Superlative degree clauses: Evidence from NPI licensing

by

Edwin Martin Howard

B.A. with Honours, Linguistics, McGill University 2009

Submitted to the Department of Linguistics and Philosophy in partial fulfilment of the requirements for the degree of
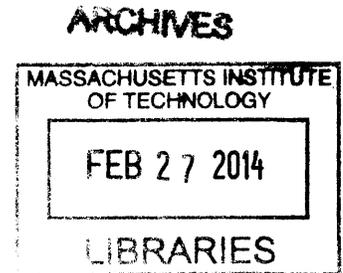
Master of Science in Linguistics

at the

Massachusetts Institute of Technology

February, 2014

Author:_____
Department of Linguistics and Philosophy, 7th February, 2014

Certified by:_____
Irene Heim
Professor of Linguistics
Thesis supervisor

Accepted by:_____
David Pesetsky
Ferrari P. Ward Professor of Modern Languages and Linguistics
Head, Department of Linguistics and Philosophy

Superlative degree clauses: Evidence from NPI licensing

by

Edwin M. Howard

Submitted to the Department of Linguistics and Philosophy on 7th February, 2014, in partial fulfilment of the requirements for the degree of Master of Science in Linguistics

Abstract

This paper concerns itself with the superlative morpheme *-est* and its ability to license Negative Polarity Items (NPIs) such as *any* or *ever*, and in particular the compositional analysis of a puzzling type of utterance which incorporates these grammatical elements. Specifically, an embedded clause structure which appears at first glance to be a restrictive relative clause modifier will be argued to be an argument of the superlative morpheme, and an analysis will be presented that brings superlatives in line with comparatives in finding its semantic restrictor in the overt syntax, in the form of a degree clause.

The first puzzle comes from an observation about the predictions of von Fintel's 1999 theory of NPI licensing when combined with those of a scope theory of ambiguities in superlative utterances: when *-est* takes scope at the VP level or higher, it fails to license the particular logical inference (Strawson-Downward Entailment) that von Fintel argues is necessary for NPI licensing. However the target data this paper focuses on seem to present counterevidence to this prediction, as they exhibit NPI licensing by the superlative within what appears to be a relative clause modifier. This puzzle evaporates under the assumption that the embedded clause in question is the degree clause argument of the superlative.

Further puzzles arise from unexpected contrasts in felicity and acceptability between minimal pairs of sentences instantiating the superlative degree clause structure. These contrasts are explained under the semantic analysis presented in the paper. It is also demonstrated that the environment where NPIs are interpreted is Strawson-Downward Entailing, in line with von Fintel 1999. NPIs are analysed here as introducing alternatives to create a nontrivial set of degree properties, which provides *-est*'s restrictor, assuming the denotation proposed in Heim 1999.

Thesis supervisor: Irene Heim, Professor of Linguistics

1. Introduction

This paper concerns itself with the superlative morpheme *-est* and its ability to license Negative Polarity Items such as *any* or *ever*, and in particular the compositional analysis of utterances such as the following:

(1)   John read the most books (that) anyone in the class has ever read.

(2)   Mary sang the loudest (that) anyone in the group has ever sung.

In (1) the NPIs *anyone* and *ever* appear within what appears at first glance to be a relative clause modifier to the noun *books*, which is itself in construction with the superlative adjective of quantity *most* (Szabolsci 1986; Gawron 1995; Hackl 2009). Since *-est* is known to license these NPIs, by dint of being a Strawson-Downward Entailing operator (von Fintel 1999; Herdan & Sharvit 2006, Gajewski 2010), it may come as no surprise that (1) is acceptable – in contrast to the ungrammatical examples in (3)a, where there is no Downward-Entailing operator (of any kind) to be found.

(3)   a. *John read some/many/three books (that) anyone in the class has ever read.

      b. John read no/few/less than three books (that) anyone in the class has ever read.

Although (2) does not appear to contain a NP for a relative clause to modify, presumably an account along these lines could be developed[1].

However, I will argue in this paper that the sketch I have just alluded to contains two flawed assumptions: the first of which is that the Strawson-Downward Entailment theory of NPI licensing necessarily predicts *-est* to license NPIs within a constiutuent it has scope over; the second is that the embedded clauses in (1) and (2) are relative clause modifiers. I will present evidence for an analysis according to which these relative clause-like elements are in fact interpreted as the complement of the superlative morpheme; the NPI elements contribute to the interpretation of such structures in specifying *-est*'s domain of comparison, while *-est* creates the SDE environment in which they are licensed.

---

[1] Such an approach would be compatible with the claims argued for in Matushansky 2008.

2. Background on superlatives

Appreciating that there is a puzzle here will involve some discussion of the readings of superlative utterances, and in particular the *est*-movement approach to certain ambiguities (Szabolcsi 1986, Heim 1999, i.a.), the extension of this approach to cover *the most NP* (Hackl 2009), and the predictions of such a theory with respect to an account of the NPI licensing properties of -*est* (von Fintel 1999; Gajewski 2010).

2.1 VP vs NP scope for the superlative morpheme

Heim 1999 discusses an ambiguity[2] detectable in utterances containing a superlative NP such as (4). Under one construal, (4) is understood as meaning that John climbed the mountain which is highest amongst all mountains under consideration; if this set comprises all mountains in the world, (4) amounts to the claim that John climbed Mt Everest. On the other possible reading, the sentence is true if John climbed a higher mountain than any other relevant individual did.

(4)   John climbed the highest mountain.   (Heim 1999)

      a. "John climbed the unique mountain whose height exceeds that of any other mountain"

      b. "John climbed a mountain whose height exceeds that of any mountain climbed by any salient alternative to John"

---

[2] According to Heim (1999, p.1, fn.1) this observation had been discussed in various sources, many of which I have not been able to obtain access. I reproduce Heim's note here in its entirety for reference:

> "This observation apparently goes back to J. R. Ross (1964) "A partial grammar of English superlatives," UPenn MA thesis. Further early references include J. Bowers (1969) "Surface structure interpretation in English superlatives," unpubl. MIT ms.; C.L. Baker (1970), anonymous review of Bowers; and R. Jackendoff (1972) Semantic Interpretation in Generative Grammar, MIT Press (p. 254). It was rediscovered in Szabolcsi (1986) and Heim (1985) "Notes on comparatives and related matters," unpubl. UTexas ms."

Of these works, I have been able to consult the last three, which are thus listed in the references for this paper.

Following Szabolsci 1986, Heim proposes an analysis of this ambiguity as resulting from different scope possibilities for the superlative morpheme -*est*. The former reading in (4)a, dubbed 'absolute' in Szabolsci's terminology, results when -*est* takes scope at the level of the NP within which it appears, whereas the latter reading in (4)b, the "comparative" or "relative" reading, comes about when -*est* takes scope outside VP.

To make sense of what it means for -*est* to have different scope possibilities, let me lay out some basic assumptions of the framework adopted in Heim (1999, 2001).

(5) $[\![$ **high** $]\!]$ = $\lambda d.\ \lambda x.$ HEIGHT $(x) \geq d$

According to (5), the gradable adjective *high* denotes a monotone function of type $\langle d,et \rangle$, which relates an individual x and a degree d if x's height is at least d. Heim assumes that the d argument is realised as a syntactic position which can be occupied by a degree-denoting expression (as in (6) for example), or can be the site where an operator on degree properties, such as the superlative morpheme, is first merged into the structure. Movement of such an operator is assumed to leave a trace of type d, which can then be interpreted as a bound variable, thus deriving a complex predicate of degrees (as represented schematically in (7)).

(6)   Ben Nevis is [[4,400 feet] high]
(7)   -est $\lambda d[$ ... [d-high] ...]

When it takes scope at NP, giving rise to the absolute reading of (4), -*est* compares members of extension of the nominal predicate *mountain* in terms of the gradable property *high*. The complement of *climb* is a definite description picking out the unique individual in the set of mountains which *high* relates to a degree greater than it does any other member of this set.

Under the relative reading of (4), -*est* moves outside the modified NP within which it appears, and takes scope above VP. -*est* then compares the verb's external argument *John* to other individuals, with respect to the gradable property derived from the meaning of the VP, that of 'climbing a d-high mountain'; this is then an instance of the schema in (7). The DP embedded within the VP which -*est* has evacuated is interpreted as an indefinite, as argued for in Szabolcsi 1986, and reflected in the paraphrase given in (4)b. The utterance is then true if, in terms of

height of mountains climbed, John is associated with a greater degree than anyone else under consideration.

2.2 _most_ as the superlative of _many_

A common analysis (Szabolsci 1986, Gawron 1995, i.a.) of _most_ in e.g. _the most books_ decomposes it into _-est_ and _many_. Hackl 2009, assuming the framework of Heim 1999, gives a gradable adjective meaning for the latter, as in (8).

(8) $[\![\text{ many }]\!] = \lambda d. \lambda x. |x| \geq d$

According to (8), _many_ returns true for a plural individual x and a degree d as long as the cardinality of x is at least d, which is to say that x contains as many as d atomic sub-parts. This analysis makes sense of the intuitive truth conditions of an utterance such as (9); this is a true assertion if John climbed more mountains than anyone else did. The paraphrased meaning is analogous to that of the relative reading of (4) in (4)b.

(9)   John climbed the most mountains.   (Hackl 2009)
         "John climbed a plurality of mountains whose cardinality exceeds that of any plurality of
         mountains climbed by any salient alternative to John"

On the other hand, Hackl argues that _the most NP_ in such examples lacks a meaning corresponding to the absolute reading. For example, the missing absolute reading of (9) would be paraphrased as in (10), and as Hackl points out, would be equivalent to _John climbed all the mountains_; clearly (9) cannot have this meaning. Gawron 1995 makes a similar observation.

(10)   !"John climbed the unique plurality of mountains whose cardinality exceeds that of any
          other plurality of mountains"   (_impossible meaning for_ (9))

Hackl suggests an explanation for the unavailability of an absolute reading for (9) in terms of the semantic ill-formedness of the definite description resulting when _-est_ takes scope at NP; for present purposes it suffices to understand that NP scope is impossible for _-est_ in _the most NP_.

2.3 The superlative as an NPI licenser

Superlative utterances are known to license the NPIs *any* and *ever*, as illustrated below:

(11)   The tallest girl to *ever* win the contest was over four feet tall.   (based on von Fintel 1999)

(12)   The longest book John (*ever*) said Tolstoy had (*ever*) written was Anna Karenina.

(Bhatt 2002, Bhatt & Sharvit 2005)

(13)   The dean praised the best student who has *any* knowledge of French.

(Herdan & Sharvit 2006)

(14)   Most men with *any* brains eat rutabagas.   (Safir 1982)

In (14) *any* is licensed within an NP headed by *most*, which is analysed by Hackl 2009 as composed of *-est+many* (Gajewski 2010).

The *est*-movement analysis has been discussed in recent work (Gajewski 2010, Herdan & Sharvit 2006, i.a.) in light of von Fintel's (1999) proposal that NPIs are only permitted in Strawson-Downward Entailing (SDE) environments; that's to say, a syntactic context that supports the truth-preserving replacement of one term by a stronger – as long as any presuppositions of the resulting utterance are satisfied[3]. This refinement to the original Fauconnier-Ladusaw generalisation on the distribution of NPIs, which appealed to garden-variety downward entailment, is proposed by von Fintel (1999) in order to deal with a variety of well-known counterexamples including *only NP*:

(15)   Only John *ever* ate *any* kale for breakfast.

(16)   a. Only John ate vegetables for breakfast.

b. Only John ate kale for breakfast.   (examples from von Fintel 1999)

---

[3] The statements of von Fintel's (1999) definition of SDE-ness (i), and the condition on NPI licensing (ii):

(i)   A function of type $<\sigma, \tau>$ is <u>Strawson downward entailing</u> (SDE) iff for all x, y of type σ such that x $\Rightarrow$ y and f(x) is defined: f(y) $\Rightarrow$ f(x). *['$\Rightarrow$' stands for 'cross-categorical entailment': relatable to '$\subseteq$']*

(ii)   An NPI is only grammatical if it is in the scope of a SDE operator.

Whereas *only John* licenses NPIs in its scope, as (15) illustrates, it does not license the downward inference from (16)a. to b., as the F-L theory requires, as it is possible for (16)a. to be true while (16)b. is not; say, if at breakfast John had eaten spinach, but no-one else ate any vegetables. However, as von Fintel observes, the inference in (16) will go through once we additionally assume the presupposition of the conclusion, that John did in fact eat kale for breakfast.

Granting that von Fintel's theory of NPI licensing is required on independent grounds, I now present how he applies it to superlatives, although for expository convenience I will be illustrating with examples taken from Gajewski 2010. First of all von Fintel points out that a superlative utterance such as (17)a presupposes the truth of a corresponding utterance without the superlative modifier.

(17)   a. Fred is the oldest linguist.

      b. Fred is not the oldest linguist.

      c. Fred is a linguist.

Both (17)a and its negation (17)b give rise to the implication that (17)c is true; this pattern, as Gajewski reminds us, constitutes a standard argument for the presuppositional status of the latter[4]. Now notice that one cannot on the basis of (18)a straightforwardly infer the truth of (18)c, where the predicate *athlete* is replaced by the stronger expression *volleyball player*; for although Erin might be tallest amongst the athletes, she could fail to be tallest among the volleyball players by not being a volleyball player.

(18)   a. Erin is the tallest athlete.

      b. <u>Erin is a volleyball player.</u>   (presupposition of (18)c)

      c. → Erin is the tallest volleyball player.

The inference goes through with the additional assumption (18)b, i.e. that she is one; and this information, as we just saw, is provided by the presupposition of (18)c.

---

[4] For further arguments and discussion of the presuppositions of superlatives, see Herdan & Sharvit 2006.

For the examples given in (11)-(14), the authors of the various cited works converge on an analysis by which -*est* is interpreted within the DP which also contains the NPI. Herdan & Sharvit 2006 and Gajewski 2010, although making different assumptions about the syntax and semantics of -*est*, each demonstrate by informal proof that -*est* generally licenses SDE inferences in such configurations.

