

Degree Operators and Scope

Irene Heim

1. Introduction

A familiar idea about gradable adjectives is that they denote relations between individuals and degrees. This is most transparent in constructions like (1), where we seem to be witnessing explicit reference to or quantification over degrees.

- (1) a. John is *six feet* tall.
b. He is *that* tall.
c. *How* tall is he?
d. *However* tall he is,

A sentence like (1a) can be analysed very simply: *tall* denotes a relation between physical objects and degrees of height, *John* and *six feet* denote respectively an individual and a degree, and the whole thing (ignoring the vacuous copula) is hierarchically structured and compositionally interpreted just like an ordinary transitive sentence. The adjective's degree argument appears to be syntactically projected, just like the individual argument, and interpreted in an analogous fashion.

Argument positions where an individual is selected may be occupied by referential DPs as well as by quantificational DPs. How about argument positions for degrees? Do we also sometimes find quantificational degree arguments in the same slots as referential ones? This indeed is a plausible diagnosis of what goes on in gradation constructions like (2).

- (2) a. He is *more than 4 feet* tall.
b. He is *taller than 4 feet*.
c. He is *taller than the bed is long*.
d. He is *as tall as you are*.
e. He is *tallest*.
f. He is *too* tall.

(2a), for example, could be paraphrased as “he is tall to *some* degree exceeding 4 feet”, with an explicit indefinite description (existential quantifier) over degrees. This kind of analysis can be given for all these examples: the underlined phrase in each example is a generalized quantifier over degrees that is generated in the same argument position occupied by the degree name *six feet* in (1a). In those cases where this phrase is superficially discontinuous, we can attribute this to an (obligatory) superficial extraposition process which does not feed LF.¹

Quantificational DPs are not always interpretable in the argument positions where they are generated, and so they move, sometimes covertly, to appropriate scope positions. There are island constraints on this movement, but within their limits, there is a certain amount of freedom. Sometimes a DP-quantifier has more than one available landing-site, and sometimes this choice gives rise to truth-conditional ambiguities (known as “scope ambiguities”). Does the analogy between individual and degree arguments extend to this phenomenon as well? Do we find ambiguities that are due to multiple scope options for degree quantifiers like the ones in (2)? This is the question I want to address in this paper.

In doing so, I will follow in the footsteps of Christopher Kennedy, who discussed the same question in his 1997 thesis. Kennedy came to a negative conclusion. He argued that the putative degree quantifiers in comparative constructions always take the narrowest possible scope – even when no known locality condition would seem to prevent them from scoping over another operator in their vicinity. He suggests that this fact is unexpected and mysterious if the constructions in question really do contain a constituent that is a degree quantifier and moves for interpretability. For this (and other) reasons he endorses a different LF-constituency (and semantics) for gradation structures, in which there is no degree-quantifier that needs to or is able to move in the first place.

My aim in this paper is modest and merely descriptive. I want to argue that the facts are more complicated than Kennedy concluded. In a limited set of environments, the multiple scope-options anticipated by the approach I have sketched do show up as truth-conditional ambiguities. Moreover, the interpretive options for various kinds of elided or implicit material give indirect evidence for two possible scope constellations even in certain cases where they truth-conditionally collapse. I will tentatively conclude that gradation constructions do contain degree quantifiers that take scope by movement, though this movement is subject to (severe) syntactic constraints. I have no proposal yet about the nature of these constraints.

Let me begin by presenting a more concrete version of the degree-quantifier analysis of comparatives that I have hinted at. Unfortunately, quite a few distinct variants are suggested by the existing literature on comparatives, and I do not have space in this paper to motivate the specific choices I make or to prove that they don’t affect my

¹ This is probably wrong. For arguments that extraposition marks scope, see Williams (1974), Guéron & May (1984), and Fox & Nissenbaum (2000). The task of integrating these authors’ insights with the results of the present paper is left for the future.

points. I will assume that gradable adjectives denote functions of type $\langle d, et \rangle^2$ which are monotone in the sense of (3). A sample lexical entry is (4).

- (3) A function f of type $\langle d, et \rangle$ is *monotone* iff
 $\forall x \forall d \forall d' [f(d)(x) = 1 \ \& \ d' < d \rightarrow f(d')(x) = 1]$

- (4) $\llbracket tall \rrbracket = \lambda d_d. \lambda x_e. x$ is tall to degree d

The adjective's degree argument is its inner argument; it is syntactically projected as sister to A and will be labelled DegP.³ In the simplest kind of case (e.g., (1a,b)), the DegP has a meaning of type d and is straightforwardly interpretable *in situ* (by functional application). In the more complicated cases which we will mostly be concerned with, the DegP is a generalized quantifier over degrees (type $\langle dt, t \rangle$). This is the case for complex DegPs headed by *-er*, *as*, *too*, etc., e.g., *-er than 6 feet*, *-er than Mary is tall*, *5 inches -er than that*. I will abstract away as much as possible from issues concerning the internal structure and composition of these complex DegPs. Until we get to sections 2.4 and 3 below, we can limit ourselves to examples in which the phrase after *than* rigidly designates a degree. The kinds of meanings I will assume for this case are defined in (5), assuming the definition of "maximum" in (6).

- (5) a. $\llbracket -er \text{ than one foot} \rrbracket = \lambda P_{\langle d, t \rangle}. \max(P) > 1'$
 b. $\llbracket exactly \text{ two inches } -er \text{ than one foot} \rrbracket = \lambda P. \max(P) = 1' + 2''$

- (6) $\max(P) := \iota d. P(d) = 1 \ \& \ \forall d' [P(d') = 1 \rightarrow d' \leq d]$

Being of type $\langle dt, t \rangle$, these complex DegPs cannot be interpreted *in situ*, but must move for interpretability to a position above the adjective's subject (not necessarily above the surface subject, if there are lower covert subjects). The movement leaves a trace of type d and creates a λ -abstract of type $\langle d, t \rangle$, which makes a suitable argument to the DegP. Here is a sample derivation.

- (7) John is taller than 4 feet.

SS minus extraposition: John is $[AP [DegP -er \text{ than } 4 \text{ ft}] \text{ tall}]$

LF: $[DegP -er \text{ than } 4 \text{ ft}]_1$ John is $[AP t_1 \text{ tall}]$

interpretation:⁴ $\max\{d: \text{tall}(j, d)\} > 4'$

² I treat degrees as a separate basic type (label "d").

³ This is not the DegP of Abney, Kennedy, and other proponents of the "DegP hypothesis". For them, AP is an argument of Deg, whereas my DegP is an argument of A.

⁴ Abbreviations: $\text{tall}(x, d)$: = x is tall to degree d . j : = John.

So *John is taller than 4 feet* is true iff the maximal degree to which John is tall exceeds 4 feet.

2. DegP-Scope and Truth-Conditions

In the syntax and semantics just presented, I deliberately highlighted a close analogy between DegPs and DPs. In both categories, we find referring (and bound variable) phrases with denotations of a basic type (d or e) as well as quantificational phrases of the corresponding generalized quantifier type ($\langle dt, t \rangle$ or $\langle et, t \rangle$). And in both, the quantificational phrases can move at LF, leaving behind basic-type traces and creating complex λ -predicates. In the case of DPs, this movement is known as QR (Quantifier Raising) and is standardly assumed to be the source of truth-conditional ambiguities. If DegPs are so similar in their LF-syntax and semantics, shouldn't we find analogous scope-ambiguities due to multiple choices of landing-sites for DegP-movement? For example, when a DegP originates in the surface-c-command domain of a quantificational DP, negation, or intensional predicate, it should be able to scope either below or above that item – at least as long as there is nothing in its path that is known to create islands for movement. In the present section, we will try to test this prediction by comparing the range of truth-conditionally distinct readings that we can generate by DegP-movement in our system with the readings that are actually possible.⁵

1.1 Caveat: Systematic Equivalences and Anomalies

Not every scope ambiguity is a truth-conditional ambiguity. For example, when *every boy* scopes above *every girl* in the sentence *Every girl saw every boy*, the truth-conditions are just the same as when it scopes below it. The methodological implications of this fact are familiar. If we want to study the properties of QR, we must choose our examples judiciously. The fact that this particular sentence has only one interpretation doesn't show that objects can't move over subjects.

