

Against ‘Long’ Movement of the Superlative Operator

Yael Sharvit and Penka Stateva
University of Connecticut

It is a matter of considerable debate whether degree operators are interpreted in their base position or in some higher position. Kennedy (1997) has shown that degree operators (e.g., the comparative operator) do not interact scopally with quantified expressions. On the other hand, Heim (1999) and Stateva (to appear) have presented evidence that the superlative operator (as in *the highest grade*) interacts scopally with intensional predicates. This paper argues that despite the fact that the superlative operator seems to interact scopally with intensional predicates, the facts receive a better account under an *in situ* analysis, rather than a movement analysis, of the operator. This point will be made by (a) looking at examples where a superlative expression is embedded under a propositional attitude verb, and focussing on readings which are neither *de re* (in the strict sense) nor *de dicto* (in the strict sense); and (b) examining negative superlatives (e.g., *the least high grade*) in extensional contexts. Our conclusion will be that Kennedy’s claim that degree operator movement is highly restricted is correct.

1. An *In Situ* Analysis and a Movement Analysis of the Superlative Operator

What is the proper analysis of sentences such as (1)?

(1) John got the highest grade.

The literature recognizes (at least at the descriptive level) that (1) is ambiguous between two readings – absolute and comparative (Ross (1964), Szabolcsi (1986)). According to the absolute reading, John’s grade scores highest among the relevant grades, and (1) is appropriate, for example, in a situation where John got an A (assuming that A is the highest grade in the American grading system). According to the comparative reading, John scores highest among the relevant grade receivers, and (1) is appropriate, for example, in a situation where John got a B, Bill – a C and Mary – a D. Under the absolute reading of (1), the prominent comparison is between grades. Under its comparative reading, the prominent comparison is between grade receivers (or their achievements, rather). It is a matter of some debate whether or not this difference in the locus of comparison is actually a genuine difference in meaning (i.e., whether we have two distinct readings on our hands).

There is no dispute that the absolute reading arises from an LF where the superlative operator does not scope above *get*. But how the comparative reading comes about (and whether it should be viewed as a reading distinct from the absolute reading) is not a settled issue. According to the ‘one-reading’ view, the

absolute and comparative “readings” are not distinct readings – but rather reflect different strategies for choosing the value for the first argument of *-est* (i.e., the restriction of *-est*, whose role is explained below). According to the ‘two-readings view’, the comparative reading arises via movement of *-est* above *get* at LF, and deletion of the definite determiner.

To illustrate how these theories work, we assume the following. First, a gradable adjective such as *high* denotes a function from degrees to $\langle e, t \rangle$ -functions (Seuren (1973), Cresswell (1976), and others), with the following meaning:

- (2) For any degree d and individual x , $[[high]](d)(x)=1$ iff x is d -high

Modified nominal expressions such as *high grade*, where the modifier is a gradable adjective, also denote functions of this type (see Heim (1999) for discussion of the compositional derivation). Each function R in the *high*-class (e.g. *high*, *high grade*, *tall man*, etc.) is monotone, in the sense that the following holds for R :

- (3) $\forall x \forall d \forall d' [R(d)(x)=1 \ \& \ d' < d \rightarrow R(d')(x)=1]$

For example, *John is four feet tall* entails *John is three feet tall*.

Secondly, the meaning of *-est* is the following (essentially as in Heim (1999)), where K is a restriction on the domain of the superlative operator (a comparison set), R – a function of the *high*-class, and x – an individual:

- (4) $[[est]](K)(R)(x)$ is defined only if $x \in K$ and $\forall y [y \in K \rightarrow \exists d [R(d)(y)=1]]$;
whenever defined, $[[est]](K)(R)(x) = 1$ iff
 $\exists d$ such that $\{z \in K : R(d)(z)=1\} = \{x\}$.

The domain restriction argument is a phonetically null variable, whose value is supplied by the context. (5)-(6) illustrate how each theory handles the interpretation of (1) (we use ‘ K ’ to represent the phonetically null restriction variable, and ‘ k ’ to represent the contextually supplied value of the variable):

- (5) *In situ* theory (One-reading/two strategies):

(Heim 1999)

- a. LF: John got the $[[K-est] [high grade]]$
- b. John got the unique x such that $\exists d [\{z \in K : z \text{ is a } d\text{-high grade}\} = \{x\}]$
- c. “Absolute” strategy for determining the value of K
 $K = \{x : \exists d [x \text{ is a (relevant) } d\text{-high grade}]\}$
- d. “Comparative” strategy for determining the value of K
 $K = \{x : \exists d \exists y [x \text{ is a } d\text{-high grade} \ \& \ y \text{ is a (relevant) person} \ \& \ y \text{ got } x]\}$

-est does not move outside the DP it originates in. The choice of the domain restriction determines which of the “readings” – absolute or comparative – is prominent. If K includes a bunch of grades that are not necessarily linked to

different receivers, the prominent “reading” is the absolute. If K includes at least one grade per relevant grade receiver, the prominent “reading” is the comparative.

(6) Movement theory (two-readings):

Absolute reading

LF: John got the [[K -est] [high grade]]

Comparative reading

(Szabolcsi (1986), Heim (1985), Heim (1999))

a. LF: John [K -est] λd [got ~~the~~ d-high grade]]

b. $\exists d$ such that $\{z \in K : z \text{ got a } d\text{-high grade}\} = \{\text{John}\}$

c. $K = \{x : x \text{ is a (relevant) person} \ \& \ \exists d[x \text{ got a } d\text{-high grade}]\}$

The comparative reading is obtained by moving *-est* to a position above *get*, abstracting over its (degree denoting) trace, and deleting *the* (thus obtaining a function of the *high*-class). The external argument of *-est* is *John*, and K is a set of grade receivers, not grades.

