Remarks on comparative clauses as generalized quantifiers

Irene Heim, MIT, heim@mit.edu

Draft, January 2006

1. Larson

Quantifiers inside comparative 'than'-clauses typically are interpreted as if they scoped in the next higher clause, over the comparative adjective. For example:

(1) John is taller than every girl is.
'for every girl x: John is taller than x'

(2) John is taller than one of the girls is.
'there is an x among the girls such that John is taller than x'

As many authors have discussed, these interpretations are not expected from either a syntactic or a semantic point of view. Syntactically, 'than'-clauses ought to be scope-islands, no easier for quantifiers to scope out of than other embedded wh-clauses, such as free relatives. Semantically, 'than'-clauses appear to be definite descriptions of degrees, or perhaps universal quantifiers over degrees[1]. Either way, interpreting a quantificational DP inside them gives readings distinct from the ones we observe.[2] The data in (1) and (2) have thus been puzzling.

Larson (1988) proposed an interesting explanation. To set the stage for his idea, let us forget the clausal comparatives in (1) and (2) for a moment, and look instead at some phrasal comparatives, where 'than' introduces a mere DP rather than a clause.

(3) John is taller than every girl.

(4) John is taller than one of the girls.

A naive analysis of these sentences begins with the assumption that a comparative adjective denotes a <u>2-place relation between individuals</u>. E.g., *taller* (*than*)[3] denotes the relation that x bears to y iff x is taller than y, or more accurately:

---

[1] See Stechow (1984), Rullmann (1995), and Kennedy (1997) for overviews, as well as the references cited in these works.

[2] See Larson (1988), Rullmann (1995), Schwarzschild & Wilkinson (2002), and the earlier authors they credit.

[3] *than* will be ignored or treated as vacuous throughout the paper.

(5) $[\![taller]\!] = \lambda x_e. \lambda y_e.$ y is taller than x

This is a meaning of the same type as that of an (extensional) transitive verb. Given its semantic type, *taller* should take two individual-denoting arguments, as indeed it does in *John is taller than Mary*, whose interpretation is now straightforward. In the examples (3) and (4), we find a slightly more complicated, but all too familiar situation: Instead of the expected type-e argument, *taller* appears to be combining with a generalized quantifier (type <et,t>). There are well-known ways of dealing with this complication, and it doesn't matter for our present purposes which of them we adopt. For example, we may say that the type-mismatch is resolved by QR, so the LF of (3) is (6).[4]

(6) every $girl_4$ [John is taller than t_4]

Evidently, this is an interpretable structure, given (5), and we derive the correct truth-conditions for (3). (Similarly for (4).)

Here there was no puzzle regarding the scope of the quantified DP. If the comparative adjective denotes a relation between two individuals, a quantificational DP in one of its argument positions will scope over it – there is just no lower possible scope-site.

Back now to the clausal comparatives in (1) and (2). Their truth-conditions, as it happens, are the same ones as the truth-conditions of the phrasal comparatives (3) and (4). Why is that? Larson answers: it's because the 'than'-clauses in (1) and (2) have the very same denotations as the 'than'-DPs in (3) and (4). And why is that? It follows from a simple and familiar syntactic analysis of these clauses, namely as wh-clauses with predicate gaps. Consider (1). On the surface, *than* is followed by *every girl is*, a sentence which is missing a 1-place predicate. If this gap is taken to be the trace of movement of an invisible wh-word ("empty operator"), the structure is (7).

(7) wh_5 [every girl is t_5]

Suppose the moved wh-element is semantically vacuous, the binder-index next to it is a lambda-abstractor[5], and the trace it binds is a variable over 1-place properties (type <e,t>). Then the entire structure in (7) denotes the function $\lambda P_{<e,t>}.[\![every\ girl]\!](P)$. This, of course, is just the generalized quantifier $[\![every\ girl]\!]$ itself.

---

[4] I am not following Larson here, who instead uses a Montagovian higher type for 'taller', i.e., an intensionalized version of $\lambda \wp_{<et,t>}.\lambda y_e. \wp(\lambda x_e.$ y is taller than x).

[5] This is as in Hein & Kratzer's (1998) analysis of restrictive relative clauses.

So here is Larson's point: Even though (1) and (3) are syntactically different, in that (1) has a clause where (3) has a DP, the complements of *taller-than* in (1) and (3) are completely equivalent semantically. They both denote the same generalized quantifier (the set of all properties that every girl has). Therefore, the analysis that was given for (3) directly carries over to (1): above the level of 'than', the syntactic and semantic computations for the two sentences are identical. Whatever we did to resolve the type-mismatch between *taller* and its object in (3) – say QR – we also do in (1). So here we QR the entire clause, to obtain the LF in (8).

(8) [wh_5 [every girl is t_5]]$_4$ [John is taller than t_4]

Given the semantic identity between the QRed phrases in (6) and (8), the entire structures will be semantically equivalent as well. We have correctly derived that (1) has the truth conditions that it has. And to do this, we have not had to scope the DP *every girl* out of its *than*-clause! We scoped the clause as a whole (though only most locally[6]), but the DP inside it was interpreted *in situ*. Yet, we derived a meaning that is equivalent to the one which would result if we did extract the DP. Larson's analysis shows how this comes about, as a consequence of the internal syntax and semantics of the 'than'-clause, which involves lambda-abstraction over properties.

2. Limitation of Larson's approach

Larson's article explores various further desirable and interesting consequences of his proposal. Here I will move right on to a problem. The examples of clausal comparatives that we started out with (such as (1)) exemplified so-called "Comparative Deletion" clauses, more specifically, Comparative Deletion clauses in which the gap is the size of an adjective phrase (AP). As is well known, this is just one possible type. There is also Comparative Deletion with NP or DP-sized gaps, and there is "Comparative Sub-Deletion", in which the gap appears to be merely the size of a degree phrase. When we consider quantificational DPs embedded in one of these other kinds of comparative clauses, we still observe the same systematic pattern of apparent non-local scope. For example, take these Sub-Deletion sentences.

(9) The desk is wider than every couch is long.
'for every couch x: the desk is wider than x is long'

(10) The lamp is taller than one of the doorways is high.
'there is an x among the doorways such that the lamp is taller than x is high'

---

[6] I.e., it was scoped only to the lowest point where it is interpretable and within the minimal clause containing it. It would have been interpreted *in situ* if I had used the Montagovian higher type for the comparative adjective, as in Larson and fn.4.

It would be nice if we could predict these meanings as well without actually scoping the DPs out of their 'than'-clauses. But Larson's analysis (as he acknowledges) does not extend to these Sub-Deletion cases. The 'than'-clauses may be wh-clauses here too, but if so, their gaps do not correspond to predicates of individuals. Perhaps they correspond to degrees, or some other semantic type. Whatever it may be, after lambda-abstracting over this other type, the wh-clause will not be a generalized quantifier over individuals, and will not be able to leave an individual variable as its trace when it QRs. It is not at all clear how these examples can be compositionally interpreted if the meaning of the comparative adjective is a relation of type <e,et>.

The moral that is commonly drawn from the existence of Sub-Deleletion comparatives (and other considerations) is that comparative adjectives do not, after all, denote relations between individuals, or at least not always. On a prominent alternative approach, the surface unit of adjective and comparative ending is not even a semantic unit at all. Instead, the comparative morpheme composes with the 'than'-clause to form a kind of quantifier over degrees (the '-er'-phrase). This takes scope over the clause containing the comparative adjective, resulting in an LFs along the lines of (11) for (9).

(11) [-er than wh_1[every couch is t_1 long]]$_2$ [the desk is t_2 wide]

Both the main clause that the '-er'-phrase takes scope over, and the 'than'-clause that it contains, denote properties of degrees. The adjectives denote relations between individuals and degrees. Surface structures in which the complement of 'than' appears to be less than a (Sub-Deletion) clause, such as our examples (1) - (4), are analyzed as involving some sort of ellipsis. The basic case from the point of view of semantic analysis is the Sub-Deletion comparative. There are a variety of variants of this type of approach, and I will not dwell on them. Suffice it to say that adopting any one of them appears to be giving up entirely on Larson's explanation of the data in (1) and (2).