When it comes to quantifiers over degrees rather than ordinary individuals, this point is especially pertinent. Due to the ordered structure of the domain of degrees and the monotonicity property of adjective meanings, there are lots of cases where the relative scopes of a DegP and another quantifier are indistinguishable. To get a feel for what I am talking about, consider a comparative with a universal subject DP.

(8) Every girl is taller than 4 feet.

⁵ This whole section closely follows the reasoning in Kennedy (1997). 2.1 and 2.2 more or less just replicate his examples, arguments, and conclusions. 2.3 and 2.4 contain differing facts and conclusions, but even there I will mostly mimic his logic of argument.

If the DegP *-er than 4 ft* is free to scope either below or above the subject, we have two LFs: a low-DegP LF as in (9a) and a high-DegP LF as in (9b).

- (9) a. [every girl]₁ [-er than 4 ft]₂ t₁ is t₂ tall
 b. [-er than 4 ft]₂ [every girl]₁ t₁ is t₂ tall

On the semantics specified above, these structures get the truth-conditions in (10a,b) respectively.

- (10) a. $\forall x[\text{girl}(x) \rightarrow \max\{d: \text{tall}(x,d)\} > 4']$
 b. $\max\{d: \forall x[\text{girl}(x) \rightarrow \text{tall}(x,d)]\} > 4'$

(10a) says that each girl's (maximal) height is above 4'. (10b) is less transparent at first: We form the set of degrees *d* such that *every* girl is tall to *d*. Given the monotonicity of the tall-relation, this is precisely the set of degrees to which the *shortest* girl is tall (or the shortest girls, if two or more are tied for shortest). So the maximum of that is the (maximal) height of the shortest girl(s), and (10b) says this is above 4'. But that's the same claim as (10a)! If every girl is taller than 4', then the shortest girl is, and if the shortest girl is taller than 4', then every girl is.

Something similar happens when the subject is an existential quantifier. If we construct two LFs for (11) analogous to the ones in (9a,b), we derive the truth-conditions in (12a,b).

- (11) Some girl is taller than 4 feet.

- (12) a. $\exists x[\text{girl}(x) \ \& \ \max\{d: \text{tall}(x,d)\} > 4']$
 b. $\max\{d: \exists x[\text{girl}(x) \ \& \ \text{tall}(x,d)]\} > 4'$

In (12b), we have the set of degrees *d* such that *some* girl is tall to *d*. This turns out to be the set of degrees to which the *tallest* girl is tall (or the tallest girls in the event of a tie). So (12b) says that the (maximal) height of the tallest girl(s) is above 4'. But if some girl is taller than 4', then the tallest one is, and if the tallest girl is taller than 4', then some girl is. So again (12b) winds up being equivalent to (12a).⁶

⁶ What about infinite domains? If there were infinitely many girls, there might be no shortest ones and/or no tallest ones. In that case, the maxima referred to in (10b) or (12b) would be undefined. So my real claim, more accurately, is that pairs like (10) and (12) are equivalent whenever these maxima are defined. For my methodological point, this qualification makes no difference. I assume (see right below) that the high-DegP LF is automatically ruled out when its interpretation refers to an undefined maximum. So even if we consider universal and existential quantifiers with infinite domains, there can be no detectable scope ambiguity.

How pervasive is this kind of equivalence? Quite. It generalizes from comparatives to equatives.⁷ It generalizes from universal and existential to all monotone increasing quantifiers. And it generalizes from quantifiers over individuals to quantifiers over possible worlds. This means that we would be wasting our time trying to detect scope-ambiguities, not only in examples like (8) and (11), but in any of the following as well.

- (13) a. Every girl is as tall as John is.
 b. Some girl is as tall as John is.
- (14) a. More than 20 girls are taller than 4 feet.
 b. Most of the papers are longer than 5 pages.
- (15) a. The paper is required to be longer than 10 pages.
 b. The paper is allowed to be longer than 10 pages.

Pick (15a), for instance. *Required* is a necessity operator, a universal quantifier over the set of accessible worlds (for some contextually given accessibility relation, plausibly here a deontic one). If the DegP *-er than 10 pp* scopes *below required*, our semantics delivers the truth-condition in (16a); if the DegP scopes *above required*, we predict (16b).⁸

- (16) a. $\forall w \in \text{Acc}: \max\{d: \text{long}_w(p,d)\} > 10pp$
 b. $\max\{d: \forall w \in \text{Acc}: \text{long}_w(p,d)\} > 10pp$

(16a) transparently says that the paper is longer than 10pp in every accessible world. (16b) instructs us to determine the maximal degree to which the paper is long in every accessible world. That is its (maximal) length in those accessible worlds where it is shortest (the “minimal compliance” worlds, so to speak). This, (16b) says, exceeds 10pp – which is tantamount to (16a), since if the paper exceeds 10pp in the accessible worlds where it’s shortest, it exceeds 10pp in every accessible world, and vice versa. The reasoning is just as for *every girl* in (8). By the same token, the possibility operator *allowed* in (15b) behaves just like *some girl* in (11). I leave it to the reader to establish the pertinent equivalences for the remaining examples.

⁷ This is assuming the (standard) “at least”-interpretation of equatives, on which *Mary is as tall as John* is true (though under-informative) when she is actually taller than him. A suitable meaning-rule for equative DegPs analogous to (5a) is: $\llbracket \text{as as one foot} \rrbracket = \lambda P. \max(P) \geq 1$. As for *exactly*-equatives, see 2.2 below.

⁸ Abbreviations: p: = the paper. $\text{long}_w(x,d)$: = x is long to degree d in world w. $\text{Acc}(w)$: = the set of worlds accessible from w. The world argument of Acc is suppressed when it is the utterance world.

So what have we learned so far about DegP-movement? Nothing yet. The examples we have analysed here are consistent with everything we have assumed plus the assumption that DegP can move to any position where it is interpretable. They are also consistent with the possibility that DegP never moves beyond the lowest position where it is interpretable. We just can't tell.

Our next set of examples will still not allow us to tell how far DegPs can move, but this time for a different reason. The high-DegP LF will not be equivalent to the low-DegP LF, but it will be semantically deviant and thus ruled out by independent principles. Consider scoping a DegP over a negation, as in the second LF given below for sentence (17a).

- (17) a. Mary isn't taller than 4 feet.
 b. not [-er than 4'] Mary is t tall
 $\neg \max \{d: \text{tall}(m,d)\} > 4 \text{ ft}$
 c. [-er than 4'] not [Mary is t tall]
 $\# \max \{d: \neg \text{tall}(m,d)\} > 4 \text{ ft}$

The predicted interpretation for (17c) refers to the maximum of the set of degrees to which Mary *isn't* tall. But this set has no maximum.⁹ So if such an LF can be generated, it is a presupposition failure. Plausibly, this suffices to explain why only one reading (the one in (17b)) is attested. For all we know, (17c) may or may not be generated by the syntax; even if it is, it will be filtered out.

This observation generalizes from plain negation to other implicitly negative or monotone decreasing operators.

- (18) a. *At most two girls* are taller than 5 feet.
 b. [at most two girls]₁ [-er than 5']₂ t₁ are t₂ tall
 $|\{x: \text{girl}(x) \ \& \ \max \{d: \text{tall}(x,d)\} > 5'\}| \leq 2$
 c. [-er than 5']₂ [at most two girls]₁ t₁ are t₂ tall
 $\# \max \{d: |\{x: \text{girl}(x) \ \& \ \text{tall}(x,d)\}| \leq 2\} > 5'$

- (19) a. (Mary works 60 hrs a week, and) she *refuses* to work harder than that.
 b. she refuses [[-er than that] PRO to work t hard]
 $\forall w \in \text{Acc}: \neg \max \{d: m \text{ works } d\text{-hard in } w\} > 60\text{hrs/wk}$
 c. [-er than that] she refuses PRO to work t hard
 $\# \max \{d: \forall w \in \text{Acc}: \neg m \text{ works } d\text{-hard in } w\} > 60\text{hrs/wk}$

⁹ This way of ruling out reading c admittedly depends on the specific semantics for *-er* that I chose, which makes reference to maximality. (Maximalization in the interpretation of the *than*-clause is standard (see von Stechow 1984, Rullmann 1995b), but in the matrix clause it is not.)

