How to be Universal when you are Existential: Negative Polarity Items in the Comparative: Entailment along a Scale

ALEX ZEPTER
Rutgers University

Abstract
Fauconnier (1975a) noticed that existential quantification, if it is related to a scale endpoint, can force entailment along the scale and as such have the effect of universal quantification: assume a partially ordered set \((X, \geq)\) and a predicate \(\phi\) such that for all \(x, y \in X, x \geq y\), if \(\phi\) is true of \(x\), it is also true of \(y\); then if there exists an element \(z\) that is ordered before all other elements and \(\phi(z)\) is true, then \(\phi\) is true for all elements in \(X\).

This paper claims that the clausal Comparative licences context sensitive Negative Polarity Items (NPIs) like English \(\textit{any}\) and \(\textit{ever}\), and German \(\textit{jemals}\) (‘ever’), due to its property of enabling entailment along a scale. I argue that the licensing condition of \(\textit{any}\) and \(\textit{ever/jemals}\) is not itself the ‘semantic scope of a downward entailing function’ (Ladusaw 1979), but rather that a downward entailing function provides one appropriate kind of context that satisfies the actual condition which is part of the NPI’s meaning (following Kadmon & Landman 1993; Krifka 1994, 1995; Jackson 1995): \(\textit{any}\) and \(\textit{ever/jemals}\) are indefinites that are licensed in particularly strong statements. The new idea is to define an assertion containing an indefinite NPI as strong if the existentially quantified formula entails a particular wide scope universal statement: assume that ‘Babajaga doesn’t see any tree’ means that it is not true that there exists some tree that Babajaga sees. This entails that for \(\textit{all}\) trees, Babajaga does not see them. In general, entailment will be granted if existential quantification is either in the scope of a downward entailing function or occurs in the context of a scale, the latter enabling entailment along the scale: if ‘Babajaga is smarter than any witch’ actually means that Babajaga is smarter than one of the smartest witches, then this entails that for \(\textit{all}\) witches, Babajaga is smarter.

‘Thus, we’d better not try to reverse the relations.’
Michel Foucault, \textit{Les Mots et les Choses}

1 INTRODUCTION
Negative Polarity Items (NPIs) are context sensitive expressions that are only licensed in a specific set of contexts. The set includes all
types of negation, which is reflected in the name ‘Negative Polarity Item’, but the scope of negation is not the only licenser; other possible environments are, at least on the surface, non-negative, such as the antecedent of a conditional, or the first argument of the quantifier every. (1–5) shows some examples containing the English NPIs any and ever, as well as German jemals (which is the correspondent of English ever):

(1) *Mo has any money.

(2) Andre doesn’t see any basketball player.

(3) If he ever sees her again, he will try to show her his collection of stamps.

(4) Every student who has any thought about NPIs gets crazy.

(5) Niemand wird jemals erfahren, was ich gerade gedacht habe. Nobody will ever know what I just thought have ‘Nobody will ever know what I’ve just thought.’

The licensed environment was identified by Ladusaw (1979: 132, 1980:12,13) as ‘the semantic scope of a downward entailing function’: a function is downward entailing if it reverses entailment relations or subset relations.

(6) Ladusaw (1979: 112):
An expression \([\delta] \) with the meaning \([\delta']\) is downward entailing iff
\[\forall x \forall y \ [x \subseteq y \Rightarrow [\delta'(y) \{ \Rightarrow \text{ or } \subseteq\} \delta'(x)]]\]

For illustration, take two expressions \(\alpha\) and \(\beta\) such that the meaning of \(\alpha\), that is \(\alpha'\), is a subset of the meaning of \(\beta\), \(\beta'\).

(7) \(\alpha' = [[\text{Heather sees a cute basketball player}]]\)
\(\beta' = [[\text{Heather sees a basketball player}]]\)

\(\alpha' \subseteq \beta'\): \([[\text{Heather sees a cute basketball player}]] \subseteq [[\text{Heather sees a basketball player}]]\)
If a function $f$ (as the meaning of an expression $\delta$) takes $\alpha'$, $\beta'$, as its argument and reverses the subset/entailment relation between $\alpha'$ and $\beta'$ such that $f(\beta') \subseteq f(\alpha')$, then $f$ is downward entailing. The meaning of sentence negation is a downward entailing function (8a), and so is the quantifier every for its first argument (8b; correspondingly, all licensors illustrated in (2-5) can be identified as downward entailing functions):

(8) a. $\alpha' = [[[Heather sees a cute basketball player]]];$
$\beta' = [[[Heather sees a basketball player]]]; f = [[[ not ]]]$
$\alpha' \subseteq \beta'$; but $f(\beta') \subseteq f(\alpha')$:  
[[Heather did not see a basketball player]] \subseteq [[[Heather did not see a cute basketball player]]]
Heather did not see a basketball player. $\Rightarrow$ Heather did not see a cute basketball player.

b. $\alpha' = [[[cute basketball player]]];$ $\beta' = [[[basketball player]]]; f = [[[ every ]]]$
$\alpha' \subseteq \beta'$; but $f(\beta') \subseteq f(\alpha')$: 
[[every basketball player]] $\subseteq$ [[[every cute basketball player]]];
therefore:
Every basketball player is out there. $\Rightarrow$ Every cute basketball player is out there.

Given Ladusaw’s approach, an NPI is only licensed if it occurs in the scope of an expression whose meaning is a downward entailing function; that is, the NPI’s meaning must be part of the meaning of the function’s argument.

Furthermore, Ladusaw (1979: 71ff) argues that the NPIs any and ever are existentials. So, the meaning of any is like the meaning of the indefinite a or some, the only difference is that the former, as an NPI, is restricted by the specific licensing condition.

(9) Andre doesn’t see any basketball player.
$\neg \exists x ([BASKETBALL PLAYER(x) \land SEE(a, x)])$
$\rightarrow$ It is not true that there exists at least one basketball player such that Andre sees this player.

(10) Every professor who has ever read a qualifying paper knows what’s fun.
$\forall y [[[PROFESSOR(y) \land \exists t [y \ read \ a \ qualifying \ paper \ at \ t]]} \rightarrow \ y \ knows \ what \ is \ fun]$ 
$\rightarrow$ Every professor who has at least one time read a qualifying paper knows what’s fun.
If Ladusaw is correct, we expect that all contexts that license NPIs are downward entailing functions, as well as that any, ever and German jemals are existential quantifiers.

Here is the subject of this paper: among the contexts licensing NPIs is the clausal comparative, in which any, ever, jemals are apparently interpreted universally.

(11) Hubi is taller than any student is.
    Hubi is taller than every student (where the students can have different heights).

(12) Today it is hotter than it ever was before.
    This day is hotter than all days in the past (where the days in the past can have different temperatures).

(13) German (jemals \(\approx\) ever):
    Sie hält es länger aus als er es jemals aushalten wird.
    She tolerates it longer (sep.prefix) than he it ever tolerate will
    ‘She tolerates it longer than he will tolerate it, for all times in the future’

Is this a problem? It depends. Observing the data, one might approach them differently:

(a) Do we want to say that any and ever/jemals are NPIs when they show up in the comparative? If they are not, then the comparative context is not relevant for the question if the licensing condition of NPIs is indeed a downward entailing function and if any, ever/jemals are existentials. We know that in English, there is Free Choice (FC) any, which is first, likewise context sensitive but restricted differently, and second, always interpreted universally. Thus, maybe there is also Free Choice ever, and in German FC jemals?

(b) If any, ever/jemals in (11–13) are NPIs, then the task to account for the data can either be a burden for the analysis of the NPIs as such or for the analysis of the comparative. Does the comparative falsify the claim that the licensing condition of NPIs is a downward entailing function and that any, ever/jemals are existentials? Or, on the contrary, does the meaning of a clausal comparative contain a downward entailing function, and furthermore does its semantics account for the universal interpretation of an ‘underlying’ existential quantifier?
To choose the first line of reasoning seems unattractive, for the following two reasons:

First, *any* and *ever/jemals* are not the only NPIs that are licensed in the comparative; we can also find others, for example idiomatic expressions:

(14) Rullmann (1995: 64):

a. He told me more jokes than I cared to write down.
b. He said the sky would sooner fall than he would budge an inch.

Second, at least in German, the comparative seems to be the only context in which *jemals* is interpreted universally (compare the data in Kürschner (1983: 214–249)). Thus, claiming that *jemals* is ambiguous between NPI and FC would leave us with the unpleasant task of having to justify why the comparative is the only possible licenser for FC-*jemals*.

The second line of reasoning branches into two: if one wants to maintain all of Ladusaw’s proposal, then it is necessary to provide an analysis of the comparative that makes it downward entailing and that accounts for the universal reading of *any, ever/jemals*. If *any, ever/jemals* are NPIs but Ladusaw is only partly right, then the goal would be to revise the analysis of the NPIs as such, including the licensing condition.

