

A Note on Intensional Superlatives

Rajesh Bhatt and Yael Sharvit

University of Massachusetts at Amherst and University of Connecticut

1. What this Paper is About

This paper is about relative clauses whose “head” contains a superlative morpheme and whose main verb is intensional. The sentence in (1) has such a relative clause. We refer to these relative clauses as “intensional superlatives”.

(1) The longest book John said Tolstoy had written was Anna Karenina.

The observed ambiguity of an intensional superlative, created by the main verb, seems to disappear under certain conditions. Our main concern here is the characterization of these conditions.

Bhatt (2002) observes that sentences with intensional superlatives are ambiguous between what he calls a ‘high’ and a ‘low’ reading, as shown in (2), where the two readings of (1) are paraphrased. In the ‘high’ reading *longest* means “longest according to the speaker”; in the ‘low’ reading *longest* means “longest according to John”.

(2) ‘High’ reading of (1)

Anna Karenina is the unique x such that x is longest among $\{y:y \text{ is a book and according to what John said, Tolstoy wrote } y\}$.

‘Low’ reading of (1)

Anna Karenina is the unique x such that according to what John said, x is longest among $\{y:y \text{ is a book and Tolstoy wrote } y\}$.

To get a better sense of the difference between these two readings, let us consider a scenario where the ‘high’ reading is intuitively true and the ‘low’ reading false (Scenario A in (3)); and a scenario where the ‘low’ reading is intuitively true and the ‘high’ reading false (Scenario B in (3)).

(3) Scenario A

John: “Tolstoy wrote Huckleberry Finn, Anna Karenina and Tom Sawyer; Tom Sawyer is the longest of these.”

Anna Karenina is actually the longest among those books.

Scenario B

John: “Anna Karenina is the longest book Tolstoy wrote. He also wrote War and Peace and some other shorter books.”

War and Peace is actually longer than Anna Karenina.

Quite generally, a surface structure such as the one in (4) yields a ‘high’ reading where *R-est* and the NP that follows it are interpreted “above” the intensional verb

at LF, and a ‘low’ reading where *R-est* and the NP that follows it are interpreted “below” the intensional verb at LF (*R* stands for an expression such as *long*, which denotes a relation between individuals and degrees).¹

- (4) [*the R-est* NP *that* Subject1 $V^{\text{intensional}}$ *that* Subject2 V _] *is* DP
- a. ‘High’ LF
 [*the* λx [*R-est*_x NP λy [Subject1 $V^{\text{intensional}}$ *that* Subject2 V y]]] *is* DP
Interpretation: DP is the unique x such that x is R-est among {y:y is NP and [Subject1 $V^{\text{intensional}}$ *that* Subject2 V y]}.
- b. ‘Low’ LF (cf. Hulsey and Sauerland, to appear)
 [*the* λx [Subject1 $V^{\text{intensional}}$ *that* *R-est*_x NP λy [Subject2 V y]]] *is* DP
Interpretation: DP is the unique x such that Subject1 $V^{\text{intensional}}$ *that* x is R-est among {y:y is NP and Subject2 V y}.

Bhatt further observes that NPI-placement may disambiguate the sentence. If we put an NPI (negative polarity item) above the main verb we lose the ‘low’ reading (see (5), whose ‘low’ reading is unacceptable), and if we place it below the main verb, we lose the ‘high’ reading (see (6), whose ‘high’ reading is unacceptable).

- (5) The longest book John ever said Tolstoy had written was Anna Karenina.
 (6) The longest book John said Tolstoy had ever written was Anna Karenina.

Interestingly, there are other cases where one of the readings is unacceptable. As Heycock (to appear) notes, not all intensional verbs support a ‘low’ reading. She claims that *know* is such a verb. According to her, (7) has only a ‘high’ reading.

- (7) The longest book John knew Tolstoy had written was War and Peace.

We show below that Kadmon and Landman’s (1993) pragmatic theory of NPI licensing suggests an interesting explanation of the NPI facts. If correct, we argue, this explanation restricts the way in which we can understand the Heycock facts. In particular, we show that it is hard to reconcile a Neg-raising account of the Heycock facts with the pragmatic account of the NPI facts.

2. Disambiguation by NPI Placement

Given our assumptions about the source of the ‘low’ reading, explaining why it is absent from (5) is perhaps not so hard. After all, if the superlative is lowered at LF, the NPI is outside the scope of its licenser. What is puzzling is the fact that the ‘high’ reading is not available in (6), especially because the verb *say* does not block NPI licensing by other licensers, such as negation, as shown by the acceptability of (8) (other operators, such as the determiner *every*, do block NPI licensing; see Linebarger 1987, Chierchia 2001, and others).

- (8) John didn’t say that Bill had ever been to Paris.

So the question is, what is so special about the superlative morpheme that it cannot license an NPI across *say*.

Since the works of Fauconnier (1975, 1978) and Ladusaw (1979), many researchers have shared the assumption that NPIs such as *any* and *ever* are indefinites that must appear in the scope of downward entailing (DE) functions. One of the questions that have preoccupied those researchers is why this should be so. Kadmon and Landman (1993) have proposed that the role of an NPI is to widen the domain restriction of the indefinite it comes attached to. It is licensed only when it yields a stronger (i.e., more informative) statement, compared to the corresponding statement where the indefinite is “NPI-less”. It so happens that DE functions are the ones that may create stronger statements. For example, an NPI is licensed by negation (a DE function) as shown in (9b). Indeed, (9b) is a stronger – that is to say, more informative – statement compared to (9a), but only in contexts where it is not already established that John didn’t eat any apple member of ($[[C']]$ – $[[C]]$) (*apples* denotes an existential quantifier with a restricted domain, $[[C]]$ is the domain restriction of *apples* supplied initially by the context, and $[[C']]$ is the “widened” restriction that is there thanks to the presence of *any*).^{2,3}

- (9) a. John didn’t eat apples.
 LF: NOT [$apples_C \lambda y[John\ ate\ y]$]
 No apple $y \in [[C]]$ is such that John ate y .
- b. John didn’t eat any apples.
 LF: NOT [$any-apples_{C \supset C} \lambda y[John\ ate\ y]$]
 No apple $y \in [[C']]$, $[[C']] \supset [[C]]$, is such that John ate y .