3. Sub-Deletion clauses as generalized quantifiers over degrees

What I want to propose in this note is that there is a way of accepting the standard wisdom about the semantic primacy of the Sub-Deletion construction while at the same time preserving the key idea in Larson's story. Larson said that comparative clauses were generalized quantifiers over individuals, and that this was so because the wh-gaps in them were variables over properties of individuals. I will say that comparative clauses are generalized quantifiers over degrees, and that this is so because the wh-gaps in them are variables over properties of degrees. As you will see, this makes for a semantic explanation of the apparent non-local scope of quantifiers in 'than'-clauses which sounds very much like Larson's. But it applies to Sub-

Deletion cases too, and in fact it goes with the kind of syntax in which '-er- and the 'than'-clause form a unit and 'than' always introduces a (possibly elliptical) Sub-Deletion clause.

Let us look at a Sub-Deletion clause like the one in (9), which I take to have the following structure.

(12) wh_1 [every couch is t_1 long]

How is this supposed to be able to denote a generalized quantifier over degrees? The standard analysis for this type of structure is based on the assumption that an adjective denotes a relation between individuals and degrees, and that the trace in (12) will consequently have to be a variable over degrees (henceforth type d). Combined with the assumption that the wh-operator is vacuous and its movement induces lambda-abstraction, we thus wind up with a meaning for the whole wh-clause which is of type <d,t>, i.e., a (1st-order) property of degrees. This is a lower type than I am aiming for. I want a generalized quantifier over degrees, i.e., <dt,t>. You may imagine ways of getting to this type by doing something fancier than mere lambda-abstraction at the edge of the clause[7]; but this is not what I have in mind. Rather, I say that the lambda-abstract has type <dt,t> because the type of the abstracted variable (the wh-trace) is already <d,t>! This in turn means that I must posit a non-standard type and meaning for adjectives: they denote relations between individuals and properties (or sets) of degrees. Concretely, my lexical entries look like this:

(13) $[\![\textit{long}]\!] = \lambda D_{<d,t>}.\ \lambda x_e.\ x\text{'s length} \in D$

This is a meaning of type <dt,et>, encoding the relation which an individual bears to a set of degrees iff the individual's length is an element of this set.

Before I proceed to show how the whole LF-syntax and compositional semantics of the sentence will unfold from this lexical entry, let me say something about it. I have used a "length-of" function in the metalanguage here, which is a function from individuals to degrees. More generally, I assume an ontology that contains degrees of various sorts, and various measure-functions (generally partial) which map individuals to degrees. One of them is the length-function, whose domain consists of physical objects with suitable shapes[8], and whose range is a

---

[7] E.g., the wh-clause might be thought of as more like a free relative than a restrictive relative, and thus incorporate the equivalent of a definite or universal determiner. There are many precedents for this in the literature, but it is not the route I take here.

[8] A complicated matter to explicate in more substantive terms. Also, disregarding temporal sense of 'long'.

sort of degrees which we also call lengths[9]. This much is standard and is also assumed where adjectives are given meanings of the simpler type <d,et>. The difference is in how the measure function is employed in the definition of the denotation. The more familiar analyses would have one of the following entries instead:

(14) $[\![long]\!] = \lambda d_d.\ \lambda x_e.$ x's length = d

(15) $[\![long]\!] = \lambda d_d.\ \lambda x_e.$ x's length ≥ d

I will return to the choice between these options in a later section and will end up modifying the claim that (13) really is the lexical meaning. But for the time being, this is the proposal.

I credit this new meaning for adjectives to Schwarzschild & Wilkinson (2002). What these authors actually say is a little different. Their emphasis is on the notion of an <u>interval</u> of degrees. An adjective, they say, relates individuals and intervals on a scale. For example, *tall* expresses the relation that x bears to those intervals on the height-scale which contain x's height. But what are intervals but certain sets of points? The difference between Schwarzschild & Wilkinson's semantics and the one I use here is just that their adjective-denotations impose a further condition on their arguments: these can't be arbitrary sets of points on the degree scale, but must be intervals (i.e., sets which satisfy the condition that for any two points they contain they also contain all points in between). The result of trying to apply an adjective to a non-interval set of degrees would presumably be undefined. As it turns out, in the actual examples which I will analyze or discuss below, the sets that adjectives apply to always happen to be intervals. So I might as well have adopted the interval-based entries of Schwarzschild & Wilkinson, but then again, I have no specific motivation to. Think of D in (13) as ranging over arbitrary sets or over intervals, as you prefer.

Now we need just one more thing before we can analyze a complete sentence: a semantics for the comparative morpheme *-er*. This will be the simplest meaning imaginable: a two-place relation between degrees, the greater-than relation (assumed to be given with our basic ontology, as in all degree-based theories).

(16) $[\![\text{-}er]\!] = \lambda d_d.\ \lambda d'_d.\ d' > d$

As for syntax, I am following the standard line that the 'than'-clause is the complement of the comparative morpheme. However, in my new scheme of semantic types, the two don't quite fit together as they are. *-er* wants a degree (type d) for its argument, but it's getting a generalized

[9] Although this is somewhat misleading, in that it is (crucially) the same sort of degrees to which a variety of other measure functions map their arguments, such as the width, height, depth, and distance functions.

quantifier over degrees (<dt,t>) instead. This is again the textbook type-mismatch, and amenable to the textbook remedies, such as QR. Our actual LF for example (9) then will not be quite the same as (11) above. The new non-standard types force us to QR the 'than'-clause out of its place as sister to *-er*.[10] On the other hand, we do not really have to QR the *-er*-phrase, since this now fits directly into the matrix adjective's argument position without a mismatch. So our LF is (17). (17a) shows the most important types (with vacuous elements omitted), and (17b) gives the semantic computation (which is just three lambda-conversions).

(17) $[\text{wh}_1 \text{ [every couch is } t_1 \text{ long]]}_2 \text{ [the desk is [-er than } t_2\text{] wide]}$

(a) $[\ 1 \text{ [every couch } t_{1,<d,t>} \text{ long]}_t\ ]_{<dt,t>}\ [\ 2 \text{ [the desk [-er } t_{2,d}]_{<d,t>} \text{ wide]]}_t\ ]_{<d,t>}$

(b) $[\lambda D_{<d,t>}.\ \forall x_e[\text{couch}(x) \rightarrow \text{x's length} \in D]](\lambda d_d.\ \text{the desk's width} \in [\lambda d'_d.d' > d])$
$= [\lambda D.\ \forall x[\text{couch}(x) \rightarrow \text{x's length} \in D]](\lambda d.\ \text{the desk's width} > d)$
$= \forall x[\text{couch}(x) \rightarrow \text{x's length} \in [\lambda d.\ \text{the desk's width} > d]]$
$= \forall x[\text{couch}(x) \rightarrow \text{the desk's width} > \text{x's length}]$

The meaning we have derived has 'every couch' taking maximal scope. We have achieved this, however, in Larsonian fashion, with an *in situ* interpretation of the DP *every couch*. Only the wh-clause containing this DP has been scoped, and only locally.

Once the Sub-Deletion case is dealt with, the other cases follow suit, provided the usual ways of viewing them as reducible to Sub-Deletion by either syntactic ellipsis or semantic mechanisms of type-shifting that can mimic ellipsis. Our original example (1) can be worked out as follows.

(18) John is taller than every girl is.
underlying (by VP ellipsis or Comparative Deletion):
John is tall [-er than wh every girl is t tall]
LF: $[\text{wh } 1[\text{every girl is } t_1 \text{ tall}]]\ 2[\text{John is [-er than } t_2] \text{ tall}]$
meaning (derived as in (17) above): $\forall x[\text{girl}(x) \rightarrow \text{John's height} > \text{x's height}]$

4. Schwarzschild & Wilkinson's desiderata

Schwarzschild & Wilkinson (2002) (henceforth S&W) document extensively that the phenomenon of 'than'-clause-internal quantifiers appearing to scope out is systematic and

---

[10] If there are syntactic qualms about QRing clauses, we can always fall back on Montagovian type-raising instead. We can use a variant of (16) which takes type <dt,t> instead of d for its (inner) argument, defined in the usual way: $[\![-er_*]\!] = \lambda P_{<dt,t>}.\ \lambda d'_d.\ P(\lambda d_d.d' > d)$.

pervasive, and that it affects not just DP-quantifiers, but also floated quantifiers, quantificational adverbs, modals (= quantifiers over worlds), and even sentence connectives ('and'). For example, they observe the validity of the following paraphrases.