This paper will ultimately decide on the last option. However, before we do this, the next section will take a closer look at the clausal comparative and its possible semantic analyses. That is, before we reject the strategy to show that the comparative is downward entailing, we want to know in greater detail why such rejection is justifiable and based on more than just conceptual taste. As a brief outline, below is a sketch of the ultimate proposal (given in Sections 3 and 4):

- The paper proposes a revision of the NPI-meaning and the NPI-licensing condition: the licensing condition should be directly related to the meaning. A downward entailing function will be a licenser, because it creates an environment that satisfies the licensing condition, but it is not itself ‘the’ condition (in the spirit of Kadmon & Landman 1993; Krifka 1994, 1995; Lahiri 1998).

- The NPIs *any, ever* and *jemals* are indefinites and thus their semantic structure contains (narrow scope) existential quantification. I will follow Heim (1982) in analyzing indefinites as non-quantificational expressions that just carry an open variable and undergo existential closure, though this is not crucial to the proposal. What is crucial is
that the NPIs differ from ‘regular’ indefinites like a in the following way:

- First, any and ever/jemals contain a function of type $\langle\langle e, t \rangle, \langle e, t \rangle \rangle$ that takes a set $P$ and gives back either the entire set or a subset; that way, the function can manipulate the size of the set that undergoes existential closure. Thus for example, the indefinite a tree differs from any tree in that for the former, existential closure always applies to the entire set of trees, whereas for the latter, existential closure might apply only to a subset of all trees, dependent on the choice of the NPI-function.

- Second, the choice of domain restriction is monitored by the licensing condition. In contrast to ‘regular’ indefinites, NPIs are used if we want to make a particularly strong statement (following Kadmon & Landman 1993; Krifka 1994, 1995). The licensing condition proposed by this paper is that NPIs are licensed if contained in an assertion, an expression of type $t$, that entails a wide scope universal statement about the entire set $P$.

- The NPIs are not licensed in all contexts, because even if the NPI-function has the power to restrict the domain differently, not all environments are of the right kind such that an appropriate restriction can be made. We will distinguish two types in which the licensing condition is satisfiable:

- First, in the scope of a downward entailing function, the NPI-function chooses the widest domain such that the entire statement is logically equivalent to or entails a wide scope universal statement.

- Second, in the scope of a degree quantifier and the context of a scale, the NPI-function zooms in on the domain corresponding to a scale endpoint such that existential quantification over that endpoint will entail universality by entailment along the scale (relevance of scale cf. Fauconnier 1975a): in the context of a comparison, all elements of a set can be ordered relative to the scale, and a comparative that is true for the elements ordered highest (or lowest) will be true for all other elements as well.

2 IS THE CLAUSAL COMPARATIVE DOWNWARD ENTAILING?

Is there a simple way to find out if the meaning of the comparative marker is a downward entailing function? No, for the following reason:
‘downward entailing function’ is a theory-internal concept of semantic representation. Thus, the question if a construction is downward entailing depends on the semantic representation we choose.

2.1 Degrees, maximality and entailing downwards

A detailed proposal that identifies the clausal comparative as downward entailing is Rullmann (1995), which follows Stechow (1984). The analysis is based on the concepts of ‘degree’ and ‘maximality’. In a nutshell: the comparative compares degrees of adjective-ness; for example:

(15) a. Babajaga is taller than Mo is.
     b. \( \exists d [TALL(b, d) \land d > max(\lambda d'[TALL(m, d')])] \)
        There exists a degree such that Babajaga stands in the tall-
        relation to this degree and this degree is greater than the
        maximal degree in the set of degrees \( d' \) such that Mo stands in
        the tall-relation to \( d' \).
     c. ‘-er’ translates as: \( \lambda R \lambda d' \lambda x [\exists d [R(d)(x) \land d > d']] \)
     d. ‘\([CP \, OP_i \, [IP \, Subj \, is \, d_i \, -Adj]]\)’ translates as: \( max(\lambda d'[Adj' \langle Subj', d' \rangle]) \)

In (15a,b), we compare the degree to which Babajaga is tall with the degree to which Mo is tall. Crucially, the degrees correspond to points on a scale. If Babajaga’s degree of tallness is greater than Mo’s maximal degree of tallness, then the entire comparative clause is true. In general, the clause embedded under than (I will call it from now the ‘subordinated clause’) denotes the unique maximal degree of a set of degrees (15d; cf. Rullmann 1995: 59); the set of degrees are all degrees that can make the subordinated clause true. Given the maximality/degree-analysis, the clausal comparative is downward entailing, because:

    If \( D \subseteq D' \), then \( \{x | x \text{ is more } Adj \text{ than } max(D') \} \)
    \( \subseteq \{x | x \text{ is more } Adj \text{ than } max(D) \} \)
MMD (my abbreviation)

Take two sets of degrees, \( D \) and \( D' \), one the subset of the other. The subset relation is reversed if we consider the sets of individuals whose degree of adjective-ness is greater than the maximal value of \( D \) and \( D' \)
respectively: $\text{MMD}'$ is a subset of $\text{MMD}$, because if $D'$ is a superset of $D$, then the maximal value of $D'$ can only be greater or equal to the maximal value of $D$; thus, all individuals that are in $\text{MMD}'$ are also in $\text{MMD}$.

Note that one might want to ask Rullmann what exactly is the downward entailing function that should be the meaning of the degree quantifier ($\rightarrow$ -er, more, less).¹ According to Stechow (1984: 55) and Rullmann (1995: 59), the subordinated clause denotes a unique maximal degree. But this maximal degree is not the argument of the downward entailing function: the subset relation that reverses concerns the set of degrees that make the subordinated clause true; only two sets of degrees can have a subset relation, their two respective maximal degrees never can. Hence, the argument of the downward entailing function is the degree set. Consequently, the meaning of the degree quantifier (given in 15c; cf. Rullmann 1995: 59) cannot be the downward entailing function: the argument of the degree quantifier is a single degree (the maximal degree denoted by the subordinated clause) and not the set of degrees. But, then, the downward entailing function corresponds to the expression ‘more Adj than max’ which includes a ‘part’ of the meaning of the clause embedded under than. Finally, the meaning of the argument of the function, that is, the set of degrees, also does not correspond to the meaning of the subordinated clause as such; the clause would denote a proposition, this is a set of worlds or a truth value.

Ignoring those details, the analysis does not only make the comparative downward entailing, it also accounts for the universal interpretation of any and ever/jemals. Here, the maximality operator bears fruit:

(17) The gardener is taller than any servant is.

The gardener is taller than $\text{max}(\lambda d' [\exists x [\text{SERVANT}(x) \wedge x \text{ is } d'-\text{tall}]])$

If the subordinated clause denotes the maximal value of all degrees that make the sentence ‘a servant is $d'$-tall’ true, then it denotes the degree of the tallest servant. Thus, if the gardener is taller than that one, then he is taller than all of the servants.

But wait a minute. If universality is due to the maximality operator which is built into the semantics of the comparative, then what about existential quantification of other indefinites that are not NPIs?

(18) a. For sure, the gardener is taller than some servant is.

¹ Thanks to Karina Wilkinson and Roger Schwarzschild for making me aware of this question.
b. Wir haben sogar länger gewartet als Patienten we have even longer waited than patients gewartet haben, die später kamen. waited have who later came ‘We waited even longer than some of the other patients who came later’.

If in (18), the maximality operator has also scope over the existential quantifiers that bind \( x \) of \textit{servant} and \textit{Patienten}, then (18a) would mean the same as (17), and (18b) that ‘we waited longer than \textit{all} of the patients who came later’. We don’t want that. In order to get the right readings, an analysis based on maximality has to assume that ‘true’ existential quantifiers (those that should be interpreted existentially) must take scope over the maximality operator.\(^2\)

Syntactically, the \textit{than}-clause is an island for movement such that quantifier raising shouldn’t be possible. However, as shown by Lerner & Pinkal (1991: 334ff), syntactic islandhood may not be a decisive argument if one considers a shorter movement as opposed to raising the quantifier all the way up to the top of the main clause. The more crucial problem is intrinsic to the semantic analysis: how do we explain that NPI-existentials are interpreted in situ but other existentials must be scoped out?