As the reader can verify, the b-LFs capture the (only) available readings, and the c-LFs refer to undefined maxima.

If scoping over upward monotone operators always gives equivalent readings, and scoping over downward monotone ones always implies presupposition failure, we are left with the non-monotone ones.

- (20) a. *Exactly two girls* are taller than 5 feet.
 b. [exactly two girls]₁ [-er than 5']₂ t₁ are t₂ tall
 $|\{x: \text{girl}(x) \ \& \ \max\{d: \text{tall}(x,d)\} > 5'\}| = 2$
 c. [-er than 5']₂ [exactly two girls]₁ t₁ are t₂ tall
 $\max\{d: |\{x: \text{girl}(x) \ \& \ \text{tall}(x,d)\}| = 2\} > 5'$

The maximum referred to in (20c) turns out to be well defined. In fact, the maximal degree to which exactly two girls are tall coincides with the maximal degree to which *at least* two girls are tall. So here we finally seem to have found a case that tells us something. The fact is that sentence (20a) unambiguously means (20b) and does not share a reading with the sentence *At least two girls are taller than 5 feet*. But if DegP is allowed to scope over the quantifier, as in (20c), then such a reading is generated. So the analysis makes inadequate empirical predictions unless supplemented by some constraint on DegP-movement.

This is a fine argument given the assumptions I have adopted in this paper. It may not go through, however, in the context of a more comprehensive theory which incorporates an account of polar opposition along the lines of Bierwisch (1987) and Kennedy (1997). Briefly, (20c) may be ruled out by an independently motivated prohibition against sets of degrees which fail to be initial or final segments of the scale. Not being able to go into this, I leave it open here whether non-monotone operators can tell us anything about DegP-movement.¹⁰

1.2 Exactly-Differentials, Less-Comparatives, and Kennedy's Generalization

The equivalences observed in 2.1 hold for simple comparatives (and equatives) and also for comparatives with simple differential phrases, but they no longer hold for comparatives with a differential phrase containing *at most* or *exactly* (or for equatives modified by *at most* or *exactly*). Nor do they hold when we switch from *more*-comparatives to *less*-comparatives. For instance, we saw that if the shortest girl is taller than John, then every girl is. But if the shortest girl is *exactly one inch* taller than John,

¹⁰ The set $\{d: |\{x: \text{girl}(x) \ \& \ \text{tall}(x,d)\}| = 2\}$ in (20c) fails to be an initial segment of the set of all heights whenever there are more than 2 girls, because it does not contain the degrees to which the third-tallest girl is tall. The constraint in question is needed to account for what Kennedy calls "cross-polar anomaly".

it doesn't follow that every girl is. And if the shortest girl is *less* tall than John, it also doesn't follow that every girl is. These and analogous non-equivalences will enable us to construct examples where the low-DegP and high-DegP LFs have clearly distinct yet equally non-deviant truth-conditions, and where we can thus get empirical evidence about the availability of each LF.

Starting with *exactly*-differentials, consider the sentences in (21a) and (22a) and their predicted truth-conditions for LFs with low-DegPs (b) and high-DegPs (c).

(21) a. (John is 4' tall.) Some girl is exactly 1" taller than that.

b. [some girl]₁ [exactly 1" -er than 4']₂ t₁ is t₂ tall
 $\exists x[\text{girl}(x) \ \& \ \max\{d: \text{tall}(x,d)\} = 4' + 1"]$

c. [exactly 1" -er than 4']₂ [some girl]₁ t₁ is t₂ tall
 $\max\{d: \exists x[\text{girl}(x) \ \& \ \text{tall}(x,d)]\} = 4' + 1"$

(22) a. (...) Every girl is exactly 1" taller than that.

b. $\forall x[\text{girl}(x) \rightarrow \max\{d: \text{tall}(x,d)\} = 4' + 1"]$

c. $\max\{d: \forall x[\text{girl}(x) \rightarrow \text{tall}(x,d)]\} = 4' + 1"$

The sentences are judged unambiguous, and the b-LFs correctly represent their intuitive meanings. What about the c-LFs? (21c) says, in effect, that the tallest girl is exactly 4' 1", which appears not to be a possible reading. Since it is a stronger claim than (21b), though, it is a bit hard to prove that it's not an alternate reading. (Speakers may be reluctant to call something false when it's true on another reading.) But with the universal example (22a), we don't have this methodological problem. (22c) expresses a weaker claim than (22b). So we can construct scenarios in which c is true while b is false: just imagine that the shortest girl is exactly 4' 1" but some other girls are taller. If c were a possible reading of the English sentence, then speakers should sometimes be willing to judge it true in this situation. But it clearly is false. So (22c) cannot be a grammatical LF. The DegP apparently cannot scope over the quantificational DP.

Test cases with *less* lead to the same conclusion. Let us first make explicit the meaning of *less*-headed DegPs. The obvious rule is (23), which differs from the rule for *-er* just in the reversal of the >-relation.

(23) $\llbracket \text{less than one foot} \rrbracket = \lambda P \langle d, t \rangle. \max(P) < 1'$

Now consider low-DegP and high-DegP derivations for the sentence in (24a).

(24) a. (John is 4' tall.) Every girl is less tall than that.

b. [every girl]₁ [less than 4']₂ t₁ is t₂ tall
 $\forall x[\text{girl}(x) \rightarrow \max\{d: \text{tall}(x,d)\} < 4']$

c. [less than 4']₂ [every girl]₁ t₁ is t₂ tall
 $\max\{d: \forall x[\text{girl}(x) \rightarrow \text{tall}(x,d)]\} < 4'$

(24b) transparently captures the intuitive meaning of the sentence. (24c) says that the shortest girl is less than 4' tall, which is a weaker claim than b and clearly not an available reading. Again, we must conclude that DegP doesn't scope over DP.

This holds not only for DPs in subject position. (25) has a universally quantified object, but still the DegP seems unable to scope over it.

(25) (The frostline is 3 and a half feet deep.) Mary set every post exactly 2 feet deeper than that.

If *exactly 2' -er than that* took scope over *every post*, the sentence would be true if Mary set just the least deep post exactly 2' below the frostline, but she set the other posts deeper. Intuitively, (25) is false in this case. Even in (26), where the quantified DP is an internal argument and to the right of the graded adjective, the DegP cannot scope over it.

(26) (John gave every candidate an A.)

?Mary was less impressed with every candidate than that.

This cannot be true just because the candidate who impressed Mary the least didn't impress her enough to get an A. The generalization appears to be the one in (27), which I will refer to as "Kennedy's generalization".

(27) If the scope of a quantificational DP contains the trace of a DegP, it also contains that DegP itself.¹¹

The most reasonable diagnosis at this point is that (27) holds generally, not just for the examples in this section, where DegP-scope happens to affect truth-conditions. The

¹¹ This formulation is intended to allow DPs that take scope *inside an argument* of the graded adjective. In (ia,b), the underlined DPs can have narrower scope than DegP, but since they don't intervene between the DegP and its trace, they are not counterexamples to Kennedy's generalization.

- (i) a. *Every student* showed up less often than that.
 'it happened less often than that that every student showed up'
 b. *An earthquake* is more likely than a snowstorm.

The only potentially genuine counterexamples to (27) that I could come up with involve indefinite complements to adjectives of distance, as in (ii).