(19) Hubert was taller than most of the others were.
for most x other than Hubert: Hubert was taller than x

(20) Hubert was taller than exactly two of the others were.
for exactly two x other than Hubert: Hubert was taller than x

(21) The suit cost more than they had each paid in taxes.
for each x among them: the suit cost more than x had paid in taxes

(22) It is warmer here today than it usually is in New Brunswick.
for most times t: it is warmer here today than it is at t in New Brunswick
(= it is usually colder in New Brunswick than it is here today)

(23) It is warmer today than it might be tomorrow.
there is a possible world w such that it is warmer in the actual world today than in w tomorrow
(= it is possible that it will be colder tomorrow than it is today)

(24) George is richer than his father was and his son will be.
George is richer than his father was, and he is richer than his son will be.

They also point out that this pervasiveness militates against the temptation to analyze sentences like (1), (2), (19), and (20) as involving a mechanism which actually scopes the DP out of its clause (such as long-distance QR). While DPs are known to have a certain degree of scopal mobility, this is unheard of for the Q-adverbs, modals, or floated quantifiers in (21) - (23), and one can't even see how it could apply to the sentence-conjunction 'and' in (24).

Our Larson-inspired analysis handles all these examples successfully. Everything in the 'than'-clauses can be interpreted *in situ*, yet the predicted meanings conform to the paraphrases. For some of the examples, where there are temporal and modal operators, we need an appropriately intensionalized version of the current extensional implementation, but this is a routine amendment needed anyway. Just for concreteness, assume that all predicates have additional argument slots for times and worlds, filled by appropriately typed variables at LF, and bindable by QAdverbs, tenses, and modal verbs. An LF for (22) may then look roughly as in (25). (I use *t* for variables over times, *w* for worlds, and 'i' and 's' as labels for the corresponding semantic types.) Simplifying matters a bit and making various arbirtary (and hopefully irrelevant) choices, I have *usually* binding a time-variable that's sister to the embedded (elided) *warm*, and *today* sitting directly in the corresponding slot of the matrix adjective. I also gave *warm* a place argument, filled respectively by *here* and (*in*) *New Brunswick*. The adjective's lexical entry should be as in (26) to accommodate this argument structure.

(25) [wh$_1$ [usually$_t$ [it is t_1 warm (at) t in NB]]]$_2$ [it is [-er than t_2] warm today here]

(26) $[\![warm]\!] = \lambda t_i.\ \lambda D_{<d,t>}.\lambda x_e.$ [the temperature in x at t] $\in$ D

Sketching the semantic computation: the lambda-abstract on the right (starting with '2') will denote the set of temperatures that are below (exceeded by) the temperature here today. This set gets lambda-converted into the position of the wh-trace bound by '3', so the overall claim is that for most times t, the temperature in NB at t falls within that set, i.e., is below today's temperature here. – My readers should have no difficulty working out the remaining examples on their own.[11]

S&W also discuss an example which they take to show that allowing quantifiers to scope out from their 'than'-clauses is not just syntactically dubious but doesn't even suffice to get all the truth conditions right.

(27) Bill did better than John predicted most of his students would do.

In this example, the 'than'-clause effectively contains two quantifiers, a modal operator (quantifier over worlds) expressed by the verb *predict* and the DP-quantifier headed by *most*. What does our system predict here? If, once again, we interpret everything in the 'than'-clause *in situ*, we get the LF-configuration (28) and meaning in (29). I use an actual-world indexical @ to fill the world-slots of the predicates not under modal operators. For the gradable predicate *do well*, I employ the entry in (30), assuming that the pertinent measure function maps people to their test scores or something like that.

(28) [wh$_1$ [John predict (in) @ λ_w [most students do t_1 well (in) w]]]$_2$
[Bill did [-er than t_2] well (in) @]

(29) $\forall$w[w conforms to what John predicted in @
$\rightarrow$ for most students x: Bill's score in @ > x's score in w]

[11] It is worth noting that even Larson's original analysis already covered some of these examples, including the floated-Q case (21) and the conjunction-case (24). The cases which crucially involve time or world quantification are, as far as I can tell, not quite so straightforward in Larson's system. This is despite the fact that the implementation he gives in his paper is in the Montague Grammar framework and hence intensional to begin with. The difficulty I see (just to hint at it) is that one seems to need entries for comparative adjectives (like *warmer*) which manipulate not just one time/world parameter but two.

(30) $[\![do\ well]\!] = \lambda w_s.\ \lambda D_{<d,t>}.\lambda x_e.$ x's score in w $\in$ D

S&W express the intuition that (28) is true in the following scenario: Before the exam, John says: "Most of my students will get between 80 and 90 points." When the exam is over, Bill gets 96 points, and Alex gets 70. (29) captures this intuition insofar as this formula is indeed true in the scenario: in all the worlds conforming to John's prediction, there is a majority of students with scores between 80 and 90, hence with scores below Bill's actual score of 96. We also see that if *Bill* in (27) were replaced by *Alex*, the formula would be false, which captures another judgment reported by S&W, namely that sentence (27) with *Alex* instead of *Bill* is false in the scenario. And as S&W stress as a desideratum, the predicted truth-conditions for (27) are not equivalent to 'for most students x, Bill did better than John predicted that x would do' (which is not true in the imagined scenario, since John made no predictions about individual students).

The remarkable degree to which our analysis predicts S&W's data invites the question of how it relates to the analysis that these authors themselves develop. Is it perhaps the same analysis or a notational variant of it? A superficial reading of their paper certainly does not convey this impression. There is one key idea which indeed we share, and which I have already credited to them: to model adjective meanings as relations between individuals and sets of degrees (in their case, intervals), and in such a way that the relation holds iff the pertinent measure function maps the individual to a point contained in the set. Beyond that, however, the ingredients of their proposal, as well as the way they present them and motivate them, look rather dissimilar. A few of the differences are quickly seen to be minor or orthogonal to our common main goals.[12] Even with such differences eliminated, however, it is not transparent that we have two implementations of the same idea. Nevertheless, the predictions of the two theories systematically coincide, and with certain auxiliary assumptions, this is provably so. (The appendix gives more detail on what I have been able to prove.) For the remainder of the paper, I will take it as given that the two proposals are equivalent enough to share the same strengths and also the same weaknesses, in every respect that is relevant for their assessment as contributions to linguistic theory. To the extent that this is so, my discussion also is commentary on S&W's project.

---

[12] For example, I might use standard Montagovian technology to raise the type of my *-er* from <d,dt> to <<dt,t>,dt> or even to <<dt,t>,<<dt,t>,t>>, thus avoinding the need for QR of the 'than'-clause and bringing my type-system and LF-syntax into closer correspondence with S&W's.

5. A systematic exception: two groups of modals[13]

Example (31), repeated here from (23) above, shows a case where the quantification over worlds expressed by a modal in the 'than'-clause – here epistemic *might* – is effectively taking widest scope in the sentence.

(31) It is warmer today than it might be tomorrow.
$\exists$w $\in$ Acc: it is warmer in @ today than it is in w tomorrow
(= it is possible that it will be colder tomorrow than it is today)

We find analoguous behavior with other modals and with attitude verbs. (27) exemplified this with the verb *predict*, and here are a couple more examples.

(32) John is more polite than his secretary must have been.
John is more polite than his secretary appears to have been.
John is more polite than his secretary is reported to have been.
$\forall$w $\in$ Acc: John is more polite in @ than his secretary was in w.

This is the pattern that our analysis predicts. (The derivations proceed analogous to (25) and (28) in the previous section.)

However, not all examples conform to the prediction. Consider (33).

(33) He was more cautious than he needed to be.