The problem becomes worse where we consider other quantifiers. As shown by Schwarzchild & Wilkinson (1999: 7ff), Schwarzchild & Wilkinson (2002: 7ff), degree-analyses in general (not only those that work with ‘maximality’) need to QR most quantifier phrases in order to predict the correct meanings. The maximality-approach not only needs an obligatory rule of QR for ‘true’ existentials but also for non-monotonic quantifiers (e.g. \textit{exactly} \( x \)). It furthermore has to raise universal quantifiers, or else the comparative allows only one particular interpretation.\(^3\)

Couldn’t we simplify the problem by assuming that all quantifier phrases have to leave the \textit{than}-clause? No, because then we would

\(^2\) One could argue that \textit{some} has to escape the \textit{than}-clause precisely because it cannot occur in the scope of a downward entailing function, following Ladusaw’s classification of \textit{some} as a positive polarity item (Ladusaw 1979: 69ff). However, this doesn’t yet explain the obligatory raising of the indefinite \textit{Patienten} in (18b).

\(^3\) Take for example ‘Hubi is taller than every student is’. Here, it is natural to assume that all students have in fact different heights but Hubi is still taller than all of them. Given a representation ‘Hubi is taller than max(\(\lambda d.\forall x[\text{STUDENT}(x) \land x \text{ is } d\text{-tall}]\))’ , the set of degrees from which the maximality operator chooses the highest value contains only degrees such that all students are equally \( d \)-tall. But this means ‘all students have the same height and Hubi is taller than that’. For a more detailed discussion of the scope-problems raised by degree-analyses, see Schwarzchild & Wilkinson (1999: 5ff).
lose the NPI-explanation. Similarly, Rullmann (1995: 60ff), following Stechow (1984: 72), wants to interpret negative quantifiers in situ, in order to account for the fact that they are banned from the subordinated clause. According to Stechow and Rullmann, comparatives like ‘*Babajaga weighs more than nobody weighs’ are ungrammatical, because the maximality operator is unable to find a maximal value in the set $\lambda d[\text{nobody weighs } d\text{-much}]$. But why doesn’t the negative quantifier simply escape the subordinated clause just as universal and existential quantifiers do?

Altogether, the maximality-approach manages to identify the clausal comparative as downward entailing, but it pays a price by requiring a QR-distribution that seems hard to justify. That this is crucial to the question of whether the comparative is downward entailing becomes obvious when we look at an alternative proposal which solves the QR-problem.

2.2 Intervals and entailing upwards

In Schwarzschild & Wilkinson (1999, 2002), all quantifier phrases are interpreted in situ. This is accomplished by analysing scalar predicates not as relations between individuals and degrees, i.e. exact points on a scale, but rather as relations between individuals and intervals. Therefore, a sentence such as:

(19) Every girl in my class who had a date with Julio is rich.

= Every girl in my class who had a date with Julio is K-rich.

can be true even if the girls in question don’t have exactly the same amount of money nor degree of fortune; rather there has to be an interval K of richness that covers them all.¹

Turning to the clausal comparative, it compares intervals on the scale provided by the scalar predicate. Crucially, in the understanding of Schwarzschild & Wilkinson, the subordinated clause neither denotes a specific maximal degree nor a specific maximal interval. Rather, it denotes a predicate of intervals, that is, a function which takes an interval as its argument. Consider the following example:

(20) The butler is a lot richer than everybody else in the house is.

= The butler is a lot richer than everybody else in the house is K-rich.

¹ Imagine the girls are all daughters of millionaires, but some of those millionaires have one million, some three million and so on; we could find an interval K that covers all individuals who are millionaires such that all are millionaire-rich if covered by K, even if under closer inspection, their respective fortune differs.
I = [[a lot -er than everybody else in the house is K-rich ]] 

\[ \text{RICH}(b)\{b\text{ is covered by } I \text{ on the scale of richness; false otherwise.}\] 

First of all, it is natural to assume that all the individuals that constitute ‘everybody else in the house’ have in fact different amounts of fortune, but the butler is still a lot richer than each of them. Imagine Duke Singur, Duke Greedon and Duke Little Smart are ‘everybody else in the house’; Singur has $1,000, Greedon has $5,000, and Little Smart has $500 saved up. Then, for (20) to be true, the butler should be a lot richer than the richest of the Dukes, which is Greedon (maybe, the butler is a Duke himself, but he is so greedy that he works as a butler in order to avoid spending). Only then, would it be true that the butler is a lot richer than everybody else. But, crucially, Greedon, Singur and Little Smart do not need to have the same amount of fortune.

Now, according to Schwarzschild & Wilkinson 1999, 2002: 18ff, 26, the entire clausal comparative is true if the (set of) individual(s) denoted by the main clause subject, thus the butler in (20), is, on the scale of the scalar predicate, here the richness scale, covered by a particular interval, call it the comparative interval I. In any clausal comparative, the portion ‘differential MORE/-er than S’ denotes I. (Note that the differential is A LOT in (20)).

Hence, for (20), if the butler’s richness corresponds to a point in I, then (20) is true; false otherwise. But, how does one generally compute I? Here is a way:

In general, we can think of an interval as a collection of points. As such, we can understand the determination of the comparative interval as the (randomly operating) test of checking points on the scale provided by the scalar predicate – such that each point satisfies a particular condition, and for that reason must be part of I. The test is repeated until I is determined. Necessary and sufficient for the question of whether a point \( p \) is part of I or not is the following:

Let us work with the example ‘the butler is a lot richer than everybody else in the house is’. Thus, we pick an arbitrary point \( p \) on the scale of richness (see (21)).

\[ \text{5 The terminology stems from van Stechow (1984: 4). For Schwarzschild & Wilkinson (2002: 15ff), differentials measure parts of the scale provided by the scalar predicate. Hence, A LOT in (20) denotes a predicate of intervals, true if the interval cuts a large portion on the richness-scale.} \]

\[ \text{6 I am grateful to Roger Schwarzschild who pointed out (in pc.) the following possibility of computing the comparative interval I in a way that is, on an intuitive level, the easiest to grasp. Any error and missing transparency caused by the presentation in the text is certainly mine.} \]
Then, we find all points \( x \) below \( p \) whose distance is given by the differential. That is, in the current case, each \( x \) must be A LOT distant from \( p \). All \( x \) will constitute together an interval \( K \), which is maximal in the sense that all points outside \( K \) are not A LOT below \( p \). For \( p \) in (21), \( K \) will start at point zero of the scale, and will end maybe a little above Greedon; keep in mind that the upper end point of \( K \) must still be A LOT below \( p \). Now, understanding the \textit{than}-clause as a predicate of intervals, we can define \( \lambda K \) [everybody else in the house is \( K \)-rich]. If it is indeed true that \( K \) covers everybody else, that is, if \( K \) covers Greedon, Little Smart and Singur, then \( p \) is a point inside \( I \). If it is not true that \( K \) covers all the Dukes, then \( p \) is outside \( I \). In the current scenario, \( p \) is in \( I \). We then probe another point \( q \) in the same way, and so on.

As said, the test runs until we find the comparative interval \( I \), such that the following holds for \( I \):

(i) For every point \( p \) inside \( I \), the \textit{than}-clause-predicate is true of measure(\( p \)).

(\( \rightarrow \) measure(\( p \)) is the interval \( K \) which exhaustively comprises all points which are A LOT below \( p \); since \( K \) covers ‘everybody else’, the \textit{than}-clause is true, and therefore \( p \) is inside \( I \).)
(ii) For every point $n$ outside $I$, the \textit{than}-clause-predicate is \textit{not} true of measure($n$).

($\rightarrow$ measure($n$) is the interval $K$ which exhaustively comprises all points which are A LOT below $n$; but $K$ does not cover \textit{everybody else}, hence the \textit{than}-clause is not true, and therefore $n$ is outside $I$.)

Finally, if the butler’s richness is covered by $I$, then the comparative ‘the butler is a lot richer than everybody else is’ is true; false otherwise.

Let us take a second example, in order to make sure that we understood the general logic. It might have become clear already that the computation of the comparative interval demands multiple constructions of $K$-intervals, all of them have to verify the \textit{than}-clause-predicate. Furthermore, for the previous example in (20), I will have started in a distance of A LOT above Greedon and extend all the way up to the top of the scale. Consider now (22), and otherwise keep the outlined scenario of the three Dukes:

(22) The butler is richer than exactly two of the Dukes are.

$=$ The butler is richer than exactly two of the Dukes are $K$-rich.