A fact which has not been highlighted in the cited works, however, is that when -*est* takes VP scope, SDE inferences are not valid. To see this, observe that (19)a. does not Strawson-Entail (19)c, as the given scenario illustrates[5].

(19) a. John read the most books.
      b. John read some travel books.
           (presupposition of (19)c)
      c. !→ John read the most travel books.

	John	Bill	Sue
books read	**15**	12	8
travel books read	3	5	**6**

Scenario verifies (19)a & (19)b; falsifies (19)c

It is easy to see that, just because John read more books than anyone else, it doesn't follow that he read more books of a particular type than anyone else. Given Hackl's observation concerning *the most NP*, we know that -*est* takes VP scope in both (19)a and (19)c; see Howard (2008) for discussion of the general invalidity of SDE inferences when -*est* takes VP scope.

If we assume then that NP-scope -*est* is generally SDE while VP-scope -*est* is not, we would expect that when a superlative-modified NP contains an NPI, -*est* is forced to take scope at NP in order to license it. This would give us an explanation of the fact that (13) for example seems to lack a VP-scope reading[6].

---

[5] Thanks go in particular to Eva Csipak for helping me design this example.

[6] However for a potential counterexample, see Gajewski 2010, wherein (i) is reported to be acceptable:

   i. #/* (/Gajewski: OK) Sue owns the most books that have <u>ever</u> been banned for profanity.

Native speakers I have consulted tend to reject (i). I suspect that native speakers who find (i) okay could, due to the conflicting requirement of the NPI to be licensed, be *in extremis* assigning (i) a parse whereby -*est* takes NP scope after all. This parse would give rise to a pragmatically anomalous meaning, violating an otherwise rather robust

(20) The dean praised the best student who has any knowledge of French.　　　(=(13))

　　　a. *"The dean praised a [student who knows French] who is better than any [student who knows French] praised by any salient alternative to the dean"　　　(*VP scope impossible*)

　　　b. "The dean praised the unique [student who knows French] who is better than any other [student who knows French]"　　　(*NP scope only*)

2.4 Section conclusion

In this section so far we have seen that, on one hand, an utterance with *the most NP* cannot have NP scope for the superlative morpheme; while on the other, given von Fintel's proposal, NPIs are expected to be not licensed within VP when *-est* takes VP scope. We are now in a position to appreciate the puzzle presented by examples like (1) and (2), which I will lay out in the next section.

3.　　Puzzles

3.1　NPIs licensed where they shouldn't be

It follows from the previous section that the configuration *the most [NP ... NPI ...]* should be ruled out; however (1) is perfect.

(1)　　John read the most books (that) anyone in the class has ever read.

Given the observation that *the most* has no NP-scope reading, we expect *-est* to take VP scope in (21)a, and in view of the findings of the previous section, we thus expect NPIs to be bad within

---

presupposition of *-est*, which makes it truth-conditionally equivalent to *Sue owns all the books that have ever been banned for profanity*; see Hackl 2009 for discussion. Investigation into the factors that influence the acceptability of utterances such as (i), I will have to leave for future research; for the purposes of the current paper, I will assume that (i) has the status reported by myself and my informants, i.e., infelicitous or downgraded in comparison to examples such as (1)&(2).

12

NP. It is therefore surprising to find that in examples such as (1) *any* and *ever* are licensed, quite naturally, in what appears to be a relative clause modifier of the NP.

Rather than taking (1) as challenging the theories either of NPIs or of superlatives sketched in the preceding section, I propose that the treatment of the embedded clauses as RC modifiers is at fault. Coming up there will be some arguments to the effect that it is incorrect to analyse these embedded clauses as such.

3.2 "Mismatch" effects

But first, notice that the constructions in (21), roughly similar to those in (1) and (2), present further puzzles when we consider the minimally-contrasting variants in (22).:

(21) a. Bill wrote the most poems that anyone ever wrote.

   b. Mary sang the loudest that any soprano ever sang.

(22) a. *Bill wrote the most poems that anyone (ever) published.

   b. #Mary sang the loudest that any baritone (ever) sang.

Replacing *wrote* in the embedded clause of (21)a. by *published* renders (22)a. unacceptable. In (22)b., the subject of the embedded clause of (21)b. *any soprano* has been replaced by *any baritone*, giving rise to the odd inference that the matrix subject Mary is a baritone, which suggests that she is male. It's not immediately obvious why this inference would come about, nor why (21)a. & (22)a. would differ in acceptability under the minimal substitution.

Taking the two contrasting pairs from (21) and (22) together, we can think of the anomalous status of the (22) examples as resulting from a mismatch between the matrix clause and the embedded clause in each case: in (21)a. the main verbs are identical, but not in (22)a.; and in (22)b. the subject *Mary* doesn't "match" that of the embedded clause, unless she is a baritone. This is the line of explanation I will pursue: that some property of the constructions in (21) and (22) requires the matrix clause and the embedded clause to stand in an appropriate correspondence relationship.

13

Whatever is behind the downgrading/infelicity effects in (22), it must be a property of superlative constructions; minimally changing the examples in (22) to replace the superlative elements gives the perfect (23).

(23)   a. Bill wrote no poems that anyone ever published.

   b. Mary sang louder than any baritone ever sang.

In (23)a, *the most* has been replaced by the downward-entailing quantifier *no*, which licenses the NPIs unproblematically. In (23)b, *the loudest* is replaced by the comparative form *louder*, and NPIs are well-known (since at least Hoeksema 1986, according to von Fintel 1999) to be licensed in the *than*-clause of comparative constructions. With the requirements of *any* and *ever* satisfied, the examples in (23) are otherwise rather banal, and it is unsurprising to see that the effects in (22) do not resurface. Thus the superlative is at least partially responsible for the badness of (22).

The NPIs must also be present for these effects to appear, as can be seen by comparing (22) to the examples in (24).

(24)   a. Bill wrote the most poems that one of the magazines published.

   b. Mary sang the most songs that the baritone (also) sang.

Appropriate contexts for the examples in (24) would be competition scenarios: *"Bill/Mary won a prize because... "*. (24)a. is a variant on (22)a. that contains no NPIs, the NPI indefinite subject of the embedded clause replaced by the plain partitive indefinite *one of the magazines*, and presents a sharp improvement. In (24)b. the embedded clause has a definite *the baritone*, and does not trigger the anomalous implication that Mary is male, thus illustrating that this "matching" effect is not observed in the absence of an NPI. Thus the badness of the examples is (22) is somehow forced by the co-occurrence of the superlative and NPIs in the embedded clause.

Finally, note that the effects are not forced when the superlative is one that allows an NP scope construal, as discussed in Section 2; *longest poem* and *oldest song* don't give rise to the effects observed with (22):

(25) a. Bill wrote the longest poem that anyone (ever) published.

b. Mary sang the oldest song that any baritone (ever) sang.

So the mismatch effects observed in (22) are triggered only in superlative constructions when the embedded clause contains NPIs and the superlative must take VP scope.

3.3 Trivial Contribution

In this subsection I sketch some arguments that the embedded clauses in (21) and (22) should not be analysed as relative clause modifiers. First of all, note that in contrast to these examples, the embedded clause in (23)a. certainly should be treated as a restrictive RC. It should not be controversial that checking the truth conditions of (23a) requires examining the intersection of the set of things Bill wrote, the set of things that are poems, and crucially, the set things that were published by a magazine; the last restriction then being the contribution of the – genuine – restrictive relative clause. Equally, determining the truth of (24)a. depends on inspecting the same set intersection, this time to in order to compare its cardinality with those of sets of published poems written by Bill's competitors; (24)a. then turns out to be true just in case none of those sets is as populous as the former.

Thus in the case of both (23)a. and (24)a., we want to analyse the NP denotation as restricted by the relevant relative clause modifier[7]. What are the grounds for giving (21)a. and (22)a. a distinct treatment? Observe that, under the uniform treatment, the paraphrase just given for (24)a. would apply approximately well to (22)a.; thus, both should mean the same thing –

---

[7] Note also that (24) and (23)a. can be modified by substituting *which* for *that* without affecting the interpretation:

  (i) *cf (23)a.*   Bill wrote no poems which/that anyone ever published.

  (ii) *cf (24)a.*   Bill wrote the most poems which/that one of the magazines published.

  (iii) *cf (24)b.*   Mary sang the most songs which/that the baritone sang.

This in contrast to the examples in (21), which are not acceptable under the substitution:

  (iv) *cf (21)a.*   *Bill wrote the most poems which anyone ever wrote.

  (v) *cf (21)b.*   *Mary sang the most songs which any soprano ever sang.

This differential behaviour with respect to *which* provides a further argument for rejecting the relative clause modifier hypothesis for (21). See footnote 17 for further discussion.

contrary to fact, since the latter is apparently uninterpretable. I take the unexpected "mismatch" effects to suggest strongly that something else is going on here.

Furthermore, consider what contribution the embedded clause would be making in (21)a. under a parse where it is interpreted as an RC, as was pointed out to me by B. Schwarz (p.c.):

(21)a. Bill wrote the most poems that anyone ever wrote.

If the embedded clause is interpreted as a relative clause, then its adjunction to the nominal predicate *poems* restricts the extension of the NP to include only poems that have been written. (The utterance would then be the claim that Bill wrote more of *those* poems than anybody else did.) But assuming that the predicate *poems* is only true of entities that have already been written, it's not clear what additional information such a relative clause could be contributing here. The extension of the modified predicate is bound to be identical to that of the unmodified NP, i.e. $[\![poems]\!]$ = $[\![poems\ that\ anyone\ ever\ wrote]\!]$; the addition of the RC makes no further restriction[8]. Thus on the restrictive RC analysis of the embedded clause in an example such as (21)a, we would anticipate the contribution to truth conditions of the embedded clause to be as trivial as in (26), which seems to be rather odd in comparison.

(26)   ?#Bill wrote every poem that anyone ever wrote.

Finally, the question alluded to in the introduction about (2) still remains: what is the attachment site of the embedded clause in a structure which has no obvious nominal for a RC to modify?[9]

(2)   Mary sang the loudest (that) anyone in the group has ever sung.

---

[8] Furthermore, consider the following modification to (21)a., suggested to me by M. Hackl (p.c.):

(i)   Bill *will* write the most poems that anyone ever wrote.

Since *write* is a verb of creation, surely the objects its internal argument picks out do not exist at the time of utterance. The putative RC, on the other hand, has past tense and would pick out entities that have already been written at the utterance time. Following this reasoning, the denotation of the putative NP *[poems that anyone ever wrote]* under the RC modifier analysis arguably describes the empty set. (i) certainly doesn't have this anomalous meaning, although I don't have anything to say about the past tense in the embedded clause.

[9] See Matushansky 2008 for discussion of the viability of analysing adverbial superlatives as containing a null N.

16

Independently of any claim about the role these embedded clauses *do* play, it should be clear that an analysis of them as restrictive relative clauses is highly suspect.


3.4 Section conclusion

In fact I have a suggestion to make about what the embedded clauses are doing in these superlative constructions, but before I set out my proposal, let us take stock of what we have learned and what an analysis of these constructions has to achieve. An analysis has to account for the NPI licensing puzzle; i.e. how VP-scope *-est* can license NPIs despite the fact that VP-scope *-est* apparently lacks the necessary SDE property. An analysis also must account for the unexpected infelicity of the examples in (22), which I suggested can be characterised as a "mismatch" between the material in the matrix clause and the embedded clause. And the analysis needs to say something about what contribution to meaning the embedded clause makes.

4. Superlative degree clauses

4.1 Comparison with comparatives

I have endeavoured up to this point to make all my observations independent of a specific proposal about the lexical semantics of the superlative. In this section I will make some more concrete assumptions, which will allow me to present some more pertinent facts about these embedded clause superlative constructions, and eventually motivate the analysis I want to propose.

For the moment I will just advertise in broad-strokes the shape that my proposal will take: I will propose to account for the tight relationship between the embedded clause and the matrix by analysing both as arguments of the superlative morpheme. Thus the LF of e.g. *Mary sang the loudest that anyone ever sang* will be something like this:

est

that anyone ever sang _

Mary sang d-loud
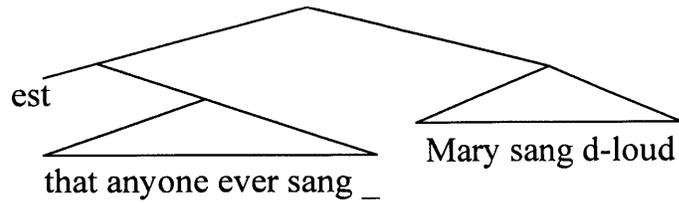
(27)

According to (27) then, *-est* first combines with the embedded clause, and then with the matrix clause denotation. This structure is familiar from common analyses of other degree constructions, most pertinently *-er... than...*, and *as... as...* , exemplified in (28); the schematic structure I have in mind is in (29).

(28) a. John read more books than there are in the library.

      b. John read as many books as Mary did.

DEG- OP

than / as ... <DEG EXPRESSION> ...

John read d-many books
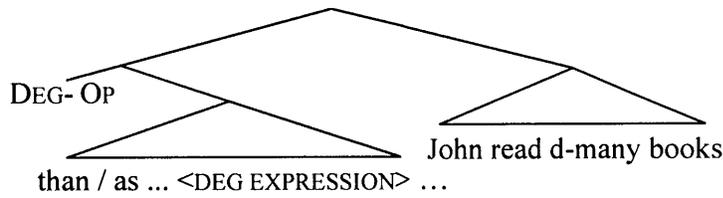
(29)

Adopting a terminology of which the first mention I have found is in Reed & Kellogg 1900, I will be using the expression "degree clause" for such embedded clauses headed by *than* or *as* in degree constructions. The analysis according to which the degree clause forms, at the relevant level of representation, a constituent with the degree morpheme dates back to at least Bresnan 1973. The assumption that the constituent thus formed is a quantificational phrase invites a semantic analysis that treats the degree clause and the matrix clause as denoting properties of degrees and the degree morpheme as an operator on them.