Can we decide between the two approaches? There are arguments in favor of both theories, and we will not review all of them here (the interested reader is referred to Heim (1999) for discussion). Instead, we will focus on what we find to be the most interesting and compelling argument in favor of the movement theory. This argument comes from examples where the superlative expression is in the (surface) scope of an intensional verb. Such examples (discovered in Heim (1999)) sometimes give rise to an intermediate – neither *de re* nor *de dicto* – reading of the superlative expression. We call these readings ‘upstairs *de dicto*’ readings.

2. ‘Upstairs *De Dicto*’ Readings – a Challenge for the *In Situ* Theory

Consider the following sentence:

(7) John needs to get the highest grade.

This example has an obvious *de re* reading and an obvious *de dicto* reading (in fact, if one believes that the absolute and comparative readings are genuinely two distinct readings, then (7) has two *de re* and two *de dicto* readings, but we will assume momentarily that the absolute and comparative “readings” are derived from the same LF, using different strategies for restricting the domain of *-est*, cf. (5)):

De re

(8) a. John needs₀ λ_1 [PRO to get₁ [the [[K -est] [high-grade₀]]]]

b. In all worlds w compatible with John’s needs in the actual world, he gets the actual grade that is higher than any other relevant actual grade.

De dicto

(9) a. John needs₀ λ_1 [PRO to get₁[the [[f_1 -est] [high-grade₁]]]]

- b. In all worlds w compatible with John's needs in the actual world, he gets the grade in w that is higher than any other relevant grade in w .

The existence of the *de re* and *de dicto* readings is unsurprising and does not shed any new light on the debate. But (7) has an additional, unexpected reading – ‘upstairs *de dicto*’ – which is the one we are interested in here. This reading is brought out in scenarios such as the following. Suppose we just conducted a survey among students about the grades they need to get in order to meet their school's requirements. (10) lists the elicited answers:

(10) Scenario I

Mary: “I need to get a C on the Math test.”

Bill: “I need to get a B- on the Math test.”

John: “I need to get a B+ on the Math test.”

(11) illustrates what goes on, according to this survey, in the worlds compatible with the needs of each of the relevant individuals: in each of the worlds compatible with her needs (her “need” worlds), Mary gets a grade which is not lower than C; in each of his “need” worlds Bill gets a grade which is not lower than B-; and in each of his “need” worlds John gets a grade which is not lower than B+:

(11) <u>Student</u>	<u>“Need” worlds/grades</u>			
a. Mary	<u>w1</u>	<u>w2</u>		<u>w15</u>
	C	C+	...	A
b. Bill	<u>w225</u>	<u>w226</u>		<u>w237</u>
	B-	B	...	A
c. John	<u>w339</u>	<u>w340</u>		<u>w357</u>
	B+	A-	...	A

Given this state of affairs, we can report the results of our survey by uttering (7). Clearly, the meaning expressed by (7) in these circumstances is neither the one corresponding to the *de re* reading of (7) (since there is no particular actual grade that John needs to get), nor the one corresponding to its *de dicto* reading (because John did not mention and did not express any necessity regarding the “height” of the grades received by the others). How can we account for this meaning then?

Heim considers and rejects two *in situ* solutions, as shown in (12). The idea here is to try to exploit the variable which denotes the comparison set (for simplicity, let us assume that the definite determiner optionally deletes):

- (12) a. John needs₀ λ_1 [PRO to get₁ [(the) **K**-est [high-grade₁]]]
 b. John needs₀ λ_1 [PRO to get₁ [(the) **f**₁-est [high-grade₁]]]

Let us start with (12a). Which grades should we “cram into” K ? Suppose that K contains all the grades that either John, Bill, or Mary got in all the worlds

compatible with their needs. Clearly, this would not give us the meaning we are after because some of these grades will be very high (given that each of the individuals has set a lower limit, but not an upper limit, for the grade he/she needs to get). And this means that for (12a) to come out true (i.e., to guarantee that John's grade is the highest), in all the worlds compatible with his needs, John will have to get an A. On the other hand, the way we understand (7), given that John's lower limit is B+, there are bound to be some worlds compatible with his needs where he gets grades of this "height". But if K contains all grades received by everyone in all of their "need" worlds, in those worlds where John gets a grade of the lower limit, his grade will not come out highest, and (12a) will be false.

What about trying to relativize the comparison set to worlds, as implied in (12b)? This would not help matters much. If we simply impose that for each of John's "need" worlds he gets a grade higher than Mary or Bill, we obtain the regular *de dicto* reading, and as we have already said, in our scenario John doesn't care whether his grade is higher or lower than the grades of the others. Another option is to have the function f collect, in each of the worlds compatible with John's needs, grades that are of exactly the lowest degree possible for each individual in the set {John, Bill, Mary}. This will not reflect the right meaning either, because John – like Mary and Bill – has only expressed the lower limit of his needs, so among the worlds compatible with his needs there are bound to be worlds where he gets grades higher than B+.