$I = [[$ \textit{some -er than exactly two of the Dukes are K-rich }$]]$

Intuitively assessed, we can say that $I$ will now start slightly above Singur and will stop slightly below Greedon. But let us go stepwise as before. Once more, we pick an arbitrary point $p$ on the scale of richness; for sake of simplicity, consider another time $p$ as in (21). Then, we find all points $y$ below $p$ whose distance is given by the differential. In (22), there is no overt differential, but be aware that in the absence thereof, Schwarzschild & Wilkinson (2002: 16f, 19) assume the presence of an implicit (existential) differential, to be written as SOME; the length of SOME is contextually specified. Thus, each $y$ must be at least SOME below $p$. All $y$ will constitute together an interval $K$, which is the largest we can construct; all remaining points on the scale, outside $K$, are not at least SOME below $p$. Now, with (22) as our comparative at stake, then for $p$ in (21), $K$ will start at point zero of the scale, and will end slightly below $p$; here, the upper end point of $K$ must still be at least SOME below $p$ (like all other points in $K$). This time, the \textit{than}-clause gives us $\lambda K$ [exactly two of the Dukes are $K$-rich]. If it is true that $K$ covers exactly two of the Dukes, then $p$ is a point inside $I$. If it is not true, then $p$ is outside $I$. In the current scenario, $p$ is \textit{not} in $I$. We continue probing points, until we have determined the entire comparative interval $I$, such that the following holds for $I$: 
For every point \( p \) inside \( I \), measure \((p) = \text{the interval } K \text{ which exhaustively comprises all points that are at least SOME below } p \) covers exactly two of the Dukes, and therefore verifies the than-clause-predicate.

For every point \( n \) outside \( I \), measure \((n) = \text{the interval } K \text{ which exhaustively comprises all points that are at least SOME below } n \) does not cover exactly two of the Dukes, and therefore falsifies the than-clause-predicate.

Finally, once again, if the butler’s richness is covered by \( I \), then the comparative ‘the butler is richer than exactly two of the Dukes are’ is true; false otherwise.

Before Schwarzschild & Wilkinson’s ultimate formula of the clausal comparative is given in (23) below, the following should be noted. Above, we were constructing the comparative interval by checking points, and collecting points that are \( \text{DIFF} \) (\( = \text{distance given by the differential} \) below them, and gathering the latter into K-intervals. Now, just as well, we could in fact describe the comparative interval by referring only to intervals and parts thereof.\(^7\)

That is, we can say: \([\text{DIFF more/-er than } S]\) denotes the comparative interval \( I \), such that for each non-empty subpart \( I’ \) in \( I \) (\( \sim \text{for each point in } I \) holds: \( K \), which is the largest interval such that each subpart of \( K \) (\( \sim \text{each point in } K \) is \( \text{DIFF} \) below \( I’ \), verifies the than-clause predicate:

\[
\text{Mn}(\mu I’[\text{Sub}(\mu K’[\text{DIFF}(I’-K’)])])
\]
- \( \mu I’[\text{Sub}(\mu K’[\text{DIFF}(I’-K’)])] \) is the comparative interval \( I \), such that for each non-empty subpart \( I’ \) of \( I \), \( I’ \) holds:
  
  The maximal interval \( K \) consisting only of parts \( K’ \), such that \( K’ \) is \( \text{DIFF} \) distant from \( I’(K = \mu K’[\text{DIFF}(I’-K’)]) \),
  
  \( K \) verifies the than-clause predicate \( \lambda K \ [x \text{ is } K-\text{ADJ}] \).

- \( \text{Mn}(\mu I’[\text{Sub}(\mu K’[\text{DIFF}(I’-K’)])]) \) is true iff for ‘\( x \) is \( \text{DIFF more ADJ} \) than \( S’ \), \( \text{ADJ}(x)(\mu I’[\text{Sub}(\mu K’[\text{DIFF}(I’-K’)])]) \) is true.

For more details on the Interval analysis, see Schwarzschild & Wilkinson (1999, 2002). For our purposes, it is at last sufficient to

\(^7\) Schwarzschild & Wilkinson’s original proposal avoids the reference to points entirely, since such points do lastly not exist within the Interval analysis. Thinking in points just helps us to find easier access to an understanding of the theory.
understand that there is no maximal interval itself denoting the meaning of the subordinated clause, nor is there a set of them. Rather, only the portion \( [[DIFF \text{ more/-er than } x \text{ is } K-\text{ADJ}]] \) equals an interval, which is the comparative interval we have labeled I above. Furthermore, I itself is a saturated function. That is, the degree quantifier (comparative marker) \text{ more/-er} denotes a function, which takes two arguments, one the \text{ than}-clause, the other the differential; saturation of the function gives us I (cf. Schwarzschild & Wilkinson 1999: 22, Schwarzschild & Wilkinson 2002: 27). But, crucially, the two respective arguments of the degree quantifier, the \text{ than}-clause and the differential, denote functions themselves, functions that take an interval as their respective argument. Thus, the \text{ than}-clause does not denote an interval itself, neither maximal nor any other, rather, it denotes a predicate of intervals.

Going back to the topic at stake, what about our question if the meaning of comparative marker is a downward entailing function? According to Schwarzschild & Wilkinson (1999: 24ff), the Interval analysis which interprets all quantifiers \textit{in situ} does not identify the comparative as downward entailing. Take two clausal comparatives that contain quantifier phrases and there is an entailment relation between the two subordinated clauses:

\begin{align*}
(24) \ a. \ & \text{Babajaga is much richer than exactly seven of my relatives were.} \\
& b. \text{Babajaga is much richer than at least four of my relatives were.} \\
& a'. \ B_1 = [\text{much -er than exactly seven of my relatives were } K\text{-rich}] \\
& b'. \ B_2 = [\text{much -er than at least four of my relatives were } K\text{-rich}] \\
(25) \ a. \ & \text{Sub1: ‘Exactly seven of my relatives were } K\text{-rich’} \\
& b. \text{Sub2: ‘At least four of my relatives were } K\text{-rich’}
\end{align*}

Sub1 entails Sub2, because if exactly seven of my relatives were \( K\)-rich, then at least four of my relatives were also \( K\)-rich (an interval \( K \) that covers the richness of exactly seven will also cover the richness of at least four). However, ‘Babajaga is much richer than exactly seven of my relatives were \( K\)-rich’ still entails ‘Babajaga is much richer than at least four of my relatives were \( K\)-rich’. Given the interval analysis, that (24a) entails (24b) follows from the subpart-relation between the two intervals \( B_1 \) and \( B_2 \) (which have to cover Babajaga’s richness in order to verify the corresponding comparatives): \( B_1 \) is a subpart of \( B_2 \) such that if Babajaga is covered by the former, she is also covered by the
latter (which means if (24a) is true, (24b) is true too; for details and proof, see Schwarzschild & Wilkinson 1999: 25). But then, assuming Schwarzschild & Wilkinson, is the clausal comparative in the end upward entailing? Comparing the entailment relations between the subordinated clauses and the entire comparative clauses, and interpreting quantifier phrases in situ, the comparative is apparently not downward entailing. But wait—what about subordinated clauses without quantifiers? According to Rullmann (1995: 65), comparatives like the following support the claim that the comparative is downward entailing, and thus call for an analysis that captures this (similar examples can be found in Linebarger (1987: 378), who follows Hoeksema’s (1983) analysis of a downward entailing clausal comparative).

(26) a. He runs more often than Mo swims.
   b. He runs more often than Mo swims the channel.

   a'. Sub1: ‘Mo swims d/K-often’
   b'. Sub2: ‘Mo swims the channel d/K-often’

(26a) entails (26b); if we ask Rullmann, this fact follows under the assumption of maximality and degrees: the maximal degree that verifies ‘Mo swims d-often’ can only be greater or equal to the maximal degree verifying ‘Mo swims the channel d-often’. Thus, if the frequency-degree of his running is greater than the former, it will be greater than the latter. However, entailment also follows if we assume intervals instead of degrees: by the same token, an interval that covers the frequency of Mo’s swimming can only be equal to or higher on the scale than an interval covering Mo’s swimming through the channel; in either case, the interval denoted by ‘[more than Mo swims K-often]’ will then be a subpart of the interval denoted by ‘[more than Mo swims the channel K-often]’ such that if his running is covered by the former, it will be covered by the latter.9

So far, so good; but is it enough to show that (26a) entails (26b)? Recall that a function is downward entailing if it reverses the

---

8 Try to grasp the necessity of the subpart relation on an intuitive level: The interval B1 (which has to cover Babajaga in order to yield truth of the comparative) must start on the scale of richness above the richness of exactly seven relatives, these seven having the smallest fortune of all of my relatives, and has to end below the 8th richest. B2, on the other hand, has to start above the richness of the four poorest relatives, and will extend all the way up to the top of the scale. Therefore B1 (independent of how far it extends) must be a subpart of B2.

9 The comparative interval denoted by [more than Mo swims K-often] is necessarily a subpart, since it necessarily starts higher on the frequency scale, while at the same time both comparative intervals extend up to the top of the scale.
entailment direction of its arguments. Given the maximality analysis, the comparative reverses the subset relation of the sets of degrees that verify the subordinated clauses. Thus for (26), the set \( D_1 \) containing all degrees that make ‘Mo swims \( d \)-often’ should be a superset of the set \( D_2 \) containing all degrees that verify ‘Mo swims the channel \( d \)-often’; then, the comparative function would reverse the relation, since ‘\( \{ x \mid x \text{ runs more often than max}(D_1) \} \)’ would be a subset of ‘\( \{ x \mid x \text{ runs more often max}(D_2) \} \)’. Therefore, what we need to check is if \( D_1 \supseteq D_2 \).

