause, is not an attested reading of this sentence (puzzle 1). Similarly, the intervention reading (11b) is missing in examples with a comparative clause modified by *exactly*.

(11) Every girl is exactly 1 inch taller than Sam is.

    a. Reading 1: $QP_i > CC_j > t_i > t_j$

        $\forall x[\mathbf{girl}(x) \rightarrow \mathbf{max}(\lambda d.\mu_{height}(x) \geq d) = \mu_{height}(\mathbf{Sam}) + 1'']$

        ✓ "For each girl x, x's height is exactly 1 inch greater than Sam's."

    b. Reading 2: $CC_j > QP_i > t_i > t_j$

        $\mathbf{max}(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mu_{height}(x) \geq d]) = \mu_{height}(\mathbf{Sam}) + 1''$

        ✗ "The *shortest* girl is exactly 1 inch taller than Sam is."

The relational semantics predicts a similar ambiguity in *wh*-sentences such as *How tall is every girl?*: the *wh*-expression should be able to take scope either above

---

2 To shorten the formulas, I've gone ahead and simplified $\mathbf{max}(\lambda d.\mu_{height}(\mathbf{Sam}) \geq d)$ to the logically equivalent $\mu_{height}(\mathbf{Sam})$.

or below the QP. Again, the intervention reading paraphrased "How tall is the shortest girl?" does not appear, and the only available reading has the quantifier scoping above the *wh*-operator, giving rise to a pair-list question.

3 Strong modals

One possible response to the absence of intervention readings is to adopt a semantics that does not generate ambiguities by eliminating the need for the comparative clause to move. Kennedy (1997), for instance, gives a semantics for comparatives which generates the right truth-conditions for (9) while keeping the comparative clause *in situ*. However, Heim (2001) shows that movement of the comparative clause is indeed needed in order to account for cases in which the predicted ambiguities do arise with strong modals such as *must*, *have to*, and *required*.

For instance, if we plug in the relational semantics just described into the ambiguous modal example (3) — assuming, as usual, that *have to* is a universal quantifier over some set of accessible worlds **Acc** — both readings are correctly predicted. (I write $\mu_{length}^{w}(x)$ for "the length of x in world w," and abbreviate the denotation of the description "the final version" as **FV**.)

(12) [This draft is 20pp. long.] The final version has to be less long than that.

    a. Reading 1: *have to* $> CC_j > t_j$

       $\forall w \in \textbf{Acc}[\textbf{max}(\lambda d.\mu_{length}^{w}(\textbf{FV}) \geq d) < 20pp.]$

       "In every accessible world, the final version is shorter than 20pp, i.e., the *maximum* length is less than that."

    b. Reading 2: $CC_j > $ *have to* $> t_j$

       $\textbf{max}(\lambda d.\forall w \in \textbf{Acc}[\mu_{length}^{w}(\textbf{FV}) \geq d]) < 20pp.$

       "The greatest length d s.t., in every accessible world, the final version is at least d-long is less than 20pp.; i.e., the *minimum* length is less than that."

Both of these readings are possible, unlike in the examples with quantificational DPs that we saw above. The ambiguity also arises with comparatives modified by an *exactly*-phrase:

(13) [This draft is 9pp. long.] The final version has to be exactly 1 page longer than that.

    a. Reading 1: *have to* $> CC_j > t_j$

       $\forall w \in \textbf{Acc}[\textbf{max}(\lambda d.\mu_{length}^{w}(\textbf{FV}) \geq d) = 10pp.$

       "In every accessible world, the final version is exactly 10pp. long."

b. Reading 2: $CC_j > $ *have to* $ > t_j$

$\mathbf{max}(\lambda d.\forall w \in \mathbf{Acc}[\mu_{length}^w(\mathbf{FV}) \geq d]) = 10pp.$

"The greatest length d s.t., in every accessible world, the final is at least d-long is exactly 10pp..; i.e., this is the *minimum* requirement."

Similarly, the relational semantics correctly generates the "minimum" reading in the *wh*-question version *How long is does the final version have to be?*, i.e., "What's the length s.t. it is at least that long in all accessible worlds?"

The fact that intervention readings are attested with *have to* and other strong modals means that we cannot eliminate the need for the comparative clause to take scope. But it leaves us with two problems: we still don't know what *does* prevent the intervention readings from surfacing with quantifiers, and we don't know what accounts for the difference in behavior between quantifiers and modals.