To sum up, it seems extremely hard, if not impossible, to analyze 'upstairs *de dicto*' readings within standard assumptions regarding the interpretation of noun phrases in intensional contexts. On the other hand, and as Heim shows, the movement theory provides an immediate solution. If we scope the superlative operator above the matrix verb and below the matrix subject (and delete *the*), we obtain precisely the truth conditions we are after:

- (13) a. John **K**-est $\lambda d[\text{needs}_0 \lambda_1[\text{PRO to get}_1 \text{the d-high-grade}_1]]$
 b. $\exists d$ such that $\{z \in K : \text{for all worlds } w \text{ compatible with } z\text{'s actual needs, } z \text{ gets in } w \text{ some d-high grade in } w\} = \{\text{John}\}$.
 c. $K = \{z : z \text{ is a (relevant) person} \ \& \ \exists d[z \text{ needs to get a d-high grade}]\}$

(13) reflects precisely the state of affairs in (11). So it looks like 'upstairs *de dicto*' cases provide a very strong argument in favor of the view that the comparative reading (a) should indeed be viewed as a reading distinct from the absolute; and (b) is obtained by movement of the superlative operator above the matrix VP. However, we will now question these conclusions. In the next section we propose an alternative *in situ* analysis of the 'upstairs *de dicto*' reading of (7), which does not run into the problems that the two attempts in (12) run into. In section 4 we present an argument against the movement analysis.

3. An Alternative *In Situ* Solution

We propose that the superlative operator never moves out of its host DP. Rather, it may be interpreted as a property. For example, in (7), *the highest grade* has the option of being interpreted as an individual (giving rise to the *de re* and *de dicto* readings of (7)), or as a property (giving rise to an ‘upstairs *de dicto*’ reading), in which case the following LF is the relevant one (for simplicity, we omit the domain restriction of *-est*):

$$(14) \quad \text{John needs}_0 \lambda_2[[\text{the-}J \dots \lambda_3[\text{est}][\text{high grade}_3]] \lambda t[\text{PRO to get}_2 t]]$$

$[\text{the-}J \dots \lambda_3[\text{est}][\text{high grade}_3]]$ is of type $\langle s, \langle e, t \rangle \rangle$. It moves (locally) because *get* takes an individual as its internal argument (alternatively, *get* undergoes type-shifting).

To obtain the property interpretation of *the highest grade*, we assume that the determiner *the* is cross-categorical (see, for example, Jacobson (1994)), and that it may apply to a set of properties to yield a unique property. Like any determiner, it comes with a variable which restricts its domain (as is commonly assumed – e.g., von Stechow (1994)). In addition, we assume the following type-shifting operation:

$$(15) \quad \lambda P'_{\langle s, \langle e, t \rangle \rangle} \lambda P_{\langle s, \langle e, t \rangle \rangle} [\forall w \in \mathbf{W}^* [P(w) = P'(w)]]$$

The function in (15) takes two properties and yields True just in case, in all worlds belonging to the contextually supplied set \mathbf{W}^* , these two properties have the same extension. We view this operation as an extension of Partee’s (1987) IDENT.¹

The property interpretation of *the highest grade* proceeds like this. The world index of *high grade* is abstracted over below *the* and above *-est* (see (14)). Next, IDENT is applied to the resulting expression (which denotes the property of being highest grade):

$$(16) \quad \lambda P' \lambda P [\forall w \in \mathbf{W}^* [P(w) = P'(w)]] (\lambda w_3 [\text{est}(\text{high-grade}_{w_3})]) \\ \Rightarrow \lambda P [\forall w \in \mathbf{W}^* [P(w) = \text{est}(\text{high-grade}_w)]]$$

The result in (16) is a set of properties to which *the* is applied, yielding a unique (contextually relevant) property:

$$(17) \quad \text{the}(\lambda P [P \in J \ \& \ \forall w \in \mathbf{W}^* [P(w) = \text{est}(\text{high-grade}_w)]])$$

“the unique property P which is a member of J and which in each world in \mathbf{W}^* has the same extension as the property of being highest grade”

J is the set of properties {‘be a B+ grade’, ‘be a B- grade’, ‘be a C grade’, ...} made salient by the context (Scenario I, (10)-(11) above). \mathbf{W}^* is a salient set of worlds where the properties ‘be a B+ grade’ and ‘be highest grade’ have the same extension. Given the results of our survey, we can characterize \mathbf{W}^* as the set in (18) (w_0 is the actual world):

- (18) $\{w: \text{for all } x \in \{\text{John, Bill, Mary}\} \text{ } x \text{ gets in } w \text{ one grade only, of the lowest possible height according to } x\text{'s needs in } w_0, \text{ and the grades that John, Bill, and Mary get in } w \text{ are the only grades in } w\}$

In other words, based on Scenario I, W^* contains worlds where the needs of everyone are minimally satisfied (ignoring any other needs they may or may not have). Given these values for J and W^* , *the highest grade* (in its interpretation in (17)) is precisely the property of being a B+ grade. After combining *the highest grade* with $\lambda t[\text{PRO to get } t]$ via the appropriate operations, we get (19) (blurring the distinction between object- and metalanguage somewhat):

- (19) John needs in w_0
 $\lambda w' [\exists x [\text{the}(\lambda P [P \in J \ \& \ \forall w \in W^* [P(w) = \lambda z [z \text{ is the highest grade in } w]])](w')(x) \ \& \ \text{John gets } x \text{ in } w']]$

So given the context resulting from our survey (i.e., given that the property in (17) denotes in this context the 'be B+' property), *John needs to get the highest grade* and *John needs to get a B+* amount to the same thing.

Clearly, the distribution of such definite descriptions of properties is not free (see Sharvit and Stateva (in preparation) for discussion of the constraints that govern the distribution of definite descriptions of this kind). But assuming that this *in situ* proposal is adequate, we find ourselves again in a position where we have to ask whether we can distinguish between the predictions of the movement theory and the *in situ* theory. We believe that there is a reason to prefer the *in situ* approach, which has to do with the fact that the movement analysis encounters a problem when a negative superlative operator is embedded in an extensional context. Section 4 discusses this problem.

4. Problems with the Movement Theory

In this section, we consider data that can help tease the two theories apart. We show that the movement theory makes wrong predictions for a set of data that involves negative superlatives, while the *in situ* theory handles these data better.