Intriguingly, \( D_1 \) must be a superset of \( D_2 \) only if we assume ‘at least’-readings; that is, if ‘Mo swims \( d \)-often’ means ‘Mo swims at least \( d \)-often’. There is no superset relation if it means, as intended by Rullmann, ‘Mo swims exactly \( d \)-often’.

Just imagine a model in which Mo swims the channel exactly two times a week, but exactly six times a week in total. In this case, \( D_1 \) of Sub1 is \( \{ 6 \} \), and \( D_2 \) of Sub2 is \( \{ 2 \} \), with \( D_2 \not\subseteq D_1 \). In general: if there is no subset relation to start with, there is also no reversal of that subset relation, and thus (26) doesn’t show that the comparative is downward entailing.

At this point, we might want to pause and ask the following question: even if a downward entailing function is a theory internal concept of semantic representation, is there really no theory-neutral way to find out if the comparative is downward entailing? Obviously, on the one hand, in order to determine if a function is downward

\[ \lambda d' [\text{Mo swims } d \text{-often}] = \{ 0, 1, 2, 3, 4, 5, 6 \} \]

\[ \lambda d' [\text{Mo swims the channel } d \text{-often}] = \{ 0, 1, 2 \} \]

Note though that such recognition might jeopardize Rullmann’s analysis of less-than-comparatives just as well (the intent to capture them being the reason for why he must reject the ‘at least’-approach in the first place). That is, Rullmann’s (1995: 77) translation of ‘John is less tall than Ben is’ is:

(i) John is less tall than Ben is = \( \exists d [\text{TALL}(j, d) \land d < \text{max}(\lambda d' [\text{TALL}(b, d')]) ] \)

Imagine that John is exactly 1.90 m tall, thus \( d = 1.90 \text{ m} \), and assume that, by monotonicity, \( \text{TALL}(j, 0.50 \text{ m}) \) is true as well. Now, if Bill is only 1.60 m, then Ben is in fact less tall than John, and not vice versa. Still, there exists a degree \( d \), such that John is \( d \)-tall, and this degree is smaller than the maximal degree to which Bill is tall. This wrongly predicts that (i) should be true in the current scenario.
entailing, we have to know what is the function and what is the argument; but, on the other hand, we’ve just said that ‘downward entailing’ must correspond to a reversal of entailment relations (subset relations respectively). Why can’t we simply assume that the subordinated clause should be the argument, and then compare the entailment relations between different subordinated clauses with the entailment relations of the corresponding comparative-clauses? Wouldn’t this give us a method to determine if the comparative is downward entailing independent of the particular analysis we adopt?

We have just seen two examples for why this is difficult. First, if one theory assumes that most quantifiers have to leave the *than*-clause, then the demonstration that with quantified subordinated clauses, there is no reversal, is either irrelevant or not theory-neutral (since it assumes that quantifiers are interpreted in situ). On the other hand, if there seems to be an entailment relation between two comparative clauses lacking quantifiers, but, under closer investigation, there is no entailment relation between the two corresponding subordinated clauses, or better, entailment depends on which theory we assume, then there is no ‘theory-neutral’ way to decide either. To put it the other way around: just in order to determine the entailment relation between two subordinated clauses, we already have to choose a semantic representation; but for this, we have to decide if we want intervals or degrees, an ‘at least’- or ‘exactly’-approach, and so on. Finally, we need to be cautious about what we mean with ‘subordinated clause’. Notice that the clause embedded under *than* has a different meaning compared with its occurrence as a main clause:

\[(27)\]

a. The butler is attractive.

b. The gardener is more attractive than the butler is.

\((27a)\) is only true if the butler *is* in fact attractive, more attractive than the average of person. But \((27b)\) can be true even if the butler is in fact quite unattractive; and even if the gardener is in fact rather unattractive, all that matters is that he looks still better than the butler does.

The general lesson: there is no meaning of the subordinated clause independent of the comparative (if we want to push it: 'the meaning of the subordinated clause only *exists* in the comparative'); and in turn there is no meaning independent of a specific analysis of the comparative.

One last attempt to introduce ‘theory-neutral’ evidence: Rullmann (1995: 65ff; following Stechow 1984: 72) highlights that comparatives which embed the modal *can* directly support the claim that the
semantics of the comparative contains a maximality operator, which in turn causes downward-entailing-ness:

(28)

\[ \begin{align*}
& a. \text{Obelix is stronger than a soldier can be.} \\
\Rightarrow & b. \text{Obelix is stronger than a Roman soldier can be.}
\end{align*} \]

If can only introduces existential quantification over possible worlds, then why do we want to interpret (28) as ‘Obelix is stronger than a (Roman) soldier maximally can be’ such that (28a) entails (28b)? Similar to Stechow’s/Rullmann’s approach to existential NPIs, analysing the degree quantifier as the locus of maximality seems to provide the right interpretation. Just as any and ever/jemals generally obtain a universal reading, so must Obelix be stronger than soldiers in all possible worlds. However, as already noticed by Schwarzschild & Wilkinson (1999: 29), this does not necessarily tell us something about the semantics of the comparative but could also be a sign for a negative polarity reading/variant of the modal can. In the latter case, can should be treated on a pair with other NPIs, and rather than supporting entailment reversal, (28) would strengthen the question of why, in the comparative, existential NPIs can receive a universal interpretation.

Altogether, there is no ‘theory-neutral’ way to determine if the comparative is downward entailing or not. Rather, we have to decide which theory we want to adopt and then see if it analyses the meaning of the comparative marker as a downward entailing function or not.12

The final result of this section: the question of whether the comparative is downward entailing or not depends on the specific analysis we adopt. If we want to assume Stechow (1984) and Rullmann (1995) such that we can maintain the claim that the comparative is downward entailing, then we have to accept the difficulties of identifying the function and its argument, and we have to face the problem of explaining why most quantifier phrases must leave the than-clause but some must stay in situ. On the other hand, if we want to adopt Schwarzschild & Wilkinson (1999, 2002), then, the meaning of

\[ \text{12 Besides comparing the entailment relations of quantified subordinated clauses with those of the corresponding comparative clauses, we haven’t said much about the question of whether the interval analysis provides a function whose argument is indeed the subordinated clause, or if it analyses some other function such that we could test whether this function is upward or downward entailing. To take a short cut: If the meaning of the comparative is indeed } \text{Mn}([\alpha]\text{F}[\text{Sub}[\mu K'[\text{DIFF}([f-K])]]]), \text{ then no part of the meaning seems to correspond to a downward entailing function. It is possible to define a downward entailing function as part of the computation; it is even possible to show a reversal in the part-of-relation of the intervals that are involved in the verification process of two comparative clauses that are related by entailment. But all those ‘semi-victories’ don’t give us the home run: Ladusaw’s original proposal was that the NPI’s meaning has to occur in the semantic scope of a downward entailing function, where this downward entailing function is the meaning of some expression } x. \text{ Given that in its strict definition, the comparative, under the Interval analysis, doesn’t fit.} \]
the comparative marker is not a downward entailing function. Drawing
the loop back to where we started, this also means that all existential
quantifiers stay in situ and are interpreted existentially. Hence, the
Interval analysis has to face the following questions:

(a) If NPIs are licensed in the clausal comparative but the meaning
of the comparative marker is not a downward entailing function,
what is the actual licensing condition of NPIs?
(b) If the NPIs any and ever/jemals are existentials, why are they
interpreted universally if embedded in the comparative?

No matter which theory we choose, the current carpet of explanations
seems too small to cover the room. However, considering the
overall difficulties of squeezing the comparative into the picture of a
downward entailing function, one might want to take this as evidence
that NPI-licensing involves more than that. We are now ready to ask
this question.

3 NPIS ARE FUNCTIONS THAT RESTRICT THE DOMAIN
OF (EXISTENTIAL) QUANTIFICATION

If we step back for a second: Isn’t it actually odd to assume
that all that distinguishes the NPIs any and ever/jemals from other
indefinites/existentials is the fact that they have to occur in the scope of
a downward entailing function? How is a downward entailing function
related to the meaning of the NPIs? Shouldn’t we ultimately prefer a
theory which answers this question?

Kadmon & Landman (1993), and Krifka (1994, 1995) suggested that
NPI any, as opposed to ‘regular’ indefinites like a, is used whenever
we want to make a stronger statement. Consequently, the licensing
condition of any spells out under which condition the statement
containing any counts as a stronger statement. Under this view, a
downward entailing function creates a good context in order to satisfy
licensing, but it is not itself ‘the’ condition. Let us take a closer look at
Kadmon & Landman’s analysis and see if it can be applied to the clausal
comparative.