4 Heim's account

In the paper which introduced the puzzles discussed here, Heim (2001) suggested framing an answer to these questions in terms of a syntactic constraint on LF-representations, calling it "Kennedy's constraint." This has been picked up in subsequent work (e.g., Bhatt & Pancheva 2004; Takahashi 2006) and dubbed the "Heim-Kennedy constraint."

(14) **Heim-Kennedy constraint**: Filter out LFs in which a quantificational DP takes scope between a degree operator and its trace, i.e. $*DegOp_i[\dots QDP \dots t_i \dots]$.

In Heim's paper (14) seems to be intended more as a placeholder for a theoretical account than as a serious explanation. But since subsequent work has treated this constraint as a *bona fide* theory, let's pause to consider why it is not a satisfactory explanation of the phenomena under discussion.

There are at least three serious problems with treating the constraint in (14) as an explanation of intervention effects. First, it is stipulative: there is no independently motivated syntactic problem with the LFs that it applies to. Second, it predicts the asymmetry between quantificational DPs and modals simply by fiat; no syntactic or semantic explanation is offered of *why* the principle would not also filter out readings like (12b). But most problematically, there are clear counter-examples involving both comparatives and *wh*-questions. (15) counter-exemplifies the constraint using the silent degree operator **Op** inside the comparative complement, and is similar to examples discussed in von Stechow 1984. (16) does the same using an amount *wh*-question.

(15) Mary lifted more than a BOY did.

"She lifted more than the *strongest* boy." $(\mathbf{Op}_i > \exists > t_i)$

(16) [The strongest girl lifted 100 pounds.] OK, but how much did a BOY lift?

"What is the greatest d s.t. a boy lifted d-much, i.e., how much did the strongest boy lift?" $(wh_i > \exists > t_i)$

(17), noted by Heim herself, even violates the proposed constraint twice.

(17) Jaffrey is closer to an airport than it is to a train station.

$\mathbf{max}(\lambda d . \exists x[\mathbf{airport}(x) \wedge \mathbf{close}(x)(\mathbf{Jaffrey})(d)]) >$
$\mathbf{max}(\lambda d . \exists x[\mathbf{train\ station}(x) \wedge \mathbf{close}(x)(\mathbf{Jaffrey})(d)])$

"The *closest* airport is closer than the *closest* train station."

(Both $CC_i > \exists > t_i$ and $\mathbf{Op}_j > \exists > t_j$)

The constraint in (14) would incorrectly rule out all three of these examples, because the existential quantifier must take scope between the comparative clause/*wh*-operator and its trace in order to generate the correct truth-conditions. All of the counter-examples that I am aware of crucially involve quantifiers with *existential* semantics (a fact which will be important shortly).

In light of this data, it seems that the correct generalization must be something more subtle than modals vs. quantifiers. Indeed, the fact that some existential quantifiers can intervene suggests that we should be looking for an explanation in the semantics of these expressions, not their syntax.

5   Resolving Puzzle 1: Restrictions on semantic operations

In fact, a semantic account of the difference between universal and existential quantifiers in their scope-taking abilities was proposed already in Szabolcsi & Zwarts 1993 (henceforth, SZ). As noted previously by Lassiter (2010a,c), SZ's theory of weak islands provides an account of restrictions on universal interveners in comparatives as well as the availability of existential interveners. I will argue below that it also suggests an account of the quantifier/modal asymmetry (puzzle 2) which dovetails neatly with a recently developed theoretical perspective on modality.

SZ are concerned with weak islands generally, but we will focus on their account of weak islands triggered by amount *wh*-expressions, illustrated in (18).

(18) a. How tall is Bill?
     b. # How tall isn't Bill?
     c. # How tall is no player?

SZ's crucial idea is to frame the explanation of these data in terms of constraints on semantic operations: "[t]he weak island effect comes about when the *wh*-phrase should take wide scope over some operator but is unable to." According to them,

some domains have a structure less rich than the Boolean algebra generated by e.g. the domain of individuals. In a poorer structure, certain semantic operations may be systematically unavailable. The empirical claims, then, are (a) that individual *wh*-questions should be exempt from weak island effects and (b) that "island-sensitive *wh*-phrases ... do not range over individuals, but over elements of *partially ordered* domains."

From this perspective, weak island phenomena and intervention constraints more generally are a "window into semantic composition," allowing us to use facts about which operators expressions of a particular type do and do not interact with to discern the underlying structure of the domain in which they denote. For example, if expressions of some grammatical category C systematically refuse to combine with negation or other expressions whose denotations make use of the complement operation, this suggests that expressions of category C denote in a lattice which is not closed under complements.