If negation is absent from (9a,b), the structure with *any* does not yield a stronger statement, compared to the one without *any*, regardless of whether or not it is established that John ate any relevant apples.

Some NPI licensors are merely “Strawson DE” (so labeled in von Stechow 1999). S1 Strawson-entails S2 iff S1 plus the semantic presuppositions of S2 entail S2. *Only* is an example of such a function. It licenses NPIs (see (10b)), and it is Strawson DE, as shown by the fact that *Only John ate apples* entails *Only John ate green apples* provided the semantic presupposition of the latter (the one contributed by *only*, namely, that John ate green apples) is true. Indeed, (10b) is more informative than (10a), but only if (the semantic presupposition of (10a) is part of the common ground and) it is not part of the common ground that no individual distinct from John ate any apple member of ($[[C']]$ – $[[C]]$).

- (10) a. Only John ate apples.
 LF: $only\ John\ \lambda x[apples_C \lambda y[x\ ate\ y]]$
 Assertion: For all $x \neq John$, no apple $y \in [[C]]$ is such that x ate y .
 Presupposition contributed by *only*: John ate some apple member(s) of $[[C]]$.
- b. Only John ate any apples.
 LF: $only\ John\ \lambda x[any-apples_{C \supset C} \lambda y[x\ ate\ y]]$

Assertion: For all $x \neq \text{John}$, no apple $y \in \llbracket C' \rrbracket$, $\llbracket C' \rrbracket \supset \llbracket C \rrbracket$, is such that x ate y .

If *only* is absent from (10a,b), the structure with *any* does not give rise to a stronger statement, compared to the one without *any*.

The superlative operator is also Strawson DE, as shown by the fact that *War and Peace is the longest book* entails *War and Peace is the longest serious book*, provided the semantic presupposition of the latter (the one contributed by *longest*, namely, that *War and Peace* is a serious book) is true (see von Stechow 1999, Herdan and Sharvit 2005, and others).⁴ Indeed, (11b) (the LF of *The longest book Tolstoy ever wrote was War and Peace*) yields a more informative statement than (11a), but only in contexts (where the semantic presuppositions of (11a) are established and) where it is not established that $\llbracket T \rrbracket$ (the restrictor of PAST, the time stretch initially supplied by the context) exhaustively covers Tolstoy's productive years as a book writer.

- (11) a. *the λx longest_x [book λy [PAST_T λt [Tolstoy write_t y]]] was WP*
 Assertion: WP is longest in $\{y: y \text{ is a book and there is a past time } t \text{ in } \llbracket T \rrbracket \text{ such that Tolstoy writes } y \text{ at } t\}$ (=C).
 Presupposition contributed by *longest*: $WP \in C$.
- b. *the λx longest_x [book λy [ever-PAST_{T, \supset T} λt [Tolstoy write_t y]]] was WP*
 Assertion: WP is longest in $\{y: y \text{ is a book and there is a past time } t \text{ in } \llbracket T' \rrbracket, \llbracket T' \rrbracket \supset \llbracket T \rrbracket, \text{ such that Tolstoy writes } y \text{ at } t\}$.

If *longest* is missing from (11a,b), the structure with *ever* does not give rise to a stronger statement, compared to the one without *ever*, because the existence presupposition contributed by *the* guarantees that these “*longest-less*” statements Strawson-entail each other. But the presence of *longest* doesn't guarantee strengthening. Crucially, when it is an established fact that $\llbracket T \rrbracket$ exhaustively covers Tolstoy's productive years as a book writer, widening $\llbracket T \rrbracket$ cannot lead to widening the set of books written by Tolstoy, and (11b), with *ever*, is not more informative than (11a). In such contexts, (11b) is predicted to be unacceptable. For example, in a context where $\llbracket T \rrbracket$ is the time stretch between Tolstoy's eighteenth and thirty-fifth birthdays, and it happens to be part of the common ground that he didn't write any book before he turned eighteen or after he turned thirty-five, it is pointless – in fact, disallowed – to use *ever* because we would not be conveying anything that is not conveyed by the “*ever-less*” (11a).

With this in mind, let us now go back to the disambiguation effects of *ever* discovered by Bhatt ((5) and (6) above). The widening+strengthening approach suggests a way to account for these effects, based on the constraints that should be placed on the context in each case to make strengthening possible. To see how, let us first set aside the (bad) ‘low’ reading of (5): *ever* is not in the scope of *longest*, so there is no hope of strengthening. The readings we are interested in are these: the (good) ‘high’ reading of (5), the (good) ‘low’ reading of (6), and the (bad) ‘high’ reading of (6). Their LFs are given in (12), (13), and (14) respectively.

- (12) *the* $\lambda x[\textit{longest}_x \textit{book} \lambda y[\textit{ever-PAST}_{T \supset T} \lambda t[\textit{John say}_t \textit{Tolstoy wrote } y]]]$
was AK
- (13) *the* $\lambda x[\textit{John said longest}_x \textit{book} \lambda y[\textit{ever-PAST}_{T \supset T} \lambda t[\textit{Tolstoy write}_t y]]]$
was AK
- (14) *the* $\lambda x[\textit{longest}_x \textit{book} \lambda y[\textit{John said ever-PAST}_{T \supset T} \lambda t[\textit{Tolstoy write}_t y]]]$
was AK

Notice the similarity between (11b) and (12)/(13): *ever* and *book* are not separated by an intensional verb. (12), where *book* and *ever* are above *say*, gives rise to strengthening only in contexts where it is not established that $[[T]]$ covers all the times where John says about an actual book that it is by Tolstoy. In (13), where *book* and *ever* are below *say*, John has to have the “right” beliefs about $[[T]]$ – the beliefs that are required for strengthening (we assume that *say* is used here in its “non-quotational” sense, and that John at least acts as if he believes what he says). (13), then, gives rise to strengthening only in contexts where the following is not an established fact: that John believes (or speaks as if he believes, rather) that $[[T]]$ exhaustively covers Tolstoy’s productive years as a book writer. For example, if $[[T]]$ is the time stretch between Tolstoy’s eighteenth birthday and his death and $[[T']]$ is his entire life (and John knows this), the common ground has to entertain the possibility that some world compatible with what John believes is a member of $\{w:\textit{Tolstoy wrote some book(s) in } w \textit{ before he turned eighteen}\}$ (i.e., it has to entertain the possibility that John says/believes something like: “I’m not sure Tolstoy did all his book writing (on or) after he turned eighteen”, or even: “Tolstoy did some of his book writing before he turned eighteen”).