Of such semantic analyses there is an abundant choice; the one in (30) is due to Seuren 1973.

(30)   Let P,Q be degree properties. Then, [[er]](Q)(P) iff ∃d[P(d) & ¬ Q(d)]

The LF in (31) is an example of how the schema in (29) applies to one of the examples in (28). The degree morpheme *-er*'s second argument is supplied by the matrix clause, its first by the

18

than-clause. Note that the latter is assumed to contain an elided gradable expression *many* which matches that in the matrix clause.

(31)   [er [λd. λw. there are d-many books in the library in w]] [λd'. λw'. John read d'-many books in w']

According to the denotation of the comparative degree morpheme *-er* in (30), (28)a's truth conditions can be paraphrased as, "there's a degree d such that John read d-many books and there aren't d-many books in the library". For these truth conditions to work we make the common assumption that the appropriate degree properties are downward monotone (Gawron 1995), i.e. if a degree d has P then every degree less than d has P. This can be made to follow from an appropriate lexical semantics for gradable expressions; e.g. Hackl's (2009) denotation for *many* in (8) has this property.

While *-er/more* (and *as*) says something about the relationship between its two arguments, *-est* can be seen as making a universal claim. Thus the denotation that Heim 1999 proposes for *-est*, given in (136), can been seen as a generalisation of Seuren's proposal for *-er*[10] to incorporate universal quantification over a relevant domain.

(32)   Let P be a degree property and A be a set of such properties. Then,

[[est]](A)(P) is defined only if $P \in A$ & $\exists Q[Q \in A$ & $P \neq Q]$

If defined, [[est]](A)(P) iff $\exists d[P(d)$ & $\forall Q \in A[Q \neq P \rightarrow \neg Q(d)]]$

According to (32), *-est*(A)(P) is true just in case there is a degree that has property P[11] and has no other property in A. Due to monotonicity, this implies that P is "bigger" than all its distinct alternatives; the set of degrees that have P must be a proper superset of the set of degrees that have Q, for each Q in A distinct from P.

---

[10] Thanks to Irene Heim p.c. for making this suggestion.

[11] I'm using the expression "has the property P" to stand for "is mapped by P to a function that maps the world of evaluation to True".

Seen from the perspective of the preceding discussion it makes sense – to skip ahead to the denouement – to assume that *-est*'s domain set argument is provided by its sister in (27), the embedded clause *that anyone ever sang*. If correct, this assumption promises to bring the syntax of superlatives in line with comparatives and equatives and could be favoured over analyses that assign these degree constructions disparate treatments. However before leaping to the conclusion that (27) is appropriate and desirable, it behoves me to justify the choice on the grounds that it delivers the correct meanings for the utterances it purports to characterise. This task will require getting clearer on what those truth conditions are, and that will involve a detailed explication of the phenomena Heim set out to capture with the proposal of (32), that of its focus-sensitivity, the discussion of which I have hopefully succeeded in putting off in a natural-seeming way until the next subsection.

I should also point out at this point that Heim's denotation in (32) – devised such as to make the superlative sensitive to focus, along the lines of Rooth's 1992 analyses of other phenomena – is not the one which most who are familiar with the superlatives literature will associate with Heim 1999. This paper is typically invoked for the analysis according to which *-est* is an operator over degree functions of type $\langle d, et \rangle$, which she discusses in the principle part of her manuscript.

Finally, the assumption of the structure in (27) will deal with the *desiderata* listed at the end of the previous section, as I will show that the putative "matching" constraint is a consequence of conditions that *-est* places on its arguments, and it will fall out naturally, given the semantic analysis that I assume, that the NPIs are licensed in an SDE environment.

4.2 *est* as a Focus-sensitive operator: Heim 1999

Consider the following scenario:

> Mary is a member of a large amateur musical theatre group, and is so desperate to get a part in the next production that she is auditioning for multiple roles. She thinks that a good way to be noticed in the various group auditions which are being held over the course of the day is to

sing loudly, but she is also aware that after a certain level of intensity, her ability to sing accurately and hold a note deteriorates the louder she sings.

Given this context, I think it is clear that example (33) is ambiguous:

(33)   Mary sang the loudest at eleven am.

Under one construal, (33) is true if, at the 11am audition, Mary sang louder than all the other performers at that audition. On another reading, (33) is true if Mary sang louder at 11am than she did at any other time that day. Intonation can help disambiguate (33); (34)a – where *Mary* is stressed – can favour the first reading, in which Mary is compared to other singers, while (34)b – where *11am* is stressed – can make salient the reading where the intensity of Mary's singing at 11am is compared to her singing at various other times.

(34)   a. MARY sang the loudest at eleven am.
          b. Mary sang the loudest at ELEVEN AM.

Heim (1999) uses the observation that prosody can help disambiguate superlative utterances to motivate an analysis according to which *-est* is a focus-sensitive operator. She proposes the LF in (35) for (34)a (and the equivalent reading of (33)); the subscript "F" indicates focus-marking on the constituent [*Mary*].

(35)   a. [est-A] [[$\lambda_{1,d}$ [Mary]$_F$ sang t$_{1,d}$-loud at eleven am]~A]

          b.
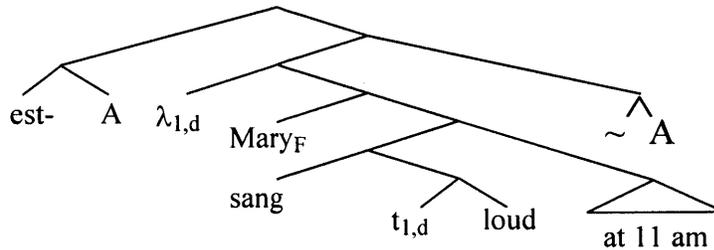
As we saw in the denotation for *-est* in (32), Heim assumes that *-est* takes two arguments in the form of a degree property P and an alternative set of degree properties A. In (35), *-est*'s sister *A* is a silent pronoun denoting such a set of degree properties, the value of which is fixed by the context, while the denotation of the rest of the utterance supplies *-est*'s P argument. The

contribution of the right-adjoined constituent [~A] I will return to in a moment, contenting myself for now to mention that it is assumed to denote a function that returns the value of its input, and we can safely ignore it temporarily. The constituent [est-A] is base-generated as the sister of the gradable adverb *loud*, and moves to take sentential scope, introducing a lambda operator binding its trace. This trace is then interpreted as a variable over degrees, which saturates locally the degree argument of the gradable adverb. I assume that the combination of *d-loud* with a verb phrase meaning returns the latter with the additional condition that the event it describes have a loudness of at least d; the constituent in (35) headed by the lambda abstractor then denotes the property that a degree d has if Mary sang at least as loud as d at 11am.

Finally, it remains to determine what the composition of the alternative set A is – bearing in mind that the truth conditions to be delivered by (35) depend on comparing Mary to other singers. Heim applies the machinery proposed in Rooth 1992 that allows focus to play a role in signalling the choice of the alternative set A, via the focus interpretation operator '~'. Rooth's squiggle operator takes two arguments, and its contribution is to introduce the presupposition that its first argument, denoted by a pronoun whose value is supplied by context, is a subset of the focus-semantic value of its second; as mentioned previously it returns the truth-conditional component of its second argument unchanged. In (35), this second argument is a degree property and its focus semantic value is a set of degree properties; generally the focus semantic value of an expression ϕ is the set of objects that is obtained by replacing the focused elements within ϕ with expressions of the same type as the focused element. In the case of (35), where *Mary* is focus-marked, the focus-semantic value of [~A]'s sister can be written as in (36)a, a form of set notation, or in terms of the characteristic function of this set, as in (36)b.

(36)   Focus semantic value of *[$\lambda_{1,d}$ [Mary]$_F$ sang $t_{1,d}$-loud at eleven am]* :

    a. { $\lambda d.\lambda w.$ x sang d-loud at 11am in w | x $\in$ D$_e$ }

    b. $\lambda P. \exists x \in D_e$ [P = $\lambda d.\lambda w.$ x sang d-loud at 11am in w]

A key element to notice about the machinery deployed here is that the set is obtained by substituting an existentially-bound variable in for *Mary*; this point will be useful in the analysis I will develop in the rest of this paper.

22

The squiggle operator's first argument in (35) is thus required, on pain of presupposition failure, to be a subset of (36), i.e. a set of degree properties. As mentioned above, *-est*'s contextually-supplied first argument A is also required, by the denotation in (32), to be a set of degree properties; thus *-est*'s A argument would be a salient choice for squiggle's first argument, conveniently also named A in (35), as in this representation they are identical[12].

Since A is only a subset of (36), the intuitive truth conditions of (34)a are only predicted to depend on a salient subset of the domain of individuals, specifically those members of Mary's theatre group who were auditioning with her, not every individual in D_e who happened to have been singing at 11am on the day in question. This effect of context on the value of *-est*'s A argument can be modelled by restricting the domain of the existential to a set of contextually-salient individuals C[13].

(37)   Appropriate value for A in (35):

   a. $\{ \lambda d.\lambda w. \ x$ sang d-loud at 11am in w $\mid x \in D_e \cap C\}$

   b. $\lambda P. \ \exists x \in D_e \cap C \ [\ P = \lambda d.\lambda w. \ x$ sang d-loud at 11am in w $]$

As can be easily seen, the set in (37) is bound to be a subset of (36), since both are sets of the same kind of objects, but (37) has more restrictive conditions on membership than (36); this point will also be crucial later on.

With all these pieces in place, we can now see how assuming the LF in (35), the denotation for *-est* in (32) and the value for A in (37) derives the intuitive truth conditions for (34)a. (and the relevant reading of (33)). Suppose the only people who participated in the 11am audition were Mary, Lee and Kim. We set C = {Mary, Lee, Kim}, and A = { $\lambda d.\lambda w.$ Mary sang d-loud at 11am in w, $\lambda d.\lambda w.$ Lee sang d-loud at 11am in w, $\lambda d.\lambda w.$ Kim sang d-loud at 11am in w }. Now

---

[12] However it is important to recognise that there is nothing necessary about *-est* and squiggle's arguments being coindexed/identical, and there is ample discussion in the literature of cases where *-est* fails to associate with a focus in its scope. Thanks to Irene Heim (p.c.) for making me aware of my misunderstanding of this point in an earlier draft, and to von Fintel (1994) for forcing the point home about what the predictions of Rooth's system actually are.

[13] To improve clarity throughout the rest of the paper I will omit C in representations where it isn't crucial.

suppose that the maximum intensity of Mary's singing was 100dB at the 11am audition, whereas Kim's and Lee's only reached 90dB and 95dB respectively (further suppose that these are realistic figures for human singing intensities). Then any degree d' corresponding to an intensity i such that 95dB $< i \leq$ 100dB will verify (35), since d' has the property that Mary sang d'-loud at 11am, but has neither of the other properties in A. If however Lee had also sung at 100dB, then there would be no such degree, as any degree d" that has the property that Mary sang d"-loud also has the property that Lee sang d"-loud, and so (35) returns False[14].

The last example highlights an important aspect of Heim's proposal, which is that -*est* quantifies over degree properties, i.e. not sets of degrees, but intensions thereof[15]. Suppose that the denotation for -*est* were the following:

(38)　Let f be a degree function of type $\langle d,t \rangle$ and A be a set of such functions. Then,

$[\![$ est $]\!](A)(f)$ is defined only if $f \in A$ & $\exists f'[f' \in A$ & $f \neq f']$

If defined, $[\![$ est $]\!](A)(f) = 1$ iff $\exists d[f(d) = 1$ & $\forall f' \in A[f' \neq f \rightarrow f'(d) = 0]]$

Then in the second version of the scenario just given, where Mary and Lee both sang at 100dB, (35) would return True, contrary to intuitions. If Mary and Lee sang at exactly the same intensity, then the sets $\lambda d.$ *Mary sang d-loud* and $\lambda d.$ *Lee sang d-loud* are identical, including all and only the degrees corresponding to intensities up to and including 100dB. Since these two sets are identical, the only $f' \in A$ such that $f \neq f'$ is $\lambda d.$ *Kim sang d-loud at 11am*; and since this set is a proper subset of -*est*'s f argument, (35) maps to True. Heim's actual proposal does not make

---

[14] This might not actually be exactly the correct prediction, since some speakers report the judgement of the utterance with respect to this scenario as not false but infelicitous. Further investigation is required here; the point remains that, as discussed in the following paragraph, the degree property analysis avoids the incorrect prediction that (34)a would be judged True in the scenario.

[15] This is a rather sloppy way of talking. The intension of a degree function, according to the system laid out in Heim & Kratzer (1998, ch.12), would be a function of type $\langle s,dt \rangle$, whereas the functions I am using here are Shönfinkelised the other way round, in that they are of type $\langle d,st \rangle$. This assumption, of little or no consequence to the discussion here, that it is last argument of a given intensional expression which is of type s will simplify matters somewhat when I get down to the compositional nitty-gritty later on.

24

this incorrect prediction, since the only way for the degree <u>properties</u> $\lambda d.\lambda w$. *Mary sang d-loud in w* and $\lambda d.\lambda w$. *Lee sang d-loud in w* to be identical is if that $\forall w \forall d$, Mary sang d-loud in w iff Lee sang d-loud in w, or in other words that it is a <u>logical necessity</u> that Mary and Lee sang as loud as each other; this condition holds only if Mary and Lee are identical.

Let us now briefly consider the derivation of the truth conditions of the other reading of (33), (34)b with focus on *eleven am*. Suppose that there were three auditions that Mary participated in: one held at 10am, one at 11am, and one at 1pm. Then we intuitively agree to the utterance in (34)b if for example Mary sang at 95dB at the second audition, but only reached 90dB at the other two; on the other hand, if Mary had also reached 95dB at the 10am audition, we reject (34)b.

Applying the interpretive procedures developed so far, we assign the LF and value for A given in (39).