SZ argue in this way that the infelicity of negation in degree questions (18b) suggests that $\mathbf{D}_d$, the domain of degrees, is not closed under complement: the semantics is unable to find the greatest degree d s.t. Bill is not d-tall, in effect because it is unable to compute the expression $\lambda d.\neg\mu_{tall}(\mathbf{Sam}) \geq d$. This predicts that, if computing the meaning of a sentence S requires applying the complement operation directly to a set of degrees, this expression will be infelicitous.

SZ argue similarly that the unavailability of the (b) reading of (19) (= (2)) indicates that the domain of degrees is not closed under the *meet* operation ($\approx$ intersection), which implements universal quantification.

(19)   How tall is every girl?

    a. ✓ "For each girl x, please tell me: how tall is x?"　　($every > wh_i > t_i$)

    b. ✗ "How tall is the shortest girl?"　　　　　　　　　　($wh_i > every > t_i$)

Essentially, the (b)-reading of the question asks us to find, for every girl x, the degrees of height to which x is tall, and then to intersect the resulting sets and locate the maximum of this set. If the lattice of degrees is not closed under meets, though, this operation cannot be performed, and the (b)-reading is semantically anomalous.

On the other hand, SZ argue that the lattice of degrees is closed under the *join* operation, which implements existential quantification. This predicts that "maximum" readings of *wh*-questions with existential quantifiers ($wh_i > \exists > t_i$) should be available, while "minimum" readings with universal quantifiers ($wh_i > \forall > t_i$) should not. As pointed out by Lassiter (2010a,c) and in (16) above, this prediction appears to be correct: existential quantifiers can intervene.

If SZ's generalization is correct, the explanation extends immediately to the data that puzzled Heim: the unavailable reading of (1)/(10) is not available because, like (19b), computing it would require intersecting the sets of degrees of heights of the

various girls, i.e. taking meets in the domain of degrees. The relevant portion of this example is repeated in (20).

(20) Every girl is less tall than Sam is.

$CC_j > QP_i > t_i > t_j$

$\mathbf{max}(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mu_{height}(x) \geq d]) < \mathbf{max}(\lambda d.\mu_{height}(\mathbf{Sam}) \geq d)$

✗ "The greatest height d s.t. every girl is at least d-tall is less than Sam's height; i.e., he is taller than the *shortest* girl."

This explanation also extends to the asymmetry between existential and universal quantification which we observed earlier in Heim's example (17): finding the greatest degree of closeness such that there *is* an airport that close requires only taking joins, which is not affected by the proposed constraint.

(21) Jaffrey is closer to an airport than it is to a train station.

$\mathbf{max}(\lambda d.\exists x[\mathbf{airport}(x) \wedge \mathbf{close}(x)(\mathbf{Jaffrey})(d)]) >$

$\mathbf{max}(\lambda d.\exists x[\mathbf{train\ station}(x) \wedge \mathbf{close}(x)(\mathbf{Jaffrey})(d)])$

✓ "The *closest* airport is closer than the *closest* train station." ($CC_i > \exists > t_i$)

Independent corroboration for a ban on taking meets, but not joins, in degree constructions comes from an asymmetry between conjunction and disjunction in comparative complements. The comparative with *or* (implemented by join) is ambiguous, with a reading where the disjunction is interpreted *in situ* yielding the "maximum" reading in (22a). However, the corresponding sentence with *and* (implemented by meet) lacks the "minimum" reading which we would expect from the *in situ* interpretation of the conjunction in (23a).

(22) Sam is taller than Bill or Sue.

    a. ✓ *In situ*:

        $\mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Bill}) \geq d \vee \mu_{tall}(\mathbf{Sue}) \geq d)$

        "He is taller than the taller of Bill and Sue; i.e., he is taller than Bill, and he is taller than Sue."

    b. ✓ **Wide-scope** *or*:

        $\mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Bill}) \geq d)$
        $\vee\ \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Bill}) \geq d)$

        "He is taller than one or the other." (e.g., with continuation "... but I don't remember which.")

(23) Sam is taller than Bill and Sue.

    a. ✗ *In situ*:

$$\mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Bill}) \geq d \wedge \mu_{tall}(\mathbf{Sue}) \geq d)$$

"He is taller than the shorter of the two, i.e., he is taller than at least one of them."

b. ✓ **Wide-scope** *and*:

$$\mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Bill}) \geq d)$$
$$\wedge \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Sam}) \geq d) > \mathbf{max}(\lambda d.\mu_{tall}(\mathbf{Bill}) \geq d)$$

"He is taller than Bill, and he is taller than Sue."