In (14), just like (13), *ever* is below *say*, so John is still responsible for having the “right” beliefs about $[[T]]$ – the beliefs needed for strengthening. But unlike (13), *book* is above *say*. Consequently, (14) gives rise to strengthening only in contexts where the following is not an established fact: that John entertains the possibility that $[[T]]$ exhaustively covers Tolstoy’s productive years as a writer of things that happen to be actual books. The common ground has to entertain the possibility that for at least some actual book(s) *y*, every world compatible with what John believes is a member of $\{w:\textit{Tolstoy wrote } y \textit{ in } w \textit{ before he turned eighteen}\}$ (i.e., that John says/believes something like: “Tolstoy wrote *y* before he turned eighteen”). It seems to us that this constraint is a relatively hard constraint for speakers to place on the context. We are not merely requiring it not to be an established fact that John entertains the possibility that $[[T]]$ covers Tolstoy’s productive years as a writer of “something” (anything, including notes to his wife), but specifically as a writer of things that are (accidentally) books. The constraint we are placing on the context doesn’t provide John with the property with which he may describe, to himself or to others, what kind of things Tolstoy wrote “outside” of $[[T]]$ (and it’s not that John doesn’t know that these objects are books – he may well know this – it’s that he doesn’t “know” that *book* is the relevant restriction). This, we suggest, is the reason why informants typically reject the ‘high’ reading of (6) (even though there is nothing semantically or syntactically wrong with (14)). Notice that for some (though admittedly not all) informants, the ‘high’ reading of *The longest thing John said Tolstoy had ever*

written was this book is significantly better than the ‘high’ reading of (6). This is expected, because *thing* is a relatively weak restriction. In fact, in contexts where John is not imagining things that do not actually exist, *thing* is a vacuous restriction, and we merely require that the common ground entertain the possibility that for some *y* or other, every world compatible with what John believes is a member of $\{w:\text{Tolstoy wrote } y \text{ in } w \text{ before he turned eighteen}\}$.

What remains to be explained is the acceptability of (8), which, we assume, has the following LF.

(15) NOT *John said ever*-PAST_{T↔T} $\lambda t[\text{Bill } be_i \text{ to Paris}]$

(15) leads to strengthening only if it isn’t an established fact that according to John, Bill doesn’t visit Paris at any time in ($\llbracket \mathbf{T}' \rrbracket - \llbracket \mathbf{T} \rrbracket$). This, we think, is not a hard constraint to place on the common ground. The difference between (14) and (15) is, then, syntactic/pragmatic. In (14), where *book*, which is part of the head of the relative clause, is above *say*, John is not given easy access, so to speak, to the restriction that would help us attribute to him the “right” beliefs about $\llbracket \mathbf{T} \rrbracket$. In (15), where there is no relative clause, no such problem arises.

To sum up, merely requiring *ever* to be in the scope of a (Strawson) DE function is not enough, because this doesn’t provide a reason why *say* should block licensing in the ‘high’ reading of (6) (especially in view of the acceptability of (8)). The pragmatic approach suggests a way to account for this, by giving us the tools we need to talk about what happens inside the environment commanded by the (Strawson) DE function.

In the next section, we discuss the facts discovered by Heycock regarding the effect of the choice of the main predicate on the availability of a ‘low’ reading. We show that her account is hard to reconcile with the pragmatic theory of NPI-licensing. In addition, it sometimes makes wrong predictions.

3. Other Constraints on ‘Low’ Readings

3.1. Verbs that Block a ‘Low’ Reading

So far we have looked at intensional superlatives with *say*. There are other verbs that induce the same kind of ambiguity (e.g., *believe*, *hope*), as the reader can verify by constructing appropriate examples and scenarios. Interestingly, some verbs do not give rise to such an ambiguity, as observed in Heycock (to appear). More specifically, Heycock claims that such verbs support only a ‘high’ reading.

An example of such a verb is the factive *know*. According to Heycock, (16) has only a ‘high’ reading (w_0 is the actual world).

(16) The longest book John knew Tolstoy had written was War and Peace.
Predicted ‘high’ reading (accepted by speakers, according to Heycock)

WP is the unique x such that x is longest among $\{y:y \text{ is a book in } w_0 \text{ and Tolstoy wrote } y \text{ in } w_0 \text{ and for all worlds } w \text{ compatible with what John believed in } w_0, \text{ Tolstoy wrote } y \text{ in } w\}$.

Predicted ‘low’ reading (rejected by speakers, according to Heycock)

WP is the unique x such that x is the longest book Tolstoy wrote in w_0 , and for all worlds w compatible with what John believed in w_0 , x is the longest book Tolstoy wrote in w .

Although we agree that speakers find the ‘low’ reading of (16) unacceptable, we think that this is in fact not so easy to show.⁵ The problem is that in order to rule out a ‘low’ reading, we have to construct an example that speakers judge as “bad”. This means that we have to construct an example where the ‘high’ reading is not possible either, but for reasons that are independent of the intensionality of the superlative relative clause (otherwise speakers may judge the sentence as “good” by resorting to its ‘high’ reading, but we wouldn’t be able to tell for sure why they judge the sentence that way or which reading they pick).

So here is what we did. It is well known that the use of superlatives requires a relatively “large” comparison set. For example, it would be odd to utter *John is the tallest student* in a situation where John is the only student. So we came up with a scenario (Scenario C in (17)) where the ‘high’ reading is, strictly speaking, true – but odd due to the fact that the comparison set is a singleton. It turns out that speakers judge (16) as “bad”. This means that indeed, as implied by Heycock, the ‘low’ reading is also unavailable.

(17) Scenario C

John once read all of Tolstoy’s books. Recently, he admitted to himself: “The longest book Tolstoy wrote is War and Peace. But I can’t remember what else he wrote.”