(39) [est-A] [[$\lambda_{1,d}$ Mary sang $t_{1,d}$-loud at [eleven am]$_F$]~A]

　　A = { $\lambda d.\lambda w$. Mary sang d-loud at t in w | t $\in$ $D_i \cap C$ }

Given the denotation in (32), this LF predicts that (34)b is judged true only if the intensity of Mary's singing at 11am reached a degree that it did not reach at any other time under consideration. These alternative times are indirectly quantified over through the makeup of A, which is restricted here by intersecting the domain of time intervals with the set of contextually salient things C. Just as in evaluating (34)a we only consider a limited subset of the domain of individuals, the question of how loudly Mary sang a week or a month previous is not relevant in evaluating (34)b; so C must at least exclude intervals that are not part of the day of the auditions[16].

---

[16] I will also assume that overlapping time intervals are excluded from C, or at least those that include the interval mentioned in the utterance property. Suppose that C contains *t'*, an interval that includes the 11am-interval. Then *A* additionally contains the property $\lambda d.\lambda w$. *Mary sang d-loud at t' in w*; thus the degree corresponding to 95dB also has this property in the scenario described, and (34)b is incorrectly predicted to map to false, since there is no degree

4.3 Deriving the presuppositions of superlative utterances

Heim's proposed denotation, repeated in more articulated form in (40), includes the definedness conditions in (a): (i) that -*est*'s P argument be a member of the alternative set A, and (ii) that there be at least one other member of A distinct from P.

(40) Let P be a degree property and A be a set of such properties. Then,

    a. ⟦ est ⟧(A)(P) is defined only if   (i) $P \in A$;

                                                   (ii) $\exists Q[Q \in A \ \& \ P \neq Q]$.

    If defined,

    b. ⟦ est ⟧(A)(P) iff $\exists d[P(d) \ \& \ \forall Q \in A[Q \neq P \rightarrow \neg Q(d)]]$

Since A is a pronoun whose value is supplied by the context, these restrictions on its makeup serve to derive the natural inference one makes on hearing an utterance such as (41)a., that Mary and at least one other individual are relevant in determining who sang the loudest at 11am; in the notation used here, this means that Mary and at least one other are in the set of contextually-salient individuals C. For reference and illustration, the values for the example used in the previous subsection are listed in (41).

(41)   a. Mary sang the loudest at eleven am.

      b. [est-A] $\lambda_{1,d}$ [Mary sang $t_{1,d}$-loud at eleven am]

      c. A = { $\lambda d.\lambda w.$ x sang d-loud at 11am in w | $x \in C$ }

      d. C = { Mary, Lee, Kim }

Scholars of Rooth 1992 might have noticed that there is potentially some redundancy in Heim's system for deriving the meaning of superlative utterances, due to an overlap between the definedness conditions in (40)a and those that Rooth assumes for ~. Squiggle also takes two arguments P,A and introduces the presuppositions that both P and some Q distinct from P are both elements of A. Given the adoption of the generally applicable squiggle operator to interpret

---

that has the utterance property and no other property in *A*. Hackl 2009, although assuming a quite different semantics for -*est*, relies on a similar stipulation in order to derive the proportional reading of *most NP*.

focus in superlative utterances, one might speculate as to whether it would not be preferable to leave ~ to do all the work of policing the content of A as related to P, and simplify the denotation of -*est* by stripping its definedness conditions.

Irene Heim (p.c.) suggests a way to show that the requirement (40)a(i) – that $P \in A$ – is still enforced in the absence of an appropriate focus for squiggle to interpret. When it appears within the context of (42)'s sequence of utterances, (41)a naturally receives a interpretation corresponding to the LF and value for A in (41)b & c, but additionally – and importantly – gives rise to the inference that Mary is a soprano:

(42)   All the sopranos that auditioned were impressive. But Mary sang the LOUDest at 11am.

This example, based on Heim's ex. 9 (1999, p.6), is pronounced with the highest prominence on the superlative adverb *loudest*, as indicated by capitalisation of the portion corresponding to its stressed syllable. Presumably, given the prominence on the adverb, there is no focus on the subject *Mary* for a [~A] coindexed with -*est*'s domain argument to interpret (Heim 1999, p.6), and a focus on the adverb does not affect the interpretation of the superlative. So we can't rely on squiggle to enforce the membership in A of [est-A]'s sister property in (42).

If we can find evidence for this requirement nonetheless being active in the interpretation of (42), then we have a kind of argument for granting Heim's definedness condition (40)a(i). Some indirect evidence of this kind comes from the inference about Mary being a soprano. First of all, we can see that this inference has the status of a presupposition based on the fact that (43)a but not (43)b forms a felicitous response to (42).

(43)   a. Hey, wait a minute! I didn't know that Mary was a soprano.

b. #Hey, wait a minute! I didn't know that Mary sang louder than the others.

The *Hey, wait a minute!* test (von Fintel 2004) is designed to pick out presupposed from asserted content. Whereas (43)a can be used to mount a legitimate challenge to (42) by questioning its speaker's taking for granted that Mary is a soprano, (43)b shows in contrast that the corresponding move cannot be felicitously made targeting the asserted component of the

27

superlative utterance in (42). This contrast indicates that it is indeed a presupposition of (42) that Mary is a soprano.

The presuppositional effect can be modelled in terms of the notation adopted here by allowing C in (41)c to take its value from the referent introduced by the expression *the sopranos that auditioned* used in the first utterance in (42); we can say that the sequence somehow makes salient the singing properties of the members of this set:

(44)   C = {x | x is a member of the set contextually salient sopranos that auditioned}

Once (44) goes in the value for C in (41)c, the truth conditional component of *-est*'s meaning in (40)b guarantees that (42) depends on comparing Mary to sopranos. However, it doesn't follow that Mary herself is a soprano without the additional assumption of the definedness condition (40)a(i) that the sister of [est-A] be a member of A.

As defined in (41)c, A is the set of properties that a degree has just in case x sang d-loud at 11am, where $x \in C$ as defined in (44). I assume the LF in (41)b for the 2nd utterance in (42); let P' be the property denoted by the sister of [est-A] in (41)b, i.e. P'= [$\lambda d.\lambda w$. Mary sang d-loud at eleven am in w]. According to the condition in (40)b, the superlative utterance in (42) is then true just in case there is a degree of loudness of Mary's singing that has no property in A, which means that so soprano sang that loudly. But nothing follows from that about Mary's range.

However, it does follow that Mary is a soprano once we assume (40)a(i). For P' $\in$ A amounts to there being a Q $\in$ A such that Q=P'; for any Q, Q=P' just in case for worlds w and degrees d, Q(d)(w) = P'(d)(w). Given the conditions on membership of A and the identity of P', if P'=Q for some Q $\in$ A, then it follows that $\forall w \forall d$, Mary sang d-loud at 11am in w iff x sang d-loud at 11am in w, for some x $\in$ C. This means that it is necessarily the case that x and Mary sang exactly as loud as each other at 11am, because there are no worlds that disagree on it. This condition can, I submit, only be satisfied if x is identical to Mary. And since x $\in$ C, x is a soprano, therefore Mary is a soprano.

The result of the *H!WaM* test is predicted if (41) only has a truth value when Mary is a soprano, and given the above reasoning, this result follows from the assumption of (40)a(i).

Although the argument for adopting it is indirect, in that I have not attempted to rule out any alternative explanation, it is useful to have found some evidence for *-est*'s definedness condition that is independent of the superlative degree clause construction, as the analysis I will propose of the puzzles this construction presented in subsection 3.2 will depend crucially on the assumption of (40)a(i). The reasoning that obtains the result of predicting (42)'s presupposition will also come in useful, as the logic of the proposed explanation of the puzzling data will be the same.

We still have some loose strands to tie up however, as (40) does not suffice to predict all the presuppositions of superlative utterances discussed so far. With respect to an example such as (41), it leaves open the question of whether Mary or anybody else actually sang at 11am: neither P nor Q are required by (40) to be <u>nonempty</u> degree properties. Recall from the discussion in subsection 2.3 that an example such as *Fred is the oldest linguist* is known to presuppose that Fred is a linguist; in the same way, it should be uncontroversial that an utterance such as (41)a presupposes that Mary sang. We can generate this presupposition if we add to (40) the condition that there be at least one degree with the property that Fred is that old.

Similarly, Hackl 2009 suggests the odd or ironic nature of (45) when uttered by a person who has no more than one mother is evidence that *-est* presupposes that its domain of alternatives is nontrivial.

(45)   #You are the best mother I have.   (Hackl 2009, attributed to Danny Fox)

We can derive the relevant presupposition with the condition, already part of (40), that there be at least one alternative of the form $\lambda d.$ *x is a d-good mother of the speaker* such that x ≠ the addressee, once we additionally require that there be at least one degree that has this property.

Putting these additional requirements together with (40), we obtain the following, slightly more complex version of Heim's proposal (see also Gajewski 2010).

(46)   Let P be a degree property and A be a set of such properties. Then,

     a. ⟦ est ⟧(A)(P) is defined only if   (i) $P \in A$;

                                       (ii) <u>$\exists d[P(d)]$</u> ;

                                       (iii) $\exists Q[Q \in A \ \& \ P \neq Q \ \& \ \underline{\exists d'[Q(d')]}$].

If defined,

b. ⟦ est ⟧(A)(P) iff $\exists d[P(d) \ \& \ \forall Q \in A[\ Q \neq P \rightarrow \neg Q(d) \ ]]$

Although (46) is the final version I am assuming throughout the rest of the paper, I will be omitting the existential conditions where they are not relevant, and employing the more streamlined version in (40) as a shorthand.

Summarising this section so far, we have seen how Heim's (1999) analysis of the superlative morpheme accounts for the disambiguating effect of focus on superlative utterances, and how its definedness conditions derive the presuppositions of superlative utterances. Along the way I made some remarks about the structure of the set of alternatives that *-est* quantifies over, which will become important in my analysis of the superlative embedded clause construction.

4.4   The superlative degree clause fixes the domain of quantification

We have seen how the string in (47) can be used to make a claim about Mary's singing compared to other singers and to other times, and how the analysis of Heim 1999 can capture these different meanings.

(47)   Mary sang the loudest at 11am.

With an analysis of these simpler cases as in hand, let us turn now to the analysis of the superlative embedded clause construction. As a first observation, consider how altering the content of the embedded clause in (2), repeated here as (48)a, affects the interpretation of its variants in (48)b & c.

(48)   Mary sang the loudest that …   a. anyone in the group has ever sung.

                                       b. she has ever sung.

                                       c. anyone in the group sang (that day).

The truth conditions of (48)a (= (2)) depend on comparing the loudness of Mary's singing not only to her co-auditioners, but to all members of the group at all past times. In (48)b however,

30

where *anyone in the group* is replaced by *she*, the truth conditions depend on Mary's singing on audition day to her own singing on other days. Significantly though, the truth of (48)b doesn't depend on the loudness of other singers; Kim and Lee could both have sung louder than Mary all day, and (48)b would still be judged true, just as long as Mary beat her own personal record for loudness. On the other hand, the truth conditions of (48)c, without *ever* but where the NPI subject is retained, depend on comparing Mary to other singers, but here Mary doesn't need to outdo historical performances.

The contrast in meaning between the examples in (48) then is strongly suggestive of a role for the embedded clause in setting the domain of comparison for *-est*. When the subject of the embedded clause is the NPI, *-est* compares the matrix subject to other individuals, but not when the subject of the embedded clause is the plain pronoun. And when *ever* is present in the embedded clause, other times enter into the comparison, which become irrelevant when *ever* is omitted.

Another pair of examples which illustrates the same effect, but perhaps even more strikingly, is to be found in (49) where a *the most*-NP is one of two internal arguments of the matrix clause. This construction permits the comparison of two different positions for the same NPI – *anyone* – in one structure.

(49) (Over the course of his lifetime...)
      a. Bill sent Chris the most letters that anyone sent him.
      b. Bill sent Chris the most letters that he sent anyone.

At issue here is the number of letters that Bill sent to Chris relative to other amounts of letters sent. However, in (49)a., the relevant amounts are those that other individuals sent to Chris, while in (49)b, we're concerned with the amounts Bill sent to other recipients. So in the first case, where *anyone* is the embedded subject, the matrix subject *Bill* is compared, while in the latter, *anyone* is the embedded dative object and the corresponding matrix argument *Chris* is compared.

The effect on meaning of the position of *anyone* in (49) is echoed by the parallel examples in (50), where the position of focus in the matrix clause has a corresponding effect. (50)a, with focus on the subject favours an interpretation where Bill is compared to senders, while focus on the indirect object in (50)b favours comparison of Chris to other send-ees.
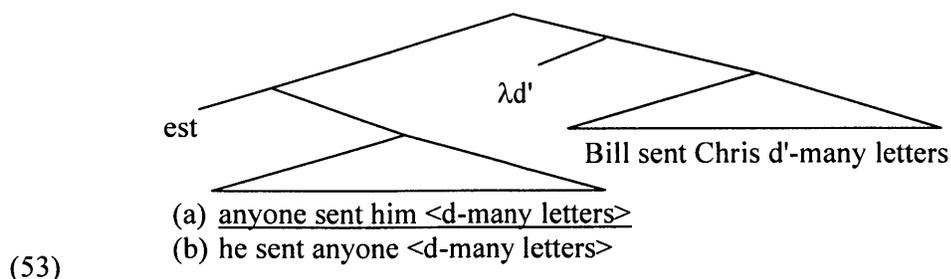
(50) a. BILL sent Chris the most letters.

b. Bill sent CHRIS the most letters.

According to the analysis in 4.2, we model the effect of focus in (50) by means of assuming the LF in (51), with focus marking either on a. *Bill* or b. *Chris*, and the values for A in (52)a & b respectively.