A note on the theoretical status of the foregoing before we proceed. SZ do not give a compositional implementation of their theory, and we could adopt either of two attitudes toward this account. First, we could think of SZ's account of weak islands as a genuine theoretical explanation of these data, with compositional implementation to be determined. Second, we might think of SZ's account simply as a set of empirical generalizations, like Heim's LF constraint, but with better empirical coverage. For our purposes it does not matter how we think of the theory, as long as we take the crucial prediction on board: operators whose semantics make use of *meets* — such as universal quantification and conjunction — cannot apply directly to degree expressions. If this constraint holds generally, as our data suggests it does, we have an account of the missing readings in (1) and (2), and so of puzzle 1 — as well as the fact that intervention readings appear with existential quantifiers but not with universal quantifiers (17 and 16).

6   Resolving Puzzle 2: Modals are Degree Expressions, Not Quantifiers

At first glance, the account of puzzle 1 that we have arrived at only makes things worse for puzzle 2, the availability of intervention readings with strong modals such as *must, have to, need*, and *require*. The reason is that these expressions are virtually always treated as universal quantifiers over some set of "accessible" worlds. For instance, *Sam is required to leave* is true if and only if, in all accessible worlds, Sam leaves.[3] Of course, the semantic constraint on taking meets applies equally in this case, and so we would predict wrongly that intervention readings should be unavailable with any modal which is implemented using a $\forall$-quantifier.

(24)   a. Every girl is less tall than Sam is.

$$\mathbf{max}(\lambda d.\forall x[\mathbf{girl}(x) \rightarrow \mu_{height}(x) \geq d]) < \mu_{height}(\mathbf{Sam})$$

---

3 See Portner 2009 for an excellent survey of theories of modality along these lines, focusing in particular on Kratzer 1991, the standard account in linguistics. Note that different theories will have different ways of determining the composition of the set of accessible worlds; for our purposes, this doesn't matter, since it is the appearance of the universal quantifier in the definition that creates trouble.

✗ "The shortest girl is shorter than Sam."

b. This draft is 20pp. long. The final version has to be less long than that.

$$\mathbf{max}(\lambda d.\forall w[\mu^w_{length}(\mathbf{FV}) \geq d]) < 20pp.$$

✓ "The minimum length is less than 20pp."

The solution that I will argue for is hinted at already in Szabolcsi & Zwarts's (1993) paper: noticing that the "minimum" reading is available in *wh*-questions with strong modals (e.g., (4)), they suggest that "the scopal properties of these verbs are not Boolean in nature." This amounts to proposing that the standard account of these modals is incorrect, and that the modals which generate intervention readings are not universal quantifiers of any stripe.

While Szabolcsi & Zwarts (1993) do not propose an alternative theory of strong modals to back up their suggestion, a number of authors have recently proposed *degree-based* accounts of the semantics of various modals and intensional verbs, including deontic (Goble 1996; van Rooij 1999), bouletic (Levinson 2003; Villalta 2008), teleological (Lassiter 2011b), and epistemic (Yalcin 2005, 2007, 2010; Swanson 2006; Lassiter 2010b). Lassiter (2011a) summarizes and expands on the arguments for this approach, and articulates a framework in which modals quite generally are degree expressions rather than quantifiers. Here we will focus on deontic and bouletic modals, which, on the degree-based approach, serve in one way or another to relate propositions to points on a scale of goodness/desirability and to compare them to a threshold value, more or less as gradable adjectives do in degree-based theories of their semantics.

There are numerous detailed arguments for this approach which I will not review here, e.g. the fact that gradability is widespread in the modal domain; see Lassiter 2011a for extensive discussion. (The authors cited above uniformly identify the notion of goodness/desirability at play as some variety of *expected utility*. This has important empirical and theoretical consequences, but won't play a major role here; as far as the issues under discussion are concerned, you can substitute an intuitive notion of "goodness" or some other formal notion without affecting anything.)

What is important for us about degree-based theories of modality is that they do not treat modals as quantifiers over possible worlds — and in particular, they do not treat the strong modals *must, have to, need*, and *require* as universal quantifiers, which would be implemented using the meet operation. As a result, degree-based theories of modality predict that restrictions on semantic operations such as those proposed by Szabolcsi & Zwarts (1993) should not apply to modals. This prediction is welcome, at least as far as this class of modals is concerned, because it means that we may be able to generate the "minimum" reading of (24b) without violating the general principles discussed in the last section.