The ‘high’ reading is odd, probably because $\{y:y \text{ is a book actually written by Tolstoy and John knows that Tolstoy wrote } y\}$ is a singleton set, so the sentence is almost trivially true. But the ‘low’ reading is also unavailable, and the reason for this is not obvious (the comparison set for John is not a singleton set; he knows that Tolstoy wrote more books, he just can’t remember which ones they are). Notice that a ‘low’ reading is still acceptable with *say*, and this is why (18) is fine in a similar scenario (Scenario D in (19)).

(18) The longest book John said Tolstoy had written was War and Peace.

(19) Scenario D

John says: “The longest book Tolstoy wrote is War and Peace, but I can’t remember what else he wrote”.

Heycock claims that all and only Neg-raising predicates support a ‘low’ reading. We refer to this as Heycock’s Observation, and we believe that it cannot easily co-exist with the pragmatic account of NPI licensing (section 2). We show this in 3.2. Furthermore, and perhaps more importantly, in subsection 3.3 we

show that Heycock’s Observation is empirically inadequate, as it overgeneralizes (*say*, for example, does not behave according to Heycock’s Observation, as the readers have probably noticed, and as Heycock herself admits).

3.2. A Possible Explanation of Heycock’s Observation

From the fact that not all verbs support a ‘low’ reading, Heycock concludes that Bhatt’s reconstruction account of ‘low’ readings is wrong, and that *-est* is always external to the relative clause.⁶ She proposes that since superlatives have a built-in negation (if John is tallest, it means that the others are not as tall as John), whenever the verb is Neg-raising, this negation will “reach into” the embedded clause, and this is how ‘low’ readings come about.

Heycock doesn’t offer a spelled-out execution of her idea. A concrete suggestion for executing her idea can be found in Hulsey and Sauerland (to appear), who rely on ideas from Bartsch (1973), Heim (2000) and others, regarding the semantics of Neg-raising verbs. We now show that this particular execution of Heycock’s idea cannot easily co-exist with the approach to NPI-licensing outlined in section 2. Let us sketch this idea using the Neg-raising verb *believe*, which supports ‘low’ readings. The account relies on the suggestion that *believe* presupposes that the subject either believes that the complement is true or believes that the complement is false. This accounts for the ‘low’ negative inference of negated Neg-raising verbs.

(20) John doesn’t believe that Mary left.

Presupposition: Either John believes Mary left or he believes she didn’t.

Assertion: It isn’t the case that John believes that Mary left.

Conclusion: John believes that Mary didn’t leave.

The LF that is assumed for the ‘low’ reading of *The longest book John believed Tolstoy wrote is Anna Karenina* is (21), where the argument of *-est* is a complex relation between degrees and individuals. The semantics assumed for *-est* is this: it takes a relation R between degrees and individuals, and an individual x, and requires x to be R-er than any other relevant individual.

(21) *the -est* $\lambda d \lambda x$ [John believed Tolstoy wrote a d-long book x] is AK

In addition, it is assumed that relations such as *long* are monotone (for example, if for some degree d Anna Karenina is d-long, it is also long to any degree $d' < d$). Accordingly, (21) says that there is a degree d such that John believes that Anna Karenina is d-long, but he doesn’t believe of any other member of the comparison set that it is d-long. Because of the presupposition of *believe*, it follows that for every x that is distinct from Anna Karenina, John believes that x is shorter than Anna Karenina. This is consistent with the intuitive meaning of the ‘low’ reading. However, if we switch to a non-Neg-raising verb we cannot rely on a similar presupposition, and we get a reading that doesn’t guarantee the same inference.⁷

What about NPI-placement? The widening+strengthening theory does not obviously rule out a ‘low’ reading with a ‘high’ *ever*. Consider (22), which corresponds to (5), and (23), which is the LF of its (unacceptable) ‘low’ reading.

- (22) The longest book John ever believed Tolstoy wrote is Anna Karenina.
 (23) *the –est* $\lambda d \lambda x$ [*ever-PAST*_{T¹→T} λt [*John believe_t Tolstoy wrote a d-long book x*]] *is AK*

Notice that *ever* is c-commanded by *–est* but not by *believe*. The problem, then, may be described as follows. To enable strengthening, the context has to be such that it is not an established fact that $[[T]]$ covers all the times at which John believes about something that it is a book of some length by Tolstoy. It seems to us that this is not a significantly harder constraint than the following: that it is not an established fact that $[[T]]$ covers all the times at which John believes about an actual book of some length that it is by Tolstoy (associated with the ‘high’ reading of (22)). In both cases *ever* is above *believe* and John is not required to have any beliefs about $[[T]]$ itself. The unwelcome prediction is that both the ‘high’ and the ‘low’ readings have the same status. Notice that the reconstruction analysis of ‘low’ readings given in (4) doesn’t generate (23) (see Endnote 1).

It is possible, of course, that the pragmatic theory of disambiguation by NPI-placement outlined in section 2 is wrong, but there is no obvious alternative to it within the Neg-raising account. Moreover, it seems to us that there are independent reasons to suspect that the Neg-raising account is not on the right track, because some non-Neg-raising verbs seem to support ‘low’ readings, and it is not clear that all Neg-raising verbs support such readings. Let us look at a number of relevant cases.

3.3. The Empirical Coverage of Heycock’s Observation

In accordance with Heycock’s Observation, *believe* and *want* (which support ‘low’ readings) are Neg-raising; *know*, *manage* and *need* (which do not support ‘low’ readings; see section 4 below) are not Neg-raising.

However, some verbs – such as *hope*, *agree* and *certain* – that support ‘low’ readings are not Neg-raising.

- (24) The longest book John hopes he will (ever) have to read is Anna Karenina.

(24) is accepted by speakers in a situation where John mistakenly believes that Anna Karenina is longer than War and Peace, he knows he will have to read them both, and hopes that at no time will the set of books he has to read include a book longer than Anna Karenina. This has to be a ‘low’ reading because War and Peace is actually longer than Anna Karenina.

- (25) The longest book John and Mary agree Tolstoy (ever) wrote is Anna Karenina.

(25) is accepted by speakers in a situation where John and Mary, independently of each other, hold the following belief: Tolstoy wrote War and Peace and Anna Karenina, and the latter is the longest book he wrote.

(26) The longest book John is certain Tolstoy (ever) wrote is Anna Karenina.

(26) is accepted by speakers when John expresses his firm belief that Tolstoy wrote Anna Karenina and War and Peace, and the former is the longest among Tolstoy's books. Yet these verbs are not Neg-raisers, as (27-29) show.