(51) [est-A] [[$\lambda_{1,d}$ Bill$_{\{a. F\}}$ sent Chris$_{\{b. F\}}$ t$_{1,d}$-many letters]~A]

(52) a. A = {$\lambda d.\lambda w$. x sent Chris d-many letters in w | x$\in D_e \cap C$ }

b. A = {$\lambda d.\lambda w$. Bill sent x d-many letters in w | x$\in D_e \cap C$ }

The presupposition introduced by [~A] in (51)a makes (52)a an appropriate choice for -est's co-indexed alternative set argument, as it is a subset of the focus value of [~A]'s sister. Since the sister of [est-A] denotes the property of degrees $\lambda d.\lambda w$. *Bill sent Chris d-many letters in w*, (51)a comes out as True, as desired, just in case this property is the biggest one in A, i.e. no-one sent more letters to Chris than Bill sent to Chris. And since [est-A]'s sister in (51)b doesn't denote a superset of (52)a, the latter can't co-refer with -est's alternative set argument, whereas choosing (52)b in this case gives the right result. But this is exactly the analysis we need for (49).

My proposal to account for these effects, and ultimately the various observations described in the preceding, is to analyse the embedded clause in as denoting -*est*'s restrictor, the alternative set of degree properties *A*. Thus while the matrix clause in (49) is assumed to denote the same degree property as [est-A]'s argument in (51), the embedded clauses are to be interpreted as denoting the respective sets of degree properties in (52); the syntactic form at the appropriate level of representation for (49), with the embedded clause integrated as the sister of -*est*, is sketched in (53).

λd'

est

Bill sent Chris d'-many letters

(a) <u>anyone sent him <d-many letters></u>
(b) he sent anyone <d-many letters>

(53)

The structure in (53) shows a little more of the relevant detail than that in (27), which I initially suggested in 4.1 for its similarity to the LFs assumed for comparative and equative constructions. I will henceforth refer to the clausal constituent in superlative embedded clause constructions as *-est*'s degree clause argument, in analogy to the *than-* or *as*-clauses that (53) evokes. Continuing the analogy, I assume that the degree clause contains an elided constituent containing the same gradable expression as the matrix clause[17]. The denotation for *-er* from (30) was such that it required its *than*-clause to denote a property of degrees, whereas Heim's generalisation of this analysis to *-est* requires a set of such properties. I propose that the NPI indefinites participate in the creation of a nontrivial set in *-est*'s complement by introducing alternatives, in a way more than a little reminiscent of the role of *wh*-words play in a Hamblin semantics for questions, or Kratzer & Shimoyama's (2002) treatment of *irgendein* indefinites.

---

[17] A potentially even better analogy for a syntactico-semantic analysis can be made with the *amount relatives* of Carlson 1977 (ARs; also known as *degree relatives*: Heim 1987; Grosu & Landman 1998; Herdan 2008), which he notes share certain properties with clauses embedded under superlatives. These include the possibility of a relativisation gap in the existential *there be* construction, as illustrated in (i), and the impossibility of *wh*-relativisation, as illustrated in (ii):

    (i)   I took with me the *(newest) book that there was on the table.   (Herdan 2008: ex.10b.)

    (ii)   Coach Hayes put the best players {that, *which} he could into the game.   (Carlson 1977: ex.72a.)

The second diagnostic applies straightforwardly to the constructions under investigation here:

    (iii)   *Mary sang {the loudest / the most songs} **which / how** anyone ever sang.

The impossibility of a *wh*-pronoun in (iii), as opposed to *that* as in e.g.(21)b, suggests it is reasonable to assume, at some level, a uniform treatment for degree relatives and superlative degree clauses. While I have nothing to say about the constructions discussed in the cited works, I adopt their consensus view that these clauses contain the ellipsis of at least a gradable expression, and often a larger constituent.

33

For another illustration of the role NPIs play within my proposal, consider the (very) partial analysis of the examples in (48):

(54) Mary sang the loudest ...

    a. ⟦ that anyone in the group has ever sung ⟧ = $\{\lambda d.\ \lambda w.\ x$ sing d-loud at t in w | $x \in D_e \cap$

        $\{y$ | y is a member of Mary's musical group$\}, t \in D_i \}$

    b. ⟦ that she has ever sung ⟧   = $\{\lambda d.\ \lambda w.$ Mary sing d-loud at t in w | $t \in D_i \}$

    c. ⟦ that anyone in the group sang ⟧ = $\{\lambda d.\ \lambda w.\ x$ sing d-loud in w | $x \in D_e \cap \{y$ | y is a

        member of Mary's musical group$\} \}$

The examples in (54) illustrate how the composition of the alternative set A varies according to the NPIs in the degree clause. When *anyone in the group* is subject of the embedded clause, as in (54)a & c, then the degree properties in A will vary with respect to the identity of the singer, with the possible singers being drawn from the set of members of Mary's musical theatre group. This restriction is modelled via intersection of this set with the domain of individuals, as shown in the relevant set descriptions on the right-hand side of the pipe ("|"); here I leave implicit the additional restriction to salient things via intersection with C. When *ever* is present, as in (54)a & b, the degree properties in A will vary along the dimension defined by singing times. Thus the truth conditions of each of the utterance depend variously on making comparisons in these aspects: when *ever* is present as in (54)b, we compare times; when *anyone in the group* is present as in (54)c, we compare singers; and when both are present as in (54)a, we compare both. Hence these values for A will deliver the relevant meanings observed for these examples, as discussed with respect to (48).[18]

---

[18] A potential wrinkle for the analysis was brought to my attention by Adam Albright (p.c.): the analysis presented here for (54)a predicts -*est*'s presupposition, due to (46)a(iii), that there be more than one member in A will be satisfied when no-one other than Mary sang, because A will still contain multiple properties of the form [$\lambda d.\lambda w.$ Mary sing d-loud at t in w], as long as there are multiple Mary-singing time intervals t in C. This prediction does not correspond to intuitions, as (54)a is judged infelicitous when Mary is the only singer. I will have to leave this issue to future research.

I will defer detailed discussion of the compositional mechanisms I adopt to derive these alternative set meaning for the superlative degree clause until Section 5. The remainder of section 4 is devoted to showing with how the superlative degree clause analysis deals with the puzzles I set out solve. I will show in subsection 4.6 that the position in which the NPIs are interpreted in the degree clause structure is a SDE environment, thus predicting them to be licensed there; in the next subsection I will demonstrate how the mismatch effects of 3.2 are to be captured. The problems arising from the treatment of the embedded clause as a restrictive relative, such as the puzzle of the trivial contribution, need not trouble us any further, however; as soon as we adopt the superlative degree clause analysis, the question of why these embedded clauses do not behave like restrictive modifiers no longer arises.

4.5 Problems solved by the degree clause account #1: Mismatch effects

With the assumption of Heim's denotation for *-est*, repeated here as (55), and the analysis of the superlative degree clause as denoting *-est*'s restrictor, the proposal provides us with an explanation of the curious effects noted in (21) & (22), repeated here as (56) & (57).

(55)   Let P be a degree property and A be a set of such properties. Then,

    a. $[\![$ est $]\!]$(A)(P) is defined only if   (i) $P \in A$;

                                               (ii) $\exists Q[Q \in A$ & $P \neq Q]$.

    If defined,

    b. $[\![$ est $]\!]$(A)(P) iff $\exists d[$ P(d) & $\forall Q \in A[$ Q$\neq$P $\rightarrow \neg$ Q(d) $]]$

(56)   a. Bill wrote the most poems that anyone ever wrote.

    b. Mary sang the loudest that any soprano ever sang.

(57)   a. *Bill wrote the most poems that anyone (ever) published.

    b. #Mary sang the loudest that any baritone (ever) sang.

The key element of (55) that will be doing the work for us is the definedness condition in a(i), which requires that *-est*'s P argument be a member of its A argument. Heim assumes this definedness condition as part of her proposal for *-est*, and I argued in section 4.3 that it is

independently motivated; the logic I employed in 4.3 to predict presuppositional effects will be recycled in the analysis of (56) & (57). To make the discussion manageable, I'm going to pretend that these examples don't contain *ever* and ignore time interval variables.

Consider first the pair (56)b & (57)b. According to the superlative degree clause analysis, these are assigned the LFs represented here, where (58)a = (56)b & (58)b = (57)b.

                   (a) <u>any soprano sang <d- loud ></u>

(58)   (b) any baritone sang <d- loud >
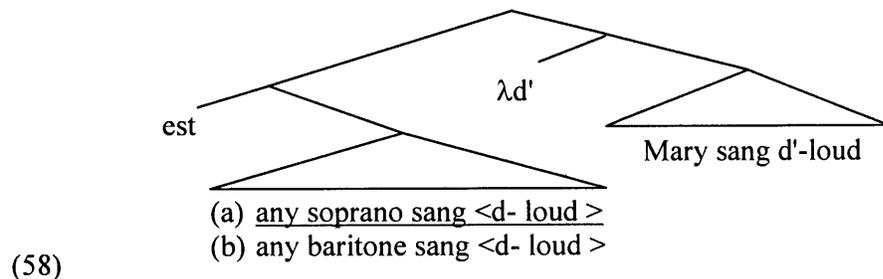
Under this analysis, *-est*'s sister in (58)a is the degree clause *any soprano sang ⟨d-loud⟩*, which denotes the set of degree properties (60)a; similarly for (58)b and (60)b.

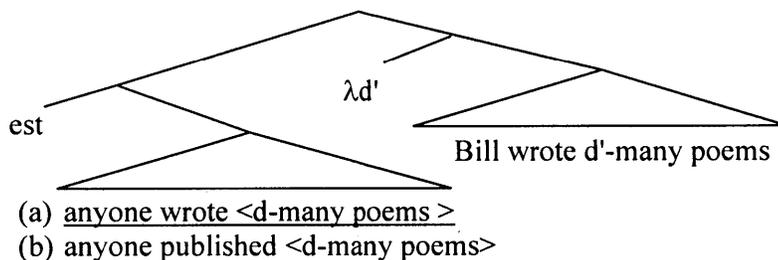(59)   Denotation of the matrix clause in (58):   [λd'.λw'. Mary sang d'-loud in w']

(60)   Denotation of the degree clause in (58): a. {λd.λw. x sang d-loud in w | x is a soprano }

                                                   b. {λd.λw. x sang d-loud in w | x is a baritone }

Here I am using e.g. "x is a soprano" to stand for the much less wieldy expression "x $\in$ $D_e \cap$ C $\cap$ {y | y is a soprano}", taking as read that the 'x's are contextually-given individuals. Applying the denotation (55), the degree clause in (58) – of which the denotations are given in (60) – is interpreted as *-est*'s A argument, while the matrix clause denotation (59) is *-est*'s P argument for both (58)a & b.

The sets in (60) have the form they do because the restrictor of the NPI indefinite, the noun *soprano* or *baritone*, goes in to restrict the set from which the 'x's of the degree property description are drawn, as we saw with *anyone in the group* in the previous subsection. So (60)a is the set of properties that a degree d has iff x sang d-loud, where the 'x's are sopranos. According to (55)a(i), (58)a is only defined if P$\in$A, i.e. (59)$\in$(60)a. For this condition to hold there has to be a soprano x such that $\forall$w$\forall$d, Mary sang d-loud in w iff x sang d-loud in w. This

36

condition is only satisfied when Mary is a soprano; thus (56)b is predicted under this analysis to give rise to the presupposition that Mary is a soprano, in accordance with observed judgements. Analogously, once we feed (59) and (60)b in to (55), we derive the prediction that (57)b presupposes that Mary is a baritone, which is the desired result, explaining the contrast in felicity with (56)b. The preceding employs an abbreviated form of the reasoning that predicted the requirement that Mary $\in$ C in the discussion of (42) in section 4.3.

By making slightly different use of the same means, the superlative degree clause proposal predicts the contrast between (56)a & (57)a. The relevant LFs are represented in (61)a & b and the assumed values for -*est*'s sister in each are given in (63)a & b respectively.

      (a) anyone wrote <d-many poems >

      (b) anyone published <d-many poems>

(61)

(62)   Denotation of the matrix clause in (61):   [λd'.λw' Bill wrote d'-many poems in w']

(63)   Denotation of the degree clause in (61): a. {λd.λw. x wrote d-many poems in w | x $\in$ C}

                                        b. {λd.λw. x published d-many poems in w | x $\in$ C }

In both (61)a & b, as in (58), the matrix clause denotation is constant, in this case (62). -*est*'s sister in (61)a is the degree clause *anyone wrote ⟨d-many poems⟩*, which denotes the set (63) a, and again for (61)b and (63)b with *published*.

Once again, -*est*'s definedness condition (55)a(i) requires that (62) $\in$ (63)a for (61)a to be defined, and similarly for the b cases. It is a straightforward application of the reasoning I have been employing to see that this condition holds, as long as Bill is in C. However, (62) is only a member of (63)b if there is an x $\in$ C such that $\forall$w'$\forall$d', x published d'-many poems in w' iff Bill wrote d'-many poems in w'. In other words, it would have to be a logical necessity that the number of poems x published = the number of poems Bill wrote. I submit that this condition is

37

impossible for any plausible common ground to support; thus (55)a(i) with respect to (60)b is impossible to satisfy.

As we will see in the next section, the licensing requirements of the NPIs force the parse in (60) and rule out an alternative parse with e.g. the embedded clause interpreted as a RC modifier of the noun. Since (60) is unavailable, there is no well-formed available structure for (57)a, and it is predicted to be pathologically unuseable.

4.6 Problems solved by the degree clause account #2: NPIs licensed

An important result of this analysis, in light of the problem for NPI licensing I identified in section 3.1, is that although VP-scope -*est* does not license SDE inferences within VP, it does support these inferences within its restrictor set A. Thus it falls out from this analysis that NPIs in a superlative degree clause are interpreted in a SDE environment, as I demonstrate here, and thus they are correctly predicted to be licensed.

Consider the pair in (64), and note that the sentences give rise to the presuppositions indicated, as predicted by the superlative degree clause analysis given the discussion from 4.3 onwards.

(64) a. John has published the most papers <u>any linguist</u> has published.

(*Presuppositions: John is a linguist, $\exists d$: John has published d-many papers*)

b. John has published the most papers <u>any syntactician</u> has published.