It remains to be shown, of course, that the specific denotations associated with strong modals in degree-based theories are able to generate the correct truth-conditions for both readings of the ambiguous sentences with strong modals. I will show that a slightly improved version of the proposal in Lassiter 2011a: §6 does indeed generate both attested readings of the crucial examples, and does so without taking meets in the domain of degrees.

The idea is this. When someone says that ϕ *has to* or *must* be the case, or that ϕ is *required*, what they are saying is that there is no relevant way of achieving $\neg\phi$ which fails to be quite undesirable/bad. Essentially, if you have to leave, there isn't any permissible way that you could stay. This characterization leaves open the possibility that ϕ itself may not be all that wonderful: all that is required is the other options are unacceptable.[4]

More formally, suppose that R is a function which assigns to a proposition ϕ (relative to a context) a set of propositions which picks out all of the contextually relevant ways that one could act to achieve ϕ in c. (We might want to restrict $R(\phi)$ to the "actionable" alternatives in the sense of Charlow (2011). I'll leave this question open for now.) So, for example, if $\{\psi_1, ..., \psi_n\}$ exhaust the relevant ways of achieving ϕ, then $R(\phi) = \{\phi \wedge \psi_1, ...\phi \wedge \psi_n\}$. θ_{bad} in the definition represents a vague threshold of goodness/desirability which is constrained to be low (minimally, below the point of indifference, the goodness of a tautology). The proposed denotation for *must*, *have to*, and other strong modals is:

(25) $[\![must/have\ to]\!] = \lambda p_{\langle s,t\rangle}.\forall q \in R(\neg p)[\mu_{good}(q) < \theta_{bad}]$

I assume that other strong modals differ only in details left open by this definition, e.g. precisely which scale is relevant (deontic goodness, teleological goodness, desirability, etc.).[5]

---

4 Here we need to pay some brief attention to the structure of goodness scales — an issue that we have avoided up to now — in order to avoid a possible pitfall. If any way of structuring the scale were allowed, we might find situations in which there are only bad alternatives, and each of them *must* be the case. But suppose that the goodness scale has the formal structure of expected utility, as Goble (1996) and Lassiter (2011a) argue. Assuming that θ_{bad} is worse than a tautology (i.e., worse than indifferent), it can be proven that, if every alternative in $R(\neg\phi)$ is worse than θ_{bad}, then ϕ is better than each of these alternatives, and indeed better than the point of indifference (cf. Jeffrey 1965 and Lassiter 2011a: §6). This has the consequence that an exhaustive set of equally good/bad alternatives always has the same degree of goodness as a tautology (and so a degree greater than θ_{bad}). Thus, for example, if I have two equally bad options A and B, both "I *must* choose A" and "I *must* choose B" will come out false by (25).

5 This is related to the definition in Lassiter 2011a, but differs in some important ways. Lassiter's version requires both (a) that p be very good and (b) that all ways of achieving $\neg p$ which have non-zero probability be fairly bad. This definition is flawed for several reasons: e.g., requirement (a) is probably too stringent, and (as von Fintel (2012) points out) the definition fails to enforce certain class-subclass relations such as the entailment from *You don't have to bring alcohol* to *You don't have*

The important point for our purposes is that, in combination with the compositional semantics for degree constructions that we are employing, this definition captures the truth-conditions of both readings of the crucial example (3)/(12), and does so without taking meets in the domain of degrees; instead, the only quantifier in the definition applies to the propositional domain, which is clearly closed under meets. Here is the example again:

(26)   [This draft is 20pp. long.] The final version has to be less long than that.

Consider first the (a)-reading, with *has to* scoping over the comparative clause. Since the modal applies to the proposition $\phi = \llbracket$*The paper is less than 20 pages long*$\rrbracket$, R applies to the negation of this proposition, returning a set of relevant ways of being 20 pages long or longer. Presumably, in this context the relevant ways to accomplish this are to be d pages long for the various lengths $d \geq 20$. So the set of propositions picked out by $R(\neg\phi)$ are

$$R(\neg\phi) = \{\{w \mid \mu_{length}^w(\mathbf{FV}) = d\} \mid d \geq 20\}$$

and the predicted truth-conditions of the first reading are:

(27) $\llbracket$*has to*$\rrbracket(\phi) = \forall\psi \in R(\neg\phi)[\mu_{good}(\psi) < \theta_{bad}]$
       "For any length $d \geq 20$, it is bad if the final version is d pages long."

This is a pretty good rendition of the intuitive meaning of the sentence under this reading: any length greater than or equal to 20 pages is impermissible, i.e., there is a maximum permissible length which is less than 20 (though the sentence doesn't say how much less than 20 it is).