- (ii) Jaffrey is closer to *an airport* than it is to a train station.

examples in the previous section, where high DegPs yielded interpretations that were either truth-conditionally equivalent to low DegPs or semantically deviant, did not provide evidence for (27). But they were, of course, consistent with it as well.

2.3 Intensional Verbs and Stateva's Argument

In the last section we only looked at DegP-scope in relation to DP-quantifiers. What about intensional verbs? The logical point is the same: when we add *exactly*-differentials or switch from *more* to *less*, we ensure non-equivalence between low and high-DegPs. Let's look at some examples and their predicted readings.

- (28) a. (This draft is 10 pages.) The paper is required to be exactly 5 pages longer than that.
- b. required [[exactly 5 pp -er than that] the paper be t long]
 $\forall w \in \text{Acc: max}\{d: \text{long}_w(p,d)\} = 15\text{pp}$
- c. [exactly 5 pp -er than that] [required [the paper be t long]]
 $\text{max}\{d: \forall w \in \text{Acc: long}_w(p,d)\} = 15\text{pp}$

(28b) says that the paper is exactly 15pp long in every acceptable world. This implies that it is not allowed to be longer than 15pp. (28c) says that the paper is exactly 15pp long in those acceptable worlds where it is shortest. This leaves open whether it might also be allowed to be longer than 15pp. The English sentence (28a) can be understood in either one of these two ways. It seems to be ambiguous, in just the way that we predict if the DegP can move either below or above the necessity operator. The same thing happens with a possibility operator.

- (29) a. (...10 pp.) The paper is allowed to be exactly 5 pages longer than that.
- b. $\exists w \in \text{Acc: max}\{d: \text{long}_w(p,d)\} = 15\text{pp}$
- c. $\text{max}\{d: \exists w \in \text{Acc: long}_w(p,d)\} = 15\text{pp}$

(29b) says that the paper is exactly 15pp in some acceptable worlds, leaving it open that it might also be allowed to be other lengths. (29c) says it is exactly 15pp long in the acceptable worlds where it is longest, which means it is not allowed to be longer than 15pp. Again, these two distinct readings are both available for the English sentence (29a).

Examples of intensional verbs with *less*-comparatives make the same point, as previously argued by Stateva (2000).¹²

¹² Stateva talks about superlatives with *least*, but the point is the same.

- (30) a. The paper is required to be less long than that.
 b. required [[less than that] the paper be t long]
 $\forall w \in \text{Acc}: \max\{d: \text{long}_w(p,d)\} < 10\text{pp}$
 c. [less than that] [required [the paper be t long]]
 $\max\{d: \forall w \in \text{Acc}: \text{long}_w(p,d)\} < 10\text{pp}$

(30b) says that the paper is less than 10pp in every acceptable world; so it's not allowed to be longer. But c merely says that it is less than 10pp in the acceptable worlds where it's shortest. c is equivalent to "the paper is not required to be as long as that". Both of these clearly distinct readings are available. Similarly for the possibility sentence:

- (31) a. (...10 pp.) The paper is allowed to be less long than that.
 b. $\exists w \in \text{Acc}: \max\{d: \text{long}_w(p,d)\} < 10\text{pp}$
 c. $\max\{d: \exists w \in \text{Acc}: \text{long}_w(p,d)\} < 10\text{pp}$

(31c), which means that the paper is not allowed to be as long as 10pp, is a possible reading for (31a).

The availability of the c-readings in these examples indicates that DegP is not, after all, always confined to its narrowest possible scope. Intervening intensional verbs seem to differ from intervening quantificational DPs in this respect, and Kennedy's generalization does not carry over to them.

Or is there a different explanation for these ambiguities? These intensional verbs are known to participate in another strange phenomenon, the so-called scope-splitting of negative and other non-monotone-increasing quantifiers.¹³ For example, *No deposit is required* means that it is not necessary to make a deposit, and *At most three attempts are allowed* means that it is not permitted to make more than three attempts. Maybe the c-readings of our comparative examples above are obtained not by DegP-movement, but by the same mechanism that splits scope in these cases. But what is that mechanism? Scope-splitting phenomena are not well understood and their analysis is controversial. One approach involves decomposition into a wide-scope negation and a narrow-scope residue, e.g., *no deposit* = NEG + *a deposit*, *at most three attempts* = NEG + *more than three attempts*, *exactly 5pp -er than that* = NEG + *more than 5pp -er than that*, *less than that* = NEG + *as as that*. But these decompositions are not all morphologically transparent (e.g., the last one involves a switch from the comparative to its dual, the equative). Another idea is to split off different wide-scope pieces in different cases, negation in some, focus-sensitive adverbs like *only*, *exactly* in others. But this would not seem to cover the *less*-cases. What I would like to suggest instead is that scope-splitting (at least sometimes) *is* DegP-movement. So I agree that the two phenomena are the same, but I question that there is a good analysis of scope-splitting

¹³ See, e.g., Jacobs (1980), Rullmann (1995a), Larson et al. (1997), and de Swart (2000).

that is a genuine alternative to DegP-scoping. Much more work is required to substantiate this suggestion. But assuming that I am on the right track with this, I tentatively conclude (with Stateva) that DegPs are able to scope over intensional verbs.

Does this hold for *all* intensional verbs? A wider range of examples reveals a mixed pattern. A few more verbs act just like *require* and *allow*, e.g., *need* and *be able*. (32a,b) are ambiguous in the same way as our examples above, and in particular permit the indicated high-DegP readings.

- (32) a. John is able to run less fast than that.
 ‘he is not able to run as fast as that’
- b. The paper needs to be exactly 5 pp longer than that.
 ‘the paper’s required minimum length is exactly 5pp longer than that’

But the verbs below behave differently in that the high-DegP readings indicated are not attested.

- (33) The paper might be less long than that.
 *‘it’s not possible for it to be as long as that’
- (34) The paper should be less long than that.
 *‘it’s not required for it to be as long as that’
- (35) The paper is supposed to be less long than that.
 *‘it’s not required for it to be as long as that’
- (36) I want the paper to be less long than that.
 *‘I don’t require it to be as long as that’

Why the difference? There are two imaginable reasons. Either the intensional verbs in (33–36) just don’t allow the DegP to move over them. Or there is something in their semantics that masks the truth-conditional effect of this movement. The second kind of explanation would be more interesting. In (33), we have an epistemic modal, and maybe the absence of the high-DegP reading can be related to a general resistance of epistemic operators against anything scoping over them.¹⁴ In (34–36), all the verbs are so-called neg-raising verbs: their outer negations are systematically understood as inner negations. Perhaps this could have something to do with why the relevant readings fail to surface, but we would need a clearer understanding of neg-raising to substantiate this.¹⁵ At this point, I am unable to spell out any concrete explanations for the unambiguity of (33–36), and it is only a hope that it will follow without specific stipulations about DegP-movement.

¹⁴ See von Stechow & Iatridou (2001).

¹⁵ In the SALT-version of this paper, I made an attempt at this, but Roger Schwarzschild showed me that it didn’t work.

2.4 DegP-Scope and De Re/De Dicto Ambiguity

Some of the older literature on comparatives took for granted that the famous Russell ambiguity in (37) was a matter of DegP-scope.

(37) John thinks the yacht is longer than it is.

When the DegP scopes below the verb *think*, it was assumed, we get the contradictory-thought reading, and the sensible reading is due to the DegP scoping above *think*. Von Stechow (1984) was the first to see clearly that this was a mistake. Analysing examples like (38), he showed that a high-DegP LF would not only be an island violation, but also misrepresent the truth-conditions.

(38) If Mary smoked less than she does, she would be healthier.

We conclude with von Stechow that the Russell-ambiguity is a matter of *de re* or *de dicto* interpretation of the *than*-clause. There is merely a one-directional connection with DegP-scope: a *de dicto* interpretation of the *than*-clause is possible only if the DegP is in the scope of the relevant intensional verb, but a *de re* interpretation is compatible with either wide or narrow DegP-scope.