4.1. Semantics for 'Least'

Since the argument against the movement theory is based on data involving negative superlative descriptions, we will briefly discuss the semantics that we assume here to interpret such expressions.

For current purposes, we adopt the semantics for *least* as in (20) (based on Stateva (to appear)). This semantics is modeled after the semantics for *-est*:

- (20) $[[\text{least}]](K)(R)(x)$ is defined only if $x \in K$ and $\forall y [y \in K \rightarrow \exists d [R(d)(y) = 1]]$;

whenever defined, $[[least]](K)(R)(x) = 1$ iff
 $\exists d$ such that $\{z \in K : R(d)(z) = 1\} = K - \{x\}$.

Least, like *-est*, gives rise to an absolute and a comparative reading. For example, (21) can be appropriate if Mary got an F (the lowest grade in the American grading system), or if she got a grade lower than every other relevant person:

(21) Mary got the least high grade.

Within the movement theory, and given our semantics in (20), (21) has two LF's. The LF in (22a) represents the absolute reading:

- (22) a. Mary got the [K-least] [high grade]
 b. Mary got the unique x such that $\exists d[\{z \in K : z \text{ is a } d\text{-high grade}\} = K - \{x\}]$
 c. $K = \{\text{grade A, grade A-}, \text{grade B+}, \text{grade B}, \dots, \text{grade C}, \dots, \text{grade F}\}$

The LF in (23a) represents the comparative reading, appropriate, for example, in Scenario II in (24):

- (23) a. Mary [K-least] λd [got the d -high grade]
 b. $\exists d$ such that $\{z \in K : z \text{ got a } d\text{-high grade}\} = K - \{\text{Mary}\}$
 c. $K = \{\text{John, Mary, Bill}\}$

- (24) Scenario II:
 Mary took one exam and got a D+;
 Bill took one exam and got a C;
 John took one exam and got a B.

In (23), *least* and its restriction K are scoped above *get*, to a position where the external argument of the superlative operator is different from the one in the LF representing the absolute reading.

4.2. A Problem for the Movement Analysis in Extensional Contexts

For some extensional contexts the movement analysis makes wrong predictions (Heim (p.c.)). Consider a scenario in which one individual is paired with more than one grade as in (25). Notice that we come across a discrepancy between our intuitions and the truth conditions derived from the LF in (23a) when we evaluate (21) with respect to Scenario III.

- (25) Scenario III:
 Mary took one exam and got a D+.
 Bill took two exams and got a C and a D-.
 John took one exam and got a B.

Since in Scenario III Bill is the person who got D-, which happens to be the lowest grade out of all grades received by the relevant people, we judge (21) as false. Our theory should predict that. However, according to (23b), for the sentence to be true it is enough to find one degree d , such that Mary didn't get a grade that is d -high but everyone else did. Contrary to what we want, this is indeed the case, because all the degrees that are higher than D+ and lower than or equal to C verify (21) according to (23b).

Unlike the movement theory, the *in situ* analysis (see (22a-b) which represent the two "readings" under the *in situ* theory, with different K 's) does not run into this problem. (22b), where K consists of grades rather than grade receivers, correctly predicts (21) to be false in Scenario III. It seems then that here, the *in situ* theory has a clear advantage. However, before deciding to reject the movement analysis entirely, let us try to save it by revising the semantics for the superlative operators while keeping the core assumptions of that theory intact.

4.3. Revising the Lexical Entry for the Superlative Operator

We saw in the previous discussion that the movement theory faces problems because it fails to consider the degrees of height of every grade that Bill got in Scenario III, where Mary's grade is "sandwiched" between the two grades that Bill got. Therefore, our first attempt to revise the semantics of the superlative operator involves universally quantifying over individual-degree pairs. The proposed semantics are given in (26) and (27) (for simplicity, the restriction K is left out):

(26) $[[\text{-est}]](R)(x)=1$ iff $\exists d[R(d)(x)=1 \ \& \ \forall y \neq x \forall d' [R(d')(y)=1 \rightarrow d > d']]$

(27) $[[\text{least}]](R)(x)=1$ iff $\exists d[R(d)(x)=1 \ \& \ \forall y \neq x \forall d' [R(d')(y)=1 \rightarrow d < d']]$

In keeping with the movement analysis, *least* scopes above *got*, yielding the LF and interpretation in (28) for (21):

- (28) a. Mary [*least*] λd [got the d -high grade]
 b. $\exists d_1$ such that Mary got a d_1 -high grade & $\forall y \neq \text{Mary} \forall d_2$ [if y got a d_2 -high grade, then $d_1 < d_2$]

What we achieved by modifying the semantics for the superlative operators is obviously not satisfactory. (21) is correctly predicted to be false under Scenario III, but we run into a more fundamental problem. For every situation which can be described with a sentence containing a superlative expression with *least* (and where our domain consists of at least two individuals) our semantics assigns that sentence truth conditions that could not be satisfied. For example, our intuitions require that under Scenario III, (29) be true, but (28b) predicts it to be false:

- (29) Bill got the least high grade.

Similarly, (21) under Scenario II should come out true. However, the prediction goes in the opposite direction in this case, too. According to (28b), (21) comes out false in Scenario II since for any degree d corresponding to the grade received by Mary (degree D+, degree D, etc.), there is at least one degree corresponding to the grades received by John or Bill that is not higher than d . These are monotonicity effects. In fact, if our semantics allowed us to consider only the degrees corresponding to the respective maximal height of each relevant grade, all the “offending” degrees would conveniently be excluded from the evaluation of (28b) under Scenario II. Let us then make a second attempt to salvage the movement theory by changing our assumptions regarding the interpretation of gradable adjectives.