For the second (intervention) reading, with *has to* taking scope between the comparative clause and its trace, the meaning is slightly more complex, but eventually works out right. To shorten the formula, define $\phi = \llbracket$*The paper is (at least) d pages long*$\rrbracket$, with a free variable d, and let $\phi[d'/d]$ denote the result of replacing the free variable d by d'. This time we are looking for R-alternatives to the negation of propositions which state that the length of the paper is *greater than or equal to* some d', i.e. $\mu_{length}(\mathbf{FV}) \geq d'$. The relevant set of ways to fulfill the negation of such a proposition are presumably to be d' pages long for the various d'' *less than d'*, i.e.,

$$R(\neg\phi[d'/d]) = \{\{w \mid \mu_{length}^w(\mathbf{FV}) = d''\} \mid d'' < d'\}.$$

On these assumptions, the predicted truth-conditions are:

---

to bring beer. Since any relevant way of not bringing alcohol is also a relevant way of not bringing beer, the revised definition fixes this problem. I've also packed Lassiter's requirement of non-zero probability into the informal notion of relevance employed here, in recognition of the fact that there are probably various ways to spell this out that we don't have a clear way to decide between.

(28) $\mathbf{max}(\lambda d'.[\![\text{\textit{has to}}]\!](\phi[d'/d])) < 20pp.$
$= \mathbf{max}(\lambda d'.\forall\psi \in R(\neg\phi[d'/d])[\mu_{good}(\psi) < \theta_{bad}]) < 20pp.$

What this says is that there is some length d' such that (a) d' is less than 20, (b) any way of being shorter than d' pages long is bad, and (c) d' is the *greatest* length that satisfies condition (b). Condition (b) ensures that we are talking about a length for which any smaller length is unacceptable. Condition (c) enforces the requirement that d' be the *largest* such length, rather than some number smaller than the minimum requirement. (Any length smaller than the minimum requirement would also satisfy (a) and (b), but we don't want *those* degrees; we're looking for the exact point which separates unacceptable lengths from acceptable ones.)

(27) and (28) show that the (slightly modified) scalar semantics for strong deontic modals captures the intuitive meaning of both readings of (26). Crucially, the scalar theory gets the right truth-conditions for both readings of these sentences without invoking universal quantification over degree sets, as the standard quantificational semantics would. It thus explains the puzzle posed by Heim. Strong modals and universal quantifiers differ with respect to their ability to take scope between a degree operator and its trace because the latter require us to apply to degrees a semantic operation (meet) which is not defined in this domain. Strong modals do not make use of the meet operation, except (harmlessly) in the domain of propositions. This result provides support for the scalar approach to deontic and bouletic modality argued for on independent grounds by Goble, Levinson, Lassiter, and others.

7 Resolving Puzzle 3: Two readings of comparatives with mid-strength modals

The final puzzle noted by Heim (2001) that we must account for is why comparatives appear to display scope ambiguities with the strong modals *must, have to, need, require* but not with the mid-strength modals *should, ought, want, supposed to*, as in the contrast between (26) and (29). Again, the standard theory of the semantics of these modals, which treats them as universal quantifiers over accessible worlds, predicts the following two readings:

(29) This draft is 20pp. long. The final version should be less long than that.

    a. ✓ "The *maximum* preferred length is less than that." (*should* $> CC_i > t_i$)

    b. ✗ "The *minimum* preferred length is less than that." ($CC_i > $ *should* $> t_i$)

On an analysis like this, it is not clear why there should be a difference between the strong and mid-strength modals with respect to intervention effects in comparatives and *wh*-questions.

In this section I will argue that the scalar semantics for these expressions — which, again, was proposed to deal with quite different problems — makes reasonable pre-

dictions in this case as well. This theory predicts that these sentences are in fact ambiguous, but that the truth-conditions of the two readings are both related to the "maximum" reading paraphrased in (30a), and neither embodies the unattested "minimum" reading (30b). Instead we have the two readings paraphrased in (30):

(30)   a. "It's better if the paper is less than 20 pages long than if it's 20 pages or more."   ($should > CC_i > t_i$)

   b. "For some possible lengths d, it would be better for the paper to be longer, but at some point getting longer is dispreferred. This point is less than 20."   ($CC_i > should > t_i$)

I'll suggest that both of these are plausible readings of (29).