3.1 NPIS create stronger statements

In a nutshell: according to Kadmon & Landman (1993), the NPI any is
an indefinite just as a is; both combine with a noun phrase such that
we get existential quantification over the set which is the meaning of the noun phrase.\(^{13}\)

The crucial difference between the NPI-NP and the ‘regular’ indefinite NP is that the former has to create a stronger statement. This is achieved by two ingredients: first, \textit{any}, contrary to \textit{a}, widens the interpretation of the noun phrase, that is, it widens the set; and second, \textit{any} is licensed only if the statement containing the NPI entails the statement that results by substituting the NPI with a regular indefinite. The core points of the analysis are given in (29):

\begin{quote}
(29) Kadmon & Landman (1993: 374) for \textit{any}:

\begin{enumerate}
\item \textit{any} CN = the corresponding indefinite NP \textit{a} CN with additional semantic/pragmatic characteristics (\(\rightarrow\) widening/strengthening) contributed by \textit{any}.
\item WIDENING: In an NP of the form \textit{any} CN, \textit{any} widens the interpretation of the common noun phrase (CN) along a contextual dimension.
\item STRENGTHENING: \textit{Any} is licensed only if the widening that it induces creates a stronger statement, i.e. only if the statement on the wide interpretation \(\Rightarrow\) the statement on the narrow interpretation.
\item LOCALITY: strengthening is to be satisfied by the ‘local’ proposition that \textit{any} occurs in.
\end{enumerate}
\end{quote}

Let us consider a concrete example to see the proposal at work:

\begin{quote}
(30) a. I don’t have any potatoes.
\textbf{b.} *I have any potatoes.
\end{quote}

Using \textit{any}, we widen the set of potatoes under consideration; that is, we consider also those potatoes which either we wouldn’t consider as relevant under normal circumstances, in a default context, or which we haven’t considered so far, in the previous context of conversation.

As a formalization, think of \(P\) as the set of potatoes considered as relevant potatoes in a default or previous introduced context; then in (30), we have \(P_{\text{ex}}\), an extended set that contains more potatoes, also borderline cases. This is widening, since \(P_{\text{ex}} \supset P\). Now, (30a) is good, because

\(^{13}\)Kadmon & Landman (1993: 357) leave the question open if indefinite NPs are in general existential quantifiers or if they just introduce open variables which get existentially closed from outside (cf. Heim 1982); whatever one assumes, it should then hold for both \textit{a} and \textit{any}. 

‘I don’t have any potatoes’ (potatoes = $P_{ex}$) 
\[ \Rightarrow \neg \exists x [P_{ex}(x) \& \text{I have } x] \]
\[ \Rightarrow \neg \exists x [P(x) \& \text{I have } x] \]

But (30b) is bad, because

‘I have any potatoes’ (potatoes = $P_{ex}$) 
\[ \Rightarrow / \Rightarrow \text{‘I have potatoes’ (with potatoes = } P) \]
\[ \Rightarrow \exists x [P_{ex}(x) \& \text{I have } x] \]
\[ \Rightarrow / \Rightarrow \exists x [P(x) \& \text{I have } x] \]

In general: Downward entailing functions are licensing environments, because widening creates a superset relation ($P_{ex} \supset P$), and the downward entailing function guarantees the required entailment relation, that is $\psi(P_{ex}) \Rightarrow \psi(P)$. On the contrary, simple affirmative contexts like the one in (30b/32) fail to create an environment that passes the licensing condition. Existential quantification over the extended set does not guarantee the truth of existential quantification over the smaller set. Now, what happens with widening and strengthening in the comparative?

(33) a. Babajaga is meaner than any other witch is.
    b. Babajaga is meaner than [$\exists x [P_{ex}(x) \& x \text{ is K-mean}]$]
    c. Babajaga is meaner than [$\exists x [P(x) \& x \text{ is K-mean}]$]

Assuming the Interval analysis, the comparative should fail to license NPI-\textit{any}. In parallel to the affirmative example in (30b/32), (33b) does not entail (33c). Why? Adding borderline cases to $P$, here the set of witches, we could add extraordinarily mean witches but also witches who are extremely ‘non-mean’, that is, extremely sweet. But then, in order to make (33b) true, Babajaga needs to be meaner than one witch, who gets arbitrarily picked out of the extended set of witches. This doesn’t necessarily entail that she will also be meaner than some witch arbitrarily pick out of a smaller set of witches.

But stop—since widening is along a contextual dimension, we could say that, in the comparative, widening always extends the set upwards along the scale; thus, in (33), the borderline cases added to the set $P$ are only witches who are extraordinarily mean, but not those who are borderline due to their extraordinary sweetness. This gives us a statement that passes the licensing condition, but it still doesn’t account for the universal reading of the NPI. Since widening of $P$ only adds witches that are meaner than the witches contained in $P$, then if you are
either meaner than some mean witch or you are meaner than some very mean witch, it is entailed that you are meaner than some mean witch. But since existential quantification can pick out any witch whatsoever from the extended set, there is no universal reading.

What can we conclude at this point? Here we have an alternative licensing condition which is attractive, because it relates itself to the meaning of the NPI and as such it provides a ‘deeper’ reason for why a downward entailing function is a licenser, but it still cannot account for the NPI-interpretation in the comparative. Where else can we get help? By looking further back?

Since Fauconnier (1975a), it has been repeatedly noticed that polarity items, and negative polarity items in particular, might refer to (pragmatic) scale endpoints. This is because with respect to general features of meaning and distribution, NPIs seem to resemble superlatives, they seem to refer to a relative extreme. Fauconnier (1975a: 373) suggests that free choice any ‘indicates a low point on an arbitrary scale’. Fauconnier (1975b: 196) notes that NPIs are generally expressions for a low point on a pragmatic scale. Cf. Lahiri (1998: 113ff), Lee & Horn (1994) argue that any is simply an indefinite plus even, pointing to the ‘absolute lowest value on the scale’. Similarly, Heim (1984) proposes that some NPIs (though not any and ever) contain in their semantics a covert even. Haspelmath (1997: 119ff) associates free choice indefinites with the feature ‘endpoint on non-reversed scale’, and negative polarity indefinites with the feature ‘endpoint on reversed scale’. Haspelmath also reminds us (p. 115; following Horn 1989:452–3, von Bergen & von Bergen 1993: 139–54) that many NPIs, if they are idiomatic expressions, denote ‘minimal units’, e.g. a red cent, a jot, a tittle. Similarly, Krifka (1994: 210) characterizes such NPIs as a red cent, lift a finger as ‘idiomatic expressions that denote ‘bottom’ elements of certain ontological sorts’.

Looking at it from a broader point of view, Kadmon & Landman’s analysis also reflects the concept of a scale endpoint: widening adds borderline cases to a set under consideration; relating this to a scale, even if any doesn’t denote directly one endpoint, any is still special precisely because it extends the scale towards a newly added endpoint.

Finally, notice the influence of the scale concept even on Ladusaw: it was Fauconnier’s insight about the relevance of scales in semantic (and pragmatic) structure which guided Ladusaw (1979: 106ff) in

---

14 A general side remark on low vs. high point: one has to be cautious with respect to the characterization of an endpoint as low or high; without clear definition, a low point can be easily turned into a high point if we just switch perspective onto the scale. Geometrically, it is an arbitrary decision what we call high and low respectively. See on this point also the discussion in section 3.2.2.
the development of his NPI licensing condition; precisely because downward entailing functions are scale reversing.

Is there a way to draw one picture, a picture that relates the concepts of scale, scale endpoint, downward entailing function and stronger statement? Here is one way to do this: Think of an alternative to define the condition that NPIs are licensed in stronger statements. Under the assumption that an NPI is indefinite such that the semantic representation contains existential quantification, then if stronger statement means the additional requirement of entailing a universal statement, we get precisely the effect that two types of contexts are adequate licensors: first, a downward entailing function, because existential quantification in the scope of a downward entailing function is logically equivalent to (or entails) wide scope universal quantification; and second, a scale context, because existential quantification over a scale endpoint can entail a universal statement by entailment along the scale (cf. Fauconnier 1975a: 361ff). The next section will elaborate upon this idea by proposing a meaning for the NPIs any, ever and jemals. We will see that if we understand the indefinite NPIs as functions that restrict the domain of existential quantification, we get the tools to explain the NPI-distribution in the clausal comparative.

3.2 The NPIs any, ever, jemals: functions that restrict the domain of existential quantification

The proposal has two crucial parts: On the one hand, we modify the idea of what it means to be licensed in a particularly strong statement; on the other hand, we revise the meaning of the indefinite NPIs such that they receive some power to influence the interpretation.