- (27) a. John and Mary do not agree that Bill left.
 $\neq \Rightarrow$ b. John and Mary agree that Bill didn't leave.
- (28) a. John isn't certain that Mary left.
 $\neq \Rightarrow$ b. John is certain that Mary didn't leave.
- (29) a. John didn't hope he would win.
 $\neq \Rightarrow$ b. John hoped he wouldn't win.

In a situation where John believes that Bill left and Mary thinks he didn't, (27a) is true but (27b) is false. In a situation where John thinks that it's possible that Mary left but he isn't sure, (28a) is true but (28b) is false. In a situation where John is reluctant about winning, (29a) is true but (29b) need not be.

Another verb that is problematic in this respect is *say*. That *say* is not a Neg-raising verb is evidenced by the fact that *John didn't say that Mary left* doesn't imply *John said that Mary didn't leave*. Heycock argues that *say* is a true exception to her observation: speakers accept 'low' readings because *say* is an evidential; for example, (30b) is the inferred interpretation of (30a) (with the superlative-like adjectival *only*), and it is also the source of the 'low' reading of (30a). But (30b) is ambiguous: *according to John* may be construed as modifying either the matrix or the embedded clause. We think that the unambiguous (30c) is what is actually inferred from (30a), and no 'low' reading is implied.

- (30) a. This is the only book that John says that Tolstoy wrote.
 b. This is the only book that Tolstoy wrote, according to John.
 c. This is the only book that, according to John, Tolstoy wrote.

Besides, the explanation doesn't carry over to *agree*, *certain*, and *hope*.

Turning to the class of Neg-raising predicates, two interesting cases Heycock discusses are *likely* and *should*, which she claims support 'low' readings.

- (31) a. It isn't likely that John will leave.
 \Rightarrow b. It is likely that John won't leave.
- (32) a. I don't think you should eat ice cream.
 \Rightarrow b. I think you should not eat ice cream.
- (33) a. This is the only book that it is likely that John read.
 b. This is the only book that you should read.

Indeed, (33a) is acceptable in a situation where the likelihood is that John reads only this book, and (33b) is acceptable in a situation where the requirements are that you read only this book. But we think it is a mistake to test these cases with *only*. *Only*, as opposed to *-est*, does not carry a pragmatic prohibition against its sister denoting a singleton set. So when speakers accept (33a,b) we cannot be sure that they are not resorting to the ‘high’ reading. The real test is with *-est*, where a reading which is trivially true when the comparison set is a singleton is rejected for pragmatic reasons. Indeed, it seems that (34a,b) do not have ‘low’ readings.

- (34) a. The tallest man Mary is likely to meet is John.
 b. The tallest man Mary should meet is John.

(34a) is odd in a situation where the likelihood is that Mary meets John and some (unspecific) shorter men. The ‘high’ reading is odd because there is only one specific man (namely, John) that she is likely to meet. Notice that in a situation where the likelihood is that she meets John, Bill, Fred and some shorter men (and Bill and Fred are shorter than John) the sentence may be accepted by speakers, but in that case the ‘high’ reading is pragmatically plausible, so we cannot be sure that speakers do not focus on that reading.⁸ Similar comments apply to (34b).

In sum, we believe that Heycock’s Observation does not provide an argument against the reconstruction analysis of ‘low’ readings. Let us consider an alternative observation, which is consistent with the reconstruction analysis.

4. The ‘Low-to-High’ Generalization

We suggest that the ‘low’ readings of intensional superlatives are subject to (35).

- (35) The ‘Low’-to-‘High’ Generalization
 The predicted ‘low’ reading of a sentence with an intensional superlative is rejected by speakers whenever it (Strawson) entails the ‘high’ reading of the same sentence.

We now illustrate how (35) works with some specific examples. (35) doesn’t cover cases of covert and overt negation (as in *The longest book John didn’t say Tolstoy had written* and *The longest book John doubted Tolstoy had written*, which do not have ‘low’ readings). We ignore these cases here for lack of space.

Among the predicates that according to Heycock do not support a ‘low’ reading are factives, implicatives, “needing” verbs (which she calls “strong deontic”), and certain possibility modals. Although we do not go over every single example of this sort, we argue that in all these cases the ‘low’ reading entails the ‘high’ reading. Let us start with (16) (with *know*).

Suppose w_1 and w_2 are the worlds compatible with John's beliefs in w_0 , and they are as described in (36) (a situation compatible with Scenario C in (17)).

- (36) w_1 : { y : y is a book written by Tolstoy} =

{War and Peace, Anna Karenina}
 w2: {y:y is a book written by Tolstoy} =
 {War and Peace, Tom Sawyer}

The ‘low’ reading of (16) is true in w0, because War and Peace is indeed the longest book written by Tolstoy in w0, and also in w1 and w2. And since War and Peace is longest in {y:y is a book written by Tolstoy in w0 and John believes in w0 that Tolstoy wrote y} (which happens to be a singleton, because John is not certain about either Anna Karenina or Tom Sawyer, but nothing hinges on this), the ‘high’ reading is true too (for a more detailed proof see Appendix, (i)).

The second class of verbs that block a ‘low’ reading is the class of implicatives (e.g., *manage*), as in the following example.

- (37) The longest book Mary managed to read was War and Peace.
Predicted ‘high’ reading (accepted by speakers)
 WP is the unique x that is longest among {y:y is a book in w0 and Mary managed in w0, despite some obstacles, to bring it about that she reads y}.
Predicted ‘low’ reading (rejected by speakers)
 WP is the unique x such that Mary managed in w0, despite some obstacles, to bring it about that x be the longest book she reads.

A scenario where the ‘high’ reading is true is given in (38), and a scenario where the ‘low’ reading would be true (but actually judged as “bad”) is given in (39).

- (38) Scenario E
 Mary read War and Peace, Anna Karenina, and Crime and Punishment.
 For each of these books, it wasn’t obvious that Mary would read it.
 War and Peace is the longest among these books.
- (39) Scenario F
 Mary read War and Peace and no book longer than that. It wasn’t obvious that War and Peace would be the longest book Mary would read.