To spell this out a bit more, I need to fill in a treatment of clausal complements to *than* (which I have avoided in the examples so far). Following standard practice, I take *than*-clauses to be derived by wh-movement of a covert operator from the degree-argument position of an adjective. The trace is interpreted as a variable over degrees. The wh-clause as a whole may be treated as a definite description of a maximal degree (von Stechow 1984).

(39) $\llbracket wh_1 \text{ the bed is } t_1 \text{ long} \rrbracket = \max \{d: \text{long}(\text{the bed}, d)\}$

(39) in conjunction with our meaning for *-er* predicts that the sentence *John is taller than the bed is long* is true iff the (maximal) degree to which John is tall exceeds the (maximal) degree to which the bed is long.

As for the *de re/de dicto* distinction, I assume that LF-representations contain explicit world-arguments for each predicate, and that the world-argument of the restrictor of a quantifier may, but need not, be bound locally. So even if we only consider narrow scope for the DegP, sentence (37) has two LFs, which differ in the choice of world-variable in the *than*-clause.¹⁶

¹⁶ The free variable *w* here stands for the utterance world (the evaluation world for the whole sentence). The idea that quantifiers can have narrow scope even when their restrictors are interpreted *de re* is widely accepted nowadays, not just for comparatives but (especially) for DP-quantifiers. Abusch (1994), for example, makes this point about an example like *If every senator were a rancher instead, I'd be happy* (which is transparently parallel to von Stechow's (38)). See also Percus (2000).

- (40) a. John thinks_w λw'[[*-er than wh it is t long_{w'}*] the yacht is t long_{w'}]
 $\forall w' \in \text{Acc}(w): \max\{d: \text{long}_{w'}(y,d)\} > \max\{d: \text{long}_w(y,d)\}$
- b. John thinks_w λw'[[*-er than wh it is t long_w*] the yacht is t long_{w'}]
 $\forall w' \in \text{Acc}(w): \max\{d: \text{long}_{w'}(y,d)\} > \max\{d: \text{long}_w(y,d)\}$

(40a) expresses the contradictory *de dicto* reading and (40b) the sensible *de re* reading. As we see here, both readings are compatible with narrow DegP-scope, and therefore, Russell ambiguities cannot be used as evidence for scopal mobility of DegPs.¹⁷

But notice that the ambiguities presented in the previous section were not Russell ambiguities. I have deliberately been using phrases after *than* that rigidly pick out degrees, so the issue of *de re* vs. *de dicto* didn't arise in the first place. Even though they are restricted to intensional contexts, the ambiguities we saw with *exactly*-differentials and *less*-comparatives cannot be explained away as involving *de re* without wide DegP-scope.

We can, of course, contrive examples in which there are both DegP-scope and *de re/de dicto* ambiguities.

- (41) The box is required to be less wide than it is tall.

We expect three possible readings, and this, I think, is what we find. (41) can mean that, in order to satisfy the requirements, the box must be taller than wide. This reading is generated by leaving the DegP low and interpreting the *than*-clause *de dicto*. (41) can also mean that the box should be less wide than the height that it actually happens to be. On this reading, the requirements don't regulate the box's proportions, only its width. We obtain this interpretation by means of a low-DegP and a *de re than*-clause. Finally, (41) can mean that the box isn't required to be as wide as its actual height. This reading has a high-DegP and (therefore) a *de re than*-clause. There can be no such thing as a high-DegP, *de dicto than*-clause reading. In this respect, (41) contrasts with (42).

- (42) The box is not required to be as wide as it is tall.

Recall that when the *than*-phrase is rigid, the high-DegP reading of *less* coincides in meaning with the negation of the corresponding equative. Given a *de re* reading of the *than* or *as*-clause, the high-DegP reading of (41) is equivalent to (42). But (42) also

¹⁷ In later work (1993, 1998), von Stechow reversed his earlier position and decided that (40b) was not, after all, sufficient to capture the sensible reading of (37). The problem he saw is that (40b) has a comparison in the complement of *think*, but intuitively the thought reported by (37) need not have a comparative content. For example, John might just be thinking "the yacht is 14' long" (when in fact it is 13' long). I don't see the force of this objection (at least not given the semantics I am assuming here). (40b) just says that in every doxastic alternative of John's, the yacht is longer than it is in the actual world. This is true if he thinks it is 14' long. Despite its form, (40b) does not actually say he has a comparative thought.

allows a reading not shared by (41), on which the *as*-clause is *de dicto*: (42) can mean that it isn't required for the box to be as wide as tall (i.e., that it is allowed to be taller than wide). These judgements may be too tenuous to build strong conclusions on. But if they are real, then we may also have an argument here against at least one version of a scope-splitting analysis, which would say that *less* is decomposed into negation and equative, and only the negation can outscope the intensional verb.

3. DegP-Scope and Syntactic and Semantic Ellipsis

3.1 Syntactic Ellipsis: VP-Deletion

A standard argument for QR (covert movement of DPs) is that it is needed to license ellipsis in examples which on the surface suffer from antecedent containment (henceforth AC). For example, the deleted VP in (43a) seems to have an antecedent (the matrix VP) which contains it and therefore is not identical to it.¹⁸ But after QR has applied, as in (43b), AC is resolved and the two VPs match perfectly.

- (43) a. I read every book *wh* that you did ~~read t~~.
 b. [every book *wh* that you did ~~read t~~] I read t

A completely analogous argument can be put forward for covert DegP-movement. In the SS of (44a), the deleted VP is contained in its antecedent, the matrix VP. If covert DegP-movement applies as in (44b), the result is a structure with no AC and properly matched VPs.

- (44) a. Mary ran faster than *wh* John did ~~run t fast~~.
 b. [-er than *wh* John did ~~run t fast~~] Mary ran t fast

The idea that DegP-movement is involved in licensing ellipsis in comparative clauses has a long tradition.¹⁹ One recent argument in its favour was presented by Wold (1995). Carlson (1975) had observed that antecedent-contained deletion (ACD) in relative clauses is degraded when the head of the relative clause is a weak indefinite, especially an existentially read bare plural.²⁰

¹⁸ I use strike-out to represent an elided VP with a certain intended interpretation.

¹⁹ See Bresnan (1973) and Sag (1976), among others.

²⁰ There is some doubt about the correctness of Carlson's descriptive generalization. According to Pesetsky (pc), at least some examples of this sort improve with the addition of *too* in the elliptical clause. This may undermine Wold's argument.

(45) *John was climbing trees that Bill was.

Diesing (1992) explained this by her thesis that existential bare plurals do not QR, but remain inside VP to get bound by existential closure. Thus they cannot get out of the VP to resolve AC, and ellipsis cannot be licensed. But (as also noted by Carlson) existential bare plurals do allow ACD in comparative clauses.

(46) a. John was climbing higher trees than Bill was.

b. John was climbing more trees than Bill was.

Diesing left this as a potential challenge for her account. Wold showed that there is no problem if AC can also be resolved by DegP-movement instead of QR.

(47) [-er than wh Bill was ~~climbing t high trees~~] John was climbing t high trees

The bare plurals in (46) can stay in VP, low enough to get caught by existential closure, while DegP moves high enough to license the ellipsis.

In the first part of this paper, we found some semantic and syntactic constraints limiting where DegPs can move. If DegP-movement is indeed a prerequisite for ellipsis licensing, these constraints should affect the range of possible ellipses and their interpretations. Let us see whether this is true.

One thing that we saw above was that DegP-movement cannot cross over a quantificational DP (Kennedy's generalization). So what will happen if we construct an example in which, on the one hand, DegP must move out of VP to resolve AC, but, on the other hand, there is an intervening DP-quantifier in that VP? (48) is a case in point.

(48) John pushed every weight higher than Mary did.

What we expect is that the only grammatical derivation will be one in which the DP *every weight* first moves out of the VP, so that then the DegP can also move out of VP without crossing over it. (It will only cross over its trace, which is fine.) So we get the LF in (49).