What we predict this to mean is that in all accessible worlds, his degree of caution is less than his actual degree of caution. Let's say the accessible worlds are those in which a certain goal is achieved, e.g., to get home without getting robbed. Then the predicted meaning is, in effect, that all the worlds in which gets home without getting robbed are worlds in which he is less cautious than he actually is. In other words, the sentence should imply that he didn't in fact get home without getting robbed. To do so, he would have <u>had</u> to be less cautious than he was. He was too cautious (for the pupose at hand). – Now, it may or may not be possible to understand the sentence (33) in this way[14], but there certainly is a salient grammatical reading that claims

[13] This section essentially replicates Schwarzschild (2004), which in turn can be seen as working out a very brief remark at the very end of Heim (2000b). (This is the handout from my presentation at SALT 10, but the relevant part did not make it into the paper published subsequently in the proceedings, Heim 2000a.) Heim (2000b) defines an operator **S&W** (the "Schwarzschild & Wilkinson Convertor"), which is the same as Schwarzschild's **Π** except for the order of its arguments. In fact, **S&W** was schönfinkeled the wrong way round to create a moveable constituent, and Schwarzschild (2004) corrected this error.

[14] Once the pragmatics is properly controlled for, English speakers do get this type of reading also. Many thanks to Noah Constant for helping me clarify this and for providing me with the right kinds of examples, such as (i).

something different or weaker: He didn't have to be so cautious. He could (also) have gotten home safely with a lesser degree of caution. There are some accessible worlds in which he was less cautious than in fact (with no implication that the actual world isn't accessible too, i.e. that he didn't in fact reach his goal). In short, the meaning predicted by our theory and the actually observed (salient) meaning contrast as follows:

(34) predicted: $\forall w \in Acc$: he is more cautious in @ than in w.
'it would have been necessary to be less cautious'
observed: $\exists w \in Acc$: he is more cautious in @ than in w.
'it would have been okay to be less cautious'

The prediction, of course, is derived on basis of analyzing *need* as a necessity operator. If *need* were a possibility operator, we would predict the other, more adequate meaning. But there is overwhelming independent evidence (from sentences not involving comparatives) that *need* is a verb of necessity rather than possibility, so this is not helpful. We must concede that we have a problem here. (33) is a counterexample to our theory.

A wider survey of examples shows there are other modal predicates, both necessity and possibility operators, which behave likewise unexpectedly from the point of view of our theory (at least on one natural reading). The intuitions about (33) can be replicated with variants such as *...than he had to be*, *... than (it was) necessary*, and *... than (it was) required*. For a possibility operator, consider *be allowed to* in (35).

(35) He charged more than he was allowed to.
He charged more than allowed.
He charged more than was legitimate.

We predict this to mean that there is some (deontically accessible, say, law-abiding) world in which he charged less than he did in fact. But the sentence is heard to make a stronger claim than that, viz., that he broke the law. He broke the law because there aren't just some law-abiding worlds in which he charged less, but rather all law-abiding worlds are ones where he charged less. (Again, our inadequate prediction is contingent on the assumption that *allow* expresses possibility, not necessity, but this is independently supported.)

(i) He was older than he needed to be (to get a kid's ticket).

As Constant observes, (i) is indeed understood to mean that in every (lawful) world in which he get's a kid's ticket, he is younger than he is in fact, as predicted by our current theory. My point here is not that such readings don't exist, but only that they are not the only ones. This is what (33) with the scenario I sketch is intended to show.

With some examples conforming to our predictions and others going against them, it would be interesting to establish a generalization of some kind, to correlate the division between the two sets with some other ways in which they behave differently. For the time being, however, I will content myself with implementing a mechanism that generates the missing readings. I will not explain why the mechanism applies only in some cases and not in all, nor why it sometimes seems to apply obligatorily.

Examples like (33) and (35) originally caught my attention because not only are they problematic for the approach I am elaborating in this paper, but they actually are handled with striking success by (at least some versions of) the more conventional approach to comparatives that has prevailed in almost all the literature (except Larson and S&W). Suppose that adjectives after all denoted relations between individuals and degrees (not sets thereof), and more specifically had meanings such as (36) (parallel to (15) above). Suppose further that *-er* had a semantics based on maximality (as in v. Stechow 1984, Rullmann 1995, and Heim 2000a).

(36) $[\![\textit{cautious}]\!] = \lambda w_s.\ \lambda d_d.\ \lambda x_e.$ x's degree of caution in w $\geq$ d

(37) $[\![\textit{-er}]\!] = \lambda P_{<d,t>}.\ \lambda Q_{<d,t>}.\ \max(Q) > \max(P)$

Given the semantic types in this set-up, we would have LFs in which the 'than'-clause can stay *in situ* but the *-er*-phrase must scope. The analysis of (33) would be this.

(38) [-er than wh_1 [need @ λ_w[he be $t_{1,d}$ cautious *w*]]]$_2$ [he was $t_{2,d}$ cautious @]

(39) max({d: his degree of caution in @ $\geq$ d}) >
max({d: $\forall$w $\in$ Acc(@):his degree of caution in w $\geq$ d})
simplifies to:
his degree of caution in @ >
max({d: $\forall$w $\in$ Acc(@):his degree of caution in w $\geq$ d})

To see what the formula in the last line says, look at the set whose maximum is being referred to. This set contains the degrees of caution which he reaches or exceeds in every one of the accessible worlds. So this is the <u>intersection</u> of all the intervals (0, his degree of caution in w) for all the accessible w. This intersection coincides with the <u>smallest</u> of these intervals, i.e., it contains just those degrees of caution which he reaches or exceeds in the accessible worlds in which he is <u>least</u> cautious. So the maximum is his degree of caution in these "least-cautious" accessible worlds. In other words, it is the <u>minimum</u> of caution that is needed to achieve the goal. The formula in (39) says that he exceeded that required minimum – which is the meaning which we apparently understand. Notice that it is equivalent to the existential claim that there is an accessible world in which he is less cautious than he is. – An analogous argument shows that this type of semantics also gets the desired meaning for (35), the *allowed*-sentences.

This, of course, is no cause for proponents of the conventional approach to celebrate. Their success with these few examples is linked directly to their failure with the many others. All cases in which the 'than'-clause contains a universal quantifier – be it the necessity modal *need to*, an attitude verb like *claim* or *predict*, or a universal DP quantifier such as *every girl* – have the same logical structure, and therefore all of them wind up with "greater-than-minimum" readings analogous to (39). E.g., *more polite than the secretary appears to have been* is predicted to mean 'more polite than the secretary was in the (epistemically) accessible world in which she was least polite', and *taller than every girl is* is predicted to mean 'taller than the shortest of the girls'. These meanings are wrong, and what is right instead are the meanings predicted by the Larsonian approach that I worked out in the previous sections (and by S&W). Putting the two types of examples side-by-side makes it very plain that what one of the approaches gets right, the other gets wrong. We can't have it both ways – unless, I suggest, we acknowledge that <u>both</u> are right in a sense, and we construct a hybrid analysis that incorporates basic ingredients of each. Here is how.

To begin, what does an adjective mean? Is 'tall' the relation 'x's height $\geq$ d' between things and degrees, or the relation 'x's height $\in$ D' between things and sets of degrees? Both, I say, but the more basic, lexical, meaning is the first. The second is the result of composing the adjective-root with an invisible operator, which (following Schwarzschild 2004) I write as **Π**.[15] More precisely, **Π** is generated in the degree-argument position of an adjective, where it combines with whatever is traditionally generated in that slot. For example, in the underlying structure of a 'than'-clause, where we used to posit a **wh** in the adjective's degree-argument position, we now generate [**Π** **wh**] in this slot. Similarly, in the main clause of a comparative, where previously we had the '-er'-phrase [**-er** + 'than'-clause], we now have [**Π** [**-er** + 'than'-clause]]. I call the complex of the **Π**-operator and its sister the "**Π**-phrase". The semantics of **Π** is as follows:

(40) $[\![\Pi]\!] = \lambda D_{<d,t>}.\ \lambda P_{<d,t>}.\ \max(P) \in D$

Let us see how this works in a simple case, like the 'than'-clause *than Mary is tall.* The AP here starts out as (41a). The *wh* moves to the edge of the clause as usual, and the **Π**-phrase undergoes a short movement (say, to the edge of AP) to resolve type-mismatch; so we have

[15] Schwarzschild means "**Π**" to stand for "p-i", which abbreviates "point to interval".