We argue that the predicted ‘low’ reading of (37) entails its predicted ‘high’ reading. We assume that *manage* takes as its arguments a property of individuals P and an individual x. It asserts that P holds of x in w0, and presupposes that there was some obstacle to that (in other words, it presupposes that not all the relevant worlds accessible from w0 are such that P holds of x). For example, *Mary managed to find her car* asserts that she found her car in w0, and presupposes that at least one world accessible from w0 is such that she didn’t find her car or that she doesn’t have a car (if, for example, it was destroyed).

What is the set of relevant worlds “accessible from w0” in this case? We assume that this is the biggest set of worlds that are minimally different from the actual world. That is to say, the same general circumstances that hold in the actual world hold in them too. In the default case, these worlds will not be different from the actual world in these respects. In our case, Mary doesn’t possess the ability to fly in them, she has the same financial means in them as she actually has, etc.

Now consider (37), and suppose w_3 and w_4 are the set of worlds accessible from w_0 , and that w_3 , w_4 and w_0 are as described in (40) (which is compatible with Scenario F in (39)).

- (40) w_3 : $\{y:y \text{ is a book and Mary reads } y\} =$
 $\{\text{War and Peace, Crime and Punishment}\}$
 w_4 : $\{y:y \text{ is a book and Mary reads } y\} =$
 $\{\text{Crime and Punishment, Anna Karenina}\}$
 w_0 : $\{y:y \text{ is a book and Mary reads } y\} = \{\text{War and Peace}\}$

The ‘low’ reading of (37) is true in this state of affairs, because not all the accessible worlds are such that War and Peace is the longest book Mary reads, or a book that she reads at all. In fact, since the set of accessible worlds is the largest set of such worlds, there is bound to be at least one world where Mary doesn’t read War and Peace at all (in our case, w_4 is such a world). And indeed, in this state of affairs the ‘high’ reading of (37) is true too, because War and Peace is longest in $\{y:y \text{ is a book in } w_0 \text{ and Mary managed in } w_0 \text{ to read } y\}$ (which happens to be a singleton, but nothing hinges on this), rendering the ‘high’ reading true. Notice that all the accessible worlds are such that the books in them are books in the actual world (we do not consider worlds where Anna Karenina is a pamphlet), and they all have the same length as they do in the actual world (we do not consider worlds where Anna Karenina is longer than War and Peace). This follows from our assumption that the accessible worlds are as similar as possible to the actual world (for a more detailed proof, see Appendix, (ii)).

Next consider “needing” verbs, as in (41).

- (41) The longest book John needs/has to read is Anna Karenina.
Predicted ‘high’ reading (accepted by speakers): AK is the unique x such that x is the longest among $\{y:y \text{ is a book in } w_0 \text{ and } \text{NEED}(\text{John})(w_0) \subseteq \{w:\text{John reads } y \text{ in } w\}\}$.
Predicted ‘low’ reading (rejected by speakers): AK is the unique x such that $\text{NEED}(\text{John})(w_0) \subseteq \{w:x \text{ is the longest book that John reads in } w\}$.

Scenario G in (42) is appropriate for the ‘high’ reading of (41), and Scenario H in (43) would be appropriate for its ‘low’ reading.

- (42) Scenario G
The requirements for passing Comp Lit I are: reading Tom Sawyer, Huckleberry Finn, and Anna Karenina. If John doesn’t pass Comp Lit I, he will be kicked out.
- (43) Scenario H
The requirements for passing Comp Lit II are: reading Anna Karenina and no book that is longer than Anna Karenina. If John doesn’t pass Comp Lit II, he will be kicked out.

We argue that the predicted ‘low’ reading of (41) entails its predicted ‘high’ reading. We assume, here too, that in the default case the members of $\text{NEED}(\text{John})(w_0)$ (i.e., the set of worlds compatible with John’s needs in w_0) are “minimally different” from w_0 . That is to say, they contain the same books as w_0 , and those books have the same lengths in them as they do in w_0 .

Suppose $\text{NEED}(\text{John})(w_0) = \{w_{17}, w_{18}\}$, and that these worlds are as described in (44) (which is compatible with Scenario H in (43)).

- (44) w_{17} : $\{y:y \text{ is a book and John reads } y\} =$
 $\{\text{Anna Karenina, Tom Sawyer, Pride and Prejudice}\}$
 w_{18} : $\{y:y \text{ is a book and John reads } y\} =$
 $\{\text{Anna Karenina, Huckleberry Finn, Emma}\}$

In this case the ‘low’ reading is true, because w_{17} and w_{18} are such that Anna Karenina is the longest book John reads. But because these are all books in w_0 as well (and they have the same lengths as they do in w_0), it follows that the ‘high’ reading is true too, because Anna Karenina is longest in $\{y:y \text{ is a book in } w_0 \text{ and John needs in } w_0 \text{ to read } y\}$ (which happens to be a singleton here, but nothing hinges on this; for a more detailed proof, see Appendix, (iii)).

The assumption that $\text{NEED}(\text{John})(w_0)$ contains worlds that are as similar as possible to w_0 is based on our intuitions regarding the following contrast.

- (45) a. John mistakenly believes he has a violin. He hopes to sell his violin.
 b. #John mistakenly believes he has a violin. He needs to sell his violin.

The oddity of (45b) is the result of the fact that the set of worlds compatible with John’s needs reflects, by and large, the speaker’s beliefs rather than John’s. The wellformedness of (45a) is the result of the fact that the set of worlds compatible with John’s desires reflects his beliefs rather than the speaker’s. In other words, when evaluating sentences with *need* we tend to look at modal bases that reflect our beliefs, as much as possible.⁹ When evaluating sentences with *believe*, *hope*, *want*, etc. we look at modal bases that reflect the subject’s beliefs (however mistaken those beliefs might be). And indeed, as we saw in section 3, *hope* is one of those verbs that support a ‘low’ reading of intensional superlatives.

So, to repeat, the generalization is that speakers reject a ‘low’ reading whenever it entails the ‘high’ reading. It is hard to characterize the class of predicates that give rise to such an entailment using well-established predicate classes. What is certain is that predicates that are not factive and have modal bases whose members tend to be very different from the actual world will support ‘low’ readings of intensional superlatives, because then the ‘low’ reading cannot, in principle, entail the ‘high’ reading. Not all such predicates are Neg-raising (e.g., *hope*, *agree on*, and *certain*, discussed in 3.3),¹⁰ and some Neg-raising predicates are not like that (e.g., *likely* and *should*, discussed in 3.3).