(49) every weight₁ [-er than wh Mary did ~~push t₁ t high~~]₂ John pushed t₁ t₂ high

(49) says that for every weight *x*, John pushed *x* higher than Mary pushed *x*. This is what the English sentence in (48) means, and indeed the only reading that it appears to have. The general prediction is that quantificational DPs must scope "out of the way" whenever they are in the path of a DegP containing an antecedent-contained elided VP. This has intricate empirical implications, which for the most part remain to be tested. For example, we predict an interference with sloppy-identity construals of pronouns in examples like (50).

(50) I planted one of my trees deeper than you did.

If we are on the right track, (50) should only allow a pragmatically odd reading: I planted one of my trees deeper than you planted that same tree. It is not obvious that this is correct, but also not easy to say what other readings are possible. More work is called for here.²¹

We also observed that DegP doesn't scope over negation and other monotone decreasing items, since the resulting readings would systematically be presupposition failures. If DegP-movement is needed to resolve ACD, it follows that (51) is ungrammatical on any reading that takes the matrix VP headed by *refuse* to be the antecedent of the elided VP.

(51) *John refused to run faster than Mary did ~~refuse to run t fast~~.

This is correct. (The only good reading for this surface string is that John refused to run faster than Mary *ran*.)²² However, it turns out that we would be making this correct prediction regardless of whether DegP-movement plays any role in ellipsis licensing. The reading indicated in (51) also involves a presupposition failure *within* the *than*-clause: there is no maximal degree *d* such that Mary refused to run *d*-fast. This in itself rules out the intended construal of the elided VP, independently of any assumptions about how ellipsis is licensed. So the judgement in (51) is merely consistent with the hypothesis that DegP-movement is needed for ellipsis licensing; it does not help support it.

With those intensional verbs that we saw a DegP *can* cross, we get ambiguity of antecedent-size, as expected. (52) is from Williams (1974).

- (52) a. My father tells me to work harder than my boss does.
 b. My father tells me to work harder than my boss does ~~work t hard~~.
 c. My f. tells me to work harder than my b. does ~~tell me to work t hard~~.

²¹ To make sense of the facts in this area, it will be necessary, for one thing, to take a systematic look at the behaviour of quantifiers inside the *than*-clause. There are well-known puzzles here, and a recent paper by Schwarzschild & Wilkinson compellingly argues that they can only be solved if we change some of the basic assumptions that I have unquestioningly taken over from the literature. Obviously, I need to make my reasoning consistent with their insights somehow, and before I manage to do this, there is reasonable doubt as to whether this paper is even asking the right questions.

²² This is a reading where the elided VP is headed by *run* and is interpreted *de re*. That is, we are saying about Mary's actual speed that John refused to surpass it. It may be marginally possible also to get a *de dicto* reading, meaning that John refused to outrun Mary. That reading is easier if the auxiliary is changed from *did* to *would*, or (even better) if we switch to the Comparative Ellipsis variant with no auxiliary at all.

- (53) a. John needs to drive faster than Mary does.
 b. John needs to drive faster than Mary does ~~drive t fast~~.
 c. John needs to drive faster than Mary does ~~need to drive t fast~~.

To bring out each of the two readings of (53), imagine two contexts. In the context for reading b, we are talking about what it takes for John to win the race. Here there is no absolute minimal requirement for the speed he needs to attain; what matters is that his speed exceed Mary's. In the context for reading c, we are discussing John's and Mary's chances of making it to Ithaca by 8 pm tonight (at the latest). Since John is currently 300 miles away from Ithaca, whereas Mary is only 200 miles away, the minimal speed for him to make it exceeds the minimal speed for her to make it. (Notice that there is no requirement here on their *relative* speeds.)

Interestingly, this ambiguity of antecedent-choice carries over to examples with verbs which we found *not* to show truth-conditionally distinct high-DegP readings, like the neg-raising verb *want*. (54) is the variant of Williams' example analysed in Sag (1976).

- (54) a. Mary's father wants her to work harder than her boss does.
 b. Mary's father wants her to work harder than her boss does ~~work t hard~~.
 c. M's f. wants her to work harder than her b. does ~~want her to work t hard~~.

This indicates that DegP can, after all, scope over these verbs, and that the reason why we didn't see truth-conditional reflexes of this scoping (when looking at *less*-DegPs) must be in their semantics. So we really need to identify the link between neg-raising behaviour and lack of DegP-scope ambiguity that I hypothesized at the end of section 2.3.

Williams and Sag used these examples to point out a correlation between ellipsis and availability of *de dicto* readings. The non-elliptical (or less elliptical) (55) allows a *de dicto* reading of the *than*-clause that is not shared by its elliptical counterpart (52c).

- (55) My father tells me to work harder than my boss tells me to.

(55) can describe a state of affairs where my father tells me: "I don't care how many hours you work, but make sure it's more than whatever your boss demands. If he says you have to work 20 hours, work at least 21; if he tells you to work 30, work 31; etc." (52a) cannot describe this scenario on any of its readings. The only reading it shares with (55) is a *de re* reading, where the embedded *tell*-clause is evaluated in the utterance world, not in the worlds conforming to what my father demands. That is, (52c) has to mean that my boss imposes a certain minimum requirement for how hard I have to work, and my father imposes another one which is higher than that. Williams and Sag took this fact to support the hypothesis that ellipsis licensing depends on DegP-scope. They reasoned on the basis of a naïve identification of *de re* and *de dicto*

with wide and narrow scope, respectively. As we saw in 2.4 above, this is not justified; there is only a one-directional implication, namely that if the DegP's scope is wide, the *than*-clause must be read *de re*. But fortunately, it is only this direction of the connection that is required for their argument. If the DegP must scope over the matrix verb in order to license the broad construal of the ellipsis, then the *than*-clause gets carried along out of the matrix verb's scope and therefore can only be read *de re*. This is just what was observed.

We can confirm the correlation between ellipsis-size and DegP-scope even more directly by changing Williams' kind of example to a variant with *less*.

(56) Mary needs to drive less fast than John does ~~need to drive t fast~~.

When the elided VP is understood to be headed by *need* (as indicated), the DegP must scope over *need*, and the truth-conditions therefore are those of a negated equative ("she doesn't need to drive as fast as he needs to drive"). This prediction is also correct.

What have we learned in this section? There is more fieldwork to be done in the area where ACD in comparative clauses interacts with Kennedy's generalization. Apart from this, it appears that our previous conclusions about DegP-movement are at least consistent with the view that ellipsis licensing relies on DegP-movement. Our last set of facts, about DegP-scope and ellipsis-size in examples with intensional verbs, shows moreover that this combination of assumptions makes some interesting correct predictions. It is difficult, however, to draw stronger conclusions at this point. We have not given any consideration to alternative hypotheses, which might account for ellipsis licensing without appeal to DegP-movement. Before we examine concrete versions of such alternative theories, we cannot know whether they will have any difficulty making the same correct predictions also.

What such alternatives might there be? That depends, for one thing, on our background assumptions about VP-ellipsis in general. We have tacitly taken for granted here that an elided VP and its antecedent must have LFs that are identical up to possible differences in indices.²³ If we consider the possibility that some less restrictive parallelism condition might be sufficient, there may be ways to establish the required parallelism without DegP-movement.²⁴ Even if we stick to the strict (near-)identity

²³ See Rooth (1992) and work based on this for more detail on the assumptions about ellipsis licensing that underlie this requirement.

²⁴ For example, it may suffice to move just the *than*-clause or the *wh*-clause after *than*:

(i) [wh John did ~~run t fast~~] Mary ran [-er than t] fast

The VPs are still not identical here, but AC is resolved. And as Kennedy (pc) points out, the existential closures of the two VPs are semantically equivalent ($\lambda x.\exists d [x \text{ ran } d\text{-fast}] = \lambda x.\exists d [x \text{ ran faster than } d]$), which suffices for ellipsis licensing according to a recent proposal by Merchant (1999). But this idea runs into a problem with differential phrases, as in *Mary ran 5 km/h faster than John did*, where the analogous equivalence fails. ($\lambda x.\exists d [x \text{ ran } d\text{-fast}] \neq \lambda x.\exists d [x \text{ ran } 5 \text{ km/h faster than } d]$.)

condition we have been assuming, it is not obvious that DegP-movement is the only way to resolve AC. What about moving the entire AP, as in (57)?