Let us assume that adjectives such as *tall* or *high* come with an understood “exactly” (cf. von Stechow (1984)). For example, *John is five feet tall* implies that John is exactly five feet tall. Accordingly, *–est* and *least* also come with an understood “exactly”. As expected, given this assumption, we face no problem in accounting for the data discussed above. To see this, consider the predictions that the movement theory makes now for (21) with respect to Scenarios II and III.

- (30) $\exists d_1$ such that Mary got an exactly d_1 -high grade & $\forall y \neq \text{Mary} \forall d_2$ [if y got an exactly d_2 -high grade, then $d_1 < d_2$]

Recall, that we need (21) to come out true under Scenario II, and false under Scenario III. Under the former, all the exact heights of the grades that the others got – namely, C and B – satisfy the condition of being higher on the scale than the exact height of the grade that Mary got (namely, D+). Therefore the sentence is predicted to be true. Under the latter scenario, however, not all the exact heights of the grades that the others got satisfy the condition of being higher than the exact height of Mary’s grade. In particular, the D- that Bill got does not, and the truth conditions are not met. Once more, a welcome result.

However, we cannot endorse these amendments to the movement theory. By neutralizing the monotonicity principle in (3) we lose the account for ‘upstairs *de dicto*’ readings, because the account depends on this principle. To see this, consider again the ‘upstairs *de dicto*’ reading of (7), repeated in (31):

- (31) John needs to get the highest grade.

By moving the superlative operator above the intensional verb we get the following truth conditions:

- (32) There is a degree d_1 such that for all worlds w compatible with John’s actual needs, John gets in w an exactly d_1 -high grade in w ; and for all others y , for all d_2 , if for all worlds w compatible with y ’s actual needs, y gets in w an exactly d_2 -high grade in w , then $d_1 > d_2$.

These truth conditions do not capture the ‘upstairs *de dicto*’ reading of (31). For

one thing, they require John to climb a mountain of the same height in all his “need” worlds. In addition, they incorrectly predict (31) to be true if John says he needs exactly B, and Mary says she needs at least B+. We conclude that neutralizing the effects of monotonicity is undesirable,² and we keep the original semantics of the superlative operator.

To summarize this section, we showed that our attempts to save the movement analysis by revising the semantics of the superlative operators failed. We were able to figure out a way to avoid the problems in extensional contexts but the price we had to pay was losing the account of ‘upstairs *de dicto*’ readings. On this we conclude that the *in situ* theory has some advantage compared to the movement theory.³

5. More on Negative Superlatives

Recall that in section 2, we showed that the movement theory has quite powerful tools to handle ‘upstairs *de dicto*’ readings. Since our goal is to argue in favor of the *in situ* theory, in this section we will show that this theory can account for a wide range of such readings equally successfully.

Negative superlatives have been shown to give rise to two ‘upstairs *de dicto*’ readings (Stateva (to appear)).⁴ Consider (33):

(33) Mary needs to get the least high grade.

The first reading becomes available in Scenario I (see Section 2, (11)), where the results of the survey can be reported by (33). In all the worlds compatible with her needs, Mary gets at least C; in all the worlds compatible with his needs, Bill gets at least B-; and in all the worlds compatible with his needs, John gets at least B+. We call this reading the ‘at least upstairs *de dicto*’ reading. To derive it under the movement hypothesis, Stateva uses Heim’s (1999) strategy of raising the degree operator to a position where it takes scope over the intensional verb. The LF and truth conditions are given in (34a) and (34b):

- (34) a. Mary [**K**-least] $\lambda d[\text{needs}_0 \lambda_1[\text{PRO to get}_{t_1} \text{ the } d\text{-high-grade}_{t_1}]]$
 b. $\exists d$ such that $\{z \in K : z \text{ needs in } w_0 \lambda w[z \text{ gets in } w \text{ a } d\text{-high-grade}_w]\} = K - \{\text{Mary}\}$

The second ‘upstairs *de dicto*’ reading of (33) comes up in a different scenario. Suppose Mary says that she needs to get a grade that is at most C; Bill says that he needs to get at most B-, and John says that he needs to get at most B+. (33) is a good report of this situation, and we call this reading the ‘at most upstairs *de dicto*’ reading. The movement hypothesis combined with one additional assumption about the morphological make-up of *least* can account for that reading, too.

Heim (1998) assumes that the comparative operator *less* can be decomposed in the syntactic component into the operator *-er* and a negation operator.⁵ One consequence of this proposal is that the degree operator and negation become movable independently of each other and can be interpreted in different positions. Stateva expands that proposal to *least*, assuming that it can be decomposed into the operator *-est* and negation. There are at least two logical possibilities to derive ‘upstairs *de dicto*’ readings under the assumption that *least* decomposes into two operators: (i) *-est* is raised above the intensional verb, and so is negation; (ii) *-est* is raised above the intensional verb but negation is interpreted in its base-generated position. The first option derives the ‘at least’ reading.⁶ The second option comes in handy for representing the ‘at most’ reading as in (35) (the negation operator turns any member of the *high*-class into its antonym).⁷

- (35) a. Mary [**K**-est] $\lambda d[\text{needs}_0 \lambda_1[\text{PRO to get}_1 \text{ a not-d-high-grade}_1]]$
 b. $\exists d$ such that $\{z \in K : z \text{ needs in } w_0 \lambda w[z \text{ gets in } w \text{ a not-d-high-grade}_w]\}$
 = {Mary}