Scalar theories of the semantics of the mid-strength modals generally analyze them as operators which compare the goodness/desirability of a proposition to the goodness/desirability of its negation or of a set of relevant alternatives. Crucially, on this semantics $should/ought/want(\phi)$ states an overall preference for ϕ, without excluding the possibility that there are some ways of achieving $\neg\phi$ that are not terrible; ϕ is simply, on balance, better than its negation or (some function of) the set of alternatives. Sticking with the simpler, negation-based version for the moment:

(31) $[\![should]\!] = \lambda p_{\langle s,t\rangle}.\mu_{good}(p) >_s \mu_{good}(\neg p)$
   "*Should p* is true if and only if p is significantly better than $\neg p$ (by the relevant metric of goodness)."

This definition predicts two readings of (29). The first is:

(32)   **Reading 1**: $should > CC_i > t_i$
   $[\![should]\!](\mathbf{max}(\lambda d.\mu_{length}(\mathbf{FV}) \geq d) < 20)$
   $= [\![should]\!](\mu_{length}(\mathbf{FV}) < 20)$
   $= \mu_{good}(\mu_{length}(\mathbf{FV}) < 20) >_s \mu_{good}(\mu_{length}(\mathbf{FV}) \geq 20)$
   "It's better if the paper's (max) length is less than 20pp. than it it's 20 pages or more."

This reading is close to the paraphrase in (30a) above, but note that it is not about a maximum preferred length. These truth-conditions leave open the possibility, for example, that some lengths less than 20 pages could be better than others; for instance 10 pages might be preferred to any other length, but 11-19 pages is acceptable as well. All that (32) requires is that less than 20 is better than 20 or more.

Turning now to the intervention reading, we derive:

(33)   **Reading 2**: $CC_i > should > t_i$
   $\mathbf{max}[\lambda d.[\![should]\!](\mu_{length}(\mathbf{FV}) \geq d)] < 20$

$$= \mathbf{max}[\lambda d.\mu_{good}(\mu_{length}(\mathbf{FV}) \geq d) >_s \mu_{good}(\mu_{length}(\mathbf{FV}) < d)] < 20$$

"There is some non-empty set of degrees D such that, for each $d \in D$, it is better on balance for the paper to be d pages or more than it is for the paper to be less. No $d \geq 20$ is in this set."

This one is more complex, but the key observation is that, unlike the reading predicted by quantificational theories of *should*, it is not about a minimum preferred length (i.e., a unique length d for which it is preferred that the paper not be shorter than d pages). Instead, the truth-conditions in (33) boil down to the following. First, there must be *at least* one length, less than 20 pages, such that being that long or longer is generally better than being shorter. (For example, this reading excludes the possibility that shorter papers are always better, so that the absolute best length is 1 page.) Second, being 20 pages or longer cannot be significantly better than being less than 20 pages; or conversely, being less than 20 pages long must be (at least) roughly as good as being 20 pages or more, and so on for 21, 22, etc.

Importantly, the second condition is nearly equivalent to the truth-conditions in (32); the difference is merely in whether we assert that being ≥ 20 pages is significantly *worse* than being < 20 (32) or that it is not significantly *better* (33). These truth-conditions are extremely difficult to distinguish in practice: Horn (1989), in particular, discusses at length similar inferences involving relative adjectives (such as *not tall* ⤳ *short*) and neg-raising inferences with what I have called mid-strength modals, as well as certain other intensional verbs. For all practical purposes, then, (33) entails (32). This suggests an explanation of why there is no detectable ambiguity: it is difficult or impossible to reliably detect ambiguities when the two readings are not logically independent (cf. Heim 2001; Szabolcsi 2010: §6.3).

The scalar semantics improves on the quantificational semantics in this respect: we can explain the absence of two sharply differentiated readings simply by noting that the scalar theory generates two endings which are truth-conditionally extremely close, and difficult to distinguish in practice. This is a considerable improvement on quantificational theories, which must rely on special syntactic or semantic restrictions to explain why the apparently unobjectionable (30b) is not observed.

8 Conclusion

It is generally assumed that strong modals such as *must* and *have to*, as well as mid-strength modals such as *should, ought*, and *want*, are restricted quantifiers over possible worlds. However, the strong modals behave differently from quantificational DPs such as *every girl* in the way that they interact with certain other operators: in the case at hand, we have seen that strong modals can scope below a degree operator, while universal quantifiers cannot. I have argued for an account of restrictions on

universally quantified DPs that extends Szabolcsi & Zwarts's (1993) theory of weak islands to comparatives, noting that this approach improves on Heim's (2001) LF constraint because it correctly allows for intervention readings with existentials.