Some of the predicates that block ‘low’ readings are existential quantifiers over worlds (unlike, for example, *need* which is a universal quantifier). With such

predicates, whenever the ‘low’ reading is rejected by speakers it Strawson-entails the ‘high’ reading. Consider (46).

- (46) The longest book that John could have read is Anna Karenina.
‘High’ reading (accepted by speakers): AK is the unique x such that x is longest among $\{y:y \text{ is a book and there is an accessible world where John reads } y\}$
‘Low’ reading (rejected by speakers): AK is the unique x such that there is an accessible world where x is the longest book John reads.

The ‘low’ reading doesn’t entail the ‘high’ reading in the strict sense. To see this, suppose the accessible worlds – w_{20} and w_{21} – are as described in (47).

- (47) w_{20} : $\{y:y \text{ is a book and John reads } y\} =$
 $\{\text{Anna Karenina in English (=AK), Tom Sawyer}\}$
 w_{21} : $\{y:y \text{ is a book and John reads } y\} =$
 $\{\text{Anna Karenina in English, Anna Karenina in French}\}$
The two versions of Anna Karenina are equally long in w_0 , w_{20} and w_{21} .

In this case, the ‘low’ reading is true (because only AK is such that in some accessible world – w_{20} – it is the longest book John reads), but it doesn't follow that the ‘high’ reading is true, even if w_{20} and w_{21} are very similar to the actual world. This is because, since the two versions of Anna Karenina are equally long, there is no unique longest book that John reads in an accessible world. However, if we take into account the presuppositions of the ‘high’ reading and assume that there is indeed a longest element in $\{y:y \text{ is a book and there is an accessible world where John reads } y\}$, the ‘high’ reading is indeed entailed (see Appendix, (iv)).

Notice that in many (if not all) of the cases where the ‘low’ reading is unacceptable, the two readings Strawson-entail each other. What can we make of this? When speakers accept a sentence of this sort (where the two readings are Strawson-equivalent) in a situation where both readings are pragmatically plausible, it is hard to tell which reading they have in mind anyway. But if we concede that an important factor in determining speakers’ perception of two readings of the same surface string as distinct is that they not be Strawson-equivalent, the distinction between the ‘high’ and the ‘low’ readings in the cases where they are Strawson-equivalent becomes blurred. This, we suggest, might be the reason why speakers reject the ‘low’ reading when the two readings are Strawson-equivalent and the ‘high’ reading is pragmatically odd.

5. Summary

We have seen that disambiguation of intensional superlatives by NPI placement cannot be explained by merely assuming that an NPI is licensed in the scope of a (Strawson) DE function. On the other hand, a pragmatic theory based on widening and strengthening does provide a useful way of accounting for the facts. We also

saw that the Neg-raising approach to the Heycock facts is incompatible with this account of NPI licensing. Furthermore, it doesn't predict certain non-Neg-raising verbs to support 'low' readings, contrary to fact. Notice that the 'Low'-to-'High' Generalization is compatible with the pragmatic explanation of disambiguation by NPI-placement, because the presence of *ever* doesn't effect the entailment relations between the relevant 'low' and 'high' readings.

Appendix

- (i) A. Assume WP is the unique x such that: (i) x is longest in BOOK ($=\{y:y$ is a book Tolstoy wrote in $w_0\}$), and (ii) for all $w \in \text{ACC}_{w_0}$, x is the longest book Tolstoy wrote in w (i.e., (16) is true on its 'low' reading).
 B. ACC_{w_0} is the set of worlds compatible with what John believed in w_0 .
 C. Given that WP is longest in BOOK and for all worlds $w \in \text{ACC}_{w_0}$, WP is the longest book Tolstoy wrote in w , then $\text{WP} \in \text{BOOK}^{\text{know}}$ ($=\{y:y$ is a book Tolstoy wrote in w_0 and for all worlds w in ACC_{w_0} , Tolstoy wrote y in $w\}$).
 D. Given that $\text{BOOK}^{\text{know}} \subseteq \text{BOOK}$, WP is longest in BOOK , and $\text{WP} \in \text{BOOK}^{\text{know}}$, it follows that WP is longest in $\text{BOOK}^{\text{know}}$. That is to say, (16) is true on its 'high' reading.
- (ii) A. Assume WP is the unique x such that: (i) x is longest in BOOK_{w_0} ($=\{y:y$ is a book in w_0 and Mary read y in $w_0\}$), and (ii) not all worlds $w \in \text{ACC}_{w_0}$ are such that x is the longest book Mary reads in w or such that Mary reads x in w (i.e., (37) is true on its 'low' reading).
 B. ACC_{w_0} is the biggest set of worlds that are similar to w_0 , but for the possible difference that Mary reads a longer book than WP or that she doesn't read WP at all.
 C. Given B, it follows that not all worlds $w \in \text{ACC}_{w_0}$ are such that Mary reads WP in w .
 D. Now consider $\text{BOOK}^{\text{manage}}$ ($=\{x:x \in \text{BOOK}_{w_0}$ and not all w in ACC_{w_0} are such that Mary reads x in $w\}$). From A, we know that WP is longest in BOOK_{w_0} . From C, we know that $\text{WP} \in \text{BOOK}^{\text{manage}}$. Since $\text{BOOK}^{\text{manage}} \subseteq \text{BOOK}_{w_0}$, it follows that WP is longest in $\text{BOOK}^{\text{manage}}$. That is to say, (37) is true on its 'high' reading.
- (iii) A. Assume AK is the unique x such that all worlds $w \in \text{NEED}(w_0)(\text{Mary})$ are such that x is longest among BOOK_w ($=\{y:y$ is a book in w and Mary reads y in $w\}$) (i.e., (41) is true on its 'low' reading).
 B. $\text{NEED}(w_0)(\text{Mary})$ is the largest set of worlds that are just like w_0 but for the possible difference that Mary's needs are fulfilled.
 C. Now consider $\text{BOOK}^{\text{need}}$ ($=\{y:y$ is a book in w_0 and all worlds w in $\text{NEED}(w_0)(\text{Mary})$ are such that Mary reads y in $w\}$). Since all the things that are books in w_0 are also books in $\text{NEED}(w_0)(\text{Mary})$ and vice-versa (from B), it follows that for all w in $\text{NEED}(w_0)(\text{Mary})$, $\text{BOOK}^{\text{need}} \subseteq \text{BOOK}_w$.