(57) [[-er than wh John did ~~run~~ fast] Mary ran t

For this to be a viable alternative, we would need appropriate interpretive mechanisms to deal with the AP-sized traces and their binding in this structure. It is not at all obvious what exactly these should be, how they would affect the reasoning in section 2, and whether this path in the end would lead to something substantively different from DegP-movement.²⁵ It is incumbent upon me to deal with these questions, but not in the space of this paper.

3.2 Semantic “Ellipsis”: Superlatives and *Too*

As we just saw, it is difficult to make a strong argument for DegP-movement on the basis of facts about VP-ellipsis, because one needs so many auxiliary assumptions. Part of the difficulty is that it is controversial just how this kind of ellipsis is licensed and how close the match between the missing VP and its antecedent has to be. To side-step this controversy, it would be useful if we could look at cases of what I will call here “semantic ellipsis”. By this I mean a construction in which there aren’t two parallel pieces of syntactic (LF) structure which stand in an anaphor-antecedent relation, but rather there is a single piece which, however, is used twice in the semantic calculation. Analyses of this kind have often been proposed as competitors for analyses in terms of syntactic ellipsis, notably for so-called Comparative Ellipsis,²⁶ but there we would be getting onto even more controversial ground. The constructions I will look at here are ones for which (to my knowledge, at least) a semantic-ellipsis analysis is the only kind of analysis on the market.

Consider the superlative.

(58) John screamed (the) loudest.

(58) can be paraphrased by a comparative: *John screamed louder than anyone else did*. The LF of the comparative paraphrase contains *two* degree predicates, one in the

²⁵ One proposal for interpreting structures like (57) is worked out in Kennedy (1997). I will discuss this on another occasion.

²⁶ See, e.g., the “direct analysis” of phrasal comparatives in Heim (1985), and similar analyses for Comparative Ellipsis as well as related constructions like Bare Argument Ellipsis and Gapping that are defended by Reinhart & Rooth (1986), Reinhart (1989), Kennedy (1997), and others. See also Lechner (1999) for critical discussion.

matrix clause and one in the *than*-clause. The superlative has only a matrix clause and still manages to convey the same meaning. The following analysis reveals how.²⁷

- (59) a. $\llbracket -est \rrbracket = \lambda R_{\langle d, et \rangle} . \lambda x . \max \{ d : R(x, d) \} > \max \{ d : \exists y \neq x : R(y, d) \}$
 b. LF for (58): John -est₁ \llbracket [scream t₁ loud] \rrbracket

Notice that R (the denotation of *-est*'s complement) is "used twice" in the calculation prescribed by entry (59a). We use it to determine one set of degrees for John, and another one for everybody else.

Recall now our example (53) (repeated here as (60a)) and compare it with the analogous superlative in (60b).

- (60) a. John needs to drive faster than Mary does.
 b. John needs to drive fastest.

The two readings we observed in (60a) are replicated in (60b). To bring out each reading, just amend the scenarios that I gave above by the addition of a third person, say Bill. Then (60b) might mean that John needs to outdo both Mary and Bill (say, in order to win the race). Or it might mean that the minimal speed that John needs (to get to Ithaca by 8 pm) is higher than the minimal speeds needed by Mary or Bill (since they are already nearer to Ithaca). In the comparative example (60a), the ambiguity clearly involves different-sized choices of antecedent for the VP-ellipsis. On the view that ellipsis licensing requires LF-identity – but not independently of this view – this implies that the two readings also differ in the scope of the DegP. In the superlative variant (60b), there is no ellipsis to resolve. So how do we generate the analogous two readings? The obvious way – and here the only one available to us – is by scoping the DegP. As the reader can verify, if we scope *-est* below *need*, as in (61a), the entry in (59a) gives us the first reading, and if we scope it above, as in (61b), it gives us the second one.

- (61) a. John₁ needs [t₁ -est₂ to drive t₂ fast]
 b. John -est₂ λ₁ [needs [t₁ to drive t₂ fast]]

Our analysis of the ambiguity in the superlative examples is parallel to our analysis for the comparative variant: in both cases, we invoke DegP-movement across the intensional verb to obtain the second reading. So what is new here, what do we learn from the superlative example that we haven't already learned from the comparative one? The new point is that the evidence for DegP-movement needn't be contingent on assumptions about ellipsis licensing. Whether or not you accept the relatively strict

²⁷ Heim (1985). In Heim (1999), I propose a variant that makes the semantic type the same as for the comparative and equative DegPs (by eliminating the external argument). This amendment would not affect the present reasoning.

parallelism conditions on VP-ellipsis that imply that DegP must be able to move above *need* in (60a), you can't get around this movement if you want to account for all the readings of (60b).

For the same methodological reason, it may be useful to look at gradation constructions with *so ... that, too, and enough*.²⁸ These also seem to involve "semantic ellipsis", in the sense that the material in the scope of the degree-operator is used twice in the semantic calculation. For example, *too tall* roughly means "taller than is compatible with certain (contextually given) goals or desires". If John is too tall, he is taller than it is acceptable for him to be tall. Notice how the clause "John (be) tall" appears twice in this paraphrase, once unmodalized to the left of *than*, and another time under a possibility operator to its right. The following entry for *too* makes this transparent in the double occurrence of P.²⁹

$$(62) \llbracket \textit{too} \rrbracket^W = \lambda P_{\langle s, dt \rangle}. \max(P(w)) > \max \{d: \exists w' \in \text{Acc}(w): P(w')(d) = 1\}$$

Consider now an example like (63).

(63) John needs (to have) too much money.

(63) has one reading where it means that what John needs is to have too much money. Imagine he is the type of person who only feels secure when he has more money than he can spend. Unless he has too much money (for him to spend), his (emotional) needs are not satisfied. But (63) also has another reading, where it means that John's financial needs are too high. Imagine that he has many financial obligations (mortgage, alimony, gambling debts, ...). He considers taking a job as a linguistics professor, but thinks better of it, since his needs regarding salary are too high. He needs too much money for him to be a linguist.

These two readings differ in the contextually supplied accessibility relation for the implicit possibility operator associated with *too*, but that is not their only difference. They also differ in the scope of *too* with respect to *need*. For the first reading, the LF in (64a) is adequate, but for the second one, it must be as in (64b).

$$(64) \text{ a. } \text{John}_1 \text{ needs } [\textit{too}_2 [\text{t}_1 \text{ to have } \text{t}_2 \text{ much money}]] \\ \forall w \in \text{Acc}_{\textit{need}}: \max \{d: j \text{ has } d\text{-much money in } w\} > \\ \max \{d: \exists w' \in \text{Acc}_{\textit{too}}(w): j \text{ has } d\text{-much money in } w'\}$$

²⁸ See von Stechow (1984) for a treatment of *too*, and Meier (2000) for a detailed recent investigation and semantic analysis of all three constructions.

²⁹ The same point still holds when we consider examples where *too* selects a *for*-infinitival. The *for*-clause supplies an explicit restrictor for the modal operator introduced by *too*, but it does nothing to spell out its nuclear scope. So we still need to use DegP's sister twice, as shown in the following rule for *too* + (*for*-)CP.