This is the desired interpretation. However, since in the previous section we rejected the movement analysis on independent grounds, we have to show that the *in situ* theory can account for these two ‘upstairs *de dicto*’ readings. We propose to use one single LF to derive both readings and manipulate the choice of the contextually supplied W^* . (36), which is reminiscent of (19), represents both the ‘at least’ and the ‘at most’ readings obtained from that LF:

- (36) Mary needs in $w_0 \lambda w'[\exists y[\text{the}(\lambda P[P \in J \ \& \ \forall w \in W^*[P(w) = \lambda x[x \text{ is least high grade in } w]])](w')(y) \ \& \ \text{Mary gets } y \text{ in } w']]$

According to (36), (33) is true if and only if in all the worlds compatible with what Mary needs there is a grade y , such that y has the property denoted by *the least high grade* (in its property-meaning), and Mary gets y . To derive the ‘at least’ reading under which each of the three grade-receivers needs to get a grade at least as high as the one they named, W^* has to be the following:

- (37) $\{w : \text{for all } x \text{ in } \{\text{John, Bill, Mary}\}, x \text{ gets the lowest grade possible according to } x\text{'s needs in } w_0, \text{ and the grades that John, Bill and Mary get in } w \text{ are the only grades in } w\}$

Since we only collect worlds that contain exactly three grades – C, B- and B+ – the property denoted by *the least high grade* in (36) is a sub-property of the ‘be a C grade’ property, namely, ‘ $\lambda w \lambda x[x \text{ is a C grade in } w \ \& \ \text{if } w \text{ is in } W^*, x \text{ is an exactly C grade in } w]$ ’.

To construct the W^* that would lead to an interpretation reflecting the ‘at most’ reading, we choose worlds which contain three grades only: one grade received by Mary which is at most C, one grade received by Bill which is above C and at most B-, and one grade received by John which is above B- and below B+.

To do so, we define for every individual x and a “need” world w a set of degrees of grade height such that according to x 's declaration in w_0 they are degrees of grades that x gets in w .

- (38) For all x and w , $NEED(x)(w) = \{d: \text{there is a } w' \text{ such that } w' \text{ is a world compatible with what } x \text{ needs in } w \text{ and } x \text{ gets a } d\text{-high grade in } w'\}$

Second, we define a ‘smaller than’ relation that holds between relevant sets of degrees as in (39):

- (39) For all $\langle x, y \rangle$ and for all w , $NEED(x)(w) < NEED(y)(w)$ iff the highest member of $NEED(x)(w)$ is lower than the highest member of $NEED(y)(w)$.

Finally, we have all ingredients necessary to characterize W^* :

- (40) $W^* = \{w: \text{each } x \text{ in } \{\text{John, Bill, Mary}\} \text{ gets one grade only in } w, \text{ whose exact height is in } NEED(x)(w_0) \text{ but not in } NEED(y)(w_0), \text{ where } y \text{ is any member of } \{\text{John, Bill, Mary}\} \text{ such that } NEED(y)(w_0) < NEED(x)(w_0); \text{ and the grades that John, Bill and Mary get in } w \text{ are the only grades in } w\}$

The property denoted by *the least high grade* then, in the context under discussion, is ‘ $\lambda w \lambda x [x \text{ is an at most C grade in } w]$ ’.

To sum up, the movement theory which successfully derives all the ‘upstairs *de dicto*’ readings has to put up with some controversial assumptions about the possibility to extract an operator out of a definite noun phrase. In addition, we saw that negative superlatives can create problems in certain extensional contexts if movement is adopted. The *in situ* theory, on the other hand, does not have these disadvantages. However, since it relies so heavily on contextually supplied information, the issue of how the choice of the values for the free variables is constrained needs to be addressed. In addition to this issue, the *in situ* proposal opens up many more questions. We discuss some of them below.

6. Some Open Questions

Due to space limitations, we are only able to address a few of the issues raised by the *in situ* proposal (see Sharvit and Stateva (in preparation) for more discussion).

6.1. “Sandwich” Scenarios and ‘Upstairs De Dicto’ Readings

There are ‘upstairs *de dicto*’ cases for which the two theories make different predictions. Potentially, these cases could point to the superior theory.

Consider a slightly more complicated scenario than the ones discussed so far. Suppose we ask people about the minimal requirements they must meet to keep their scholarships. John says: “I need to get two grades, a D+ in Math, and a

B in English”. Suppose further that Bill says that he only needs to get a C in Math, and Mary says that she needs to get an A- in English. What is then the status of (41a) and (41b)?

- (41) a. John needs to get the least high grade.
b. Bill needs to get the least high grade.

Under the ‘upstairs *de dicto*’ reading of (41a-b), the movement analysis predicts (41a) to be false and (41b) to be true as we can see from their respective truth conditions in (42a) and (42b):

- (42) a. $\exists d$ such that $\{z \in \{\text{John, Bill, Mary}\} : z \text{ needs to get a } d\text{-high grade}\} = \{\text{Bill, Mary}\}$
b. $\exists d$ such that $\{z \in \{\text{John, Bill, Mary}\} : z \text{ needs to get a } d\text{-high grade}\} = \{\text{John, Mary}\}$

The *in situ* analysis goes in the opposite direction and predicts (41a) to be true, and (41b) to be false, if W^* is the set given in (43). In (44), *the least high grade* denotes, presumably, the relevant sub-property of ‘be a D+ grade’ – ‘ $\lambda w \lambda x [x \text{ is a D+ grade in } w \ \& \ \text{if } w \text{ is in } W^*, x \text{ is an exactly D+ grade in } w]$ ’.