(*Presuppositions: John is a syntactician, $\exists d$: John has published d-many papers*)

Intuitively, the set-to-subset (since $[\![syntactician]\!] \subseteq [\![linguist]\!]$) inference from (64)a. to (64)b. goes through as long as John is indeed a syntactician, i.e. (64)b.'s presupposition is satisfied. This result is predicted by this semantic analysis: if the property in (65) is the largest set in (66) (i.e. (64)a is true) and (65) is in (67) (i.e. (64)b is defined), then (65) is also the largest set in (67), since (67) is bound to be a subset of (66); thus the truth of (64)b is guaranteed.

38

(65) λd. λw. John published d-many papers in w   -est *'s P argument in both ((64)a. & b.)*

(66) {λd. λw. x published d-many papers in w | x is a linguist}   est *'s A argument for* (64)*a.*

(67) {λd. λw. x published d-many papers in w | x is a syntactician} est *'s A argument for* (64)*b.*a.

More generally speaking, this pattern of inference is S(trawson)-valid since -*est* licenses set-to-subset inferences in its restrictor.

An important point to make clear is that replacing *linguist* by *syntactician* in (64) does in fact mean that (67) is a subset of (66), because both sets contain the <u>same kind</u> of objects: degree properties of the form *λd.λw. x published d-many papers in w*. The substitution imposes more stringent conditions of membership, but does not alter the nature of the members. I'm taking pains to belabour this point because of a confusion I foresee arising quite naturally concerning the predictions of the proposal with respect to pairs of sentences such as the following:

(68)   a. John has published the most <u>linguistics</u> papers any of these candidates has published.

   b. John has published the most <u>syntax</u> papers any of these candidates has published.

This is a variation on (19), which I gave as a counterexample to the putative claim that VP scope is SDE, and the inference from (68)a. to (68)b. is no more S-valid than in (19), as the reader can verify. Why should we be worried that this example might be any different? Consider what my proposal claims to be the analysis for (68)a.:

(69)
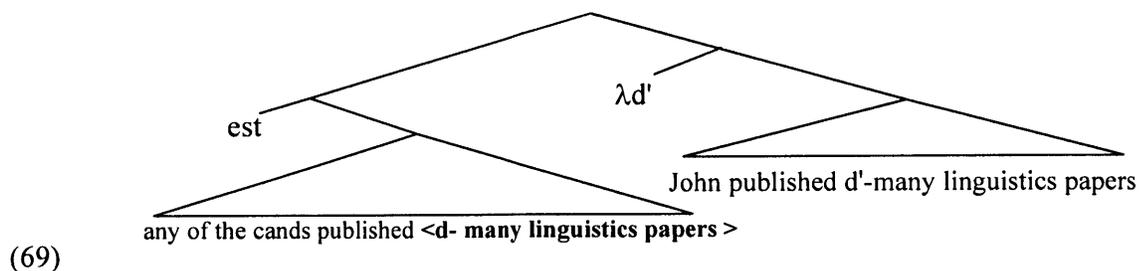
As the sketch in (69) illustrates, the structure I propose has an elided constituent in the degree clause, parallel to that in the matrix, containing the constituent *linguistics papers*. Since I am arguing that the degree clause denotes the restrictor of -*est*, where we just saw SDE inferences are licensed, shouldn't we expect to be able to S-validly infer (68)b., where the replacement of *linguistics papers* by the stronger expression *syntax papers*? Categorically, <u>No</u>.

(70) a. $\{\lambda d.\lambda w.\ x$ wrote d- many linguistics papers in w $\mid$ x $\in$ ⟦these candidates⟧ $\}$

   $= \lambda P_{dt}.\ \exists x\ [x \in$ ⟦these candidates⟧ & P$=\lambda d.\lambda w.$ x wrote d-many linguistics papers in w]

   b. $\lambda d'.\lambda w'.$ John wrote d'-many linguistics papers in w'

(71) a. $\{\lambda d.\lambda w.\ x$ wrote d- many syntax papers in w $\mid$ x $\in$ ⟦these candidates⟧ $\}$

   $= \lambda P_{dt}.\ \exists x\ [x \in$ ⟦these candidates⟧ & P $= \lambda d.\lambda w.$ x wrote d-many syntax papers in w]

   b. $\lambda d'.\lambda w'.$ John wrote d'-many syntax papers in w'

(70) shows the 1st and 2nd arguments of -*est* in (68)a in turn, whereas (71) does so for (68)b. This way of laying things out makes clear that the type of objects quantified over in each case, despite sharing some similarity in form, are quite distinct; thus the asserted contents of the two statements are incomparable, and we don't expect any entailment relationship between them[19].

Alternatively, to come at the same point from a completely different direction, notice that the position where the replacement by a subset-denoting expression **is** S-valid – for convenient reference, in (70) and (71) the spot is occupied by *these candidates* – is one that, all things being equal, is susceptible to having its monotonicity mucked about with by higher operators. In contrast, the position that *linguistics/syntax* turns up in, as illustrated in the more articulate second line of (70)a/(71)a, is sheltered by the non-monotonic identity operator.

So this analysis correctly predicts the availability of SDE inferences, and thus the licit presence of NPIs, in just the right places in the superlative degree clause constructions under investigation. Nobody should be alarmed about the result that -*est* allows NPIs in its restrictor but not in its scope, as this is precisely the pattern that is well-known to hold with other asymmetric NPI licensers such as *every* (Ladusaw 1980). This conclusion is not however intended to give the misleading impression that my analysis straightfowardly extends to make the necessary correct predictions about the SDE-ness and NPI licensing property of -*est* generally, even in the garden-variety cases discussed in Section 2. See Section 6.2 for some discussion.

---

[19] Although of course their *presuppositions* stand in an entailment relationship.
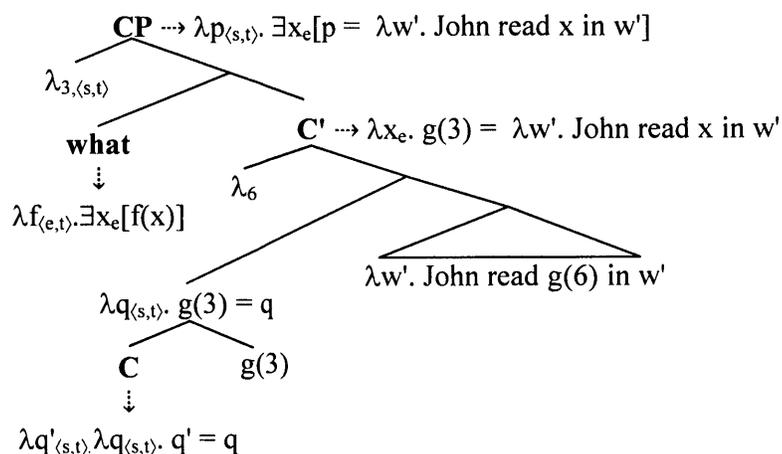
4.7 Section conclusion

To sum up the results of this section then, I have sketched a proposal for the analysis of these superlative-embedded clause constructions whereby they are interpreted as the restrictor of -*est*, under the assumption that it quantifies over a set of degree properties, thus accounting for the contribution these – otherwise mysterious – constituents make to the sentences they occur in. I furthermore illustrated how they can derive the peculiar downgradedness or infelicity these constructions bring about if there is a "mismatch" between the degree clause and the matrix. Finally, I showed how this analysis can correctly predict the licit presence of NPIs in these structures, in spite of the problem for NPI licensing I identified for VP-scope -*est* in Section 3.1. Thus the analysis sketched here meets all the desiderata outlined at the end of the last section. In the next section I discuss in greater detail the technical proposal I adopt for the derivation of superlative degree constructions, and lay out the predictions the analysis makes.

5. Deriving superlative degree clauses

5.1 A Karttunen derivation of Hamblin semantics

In this section I make a more concrete proposal for deriving a set of degree properties from the superlative degree clause. The key insight to be exploited is that the desired meaning of these elements is akin to that of the Hamblin (1973) denotation of a question; I will therefore employ a generalised version of compositional machinery that is already required elsewhere in the grammar. There are various possibilities out there on the market; I adopt the one that I am most familiar with, which is a version of Karttunen's (1977) analysis of the derivation of Hamblin denotation of a question, within an extension of the Heim & Kratzer 1998 framework (von Fintel, Fox & Heim, 2010 class notes; henceforth vF,F&H). This machinery is to a certain extent a choice of convenience, but I believe it has some advantages, in that its assumption that the NPI indefinites must undergo QR leads to welcome predictions.

41

The key point from Karttunen is the assumption of a silent morpheme – labelled '?' by Karttunen 1977, but often written as 'Q' (e.g. vF,F&H) – that turns a proposition into a (singleton) set of propositions. A *wh*-word, construed as a restricted existential, can then quantify into this set of propositions, generating the required nontrivial set. As an illustration, consider the following derivation for the Hamblin denotation of the embedded question *what John read*:

$$\text{CP} \dashrightarrow \lambda p_{\langle s,t\rangle}.\ \exists x_e[p = \lambda w'.\ \text{John read } x \text{ in } w']$$

$$\lambda 3,_{\langle s,t\rangle}$$

what
$$\lambda f_{\langle e,t\rangle}.\exists x_e[f(x)]$$

$$\text{C'} \dashrightarrow \lambda x_e.\ g(3) = \lambda w'.\ \text{John read } x \text{ in } w'$$

$$\lambda 6$$

$$\lambda w'.\ \text{John read } g(6) \text{ in } w'$$

$$\lambda q_{\langle s,t\rangle}.\ g(3) = q$$

$$\text{C} \qquad g(3)$$

$$\lambda q'_{\langle s,t\rangle}.\lambda q_{\langle s,t\rangle}.\ q' = q$$

(72)

Here the *wh*-word *what* has moved to the specifier of CP as one would expect, while its sister denotes the result of applying Predicate Abstraction to the structure that contains the interpreted trace of *what*. The interrogative head C is assumed to host the set-creation morpheme, which in this setup has a denotation of type $\langle st,\langle st,t\rangle\rangle$, and thus takes two arguments which have the type of a proposition, $\langle s,t\rangle$. It then returns True just in case its two arguments are identical. In this tree its first argument is a variable bound by the lambda operator at the root of the CP, and its second is the IP denotation. This allows the CP to denote the function which assigns True to any proposition identical to "that John read x" for some x, the existential quantification being contributed by *what*, crucially scoping outside the singleton set. The set characterised by the CP denotation is the desired set of propositions $\{\lambda w.\ \text{John read } x \text{ in } w \mid x \in D_e\}$.

A couple of notes on the details here: I found an antecedent to the formulation of the set-creation morpheme I'm using here in Beck 1996, who credits it to lecture notes of A. von Stechow. The binding of the proposition-type variable from the root of the CP is a necessary stipulation to derive the correct result. To make this sit fully comfortably within the H&K
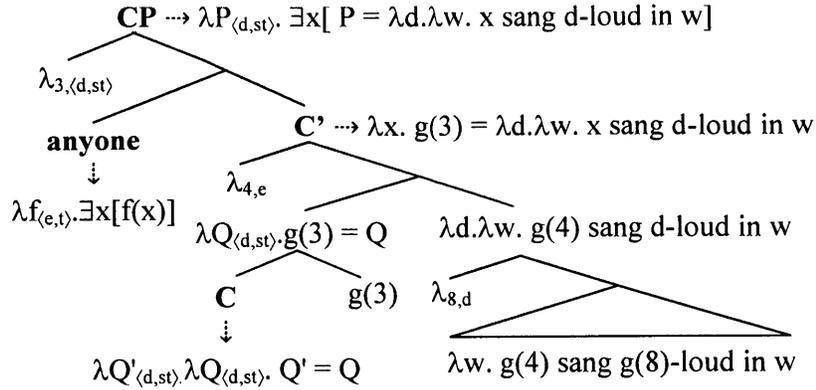
system, D. Fox (in vF,F&H) suggests assuming that the proposition-type variable is the trace of a (semantically vacuous) element. This null element is stipulated to move to adjoin to the root of the CP and introduce a lambda-binder in the usual way (similarly to the use H&K make of *PRO* in their ch.8), except that the trace is interpreted as a variable of type $\langle s,t \rangle$. D. Fox tentatively suggests a story where the requisite movements are motivated by the need to avoid type-mismatches; this is not the appropriate place to spell out such a proposal or examine its conseqences.

A particular feature of the (vF,F&H version of the) Karttunen analysis that I will be exploiting here is how it deals with a multiple *wh*-question such as *who read what*. All the *wh*-words are assumed to move to the specifier of interrogative C, whether overtly or covertly:

(73) λp [$_{CP}$ *who* λx [*what* λy [$_{C'}$ [$_C$ $\lambda q[p=q]$]] $\lambda w[x$ read y in $w]$]] =

    $\lambda p. \exists x_e \exists y_e [\ p = \lambda w[x$ read y in $w]\ ]$] =

    $\{\ \lambda w[x$ read y in $w]\ |\ x,y \in D_e\ \}$

As the reader is welcome to verify, under this analysis the semantics doesn't care about the c-command relationships between the *wh*-elements in such relatively simple constructions; the same set results regardless, which is adequate for current purposes.

Turning now to superlative degree clauses, notice that assigning a type of $\langle st, \langle st,t \rangle \rangle$ to the set-creation morpheme won't do; since we need to derive not a set of propositions but a set of degree properties, it needs to identify two variables of type $\langle d,st \rangle$. It follows that the IP denotation must be $\langle d,st \rangle$, thus we need a binder of a variable of type d on top of the proposition-denoting node, and in fact one that binds the d-argument of the gradable predicate within.