The fact that intervention readings are not ruled out with strong modals was explained by adopting recent proposals according to which strong modals are not treated as universal quantifiers over possible worlds, but as scalar operators. I argued that a slightly modified version of one recent proposal along these lines — Lassiter 2011a — predicts the right truth-conditions for both readings of comparatives with strong modals, and does so in a way which avoids the restrictions on semantic operations that Szabolcsi & Zwarts propose. Finally, I suggested that the fact that mid-strength modals do not display the expected ambiguity is due to the fact that quantificational theories of modality lead us to expect the wrong ambiguity. There are, I argued, ambiguities in comparatives with mid-strength modals, but the two readings are truth-conditionally very close and indeed, with a neg-raising step, one of these readings becomes a special case of the other.

Overall, the argument supports the scalar semantics for modality and also underscores the importance and usefulness of seeking a unified theory of intervention constraints in comparatives and *wh*-questions.

References

Bhatt, Rajesh & Roumyana Pancheva. 2004. Late merger of degree clauses. *Linguistic Inquiry* 35(1). 1–45.

Charlow, Nate. 2011. What We Know and What to Do. *Synthese* 1–33. doi:10.1007/s11229-011-9974-9.

Chomsky, Noam. 1977. On *wh*-movement. In Peter Culicover, Thomas Wasow & A. Akmajian (eds.), *Formal syntax*, vol. 132, Academic Press.

von Fintel, K. 2012. The best we can (expect to) get? Challenges to the classic semantics for deontic modals. Paper presented at the Central meeting of the American Philosophical Association, February 17. http://mit.edu/fintel/fintel-2012-apa-ought.pdf.

Goble, L. 1996. Utilitarian deontic logic. *Philosophical Studies* 82(3). 317–357.

Heim, Irene. 2001. Degree operators and scope. In Fery & Sternefeld (eds.), *Audiatur vox sapientiae: A festschrift for Arnim von Stechow*, 214–239. Akademie Verlag.

Horn, Laurence. 1989. *A Natural History of Negation*. University of Chicago Press.

Jeffrey, Richard C. 1965. *The logic of decision*. University of Chicago Press.

Kennedy, Chris. 1997. *Projecting the adjective: The syntax and semantics of gradability and comparison*: U.C., Santa Cruz dissertation.

Kratzer, Angelika. 1991. Modality. In von Stechow & Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, de Gruyter.

Lassiter, Daniel. 2010a. Explaining a restriction on the scope of the comparative operator. *University of Pennsylvania Working Papers in Linguistics* 16(1).

Lassiter, Daniel. 2010b. Gradable epistemic modals, probability, and scale structure. In N. Li & D. Lutz (eds.), *Semantics & Linguistic Theory (SALT) 20*, 197–215. CLC Publications.

Lassiter, Daniel. 2010c. The Algebraic Structure of Amounts: Evidence from Comparatives. *Interfaces: Explorations in Logic, Language and Computation* 38–56.

Lassiter, Daniel. 2011a. *Measurement and Modality: The Scalar Basis of Modal Semantics*: New York University dissertation.

Lassiter, Daniel. 2011b. Nouwen's puzzle and a scalar semantics for obligations, needs, and desires. In Neil Ashton, Anca Chereches, & David Lutz (eds.), *Semantics and Linguistic Theory 21*, 694–711. CLC Publications.

Levinson, D. 2003. Probabilistic model-theoretic semantics for *want*. In Robert B. Young & Yuping Zhou (eds.), *Semantics and Linguistic Theory (SALT) 13*, 222–239. CLC Publications.

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

van Rooij, Robert. 1999. Some analyses of pro-attitudes. In H. de Swart (ed.), *Logic, game theory, and social choice*, Tilburg University Press.

Rullmann, Hotze. 1995. *Maximality in the Semantics of* wh-*constructions*: University of Massachusetts, Amherst dissertation.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3(1). 1–77.

Swanson, Eric. 2006. *Interactions With Context*: MIT dissertation. http://mit.dspace.org/bitstream/handle/1721.1/37356/123190034.pdf?sequence=1.

Szabolcsi, A. 2010. *Quantification*. Cambridge University Press.

Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1(3). 235–284.

Takahashi, Shoichi. 2006. More than two quantifiers. *Natural Language Semantics* 14(1). 57–101.

Villalta, Elisabeth. 2008. Mood and gradability: An investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy* 31(4). 467–522.

Yalcin, Seth. 2005. A puzzle about epistemic modals. *MIT Working Papers in Linguistics* 51. 231–272.

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

Yalcin, Seth. 2010. Probability Operators. *Philosophy Compass* 5(11). 916–937.