- (43) $\{w : \text{John gets two grades in } w - \text{ a D+ and a B, Bill gets a C and Mary gets an A-; and these grades are the only grades in } w\}$
(44) a. John needs in w_0 $\lambda w' [\exists y [\text{the}(\lambda P [P \in J \ \& \ \forall w \in W^* [P(w) = \lambda x [x \text{ is least high grade in } w]])](w')(y) \ \& \ \text{John gets } y \text{ in } w']]$
b. Bill needs in w_0 $\lambda w' [\exists y [\text{the}(\lambda P [P \in J \ \& \ \forall w \in W^* [P(w) = \lambda x [x \text{ is least high grade in } w]])](w')(y) \ \& \ \text{Bill gets } y \text{ in } w']]$

Unfortunately, intuitions about the status of (41a) and (41b) in the given situation vary. Some speakers judge them both as neither true nor false. Others judge (41a) true and (41b) false, as predicted by the *in situ* theory. Given the diversity of judgments, we cannot take this to be an argument in favor of the *in situ* analysis.

6.2. *Are Superlatives Indefinite Descriptions?*

According to the *in situ* analysis, unlike the movement analysis, superlative expressions are always definite descriptions. If there is evidence favoring the view that the superlative construction can be indefinite, it might point in favor of the movement theory. The mechanism of deriving the comparative reading under the latter theory involves covert movement of the operator outside of its base position in the superlative description. Given standard assumptions about conditions on extraction, this syntactic operation is not possible if the site of extraction is a definite description (see Chomsky (1973) and Szabolcsi (1986)):

- (45) a. Who did you take a picture of?

- b. *Who did you take the picture of?
- c. Who did you take the best picture of?

As the contrast between (45a) and (45b) shows, extraction is impossible out of a definite noun phrase. The grammaticality of (45c) suggests that the superlative expression is an indefinite, thus lending support to the one of the assumptions underlying Heim's and Szabolcsi's theory of the movement theory (i.e., *the*-deletion and *-est*-extraction). Clearly, (45c) is a puzzle for the *in situ* theory.

We do not have a full answer to this problem, but a possible solution is this. To obtain the 'upstairs *de dicto*' readings, we crucially assume that the superlative expression can denote a property. If we allow such interpretation in extensional cases too (such as *John climbed the highest mountain*), we can maintain the idea that (at least sometimes) superlative expressions are in some sense indefinite, and therefore are not necessarily islands for extraction.⁸

6.3. Superlatives and Focus

Both Szabolcsi and Heim argue that only the comparative reading of superlatives arises in the presence of focus:

- (46) a. JOHN gave the hardest exam to Mary.
- b. John gave the hardest exam to MARY.

If Szabolcsi and Heim are right, then the movement theory draws a nice correlation between the presence of focus and the special mechanism (degree operator movement) for deriving the comparative reading that focus makes prominent. However, we do not believe that this is an argument against the *in situ* theory. First, contrary to Szabolcsi, Heim argues that comparative readings are not necessarily prompted by focus, so the correlation between focus and movement goes only one way. Secondly, as shown by Heim (1999), the *in situ* theory has no problem of accounting for the focus effects. Thirdly, we think that even the absolute reading (of say, *John climbed the highest mountain*, where *the highest mountain* refers to Mount Everest) can arise when the subject is focused (for example, if there is a dispute regarding who climbed Mount Everest).

How are focus effects obtained in the *in situ* theory? For example, how do we guarantee that in (46a) the comparison set contains exam givers and not exam takers? Heim proposes a LF in the spirit of Rooth (1992) and von Stechow (1994):

- (47) the [\cup K-est] [hard exam] [λx [JOHN gave x to Mary]~K]

The superlative phrase is moved, creating a λ -abstract to which the focus operator is adjoined. The focus operator introduces an anaphor which denotes a subset of the focus semantic value of the sister of $\sim K$ (the λ -abstract) which is given in (48):

- (48) {Y: $\exists y$ such that Y = λx [y gave x to Mary]}

The anaphor *K* has an antecedent in the domain restriction of the superlative operator. That domain was previously assumed to be a set of individuals, while *K* is a set of sets of individuals. To resolve this mismatch, the argument of *-est* in this case is assumed to be the union of *K*. (47) then, requires that all sets of relevant hard exams are also sets of things given by somebody to Mary.

As (49) shows, focus effects in ‘upstairs *de dicto*’ cases can be obtained in a similar way, with the relevant λ -abstract being over properties:

- (49) a. JOHN’s sister needs to get the highest grade.
 b. $\text{the-}\cup\text{J}(\lambda\text{P}[\forall w \in W^*[\text{P}(w)=\lambda x[x \text{ is highest grade}_w]]]) [\lambda\text{P}[\text{JOHN’s sister needs in } w_0 \lambda w'[\exists y[y \in \text{P}(w') \ \& \ \text{PRO gets } y \text{ in } w']]]\sim\text{J}]$
 c. $\text{J} \subseteq \{\text{Q: } \exists x \text{ such that } \text{Q} = \lambda\text{P}[x\text{’s sister needs in } w_0 \lambda w'[\exists y[y \in \text{P}(w') \ \& \ x\text{’s sister gets } y \text{ in } w']]]\}$

The constraint on *J* as specified in (49c) guarantees that John’s sister is compared to other people’s sisters (and not simply to other people in general). In order to derive the focus effects in (49) some syntactic operation must apply displacing *the highest grade* into a position where it has scope over the matrix subject. QR is indeed a likely candidate, because movement of *the highest grade* is not semantically vacuous, since it effects focus interpretation. However, QR is subject to locality restrictions that are not always clear (Reinhart (1997)). For instance, it is not at all clear whether (50) can have a reading where the embedded quantified phrase scopes above the matrix subject:

- (50) Some student needs to meet every professor.

If QR cannot apply in (50), then it should not be available for (49b) either. This, too, is left as an open problem.