43

$$CP \dashrightarrow \lambda P_{\langle d,st \rangle}. \exists x[\ P = \lambda d.\lambda w.\ x \text{ sang } d\text{-loud in } w]$$

$\lambda 3,\langle d,st \rangle$

anyone
↓
$\lambda f_{\langle e,t \rangle}.\exists x[f(x)]$

$C' \dashrightarrow \lambda x.\ g(3) = \lambda d.\lambda w.\ x \text{ sang } d\text{-loud in } w$

$\lambda 4,e$

$\lambda Q_{\langle d,st \rangle}.g(3) = Q \qquad \lambda d.\lambda w.\ g(4) \text{ sang } d\text{-loud in } w$

$$C \qquad g(3) \qquad \lambda 8,d$$

C
↓
$\lambda Q'_{\langle d,st \rangle}.\lambda Q_{\langle d,st \rangle}.\ Q' = Q \qquad \lambda w.\ g(4) \text{ sang } g(8)\text{-loud in } w$

(74)

To exploit the set-creation machinery correctly, we need to assume that an NPI such as *anyone*, like the *wh*-words in the interrogative system just outlined, denotes a restricted existential quantifier, which moves to scope over the singleton set denotation. As I concluded in the previous section, this position is SDE, and therefore in this system the NPI is forced to move there in order to satisfy its licensing conditions. This movement has to be covert at least some of the time, given the existence of nonsubject NPIs in superlative degree clauses such as (75)a (= (49)b).

(75) a. Bill sent Chris the most letters [that he sent anyone].

b. Derivation of superlative degree clause:

λP [$_{CP}$ *anyone* λx [$_{C'}$ [$_C$ $\lambda Q[P{=}Q]$] $\lambda d\lambda w$[Bill sent x d-many letters in w]] =

$\lambda P.\ \exists x_e[\ P = \lambda d.\lambda w.\ \text{Bill sent } x \text{ d-many letters in } w\ ] =$

$\{\ \lambda d.\lambda w.\ \text{Bill sent } x \text{ d-many letters in } w\ |\ x \in D_e\ \}$

This setup also allows for cases of multiple NPI-movement, similarly to multiple *wh*-movement (e.g. (73)).

(76) a. Mary sang the loudest [that anyone ever sang].

b. Derivation of superlative degree clause:

λP [$_{CP}$ *anyone* λx [*ever* λt [$_{C'}$ [$_C$ $\lambda Q[P{=}Q]$] $\lambda d\lambda w$[x sing d-loud at t in w]] =

$\lambda P.\ \exists x_t \exists t_t\ [\ P = \lambda d.\lambda w.\ x \text{ sing } d\text{-loud at } t \text{ in } w\ ] =$

$\{\lambda d.\lambda w.\ x \text{ sing } d\text{-loud at } t \text{ in } w\ |\ x \in D_e,\ t \in D_i\ \}$

44

The contribution of NPI indefinites is justified by an independently motivated theory of NPI licensing, but we might wonder to what extent the analysis laid out here represents a natural extension of the original system for question denotations. To reiterate, I needed to additionally assume a lambda operator just below [C-*t*] to bind the degree argument of the gradable expression, and a tailor-made denotation for **C** to generate sets of degree properties. This latter capacity could be made to follow from the assumption that the set creation morpheme denotes the <u>identity</u> relation, since the natural interpretation of this assumption is that it is <u>type neutral</u>[20]; then the generalisation follows automatically. A unique set-creating morpheme would in that case be able to generate any kind of set, and the necessary stipulation reduces to that of constraining its distribution, and in particular ensuring that it gets the right type of argument. To be clear about what the problem would amount to in the present case, if there was no type-*d* lambda binder right under **C**, then the entire **CP** would have a question denotation; we would end up predicting that *John knows who came* would be synonymous with **John knows that anyone came* !

However, as I already mentioned, I need some way of ensuring that the set-creating morpheme's sister is ⟨*d,st*⟩ in any case. If we further assumed that the constituent containing just the identity morpheme and its sister were base-generated in the degree argument position of the gradable expression, undergoing movement to adjoin to the proposition-denoting clause and introducing a binder over its type-*d* trace there, then everything I needed to assume in the preceding paragraph would follow. I would then have to find a way of motivating the movement to exactly that position; also, this move would mean breaking the association between the set-creating morpheme and **C**. The latter point is not a cause for concern since it does not seem particularly desirable or natural to identify the set-creation morpheme with the complementiser *that*. That's all the speculation I can offer on these issues in the present context.

---

[20] This idea comes to me via D. Fox, class notes.

5.2 Predictions of the NPI-movement analysis

A feature of this analysis is that a NPI indefinite in a superlative degree clause is required to undergo movement to scope over the set creating morpheme, in order both to derive the attested truth conditions, and to satisfy its own licensing conditions. Since this is movement operation is syntactic we should, as a default stance, expect it to be subject to well-known constraints on syntactic movement. To discover evidence of island effects in superlative degree clause constructions comes therefore as a welcome prediction of this account.

The design of this test was worked out in collaboration with D. Fox (p.c.). The relevant structure is one where the NPI appears within an island for movement, in the instance an adjunct island, whereas the degree morpheme does not. This is to make sure that a judgement of unacceptability for the relevant test sentence only diagnoses a failure of the NPI movement predicted by this analysis, and cannot be ascribed to other plausible, independent causes such as *-est* (or indeed the set creating morpheme within the degree clause, if we follow the suggestion in the preceding subsection that it undergoes movement) being unable to move freely.

Furthermore, the adverbial clause within which the NPI is to be embedded is chosen to ensure that the presence of the NPIs forces a parse for the sentence as a superlative degree clause construction; if the NPI can be licensed independently then the embedded clause containing it can be parsed as an everyday restrictive RC. Adverbial clause heads such as *without* or *before* are able to license NPIs independently, whereas *after* does not: cf. *The dean arrived before/without/*after anyone else*. Therefore in (77)a,b, the NPI *anyone* can only be licensed by being interpreted in the restrictor of the superlative.

(77)   Mary read the most books after talking to Chomsky

      a.*... that she read after talking to anyone.

        cf. Mary read the most books after talking to CHOMSKY.

      b. OK ... that anyone read after talking to him.

        cf. MARY read the most books after talking to Chomsky.

The contrast between the continuations in (77)a & b seems to me to be subtle but solid. The reading which (77)a should have, paraphrasable as "Mary read more books after talking to Chomsky than she read after talking to anyone else", is perfectly reasonable, and is available for the comparison sentence with focus on *Chomsky*. However (77)a seems to be an impossible continuation for (77), particularly in contrast to the much improved (77)b, the meaning of which corresponds to the sentence with focus on the matrix subject *Mary*. In (77)a the NPI *anyone* appears inside the adverbial clause island, as the internal argument of *talking to*, while in (77)b *anyone* appears outside the island as subject of the degree clause. The latter NPI is able to move unhindered to its scope position, whereas the former cannot do so without crossing the island – and since its licensing conditions require it to move, the structure is correctly predicted to be ungrammatical. Hence the distinction in acceptability between the two continuations receives an explanation under my assumption that NPIs have to move at LF.[21]

The contrast in (77) therefore argues for the movement derivation of superlative degree clauses I propose. It allows us to distinguish my account from a possible alternative proposal, using an approach based on Kratzer & Shimoyama (2002), to generate the Hamblin set of degree properties meaning of the superlative degree clause. The Kratzer & Shimoyama mechanisms would be able to derive an interpretation for the degree clause by leaving these NPI indefinites *in*

---

[21] There might something to learn from the comparison of the contrast in (77) with a third variant, this time with two NPIs, one inside the island and one without:

    (i)   ? Mary read the most books after talking to Chomsky that anyone read after talking to anyone.

If (i) is reliably better than (77)a, then we may have evidence for a Minimal Compliance effect (Richards 1998), with the island violation provoked by moving the lower *anyone* improving thanks to licit movement to the same position of the higher *anyone*. More investigation is required here. Certainly it seems that any contrast between (77)a. and (77)b. is obliterated in variants additionally containing *ever*:

    (ii)   Mary read the most books after talking to Chomsky that she ever read after talking to anyone.

    (iii)   Mary read the most books after talking to Chomsky that anyone ever read after talking to him.

situ; since it would not require movement operations to derive the desired interpretation, such an approach would have difficulty accounting for the contrast in (77)[22].

6.   Potential extensions, problems, remaining puzzles

6.1   Where the superlative degree clause account should extend, but doesn't obviously seem to:

Modal superlatives (Schwarz 2005, Romero 2010, 2013)

(78)   John brought the longest possible ladder.

'There's no accessible world where J. brought a longer ladder than he brought in w$_@$'

(79)   John climbed the most possible mountains.

'There's no accessible world where J. climbed more mountains than he climbed in w$_@$'

Romero independently develops an account for the data in (78) & (79) that is similar to that discussed here, using Heim's denotation for *-est* and proposing that its alternative set argument is provided by an overt constituent containing ellipsis, in this case *[possible (John climb d many mountains)]*.

The natural way to extend the present account to (78), (79) is to allow *possible* to play an analogous role to *ever*, i.e. introduce alternative accessible worlds. A proposal for the denotation of *-est*'s complement is given in (80)a:

---

[22] I think it's worth stressing that I don't expect judgements to be particularly sharp with respect to sentences of the complexity, and potential sensitivity to discourse context, of (77); but neither is my account in any particular danger if the contrast in (77) turns out to be less than robust: everything I've said up to, and a goodly portion of, Section 5 still stands, and we'd just have to countenance an alternative approach such as K&S to deriving the meaning for superlative degree clauses.

(80) a ⟦ possible ⟨~~John climbed d-many mountains~~⟩ ⟧$^{w@}$ =

$\{\lambda d. \text{John climbed d-many mountains in } w \mid w \in Acc(w_@)\}$

b. $⟦(79)⟧^{w@}$ = est($\{\lambda d.\, j$ climbed d-many mountains in $w \mid w \in Acc(w_@)\}$)($\lambda d'.\, j$ climbed d'-many mountains in $w_@$) = 1 iff

$\exists d'[j$ climbed d'-many mountains in $w_@$ & $\forall f \in \{\lambda d.$ John climbed d-many mountains in

$w \mid w \in Acc(w_@)\}[f \neq [\lambda d''.\, j$ climbed d''-many mountains in $w_@] \rightarrow f(d') = 0]$

This analysis predicts the correct truth conditions for (79), but only under the assumption that -*est* is quantifying over degree sets, not degree properties. Suppose John climbed exactly as many mountains in w_1 as he did in $w_@$; then the degree sets (81)a,(81)b are identical, and -*est* does not distinguish between them.

(81) a. $\lambda d.$ John climbed d-many mountains in w_1

b. $\lambda d.$ John climbed d-many mountains in $w_@$

If -*est* could distinguish between e.g. (81)a,b, then the predicted truth conditions would be impossibly strong: "There's no other possible world where John climbed as many mountains as he did in $w_@$".

Hence, under the assumption of quantification over degree sets, (80) correctly predicts the paraphrase in (79). However, it won't do to assume that -*est* always quantifies over degree sets, as we needed degree <u>properties</u> to derive the presuppositional mismatch effects in (57), one of the key results of this paper. At the present time I have to leave open the question of how to proceed with the unification of the superlative degree clause and the modal superlative data.[23]

---

[23] Romero 2013 was charitable enough to include my confused ramblings on this issue from an earlier draft of mine, omitting my errors of notation, in an appendix of her paper.

6.2 Accounting for NP scope superlatives & NPI licensing under this analysis

This paper set out to provide an analysis of a particular kind of superlative construction in English, being mindful of an independently motivated theory of NPI licensing. What it didn't aim to do was to provide a unified analysis of English superlatives or a complete explanation of their NPI licensing properties. This is just as well, because the proposal laid out for the target construction in Sections 4 & 5 does not obviously extend to the basic cases discussed in Section 2. Using the -*est* focus operator analysis to capture NP-scope readings for ambiguous cases like (4), of which a trivial variant is in (82), requires nonobvious assumptions, and predicting NPI licensing under it in even simple predicative utterances such as (83) is quite mysterious.

(82)   John climbed the tallest building.

a. "John climbed the unique building whose height exceeds that of any other building"

b. "John climbed a building whose height exceeds that of any building climbed by any salient alternative to John"

(83)   War & Peace is the longest book Tolstoy <u>ever</u> wrote.

(84)   Let P be a degree property and A be a set of such properties. Then,

$[\![$ est $]\!](A)(P)$ is defined only if   i) $P \in A$

ii) $\exists d[P(d)]$

iii) $\exists Q[Q \in A \;\&\; P \neq Q \;\&\; \exists d'[Q(d')]\,]$

If defined, $[\![$ est $]\!](A)(P) = 1$ iff $\exists d[\, P(d) \;\&\; \forall Q \in A[\, Q \neq P \rightarrow \neg\, Q(d)\,]\,]$

Heim's denotation for -*est*, repeated here in (84), requires that it move to a proposition-denoting node for interpretation. This is straightforward when it takes VP-scope, and I have discussed such examples throughout the current text. It's less obvious how to derive NP-scope, since we don't typically assume that NP contains a proposition-denoting node. Heim & Kratzer (1998, ch.8) point out a necessity for such a node within NP in order to derive the truth conditions of sentences such as *No owner of an espresso machine is unhappy* – the problem being that the embedded indefinite needs to be interpreted at such a node, but can't outscope *no*. They posit a semantically vacuous, phonologically null element *PRO* in an internal subject

50

position for the NP. This *PRO* moves to a position just under *no*, introducing a binder for its own trace; the interpreted trace saturates the NP's external argument creating a propositional node, but this argument position is then opened up again to be operated upon by the taller determiner. Irene Heim (p.c.) suggests that this is a possible route to take for NP-scope readings of superlative utterances under the assumption of (84).