(41b). (The subject *Mary* also moves out of the AP, but let's pretend it is reconstructed back in.) The meaning is computed in (41c).

(41) (a) Mary [**Π wh**] tall

(b) $[_{CP}$ wh$_1$ [_ is $[_{AP}$ [**Π** $t_{1,<d,t>}]_2$ [Mary $t_{2,d}$ tall]]]]

(c) $[\![tall]\!] = \lambda d_d.\ \lambda x_e.$ x's height $\geq$ d
$[\![2[Mary\ t_2\ tall]]\!] = \lambda d.$ Mary's height $\geq$ d
$[\![\Pi\ t_1]\!]^g = \lambda P_{<d,t>}.\ \max(P) \in g(1)$
$[\![AP]\!]^g = [\![\Pi\ t_{1,<d,t>}]\!]^g([\![2[Mary\ t_2\ tall]]\!])$
$= 1$ iff $\max(\lambda d.$ Mary's height $\geq d) \in g(1)$
$= 1$ iff Mary's height $\in g(1)$
$[\![CP]\!] = \lambda P_{<d,t>}.$ Mary's height $\in$ P

We see that the result is the same as before, only arrived at more indirectly. The **Π**-operator kind of mediates between the set-of-degrees variable bound by the *wh* and a basic adjective meaning that selects for degrees rather than sets of degrees. It does this by converting the set-variable into a generalized quantifier over degrees, which then can take the rest of the AP as its argument.

Clearly, if the **Π**-phrase always scoped locally in its host-AP, as it did in this illustration, the new analysis would not differ from the previous one in its predictions. But the point is that in this new more articulated structure, we have a new phrase which, at least in principle, can move to take non-local scope. The **Π**-phrase, being a generalized quantifier over degrees, could move up and be interpretable in higher positions too. This is what Heim (2000b) suggested, and Schwarzschild (2004) argued more explicitly, goes on in examples like (33), and what is responsible for their previously problematic truth-conditions. Here is the analysis.

(42) He is taller than he needs to be.
[wh$_3$ [**Π** t$_3$]$_2$ [need-@ λ_w[he be t$_2$ tall-w]]]$_1$ [[**Π** er than t$_1$]$_4$ [he t$_4$ tall-@]]
1[[**Π** er than t$_1$]$_4$ [he t$_4$ tall-@]] : $\lambda d_d.$ his actual height > d
2[need-@ λ_w[he be tall-w t$_2$]] : $\lambda d_d.\ \forall w \in Acc(@)$:his height in w $\geq$ d
[wh$_3$ [**Π** t$_3$]$_2$ [need-@ λ_w[he be t$_2$ tall-w]]] :
$\lambda D_{<d,t>}.\ [\max_d.\ \forall w \in Acc(@)$: his height in w $\geq d] \in D$
His actual height > $[\max_d.\ \forall w \in Acc(@)$:his height in w $\geq$ d]

Notice that the **Π**-phrase scopes over *need*, and that the final line mirrors (39).

To prevent massive overgeneration of unattested readings, we must make sure that **Π** never moves over a DP-quantifier, an adverb of quantification, or for that matter, an epistemic modal or attitude verb. Nor must any DP-quantifier be allowed to reconstruct down into its scope. Its capacity for scoping over other material must be severely constrained. I have nothing to say here about why this should be the case. For pertinent discussion, see in particular Heim (2000a) and Takahashi (2006).

6. Monotonicity properties and NPI-licensing

S&W point out that their analysis makes the comparative clause an upward entailing (UE) environment. The same is true for Larson's analysis and the Larson-inspired proposal in this paper. In the implementation I have adopted, this is particularly obvious, since the 'than'-clause is literally taking maximal scope.[16] Clearly, inferences such as (43) and (44) are predicted valid, given that the respective DPs are effectively interpreted with widest scope.

(43) John is taller than some professional basketball players are.
Therefore, he is taller than some professional athletes are.

(44) John is taller than most of his ancestors were.
Therefore, he is taller than some of his ancestors were.

However, this does not mean that we don't find non-UE or even downward-entailing (DE) environments further down within the 'than'-clause. Trivially, if the 'than'-clause itself contains DE operators (such as negation or universal determiners), then phrases within the scope of these will be in DE environments. But what I have in mind here is a group of more interesting cases, in which DE environments are created by configurations of operators which are unique to comparative clauses. In a nutshell, *-er* and **Π** together conspire to create DE-environments in (certain regions of) the 'than'-clause.

Notice that it would not be correct to say that **Π**-phrases in themselves are DE.[17] The following inference schema is not valid.

(45) $P_{<d,t>} \subseteq Q_{<d,t>}$
$[\![\Pi]\!](D)(Q) = 1$; i.e., $\max(Q) \in D$
Therefore, $[\![\Pi]\!](D)(P) = 1$; i.e., $\max(P) \in D$

---

[16] Had we used the Montagovian higher-typed $-er_{*}$ of note 10 instead, it would be straightforward to show that that is a UE operator.

[17] I am talking here about the monotonicity properties of the **Π**-phrase, not the **Π**-operator. We could ask about the latter as well, of course. As it turns out, the **Π**-operator is not DE either (it is even UE). But this fact is not relevant to anything, since the sister of **Π** never seems to be complex enough to host any NPIs in the first place.

There is no guarantee that if the maximum of a set is in D, then the maxima of its subsets are also in D. However, suppose we derive an LF for a comparative in which **Π** in the 'than'-clause ends up in the immediate scope of *wh*, and **Π** in the matrix-AP in the immediate scope of the 'than'-clause. ("Immediate scope" means that nothing intervenes.) An example of such an LF is (42) above, and here is the general schema.

(46) $[wh_1\ [[\mathbf{\Pi}\ t_1]\ \phi]]_2\ [[\mathbf{\Pi}\ \text{er than}\ t_2]\ \psi]$
where ϕ and ψ are constituents of type <d,t>, and t_1 and t_2 are variables of types <d,t> and d respectively

In (46), ϕ is in a DE environment, in the sense that the following inference is valid.[18]

(47) Suppose that
(i) $[\![\phi]\!] \subseteq [\![\chi]\!]$ and
(ii) $[wh_1\ [[\mathbf{\Pi}\ t_1]\ \chi]]_2\ [[\mathbf{\Pi}\ \text{er than}\ t_2]\ \psi]$ is true.
Then $[wh_1\ [[\mathbf{\Pi}\ t_1]\ \phi]]_2\ [[\mathbf{\Pi}\ \text{er than}\ t_2]\ \psi]$ is also true.

The proof proceeds by using the entries of **Π** and *-er* (and some lambda-conversion) to work out the meaning of the schematic structure. This shows (by calculations parallel to (42) above) that the claim in (47) amounts to the claim in (48).

(48) Suppose that
(i) $[\![\phi]\!] \subseteq [\![\chi]\!]$ and
(ii) $\max([\![\psi]\!]) > \max([\![\chi]\!])$.
Then $\max([\![\psi]\!]) > \max([\![\phi]\!])$.

It is clear that (48) must hold. Since $[\![\phi]\!]$ is a subset of $[\![\chi]\!]$, $\max([\![\phi]\!])$ cannot be higher than $\max([\![\chi]\!])$. Therefore, whatever exceeds the latter must exceed the former.

This result is welcome, because it arguably allows us to make sense of at least some instances of NPI-licensing in comparative clauses. A particularly interesting case is the German verb *brauchen*, which (when it embeds an infinitive) has the typical distribution of an NPI.[19] Its meaning is that of a necessity-operator, and when it occurs in a comparative clause, we get the kind of truth conditions that indicate that **Π** is scoping over it. In fact, (49) means the same thing as (33).

---

[18] It may be worth noting also that ψ is not in a DE environment (despite being in the scope of a **Π**-phrase as well). In fact, ψ is in an UE environment

[19] See also the corresponding item in Dutch (discussed by Rullmann 1995, among others), and English "auxiliary" *need* (the *need* that doesn't inflect for 3rd person present and takes a bare infinitive without *to*), as described by Potts (2000).