- D. Since all the books have the same lengths in w_0 and throughout $NEED(w_0)(Mary)$, and since AK is longest in $BOOK_w$ for all $w \in NEED(w_0)(Mary)$, it follows that AK is longest in $BOOK^{need}$. That is to say, (41) is true on its ‘high’ reading.
- (iv) A. (i) Assume AK is the unique x such that there is world $w \in ACC_{w_0}$ where x is the longest book John reads, and (ii) let AK' be the unique x such that x is longest among $\{y: y \text{ is a book in } w_0 \text{ and there is a world } w \in ACC_{w_0} \text{ where John reads } y\}$ (i.e., (46) is true on its ‘low’ reading, and the presupposition of its ‘high’ reading is true).
- B. ACC_{w_0} is the biggest set of worlds w that are just like w_0 , but for the possible difference that John does in w what he is permitted in w_0 .
- C. Suppose $AK \neq AK'$. Given B, if AK' is longer than AK , A(i) is contradicted; if AK' is shorter than AK , A(ii) is contradicted; and if they are equally long, A(i)-(ii) are contradicted. Therefore, $AK = AK'$. That is to say, (46) is true on its ‘high’ reading.

Endnotes

* For very valuable comments, we are grateful to the audiences at UConn, UMass, NYU, Semantiknetzwerk II (Nijmegen), SALT15 (UCLA), Tel-Aviv University, and Egg’05 Summer School (University of Wroclaw). Special thanks go to Gidi Avrahami, Jon Gajewski, Caroline Heycock, and Ede Zimmermann for their helpful and thought provoking feedback. All errors are of course ours.

¹ We assume here that *-est* (or even *R-est*) cannot separate at LF from the rest of the “head” (cf. Sharvit and Stateva 2002). Thus, according to this analysis, no LF where *-est* (or *R-est*) is above the main predicate and the rest of the “head” is below it can be generated (cf. (21) in 3.2).

² For simplicity, we systematically omit the variable assignment g from $\llbracket _ \rrbracket^g$.

³ Throughout this paper, the variable denoting the bigger domain is free, like the one denoting the original domain (implying they are both supplied contextually). A better alternative might be to have the variable denoting the bigger domain universally bound (along the lines suggested in Chierchia 2001). We haven’t yet worked out the details of this alternative with respect to the data discussed here.

⁴ See Herdan and Sharvit (2005) for why adjectival *only* is also Strawson DE.

⁵ The mere unacceptability of *The longest book John knew Tolstoy had ever written was War and Peace* is not sufficient to show this, as it could, in principle, be the result of a clash between the factivity of *know* and the presence of the NPI.

⁶ In Sharvit (to appear) it is also argued that *-est* is always external to the relative clause, but for reasons that are independent of the facts discussed here.

⁷ But we still predict a plausible reading. For example, we predict (16) to have a reading where WP is the unique x such that there is a degree d such that John knew that x is a d -long book, and for every $y \neq x$, John didn’t know that y is a d -long book. It is not obvious, under this view, why this reading is unavailable.

⁸ A ‘low’ *ever* with a ‘high’ *tallest man* is possible here, but given our analysis in section 2 this is because *likely* is not an attitude verb and Mary is not required to have any beliefs (specifically not about the time stretch that is widened by *ever*).

⁹ How similar these modal bases are to w0 depends on the context. In contexts where these modal bases are not similar to w0, *need* behaves more like *believe*.

¹⁰ Interestingly, Heycock observes the unacceptability of the ‘low’ reading of *The longest book that it is certain Tolstoy wrote is Anna Karenina*. We think there are two *certain*’s. Only the one used with a subject expletive has a modal base whose members are minimally different from the actual world.

References

- Bartsch, Renate: 1973, ‘ “Negative Transportation” gibt es nicht’, *Linguistische Berichte* **27**.
- Bhatt, Rajesh: 2002, ‘The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification’, *Natural Language Semantics* **10**, 43-90.
- Chierchia, Gennaro: 2001, ‘Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface’.
- Fauconnier, Giles: 1975, ‘Polarity and the Scale Principle’, *Chicago Linguistics Society* **11**, 188-199.
- Fauconnier, Giles: 1978, ‘Implication Reversal in a Natural Language’, in F. Guenther and S.J. Schmidt (eds.) *Formal Semantics and Pragmatics for Natural Languages*, 289-302. Dordrecht, Reidel.
- von Fintel, Kai: 1999, ‘NPI Licensing, Strawson Entailment, and Context Dependencies’, *Journal of Semantics* **16**, 97-148.
- Heim, Irene: 2000, ‘Degree Operators and Scope’, in B. Jackson and T. Matthews (eds.) *Proceedings of the Semantics and Linguistic Theory* **10**, CLC Publications, Cornell University.
- Herdan, Simona and Yael Sharvit: 2005, ‘Definite and Non-definite Superlatives and NPI-licensing’, ms. University of Connecticut.
- Heycock, Caroline: to appear, ‘On the Interaction of Adjectival Modifiers and Relative Clauses’, *Natural Language Semantics*.
- Hulsey, Sarah and Uli Sauerland: to appear, ‘Sorting out Relative Clauses’, *Natural Language Semantics*.
- Kadmon, Nirit and Fred Landman: 1993, ‘Any’, *Linguistics and Philosophy* **16**, 353-422.
- Ladusaw, William: 1979, *Polarity Sensitivity as Inherent Scope Relations*. Ph.D. dissertation, University of Texas at Austin.
- Linebarger, Marcia: 1987, ‘Negative Polarity and Grammatical Representation’, *Linguistics and Philosophy* **10**, 325-387.
- Sharvit, Yael: to appear, ‘Two Reconstruction Puzzles’, in C. Barker and P. Jacobson (eds.) *Proceedings of the Workshop on Direct Compositionality* (Brown University, June 2003).
- Sharvit, Yael and Penka Stateva: 2002, ‘Superlative Expressions, Context, and Focus’, *Linguistics and Philosophy* **25**, 453-504.