(85)   a. John climbed [$_{DP}$ the [PRO λ_3 [[-est-A] λ_9[$_{NP}$ t$_{3,e}$ [t$_{9,d}$-tall building]]]]]

b.
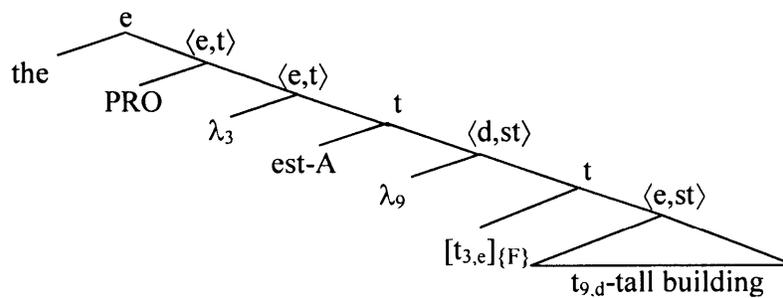
The tree for the DP denotation in (85)a. is given in (85)b., and demonstrates how all the types work out, supplying *-est* with the right kind of node to be interpreted at within NP, and yet allowing the whole definite DP to fit in to the larger structure. The lambda abstract under PRO then denotes the function that assigns True to any x just in case there's a degree d that has the property that x is a d-tall building, and no other property in A The definite DP then picks out the unique individual that this function assigns True, and the whole utterance is true just in case John climbed it.

In order to guarantee the correct meaning, the value of the alternative set A must be constrained, bearing in mind that at the end of the day we want to derive a meaning for the DP that picks out the unique building in the context which is tallest compared to other buildings. The set of alternative properties that will get the right result is {λd.λw. x is a d-tall building in w | x$\in$C}; then the definite DP is bound to pick out the tallest building in C. Under Heim's assumptions, salience and overall plausibility are primarily responsible for fixing the value of the contextually supplied domain variable (Heim 1999, p.4, fn.6), but nothing in the analysis definitively rules out, for example, a degree property such as [λd.λw. Ben Nevis is a d-tall mountain in w] being in A. Should we be worried about this? Regardless of the height of the

building John climbed, Ben Nevis is still taller.[24] This means that, regardless of the value g assigns to 3, there is no degree that has the NP-denotation property [λd.λw. g(3) is a d-tall building in w] that has no other property in A; they all have the Ben-Nevis-property, due to monotonicity. Since there is nothing that g can assign 3 that will give a property for which *est*-A will return true, the predicate passed up to the definite will be empty and violate *the*'s own existence presupposition (which I assume to be uncontroversial). Surely following plausible principles of communicative discourse would be to choose a value for a contextually-supplied variable that does not give rise to any presupposition violations; so I submit that cooperative conversational partners will simply exclude the Ben-Nevis-property in that case, and the situation described here will not arise. I conclude that the analysis adequately captures the truth conditions of the NP-scope reading of (85).

Turning our attention now to the matter of how NPIs are to be licensed in examples such as (83): According to the discussion in Section 4, in order to be licensed *ever* needs to be interpreted in *-est*'s restrictor, under the assumption of (84). However, according to the analysis just given for NP-scope *-est*, all the lexical material within NP is interpreted as *-est*'s scope argument; this is the problem in a nutshell.

Supposing that *ever* in (83) was somehow interpreted in *-est*'s restrictor and not its scope; the resulting predicted truth conditions would be something like in (86):

(86)   est(λd.λw. x is a d-long book in w & [Tolstoy wrote x at t in w] | x∈D_e, t∈D_i })

(λd.λw. W&P is a d-long book & Tolstoy wrote W&P at t' in w)

where t' is the time at which Tolstoy actually wrote W&P

I have not thoroughly investigated whether (86) adequately predicts the truth conditions for (83), and in any case, I don't understand the assumptions that would be required to derive (86) from the pieces (83) makes available. It would require computing in parallel two distinct denotations

---

[24] At the time of writing, the tallest man-made structure, at 829.8 m. (http://en.wikipedia.org/wiki/Burj_Khalifa), is still not taller than Ben Nevis, which stands at 1,344 m above sea level (http://en.wikipedia.org/wiki/Ben_Nevis).

for the same NP constituent, one which included *ever* and the other which somehow excluded it. The parallel computation suggestion is somewhat reminiscent of how the focus semantics theory invokes two different values for the same constituent, its ordinary and its focus semantic values. Perhaps some more imaginative researcher than I will come up with a way of fusing together Rooth's and Kratzer & Shimoyama's systems into a hybrid that will permit the compositional incorporation of alternative invoking expressions into the focus dimension while remaining apart from the ordinary semantic dimension; for my own part I remain sceptical.

Possibly a more promising way out of this impasse is to assume lexical ambiguity, whereby NP and VP scope readings for -est are actually the result of two different lexical entries. The VP-scope version would have the denotation and all the properties discussed in the main part of this paper, while the NP-scope version could be assigned a denotation that guarantees SDE inferences, and thus predicts NPI licensing, in the right places. The denotation von Fintel (1999) proposes has this property[25]:

(87) Let $R \in D_{<d,et>}$ be a gradable adjective denotation, $f \in D_{<et>}$ be a NP denotation, $x \in D_e$.

Then, $[\![est]\!](R)(f)(x)$ is defined only if $f(x) = 1$

If defined, $[\![est]\!](R)(f)(x) = $ True iff $\exists d:[R(d)(x) = 1 \; \& \; \forall y:[[f(y)=1 \; \& \; y \neq x] \rightarrow R(d)(y)=0]]$

It can be shown that under (87), $[\![est]\!](R)$ denotes a SDE function for any $R \in D_{<d,et>}$, i.e. *-est* is SDE on its NP argument (see Herdan & Sharvit 2006). This is promising since that's the position where we want to predict NPIs to be allowed. The lexical ambiguity approach might be a bit of a cop-out, but could make sense in a world where degree operators generally come in different-valence flavours (see e.g. Bhatt & Takahashi (2011) on comparatives crosslinguistically).

---

[25] This denotation bears only an underlying resemblance to von Fintel's actual proposal, since I have taken the liberty of translating it into the framework adopted in this paper.

An important step in the genesis of the ideas in this paper was inspired by Nouwen's (2010) discussion of (88):

(88)   The highest amount of prize money that Federer won so far is $1 million. That was for
          Wimbledon 2009.   (Nouwen 2010: ex. 29)

Nouwen observes that the money Federer has won during his career is a definite amount, but the interpretation of (88) requires dividing this amount into parts for the superlative to compare. Nouwen suggests that (88) contains a silent indefinite, *in some tournament*, which individuates the amounts *-est* quantifies over. This suggestion inspired the role I assign NPIs in superlative degree constructions. But once again, it's not clear to me that my analysis can apply to a variant on Nouwen's example with *ever*, as in (89)a. In (89)b., I offer a sketch of the truth conditions similar to that in (86), with the same reservations.

(89)   a. The highest amount of prize money that Federer has ever won is $1 million. .

          b. the λy [est({λd.λw. x is prize money $_w$ & HIGH-AMOUNT $_w$ (x) $\geq$ d

                                                    & F won x at t in w | x$\in$C$\cap$D$_e$, t$\in$C$\cap$D$_i$})

             (λd'.λw. y is prize money $_w$ & HIGH-AMOUNT $_w$ (y) $\geq$ d' & F won x at t' in w)]

                                                                                                    is $1 million

                              (where t'$\leq$[[Wimbledon 2009]])

Here it's clear more than ever that C is carrying the burden of excluding time intervals that overlap, or properly include more than one tournament, so that the definite subject doesn't end up picking out, for example, the sum that Federer won over the course of 2006 ($8.3 million, according to Nouwen). Can the von Fintel theory do better?

          c. the [est([[high-amount]])([[of prize money that Federer has ever won]])] is $1 million

             the [est(λdλy HIGH-AMT(y) $\geq$ d)(λx. x is prize money & $\exists$t$\in$C$\cap$D$_i$[F won x at t])]

                                                                                                    is $1 million

Again, the domain of intervals will have to be restricted in order to exclude e.g. the 2006 interval. On balance I think the von Fintel theory has the better chance of getting it right here, since the compositional semantics is relatively transparent (see the discussion of (86)).

6.3 Absence of mismatch effects

(90) a. John read the most books I've ever seen (anyone read).

   b. Mary sang the loudest I've ever heard (anyone sing).

Problem: The property $\lambda d.\lambda w.$ *John read d-many books in w* presumably cannot be a member of the set of properties $\{\lambda d.\lambda w.$ *I saw x read d-many books in w* $\mid x \in D_e\}$

It might be relevant that, for example, (90)b seems to presuppose that I heard Mary sing. Perhaps the presupposition is relatively easy to accommodate. If so, in what way is it easier than it was with the examples we wanted to use the unsatisfiable presupposition to exclude? Presumably the following examples will be relevant in determining the way forward here:

(91) a. ?John wrote the most poems in Davis Sq. that anyone ever wrote in Somerville.

                                                        (Norvin Richards, p.c.)

   b. ?/* John wrote the most poems in Somerville that anyone ever wrote in Davis Sq.

(92) a. ?! John won the most games of chess that anyone ever lost.

   b. ?! John sold the most cars that anyone ever bought.   (Martin Hackl, p.c.)

7.   Conclusion

I hope to have shown why superlative degree clause constructions present an interesting puzzle, and furthermore that my extension of Heim's analysis captures the key facts about the meanings of and the restrictions on such constructions, in addition to being compatible with von Fintel's generalisation on the distribution of NPIs.

8. References:

Beaver, D. I., and B. Z. Clark. 2008. *Sense and Sensitivity: How Focus Determines Meaning.* Oxford: Wiley-Blackwell.

Beck, S. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics.* 4. 1-56.

Bhatt, R. 2002. The raising analysis of relative clauses: evidence from adjectival modification. *Natural Language Semantics.* 10. 43-90.

Bhatt, R. & R. Pancheva. 2004. Late Merger of Degree Clauses. *Linguistic Inquiry.* 35. 1-45.

Bhatt, R. & Y. Sharvit. 2005. A note on intensional superlatives. In E. Georgala and J.Howell (eds.), *Proceedings of Semantics and Linguistic Theory* 15, CLC Publications, Cornell University. 62-79.

Bhatt, R. & S. Takahashi. 2011. Reduced and unreduced phrasal comparatives. *Natural Language and Linguistic Theory.* 29. 581-620.

Carlson, G. 1977. Amount relatives. *Language.* 53. 520-542.

von Fintel, K. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics.* 14. 97-148.

Gajewski, J. 2010. Superlatives, NPIs and most. *Journal of Semantics.* 27. 1-13.

Gawron, J. M. 1995. Comparatives, superlatives, and resolution. *Linguistics and Philosophy.* 18. 333-380.

Grosu, A. and F. Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics.* 6. 125-170.

Hackl, M. 2009. On the grammar and processing of proportional quantifiers: most versus more than half. *Natural Language Semantics.* 17. 63-98.

Hamblin, C. 1973. Questions in Montague English. *Foundations of Language.* 10. 41-53.

Heim, I. 1984 A note on Negative Polarity and Downward Entailingness. in: C. Jones and P. Sells. (eds.) proceedings of NELS 14, Amherst: GLSA. 98-107.

Heim, I. 1999. Notes on superlatives. Ms. MIT, Cambridge.

Heim, I. 2001. Degree operators and scope. In Féry, C., and W. Sternefeld (eds), *Audiatur Vox Sapientiae. A Festschrift for Arnim von Stechow.* 214-239.

Herdan, S. 2008. A superlative theory of amount relatives. In Chang, C. B., and H. J. Haynie (eds), *Proceedings of the 26th West Coast Conference on Formal Linguistics.* 234-242. Somerville, MA: Cascadilla.

Herdan, S., and Y. Sharvit. 2006. Definite and indefinite superlatives and NPI licensing. *Syntax.* 9. 1-31.

Hoeksema, J. 1986b. Monotonie en superlativen. In Hoppenbrouwers, C. (ed.), *Proeven van Taalwetenschaap.* Groeningen: Tabu. 38-49.

Howard, E. 2008. Superlatives, NPIs and Strawson-entailment. Undergraduate thesis. McGill University, Montreal.

Jackendoff, Ray S. 1972. Semantic Interpretation in Generative Grammar. Cambridge, MA: MIT Press.

Karttunen, L. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1. 3-44.

Kratzer, A. & J. Shimoyama. 2002. Indeterminate pronouns: the view from Japanese. In Yukio Otsu (ed.), *The Proceedings of the Third Tokyo Conference on Psycholinguistics*. Tokyo: Hituzi Syobo. 1-25.

Ladusaw, W. A. 1980. On the notion 'affective' in the analysis of negative polarity. Journal of Linguistic Research 1:1–16.

Matushansky, O. 2008. On the attributive nature of superlatives. *Syntax*. 11:1. 26–90.

Nouwen, R. 2010. Two puzzles of requirements. In Aloni, M. and K. Schulz (eds), *The Amsterdam Colloquium 2009*. Springer.

Reed, A & B. Kellogg. 1900. *A high school grammar dealing with the science of the English language, the history of the parts of speech, the philosophy of the changes these have undergone, and present usage respecting forms in dispute*. New York: Maynard, Merrill, & Co.

Richards, N. 1998. The Principle of Minimal Compliance. *Linguistic Inquiry*. 29:4. 599–629.

Romero, M. 2010. Decomposing modal superlatives. Handout of talk given at NELS 41, UPenn.

Romero, M. 2013. Modal superlatives: a compositional analysis. *Natural Language Semantics*. 21. 79-110.

Rooth, M. 1985. Association with focus. PhD. dissertation. University of Massachusetts, Amherst.

Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics*. 1. 75-116.

Sag, Ivan. 1976. Deletion and Logical Form. PhD dissertation, MIT.

Safir, K. 1982. Syntactic chains and the definiteness effect. PhD dissertation, MIT.

Schwarz, B. 2005. Modal superlatives. In Proceedings of SALT XV. Cornell: CLC Publications. 187-204.

Seuren, P.A.M. 1973. The comparative. In: F. Kiefer & N. Ruwet (eds.), *Generative Grammar in Europe*. Reidel, Dordrecht. 528-564.

Szabolcsi, A. 1986. Comparative superlatives. *MIT Working Papers in Linguistics* 8. Cambridge, MA. 245-266.

Williams, E. 1974. Rule ordering in syntax. PhD dissertation, MIT.

Wold, D. 1995. Antecedent-contained deletion in comparative constructions. Ms., MIT.

57