On the second count: Assume that the NPIs any and ever/jemals differ from ‘regular’ indefinites, because they contain a function δ that restricts the domain of a set P; that is, ∀P: δ(P) ⊆ P. Thus, δ is of type ⟨⟨e, t⟩, ⟨e, t⟩⟩, takes as an argument a set P and gives back a set δ(P). In the case of any, the argument is a set of individuals; in the case of ever/jemals, the argument is a set of time intervals.

Furthermore, assume that any and ever/jemals are indefinites in the sense of Heim (1982): the NPI carries along an open variable that gets existentially closed in the nuclear scope. The entire meaning of the NPI is thus a function of the form λP[δ(P)](x) (that is, type ⟨⟨e, t⟩, t⟩),

15 Thanks to Roger Schwarzschild for making me aware of the possible concept of a function that restricts domains. Any error is mine.
that takes a set and gives back an open proposition which undergoes existential closure.

Take as an example *any witch*: the argument of *any* is the set of witches; after lambda conversion, we get \[\delta(WITCH)(x)\].

How does \(\delta\) restrict the set \(P\)? \(\delta(P)\) has to be a subset of \(P\), that is \(\delta(P) \subseteq P\). Hence, \(\delta(P)\) can be either a proper subset of \(P\) or it can be equal to \(P\), but within this limitation, the domain restriction is generally free. However, determination of \(\delta(P)\) is dependent on the following licensing condition, given below together with the assumptions made so far.

(34) a. The NPIs *any, ever, jemals* are functions that restrict the domain of existential quantification over a set: \(\lambda P[\delta(P)](x)\)

\(\delta\) is of type \(\langle\langle e, t\rangle, \langle e, t\rangle\rangle\); \(\lambda P[\delta(P)](x)\) is of type \(\langle\langle e, t\rangle, t\rangle\), an open proposition that undergoes existential closure

b. \(\delta(P) \subseteq P\)

c. **Licensing condition**: ‘NPI \(P\)’ is licensed in an assertion \(\psi\) of type \(t\) only if

\[\ldots \exists x[\delta(P)(x)] \ldots \psi(t) \supseteq \forall x P(x) \rightarrow \ldots x \ldots \psi(t)\]

Given (34c), what is relevant for NPI-licensing is the entire assertion, where the meaning of the NPI is part of the meaning of the assertion. If the assertion is true, with \(\delta(P) \subseteq P\), then a second proposition must be true too: take the original assertion and substitute \('[\delta(P)](x)\)' by \('P(x)'\); furthermore, instead of existentially quantifying over \(x\), bind \(x\) by a universal quantifier such that the result is a wide scope universal statement about the entire set \(P\). The original assertion must entail this universal statement in order to guarantee licensing.

Why propose such a condition? Following Kadmon & Landman (1993); Krifka (1994, 1995), also Jackson (1995), it reflects the idea that NPIs are used if we intend to make a stronger statement than we did by using a simple indefinite. However, ‘stronger’ is understood differently; that is, licensing requires another type of entailment: NPI-statements, as opposed to statements containing other indefinites, must entail universal statements. That is, even if the semantic representation doesn’t contain

---

16 To define the indefinites following Heim (1982) is not crucial to the proposal; if one wants to assume that in general, indefinites are generalized existential quantifiers, then the existential quantifier should be part of the NPI’s meaning as well: defining this for *any*, we would have \(\lambda P.QB[[\delta(P)](x) \land Q(x)]\) such that the function \(\delta\) restricts the domain of one argument but not the other.

Furthermore, note that to assume a Heim-semantics requires that a nuclear scope is present under any TP; that is, we want to interpret the tense node present in all finite sentences as a proper temporal operator that imposes a tripartite structure (compare Heim (1982: 143) for the assumption that a temporal operator counts as quantifier phrase). Otherwise, it would be unclear how existential closure occurs in the comparative.
more than pure existential quantification, universal entailment is never a possible arbitrary accident. As a consequence, we get unexpected effects: an assertion containing an NPI has always the flavor that it doesn’t matter which member of the set we pick and talk about and that we can choose *any member whatsoever*, and the assertion will still be true. This differs from the usage of other indefinites: It might always be irrelevant about which member of the set we talk, but it is not necessarily the case that all members are interchangeable. In this sense, the remaining set resembles an alternative set (cf. Krifka 1994, 1995): existential closure picks one member \( x \) out of the set \( P \), but we need a context that is likewise true for all alternative members of \( P \).\(^{17}\)

What are the effects of the condition in terms of licensing? First, NPIs are not licensed in all assertions; rather they have to find an environment which has the right properties in order to satisfy the condition. Second, in order to accomplish licensing, the NPI-function can monitor the domain of quantification: either it gives back the entire domain or it ‘zooms’ in on the domain of a subset, both being in order to enable universal entailment. In this way, one should really understand the two poles of, first, the domain-restricting function, and second, the licensing condition, as two parts of the NPI’s meaning: an indefinite NPI is ‘more’ than a simple indefinite, because existential quantification has to entail universality and because the NPI has an additional tool to accomplish this.

We can distinguish two different types of appropriate environments which are paired with two different determinations of \( \delta(P) \). They fall together with the two options of subset relation: \( \delta(P) \) as a proper subset of \( P \) (\( \delta(P) \subset P \)), or \( \delta(P) \) equal to \( P \) (\( \delta(P) = P \)).

Before we look at both types, the following should be highlighted. Conceptually, the current proposal is still close to Kadmon & Landman’s approach, despite that there is a crucial distinction which lays in the modification of both the ideas of widening and strengthening. For Kadmon & Landman, the NPI is an indefinite, but it differs to others in that its denotation is slightly enhanced. The NPI widens a set (the denotation of the noun phrase). Thus, the NPI is, by its denotation, equipped to perform a particular function, which is strengthening the statement beyond what a ‘simple’ indefinite can do. Precisely for that matter, in any context which does not meet the licensing condition, we get ungrammaticality rather than just a violation.

\(^{17}\) Compare on this proposal also Dayal (1998: 473) who suggests a ‘meta’-constraint accommodating universal \( FC\text{-}any \) and existential \( NPI\text{-}any \): both must be contained in ‘a property loaded statement that applies to the whole class, not to particular members of this class’.
of a Gricean maxim (~ Quantity—performing a strong statement in a ‘weak’ context; see Kadmon & Landman 1993: 372f): the NPI cannot perform its defining function. Just the same reasoning can be applied here. The NPI, in the current understanding, is also an indefinite, with an additional meaning component. This enables the NPI to perform a function, strengthening, yielding ungrammaticality in any context which notoriously rejects such strengthening. Only, rather than widening a set, the NPI has the potential to ‘shrink’ it. Significantly, and from an intuitive perspective maybe surprising, such ‘shrinking’, where it is performed in a scale context, has a ‘reversing’ effect – it precisely leads to a particular strong statement, one in which it is entailed that the predication is true for all members of the set the NPI zooms in on. At last, how is universal entailment connected to indefiniteness, such that we can say that the NPI strengthens the statement beyond what a ‘simple’ indefinite can do? Note here that one prototypic property of an indefinite is arbitrariness (~ using an indefinite, we can express that it is about some arbitrary member of the set, not about one specific individual). Arbitrariness in its strictest understanding means that all members are lastly inter-changeable. The NPI precisely exploits this aspect of indefiniteness: By the explicit accomplishment of universal entailment, all members of the set are always de facto inter-changeable, because the predication is true for all members. Thus, there is, ultimately, the explicit conveyance of arbitrariness in its strictest sense (as opposed to an implicit conveyance by the use of other indefinites).

3.2.1 In the scope of a downward entailing function: \( \delta(P) = P \) In general, if existential quantification occurs in the scope of negation or another downward entailing function, we get a statement that is logically equivalent to a wide scope universal statement. In fact, Ladusaw (1979: 71ff) explicitly argues against an analysis of NPI any and ever as wide scope universals because of this logical equivalence, which offers, at least theoretically, a second option of semantic representation.