(49) Er war vorsichtiger, als er zu sein brauchte.
he was cautious-er than he to be needed

Given its truth conditions, (49) must have the same LF as (42), with *brauchen* in the scope of the **Π**-phrase. It is very attractive to say that this LF-configuration is also what licenses *brauchen* as an NPI, by placing it in a DE environment.

Interestingly, this type of situation arises not only when a **Π**-phrase takes non-local scope (moving over a modal). Even when it scopes as locally as possible, within the projection of its host adjective, it creates a DE environment in its scope. If this scope is very small, of course, it may be impossible to see this reflected in the distribution of NPIs. There may just not be room for an NPI in there; especially not if we assume – as I have already acknowledged that we must – that it is not allowed to reconstruct quantifiers into the scope of a **Π**-phrase. Then the existence of a DE environment is effectively undetectable. However, this is not a concern if the adjective in question happens to take arguments that are not just individuals but rather properties or propositions. What I have in mind are gradable adjectives such as *often* and *likely*, which select respectively for sets of times and sets of worlds as their arguments. These adjectives have plenty of room for all sorts of NPIs inside their arguments, and those NPIs will then be in the scope of the **Π**-phrase, no matter how low the latter scopes.[20]

Let me illustrate these abstract remarks with a concrete example (from Linebarger's thesis).

(50) Cows fly more often than he lifts a finger to help.

The measure function encoded by *often* is (roughly) a cardinality function, which maps sets of times to the number of their elements.[21] In (50), we compare the number of times at which cows fly to the number of times at which he helps. The lexical entry is something like this.

---

[20] For this reason, DE-environments in the scope of *often* and *likely* (and also *many*) are actually predicted independently of whether we posit a separate **Π**-operator or we wrap its contribution directly into the adjective-meaning, as I did before section 5 and as did S&W. In other words, if we did not have a **Π**-operator but instead a meaning for *often* of type <<u>dt</u>,<it,t>>, this wouldn't make a difference. *often* then would presumably denote the function $[\lambda D_{<d,t>}. \lambda P_{<i,t>}. |P| \in D]$. This function by itself is not DE in either of its arguments (specifically, $[\lambda P_{<i,t>}. |P| \in D]$ is not DE), but the configuration

$[wh_1\ [[t_1 \text{often}]\ \phi]]_2\ [[[er\ than\ t_2]\ \text{often}]\ \psi]$

creates a DE environment in ϕ. – In light of this observation, the discussion in S&W's paper (p. 6ff., regarding examples like Linebarger's (50)) is somewhat misleading. Their own theory, being essentially equivalent to the Larsonian theory of section 3 of this paper, does predict a DE environment in the scope of *often* in the 'than'-clause.

[21] Maybe it should be events instead of times, and maybe a proportional instead of a cardinal reading – I assume these fine points don't matter here.

(51) $[\![often]\!] = \lambda n_d.\ \lambda P_{<i,t>}.\ |P| \geq n$

The LF of (50) is as follows.

(52) $[\text{wh}_1\ [\mathbf{\Pi}\ t_1]_3\ [t_3\ \text{often}]_t$ [he lifts a finger to help (at) t]$]_2$
$[[\mathbf{\Pi}$ er than $t_2]_4\ [t_4\ \text{often}]_{t'}$ [cows fly (at) t']]

We already know from (47) that the whole 'often'-clause in the 'than'-clause constitutes a DE environment. Since *often* itself is UE and thus does not reverse monotonicity, this property is preserved and the scope of t_3 *often* in (52) is still a DE environment.[22] Plausibly, this is what makes it okay for the NPI-idiom *lift a finger* to occur in this domain. – A similar story can be told about (53).

(53) He's more likely to slow us down further than to lift a finger to help.

The gradable determiner *many* also falls in the class of items which take a degree argument along with co-arguments of a non-basic type. Following Hackl (2000), treat *many* as a quantificational determiner with the usual two arguments of type <e,t> plus a third argument for a degree. (The underlying measure function again is cardinality.)

(54) $[\![many]\!] = \lambda n_d.\ \lambda P_{<e,t>}.\ \lambda Q_{<e,t>}.\ |P \cap Q| \geq n$

This meaning is UE in both of its non-degree arguments, the restrictor and the nuclear scope.[23] Therefore both ζ and ξ in the LF-configuration below are DE environments.

(55) $[\text{wh}_1\ [\mathbf{\Pi}\ t_1]_3\ [t_3\ \text{many}]\ \zeta\ \xi]_2\ [[\mathbf{\Pi}\ \text{er than}\ t_2]\ \phi]$

We thus have no difficulty with licensing the NPIs in examples like these.

(56) More people took deductions than contributed anything to charity.

(57) More people enroll than will ever be able to finish.

(58) Many more people enroll than have a hope in hell of finishing.

7. NPIs and the scope of Π

I suggested that there are severe constraints on the scope of the **Π**-phrase. This was necessary in order to avoid overgenerating unattested readings in the many cases where

---

[22] More generally, we could have formulated the theorem in (47) without the restriction that the **Π**-phrase be in the <u>immediate scope</u> of *wh*, replacing it by the weaker restriction that any intervening operators be UE.

[23] More precisely, $[\![many]\!]$(n) is UE (for all n), and $[\![many]\!]$(n)(P) is UE (for all n and P).

quantifiers in the 'than'-clause can only be read with widest scope. On the other hand, we have shown that only that portion of the 'than'-clause which is below the **Π**-phrase is DE, whereas the portion above is UE. These two assumptions together make predictions about the distribution of NPIs. We expect them to be licensed in 'than'-clauses, but only in that lower portion which falls inside the scope of the **Π**-phrase. When there is material that the **Π**-phrase isn't scoping over, this should also limit the domain in which NPIs are licit. In this section, I will make a preliminary attempt to explore and test this prediction, but we will see that the facts do not fall neatly into place.

Traditional wisdom has it that comparatives license NPIs, with no particular restrictions on the type of NPI or its location within the clause.[24] Here is a representative list of examples that supports this impression.[25]

(59) John is taller than <u>any</u>one else is.

(60) I am busier than I <u>ever</u> was before.

(61) He is richer than you <u>care</u> to know.

(62) He stole more than he has confessed <u>yet</u>.

(63) Sie ist reicher, als er <u>auch nur</u> ahnt.
she is richer than he 'even'(NPI) surmises
'he doesn't even have an inkling how rich she is'

A confound that we have to watch out for is the existence of Free Choice (FC) readings for *any*. If (59) is an instance of FC-*any*, it doesn't need a DE environment. This, of course, raises another non-trivial and difficult question, viz., why comparatives should be able to license FC items. But be that as it may, there is reason to believe that they are. For example, we can have modifiers like those in (64).

(64) (a) John is taller than almost anyone else is.
(b) John is taller than just about anyone else is.

So I will not base any conclusions on the distribution of *any* and will focus on the other NPIs exemplified above instead.

---

[24] There are some systematic (and arguably well-understood) exceptions to this, particularly 'exactly'-differentials (Horn 1972) and 'less'-comparatives (on one of their readings) (Rullmann 1995).

[25] Examples with *any* and *ever* are all over the literature. Examples with *care* such as (61) are found in Rullmann (1995), Larry Horn reminded me of *yet*, and the example with *auch nur* in (63) was provided by Hubert Truckenbrodt.

If we want to account for the NPIs in (60) - (63), we are compelled to say that the **Π**-phrases in these examples are able to scope high enough to bring them into their scope. Let us first assure ourselves that this yields adequate semantic interpretations.