7. Conclusion

‘Upstairs *de dicto*’ readings pose a difficult challenge for an *in situ* analysis, and seem to support a movement analysis. However, given that negative superlatives in extensional contexts are better analyzed without movement, the *in situ* analysis has some advantage. Thus Kennedy’s original proposal that ‘long’ distance movement of degree operators is banned by the grammar finds additional support.

Endnotes

* For valuable comments and discussion, we thank Gidi Avrahami, Sigrid Beck, Deborah Chen, Irene Heim, Polly Jacobson, Gerhard Jaeger, Chris Kennedy, Fred

Landman, Barbara Partee, Mats Rooth, Susan Rothstein, Roger Schwarzschild, Aldo Sevi, Jason Stanley, Arthur Stepanov, Zoltan Szabo, and Anna Szabolcsi.

¹ The original IDENT applies to an individual to yield the property of being that individual.

² Heim (1998) also reaches the conclusion that intensional contexts provide an argument in favor of a theory that assumes the Principle of Monotonicity. The argument is built on data involving a comparative construction. Consider (i):

(i) You will have to be more patient than your brother (will have to be)(if you want to make it in philosophy).

Heim discusses Rullmann's (1995) proposal to analyze the comparative construction without assuming that monotonicity holds. A Rullmann-type representation of (i) will be that in (ii):

(ii) [er than wh₁ your brother will have to be t₁ patient] 2 [you will have to be t₂-patient]
 $\max\{d: \text{necessarily, you be } d\text{-patient}\} > \max\{d: \text{necessarily, your brother be } d\text{-patient}\}$

From a perspective slightly different from ours, Heim argues that monotonicity describes scalar predicates better. She argues that if people can have different unique degrees of patience in different worlds, then there is no degree to which the subject is patient (in each world) in (ii), hence the maxima of each of the compared sets of degrees is undefined. Given that Rullmann's proposal for the semantics of *-er* and our 'revised' entries for *-est* and *least* are so similar in spirit, it is not surprising that both suggestions face problems in similar contexts, namely, when intensionality is involved.

³ Similar problems arise if alternative frameworks (cf. Schwarzschild and Wilkinson (1999) where scalar predicates are viewed as relations between individuals and intervals) are combined with a 'movement' hypothesis.

⁴ This type of ambiguity in sentences with a superlative construction has a parallel phenomenon in comparatives. For details, see Rullmann (1995) and Heim (1998).

⁵ The mechanism that Heim (1998) proposes is not of decomposing, strictly speaking, since she only assumes the existence of *-er* in the lexicon.

⁶ For reasons of simplicity, here we represented the 'at least' reading by combining negation with *-est* and 'producing' the compound *least*. A logically equivalent result is achieved if we are to strictly observe the assumption that the lexicon only contains an entry for *-est* but not for *least*. Then the reading is derived by interpreting negation as a sister of VP.

⁷ The idea that both the degree operator and negation can freely move at LF leads to overgeneration. There is an alternative strategy of deriving both 'upstairs *de dicto*' readings, which is an expansion of Rullmann's proposal to use different bracketing combinations in comparative constructions (cf. Stateva (to appear)).

⁸ Szabolcsi discusses three additional contexts (relational *have*, existential constructions with *there*, and *ago* constructions) that disallow definite expressions but they admit superlative expressions.

References

- Chomsky, N. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. S. Anderson and P. Kiparsky, 232-286. New York: Holt, Rinehart and Winston.
- Cresswell, M. 1976. The semantics of degree. In *Montague grammar*, ed. B. Partee, 261-292. New York: Academic Press.
- von Stechow, Kai. 1994. *Restrictions on quantifier domains*, U. of Massachusetts, Amherst.
- Heim, I. 1985. *Notes on comparatives and related matters*. Ms., U. of Texas, Austin.
- Heim, I. 1998. *Negative decomposition and comparatives*, U. of Connecticut colloquium.
- Heim, I. 1999. *Notes on superlatives*. Ms., MIT, Cambridge.
- Jacobson, P. 1994. Binding connectivity in copular sentences. In *Proceedings of SALT IV*. CLC Publications, Cornell University.
- Kennedy, C. 1997. *Projecting the adjective: The syntax and semantics of gradability and comparison*. PhD Dissertation, U. of California, Santa Cruz.
- Partee, B. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and theory of generalized quantifiers*, ed. J. Groenendijk et al., 115-144. Dordrecht: Reidel.
- Reinhart, T. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20: 335-397.
- Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics*: 75-116.
- Ross, J. R. 1964. *A partial grammar of English superlatives*. MA thesis, U. of Pennsylvania, Philadelphia.
- Rullmann, H. 1995. *Maximality in the semantics of wh-constructions*. PhD dissertation, U. of Massachusetts, Amherst.
- Schwarzschild, R., and K. Wilkinson. 1999. *Interval semantics for scale predication*. Ms., Rutgers University, New Brunswick.
- Seuren, P. A.M. 1973. The comparative. In *Generative grammar in Europe*, ed. F. Kiefer and N. Ruwet. Dordrecht: Reidel.
- Sharvit, Y., and P. Stateva. in preparation. *Context, superlatives, and definiteness*. Ms., U. of Connecticut, Storrs.
- Stateva, P. to appear. In defense of the movement theory of superlatives. In *Proceedings of the Eastern States Conference on Linguistics 1999*.
- von Stechow, A. 1984. Comparing semantic theories of comparison. *Journal of semantics* 3: 1-77.
- Szabolcsi, A. 1986. Comparative superlatives. In *MIT Working Papers in Linguistics* 8, ed. N. Fukui et al., 245-265. MIT Press, Cambridge.