(i) $\llbracket \textit{too CP} \rrbracket^W = \lambda P. \max(P(w)) > \max \{d: \exists w' \in \text{Acc}(w): \llbracket \textit{CP} \rrbracket^{W'} = 1 \ \& \ P(w')(d) = 1\}$

- b. too_2 [John₁ needs [t₁ to have t₂ much money]]
 $\max\{d: \forall w \in \text{Acc}_{\text{need}}: j \text{ has } d\text{-much money in } w\} >$
 $\max\{d: \exists w \in \text{Acc}_{\text{too}}: \forall w' \in \text{Acc}_{\text{need}}(w): j \text{ has } d\text{-much money in } w'\}$

Acc_{need} is meant to pick out the worlds in which John's needs are met. Acc_{too} is different for each reading. For the intended interpretation of (64a), let's say it maps each w to the set of worlds which are like w in the respects that determine spending limits and in which John spends all his money. So (64a) says that John's needs are met only if he has more money than for him to have is compatible with him spending it all. For the intended interpretation of (64b), let Acc_{too} map each w to the set of worlds in which people obey the laws in effect in w (without starving) and in which John is a linguist. Then (64b) says that John's financial needs are greater than for them to be is compatible with him being a linguist. The important point for our purposes is that what is incompatible with being a linguist is not *having* lots of money, but *needing* lots of money. This is why we need to scope *too* over *need* in order to capture this second reading properly.

Due to the multiple layers of modality and concomitant context-dependency, the analysis of this type of example is rather complex. Still, I think we can isolate the contribution that is made by the degree operator's scope in relation to the intensional verb. If so, then we have another way here of detecting DegP-scope without taking a stand on licensing-principles for syntactic ellipsis.

4. Conclusion

We have probed the scopal behaviour of degree quantifiers, using direct evidence from truth-conditional ambiguity as well as more indirect evidence from the interpretation of elided material. The argumentation was incomplete and tentative in a number of places, but suppose the picture that began to emerge is right. Then DegPs are mobile, but they are not allowed to cross over quantificational DPs. This appears to be a syntactic constraint of some kind, because the prohibited scopings would not in any way be semantically deviant. And if our treatment of ellipsis licensing is right, it is not a constraint on the absolute length of movement, since non-quantificational DPs in the same domains (including QR-traces of quantifiers) can be crossed over. Rather it is a constraint that specifically targets interveners of a certain kind. The situation is thus highly reminiscent of the intervention effects in *wh*-questions studied by Beck (1996), although the precise connections remain to be explored. More data need to be investigated to decide whether the parallel really holds up. And most importantly, we need to work out a syntactic analysis in which the observed constraint has principled reasons.

Endnotes

This paper appears in almost identical form in B. Jackson & T. Matthews (eds.) *Proceedings of SALT 10*, CLC Publications, Cornell University. I had hoped to revise it substantially, but have not been able to do so in the available time. For help with the preparation of the SALT-version, I would like to thank Sigrid Beck, Rajesh Bhatt, Danny Fox, Martin Hackl, Roumi Izvorski, Chris Kennedy, Lisa Matthewson, Cécile Meier, Yael Sharvit, Penka Stateva, Arnim von Stechow, Zoltán Szabó, and Karina Wilkinson. More recently, I received very detailed and thoughtful written comments from Cécile Meier, Roger Schwarzschild, and an anonymous reviewer. They have convinced me that the present version suffers from many inadequacies, not just in matters of exposition, but also with regard to the descriptive claims, analyses, and argumentation. I have every intention of following up on what I have learned from their criticism, and I apologize to them and to Arnim that this is not reflected in my contribution to this volume.

References

- Abney, Steven (1987). *The English NP in its Sentential Aspect*. PhD dissertation, MIT.
- Abusch, Dorit (1994). The Scope of Indefinites. *Natural Language Semantics* 2. 83–135.
- Beck, Sigrid (1996). Quantified Structures as Barriers for LF Movement. *Natural Language Semantics* 4. 1–56.
- Bierwisch, Manfred (1987). Semantik der Graduierung. In: Manfred Bierwisch & Ewald Lang (eds.) *Grammatische und konzeptuelle Aspekte von Dimensions-Adjektiven*. (Studia Grammatica 16.) Berlin: Akademie-Verlag. 91–286.
- Bresnan, Joan (1973). Syntax of the Comparative Clause Construction in English. *Linguistic Inquiry* 4. 275–343.
- Carlson, Gregory (1975). Amount Relatives. *Language* 53. 520–542.
- Diesing, Molly (1992). *Indefinites*. Cambridge, MA: MIT Press.
- Fintel, Kai von & Sabine Iatridou (2001). *On the Interaction of Modals, Quantifiers, and If-Clauses*. Ms, MIT.
- Fox, Danny & Jon Nissenbaum (2000). *Extraposition and Covert Movement*. Ms, Harvard University & MIT.
- Guéron, Jacqueline & Robert May (1984). Extraposition and Logical Form. *Linguistic Inquiry* 15. 1–31.
- Heim, Irene (1985). *Notes on Comparatives and Related Matters*. Ms, University of Texas, Austin.
- Heim, Irene (1999). *Notes on Superlatives*. Ms, MIT.
- Jacobs, Joachim (1980). Lexical Decomposition in Montague Grammar. *Theoretical Linguistics* 7. 121–136.

- Kennedy, Christopher (1997). *Projecting the Adjective*. PhD dissertation, University of California, Santa Cruz [also Garland Press 1999].
- Larson, Richard, Marcel den Dikken & Peter Ludlow (1997). *Intensional Transitive Verbs and Abstract Clausal Complementation*. Ms, SUNY Stonybrook.
- Lechner, Winfried (1999). *Comparatives and DP-Structure*. PhD dissertation, University of Massachusetts, Amherst.
- Meier, Cécile (2000). *Konsekutive Konstruktionen und relative Modalität*. PhD dissertation, University of Tübingen.
- Merchant, Jason (1999). *The Syntax of Silence*. PhD dissertation, University of California, Santa Cruz.
- Percus, Orin (2000). Constraints on Some Other Variables in Syntax. *Natural Language Semantics* 8. 173–229.
- Reinhart, Tanya (1989). Elliptic Conjunctions – Non-Quantificational LF. In: Asa Kasher (ed.) *The Chomskyan Turn*. Oxford: Blackwell. 360–384.
- Reinhart, Tanya & Mats Rooth (1986). *Taking Tiger Mountain by Strategy*. Ms, Tel Aviv University & University of Stuttgart.
- Rooth, Mats (1992). Ellipsis Redundancy and Reduction Redundancy. In: Stephen Berman & Arild Hestvik (eds.) *Proceedings of the Stuttgart Ellipsis Workshop*. University of Stuttgart. 1–26.
- Rullmann, Hotze (1995a). Geen eenheid. *Tabu* 25. 195–197.
- Rullmann, Hotze (1995b). *Maximality in the Semantics of WH-Constructions*. PhD dissertation, University of Massachusetts, Amherst.
- Sag, Ivan (1976). *Deletion and Logical Form*. PhD dissertation, MIT.
- Schwarzschild, Roger & Karina Wilkinson (1999). *Interval Semantics for Scalar Predication*. Ms, Rutgers University.
- Stateva, Penka (2000). *In Defense of the Movement Theory of Superlatives*. In: R. Daly & A. Riehl (eds.) *Proceeding of ESCOL 1999*. (Distributed by Cornell Linguistics Club). 219–226.
- Stechow, Arnim von (1984). Comparing Semantic Theories of Comparison. *Journal of Semantics* 3. 1–77.
- Stechow, Arnim von (1993). Die Aufgaben der Syntax. In: Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.) *Syntax – A Handbook of International Research*. Berlin: de Gruyter. 1–88.
- Stechow, Arnim von (1998). *Some Remarks on Ch. Kennedy's Syntax and Semantics of Comparison*. Ms, University of Tübingen (presented at Workshop Syntax and Semantics of Comparatives, Berlin).
- Swart, Henriette de (2000). Scope Ambiguities with Negative Quantifiers. In: Klaus von Heusinger & Urs Egli (eds.) *Reference and Anaphoric Relations*. Dordrecht: Kluwer. 109–132.
- Williams, Edwin (1974). *Rule Ordering in Syntax*. PhD dissertation, MIT.
- Wold, Dag (1995). *Antecedent-Contained Deletion in Comparative Constructions*. Ms, MIT.