(65) $[\text{wh}_1\ [\Pi\ \ t_1]_3$ [ever before$]_t$ [I was t_3 busy (at) t$]]_2$
$[[\Pi$ er than $t_2]_4$ [I am (now) t_4 busy$]]$
$[\max_d.\ \exists t < \text{now. my busy-ness at } t \geq d] <$ my busy-ness now

(66) $[\text{wh}_1\ [\Pi\ \ t_1]_3$ [you care-@ λ_w [to know-w $\lambda_{w'}$ [he is t_3 rich-w'$]]_2$
$[[\Pi$ er than $t_2]_4$ [he is t_4 rich-@$]]$
$[\max_d.\ \exists w \in Acc_{care}(@).\ \forall w' \in Acc_{know}(w).$ his wealth in $w' \geq d] <$ his wealth in @

(67) $[\text{wh}_1\ [\Pi\ \ t_1]_3$ [he has yet confessed-@ λ_w [he stole-w t_3 much$]]_2$
$[[\Pi$ er than $t_2]_4$ [he stole-@ t_4 much$]]$
$[\max_n.\ \forall w \in Acc_{confess}(@).\ |\{x\text{: he stole-w } x\}| \geq n] < |\{x\text{: he stole-@ } x\}|$

(I left out the contribution of 'yet' in (67) and don't want to get into its messy semantics.) There are some questions about the appropriate accessibility relations and even the modal force of the verbs 'confess' and 'care_{NPI}'. But I think there are defensible answers to these that make these interpretations consistent with our reading of the examples.

What is surprising, however, is how non-local the scope of the **Π**-phrase will have to be if these analyses are right. In (65), it has scoped over the QAdv *ever*, which raises the question of why it can do this when it apparently cannot scope over *usually*, *sometimes*, *often*, *always*, as the meaning of (22) above and similar examples show (cf. S&W's observation). In (66) and (67), the **Π**-phrase appears to be coming from an embedded (though elided) clause, crossing over attidtude verbs such as *know* and *confess* which embed finite sentences. These examples thus contrast with S&W's *predict*-sentence ((27) above), where the observed truth conditions indicated that *predict* took maximal scope.

A hypothesis that comes to mind is that the scope of the **Π**-phrase is (at least partly) <u>determined</u> by the need for NPIs to get licensed. It is not that independent principles determine the possible scopes of the **Π**-phrase and that NPIs then are okay if they happen to be in the domains that can be scoped over. Rather, the **Π**-phrase is allowed to scope non-locally <u>because</u> there is an NPI that "attracts" it. In structurally identical examples that differ only by the absence of an NPI, the same wide scope is not possible. I don't know at present how to make syntactic sense of this hypothesis. Before we worry about that, however, it is appropriate to test whether the conjecture is even consistent with the truth-conditional evidence. We should examine real minimal pairs where only the presence of an NPI distinguishes one member from the other.

I have found one type of example which does seem to behave in the expected way. Consider first the German sentence in (68).

(68) Er ist höher gestiegen, als gestern ein Amerikaner gestiegen ist.
he has higher climbed than yesterday an American climbed has
meaning: 'there is an American he climbed higher than'

(68) means, I think, what Larson and S&W would predict, i.e., it expresses a widest-scope existential claim, indicating that the **Π**-phrase stays low. Now I replace 'yesterday' by the NPI *jemals* ('ever').

(69) Er ist höher gestiegen, als jemals ein Amerikaner gestiegen ist.
he has higher climbed than ever an American climbed has
meaning: 'he climbed higher than any American every did'

This sentence clearly means that no American ever made it as high as he did. We can derive this truth-condition if we scope the **Π**-phrase to the top of the 'than'-clause and interpret *jemals* and *jemand* in its scope as existential quantifiers over times and people respectively.

I have tried to construct other examples in which quantifiers of the sorts that are shown to scope wide in examples like S&W are "trapped" below an NPI. It is often difficult to muster clear intuitions about truth-conditions or even to get the examples to sound felicitous. Here is a sample.

(70) This is more than we ever have both (each) paid in taxes.

(71) I made yours bigger than I ever made every portion.
(cf. I made yours bigger than I made every other portion.)

I sometimes think that these sentences do have the meanings predicted by the scope-order **Π**-phrase > *ever* > universal quantifier, but I am not sure. The pertinent meaning of (71), for example, would be that your portion was bigger than any of the <u>minimal</u> portions on each of the previous occasions. Suppose I had a policy on each occasion as to what the minimum portion must be. Most people always just got that minimum from me, but for some I was a bit more generous. And you (the sentence on the intended reading would say) are one of those who got special treatment.

In another kind of example, however, the judgments fairly clearly do not bear out my conjecture. Take the following variations on (a simplified version of) S&W's *predict*-sentence.

(72) (a) Bill did better than John predicted that he would.

(b) Bill did better than John ever predicted that he would.

(c) Bill did better than John has yet predicted that he would.

The crucial intuition about the base-line example (72a) is this: if John makes a prediction that specifies an interval (e.g. "Bill will get between 80 and 90"), then the sentence is not true if Bill merely exceeds the lower bound of the interval, but only if he exceeds the upper bound. This shows that the **Π**-phrase is not out-scoping *predict*. If it did, we would effectively be computing the maximal score that Bill reaches or exceeds in every world compatible with John's prediction, and this would be 80 (the lower bound of the quoted interval). Now, what are the meanings of (72b) and (72c)? (72b) suggests that John made multiple predictions over some range of time. If each of these predictions specified a definite score – e.g. John says at one time "he'll get 82", at another time "he'll get 79", at yet another "he'll get 84" – then the intuition is that (72b) claims that Bill exceeded the maximum of these predicted scores (i.e., that he got more than 84). This works out fine with the **Π**-phrase scoping above *ever* and *predict*. But what if one or more of John's various predictions involved intervals? Suppose, e.g., he said "he'll get between 70 and 80" and at another time "he'll get between 75 and 85". I think that (72b) then claims that Bill got above 85, i.e., he scored above the upper bound of the union of all the predicted intervals. This, unforrtunately, is not what we predict if we have the scope-order **Π**-phrase > *ever* > *predict*. What we rather compute in that case is the maximum of the lower bounds of all the intervals (in this case, 75).

This piece of data is difficult to make sense of. *ever* unquestionably scopes over *predict*, so if **Π** is to license it as an NPI by scoping over it, it must also be scoping over *predict*. I can imagine only two types of responses to this dilemma. One is to look for a radically new theory of NPI-licensing which makes DEness irrelevant[26], or to question the NPI-status of *ever*. Perhaps there is, after all, an FC *ever* with a universal semantics, even though this is not attested anywhere outside of comparatives. The other response, which I am a bit more inclined to pursue, is to reconsider the semantics of *predict* (and attitude verbs more generally), so that the correct meanings come out even if **Π** scopes over *predict*.[27] I have no concrete idea so far about

---

[26] It is worth mentioning here that, unless **Π** scopes over it, *ever* is in an environment which not only fails to be DE, but is UE (i.e., not non-monotone) and moreover veridical, or at any rate not non-veridical. (These terms have not, to my knowledge, been formally defined for anything other than propositional operators. But by an obvious generalization of the existing definition, -er_* comes out as veridical.) Also, whatever we say about NPI-licensing, we would have to treat *ever* as a universal rather than an existential.

[27] I made a previous attempt in this direction in the SALT-proceedings version of Heim (2000a), but it was flawed. (This was pointed out to me by Roger Schwarzschild, and was the reason why I deleted this

how to make this work, but would like to conclude with an observation that indicates such a move may be called for independently of the analysis of comparatives and NPI-licensing. Suppose, as before, that John makes a prediction quoting an interval, such as "Bill will get between 80 and 90 points." Suppose further that Bill ends up getting 85 points. Knowing this, how do you assess the truth-values of the assertions in (73)?

(73) (a) John didn't predict that Bill would get 85 points.
(b) John didn't predict that Bill would do so well.
(c) Bill got 85 points. This was not predicted.

(73a) strikes me as true. John indeed didn't predict that Bill would get 85 points, because he only made a much less precise prediction. With (73b), however, I am much more inclined to judge it false. That is quite puzzling. *do so well* apparently means 'do as well as he did', and given what we presuppose in the context, this should make (73b) equivalent to (73a). (73c) seems to be somewhat in between. I can read it as true, like (73a), but I also see a way of taking it to be false. To the extent that the latter intuition is real, this is again quite puzzling.

I don't understand enough of what is going on here to draw any conclusions. But I have a suspicion that there is a connection between the unexpected readings of (73b, c) and our failure to get adequate truth-conditions for the sentences in (72) when we scope the Π-phrase over *predict*. And with this inconclusive remark I conclude this paper.

Appendix: The generalized-quantifier analysis (GQ) and S&W

(1) S&W, p. 26:
-er(P_{than})(Q_{matrix}) $\Leftrightarrow$ $[\mu I. [\mu K. Diff(I - K)] \in P] \in Q$
(a relation between two sets of intervals)

(2) *Diff* stands for the differential phrase. If there is no overt one, *Diff*(I - K) is equivalent to K > I.
Special case with null differential:
-er(P_{than})(Q_{matrix}) $\Leftrightarrow$ $[\mu I. [\mu K. K < I] \in P] \in Q$

The following addresses only the question of equivalence for examples without overt differentials. More will have to be said about the cases with differentials, especially about differentials with 'exactly'-interpretation.

definitions:

---

passage in the Stechow-Festchrift version). Not only was it flawed, it was also only meant to apply to neg-raising verbs, which *predict* is not.

(3) intervals:
A set of degrees D is an <u>interval</u> iff
for all d, d', d'': if $d \in D$ & $d' \in D$ & $d \leq d'' \leq d'$, then $d'' \in D$.
(If d'' is between two elements of an interval, it is itself in the interval.)

(4) <-relation among intervals:
$I < K$ iff for all d, d': if $d \in I$ & $d' \in K$, then $d < d'$.
(I is wholly below K.)

(5) μ-operator (S&W, p. 23)
$[\mu I.\ \phi[I]] := \iota I.\ \forall I'(I' \neq \varnothing \ \&\ I' \subseteq I \rightarrow \phi[I'])\ \&\ \forall I''(I \subset I'' \rightarrow \exists I'(I' \subset I'' \ \&\ \neg\phi[I']))$

(6) μ-operator, corrected[28]
$[\mu I.\ \phi[I]] := \iota I.\ \forall I'(I' \neq \varnothing \ \&\ I' \subseteq I \rightarrow \phi[I'])\ \&\ \forall I''(I \subset I'' \rightarrow \exists I'(I' \neq \varnothing \ \&\ I' \subseteq I'' \ \&$
$\neg\phi[I']))$
("the largest interval all of whose non-empty subintervals are ϕ")

abbreviations:

(7) $d < I$ abbreviates $\{d\} < I$.
$<I := \{d: d < I\}$
$<d := <\{d\}$
$P(I) := I \in P$

I will first remove a superficial difference in the semantic types for *-er* that GQ and S&W assume. In GQ, *er* + *than*-clause was interpreted in situ, S&W assume it to be moved to the edge of the matrix clause. Their semantics, however, effectively reconstructs it back. They could therefore just as well have assumed in-situ interpretation and the following meaning for *-er*.

An equivalent reformulation of (2):

(8) $\text{-er}(P_{than})(d_{matrix}) \Leftrightarrow d \in [\mu I.\ [\mu K.\ K < I] \in P]$

A further equivalent reformulation of (8):

(9) $\text{-er}(P_{than}))(d_{matrix}) \Leftrightarrow d \in [\mu I.\ P(<I)]$

---

[28] The prose formulation that S&W use throughout the paper suggests the definition in (6), where the second condition uses exactly the negation of the first. The absence of the non-emptiness requirement in the second clause of (5) seems not to be intended. It has the consequence that the μ-term is not always well-defined, e.g., in a situation where there exist people and *P* is the denotation of *than nobody is tall*. (Problem: the definition is met for smaller intervals than the intended one as well.)

(10) Proof that $[\mu K.\ K < I] = {<}I$:
Need to show that ${<}I$ meets both clauses in the definition of μ.
1st clause: Take $K' \subseteq {<}I$ and show that $K' < I$. Obvious.
2nd clause: Take $K' \supset {<}I$, define $K'' := K' - {<}I$, and show that not $K'' < {<}I$. By def., K'' contains a d' such that $d' \geq d''$ for some $d'' \in I$. This contradicts $K'' < {<}I$.

Now are my GQ analysis and (9) equivalent?

(11) Full equivalence would mean:
For all d_d and $P_{<dt,t>}$:
$P({<}d) \quad \Leftrightarrow \quad d \in [\mu I.\ P({<}I)]$

One direction is straightforward:

(12) If $d \in [\mu I.\ P({<}I)]$, then $P({<}d)$.

(13) Proof: Assume $d \in [\mu I.\ P({<}I)]$. I.e., $\{d\} \subseteq [\mu I.\ P({<}I)]$. By definition of μ, therefore $P({<}\{d\})$. And ${<}\{d\} = {<}d$. QED

For the other direction, I can currently prove special cases if I can assume a discrete scale:

(14) If P is mon↑, mon↓, or a conjunction of mon↑ and mon↓ properties, and if every non-empty interval contains a lowest point, then:
$P({<}d) \quad \Rightarrow \quad d \in [\mu I.\ P({<}I)]$

Proof: Assume $P({<}d)$. Define $I := \{d' : P({<}d')\}$ and show that this I fulfills (i) as well as the following three conditions required by the definition of μ:

(i) $d \in I$.
(ii) I is an interval.
(iii) $\forall I' [I' \neq \emptyset \ \&\ I' \subseteq I \rightarrow P({<}I')]$
(iv) $\forall I'' [I \subset I'' \rightarrow \exists I' [I' \neq \emptyset \ \&\ I' \subseteq I'' \ \&\ \neg P({<}I')]]$

Re (i): Follows directly from def. of I and premise that $P({<}d)$.

Re (ii)[29]: Take d_1, d_2, d_3 such that $d_1, d_3 \in I$ and $d_1 \leq d_2 \leq d_3$. It follows that $P({<}d_1)$ and t $P({<}d_3)$ and that ${<}d_3 \subseteq {<}d_2 \subseteq {<}d_1$.

1st case: P is mon↑. Then $P({<}d_3)$ and ${<}d_3 \subseteq {<}d_2$ imply $P({<}d_2)$.

2nd case: P is mon↓. Then $P({<}d_1)$ and ${<}d_2 \subseteq {<}d_1$ imply $P({<}d_2)$.

[29] This is the only part of the proof where monotonicity properties of P are used.

3rd case: $P = Q \,\&\, R$, where Q is mon↑ and R is mon↓. Then, by reasoning as in 1st and 2nd case, we get $Q(<d_2)$ and $R(<d_2)$. Thus $P(<d_2)$.

In all three cases, we have $P(<d_2)$, therefore $d_2 \in I$ by def. of I.

Re (iii): Let I' be a non-empty subset of I. Let d' be the minimal element of I'. Since $d' \in I$, we know $P(<d')$. By definition of d', we also know $<d' = <I'$. Hence $P(<d')$.

Re (iv): Let I'' be a proper superset of I. Define I' := I'' - I. (This is non-empty and a subset of I''.) Let d' be the minimal element of I'. Since I' is disjoint from I, we know $\neg P(<d')$. Therefore, since $<d' = <I'$, also $\neg P(<I')$.

References

Heim, I. (2000a) "Degree Operators and Scope," in SALT 10 proceedings and in Féry & Sternefeld (eds.) Audiatur Vox Sapentiae, Akademie-Verlag: Berlin

Heim, I. (2000b) handout of SALT 10 talk at Cornell

Kennedy, C. (1997) Projecting the Adjective, UC Santa Cruz Ph.D. thesis

Larson, R. (1988) "Scope and comparatives," L&P

Potts, Chris (2000). "When even *no*'s neg is splitsville," Jorge Hankamer's Web Fest: http://ling.ucsc.edu//Jorge/index.html

Rullmann, H. (1995) Maximality in the Semantics of Wh-Constructions, UMass Ph.D. thesis

Schwarzschild, R. & K. Wilkinson (2002) "Quantifiesr in comparatives: A semantics of degree based on intervals," NALS

Schwarzschild, R. (2004) "Scope-splitting in the comparative," handout from MIT colloquium, http://www.rci.rutgers.edu/~tapuz/pubs.htm

Stechow, A. v. (1984) "Comparing semantic theories of comparatives," JoS

Takahashi, S. (2006) "More than two quantifiers," NALS (forthcoming) and ch.4 of MIT Ph.D. thesis