(35) Andre doesn’t see a basketball player.
\[ \neg \exists x [\text{BASKETBALL PLAYER}(x) \land \text{SEE}(a, x)] \]
\[ \iff \forall x [\text{BASKETBALL PLAYER}(x) \rightarrow \neg \text{SEE}(a, x)] \]

To assert that it is not true that there exists a basketball player such that Andre sees that player is the same as asserting that for all basketball players, it is true that Andre doesn’t see them. The equivalence is true for other downward entailing functions as well; some examples:
NEGATION: \[ \neg \exists \varphi(P(x)) \] actual representation
\[ \equiv \forall x P(x) \rightarrow \neg \varphi(x) \]

CONDITIONAL: \[ \exists \psi(\ldots P(x)\ldots) \rightarrow \phi \] actual representation
\[ \equiv \forall x P(x) \rightarrow [\psi(\ldots x\ldots) \rightarrow \phi] \]

EVERY: \[ \equiv \forall y[[\psi(\ldots y\ldots \exists x[P(x)]\ldots)] \rightarrow \phi(y)] \] actual representation
\[ \equiv \forall x P(x) \rightarrow [\forall y[\psi(\ldots y\ldots x\ldots)] \rightarrow \phi(y)] \]

Now, given our NPI licensing condition, we want contexts in which existential quantification over \( \delta(P) \) entails a wide scope universal statement about \( P \). A downward entailing function will grant this entailment precisely if \( \delta \) gives back the entire set \( P \), that is, if it chooses the widest domain. For example, for ‘Andre doesn’t see any basketball player’, we get \( \delta(\text{BASKETBALL PLAYER}) = \text{BASKETBALL PLAYER} \).

(37) a. Andre doesn’t see any basketball player.
   b. \( \delta(\text{BASKETBALL PLAYER}) = \text{BASKETBALL PLAYER} \)
   c. \( \neg \exists x[\delta(\text{BASKETBALL PLAYER})](x) \land \text{SEE}(a, x)] \)
      \[ \equiv \neg \exists x[\text{BASKETBALL PLAYER}(x) \land \text{SEE}(a, x)] \]
      \[ \Rightarrow \forall x[\text{BASKETBALL PLAYER}(x) \rightarrow \neg \text{SEE}(a, x)] \]

We can say that in general, if the NPI-function is in the scope of a downward entailing function, it gives back \( P \) in order to satisfy the licensing condition: \( \delta(P) = P \).

But wait—what about Ladusaw’s arguments against analysing NPI any as a wide scope universal? Notice that the proposal outlined here agrees with Ladusaw about the NPI’s semantic representation: any and ever/jemals are existentially quantified, and \( \exists \) has narrow scope with respect to the downward entailing function. Therefore, all syntactic arguments that support a ‘narrow scope existential’-representation also support this analysis. However, what about the ‘lexical decomposition’ argument which is partly a ‘meaning’ argument? Ladusaw (1979: 76) notices that an expression like rarely, which can be decomposed into ‘USUALLY \( \neg \)’ containing negation such that NPIs are licensed within its scope, needs to embed an existential quantifier: in (38), only a representation ‘USUALLY \( \neg \exists \)’ gives a correct meaning and not ‘\( \forall \) USUALLY \( \neg \)’ . A representation using universal quantification should intersperse \( \forall \) between USUALLY and \( \neg \), which is impossible since the lexical item itself, rarely, is atomic and can’t be broken up.

(38) following Ladusaw (1979: 76):
   The IRS rarely audits anyone.
a. \( \text{USUALLY } \neg \exists x [\text{PERSON}(x) \land \text{IRS audits } x] \)

Usually it is not true that there exists someone such that the IRS audits him.

\( = \) The IRS almost always audits no one. \text{ correct meaning }

b. \( \forall x [(\text{PERSON}(x)) \rightarrow \text{USUALLY } \neg [\text{IRS audits } x]] \)

Everyone is such that it is usually not the case that the IRS audits him. \text{ wrong meaning }

In accordance with Ladusaw, we want (38a) to be the correct representation (since \( \delta(\text{PERSON}) \) occurs in the scope of negation, we get \( \delta(\text{PERSON}) = \text{PERSON} \)). The licensing condition does not demand that (38b) is the representation, it only demands that if (38a) is true, then (38b) should be true too. This is the case, (a) entails (b): if it is usually not true that there exists someone such that the IRS audits him, then if the IRS audits someone, this is an unusual case. But then, everyone is such that usually, the IRS doesn’t audit them (because if the IRS audits them, this is the unusual case). Note that (b) does not entail (a); if it did, both sentences would be equivalent and Ladusaw’s argument that (b) cannot be the actual representation would not go through. But we can think of a situation, in which (b) is true, but (a) at the same time false: imagine that everyone is such that it is usually not the case that the IRS audits him; and this is because so many people are on the list to be audited that the same individual is audited only occasionally. Still, looking at the total number of audits, the IRS audits frequently.

In parallel to \textit{rarely}, one can show that quantifiers like \textit{few} and \textit{at most} reject an analysis of \textit{any} and \textit{ever} as wide scope universals, but they don’t reject an analysis that assumes entailment of a universal statement.

(39) At most two boys got any red presents.

Imagine a situation in which it is true that only for at most two boys, there exists some red present such that a boy got that present; either because there were not many red presents at all or because most red presents were given to the girls. In that situation it is also true that for all \( x \), if \( x \) is a red present, there are at most two boys who got \( x \) (which is also true if zero boys got \( x \)).

Before we close this section, another point to address is the following: it has repeatedly been noticed that NPI-licensing can be canceled by factors that seem to be governed by locality, where this might concern both the semantic and the syntactic structure. A promising way to account for locality is to analyze a ‘trigger’ distinct from the NPI. Identifying such a trigger allows us to define locality
constraints on the structural relationship between the NPI and the trigger (see for example the Immediate Scope Constraint in Linebarger (1987: 338), or the discussion in Ladusaw (1992: 244–246)). Now, given a licensing condition that operates on the entire assertion, with the NPI being part of that assertion, doesn’t this exclude the option of talking about a trigger? Or are we forced to say that if there is a trigger at all, then the entire assertion is the trigger such that we should try to define locality constraints on that assertion? Compare, for example, Kadmon & Landman’s constraint that strengthening has to be satisfied by the ‘local’ proposition (Kadmon & Landman 1993: 373, and (29) above). An alternative is to define a trigger as follows:

(40) Definition of a trigger for NPI-licensing:
A trigger is a quantifier Q or a quantifier phrase QP, contained in an NPI-assertion such that, under maintenance of the set $\delta(P)$, substitution of Q by an alternative quantifier of the same type or pronominalization of QP does not necessarily maintain the required entailment of a wide scope universal statement.

Take two examples to see how the trigger-definition works: in (37), we had ‘Andre doesn’t see any basketball player’, with $\delta($BASKETBALL PLAYER$) =$ BASKETBALL PLAYER. The assertion contains the negative quantifier not of type $\langle t, t \rangle$. Substituting not by another quantifier of the same type can cancel universal entailment; choose for example usually:

(41) a. *Andre usually sees any basketball player.
    b. $\delta($BASKETBALL PLAYER$) =$ BASKETBALL PLAYER.
    c. USUALLY $\exists x[\delta($BASKETBALL PLAYER$)](x) \land \text{SEE}(a, x)]$
        $\equiv$ USUALLY $\exists x[\text{BASKETBALL PLAYER}(x) \land \text{SEE}(a, x)]$
        $\Rightarrow / \Rightarrow \forall x[\text{BASKETBALL PLAYER}(x) \land \text{SEE}(a, x)]$
        $\Rightarrow$ USUALLY SEE($a, x)$

Thus, for (37), we can identify a trigger which is $\neg$. Similarly, in (42) below, if $\delta(T) = T$, with $T$ the set of time intervals in the past and existential quantification over $t \in T$, then (42a) entails that for all times in the past, no man has worried a lot. This entailment can be canceled by pronominalization of kein Mann (‘no man’) but not by pronominalization of viele Gedanken (‘many thoughts’); thus, the QP kein Mann is a trigger for licensing but viele Gedanken is not.

(42) a. Kein Mann hat sich jemals viele Gedanken gemacht. 
    no man has himself ever many thoughts made
Enhancing the proposal by the definition of a potential trigger gives us the possibility of restricting licensing by referring to the local relationship between the NPI and the trigger. One such example of a concrete locality constraint will be discussed at the end of the next section, concerning licensing in the context of a scale. As for the context of a downward entailing function, we just want to notice that in the examples above, it is unsurprisingly the downward entailing function that can be identified as a trigger. However, to highlight the point once more: the downward entailing function, and more generally the trigger, is not the licensing condition. The trigger just provides the appropriate environment for satisfaction of the actual licensing condition: an assertion containing an NPI has to entail a wide scope universal statement. Occurring in the scope of a downward entailing function, together with the NPI’s choice for \( \delta(P) = P \), is one way to accomplish this.

We have seen that universal entailment is trivially satisfied if the actual representation, which contains existential quantification, is logically equivalent to a wide scope \( \forall \) representation. This is true for most cases where the existential has narrow scope under a downward entailing function. However, we also discussed that logical equivalence is not a must for licensing; pure entailment is enough. Let us now look at the second type of adequate context, where we get \( \delta(P) \subset P \).

3.2.2 In the scope of a degree quantifier: \( \delta(P) \subset P \) In the context of a scale, entailment of a universal statement is accomplished by ‘entailment along the scale’. As noticed by Fauconnier (1975a), existential quantification over the endpoint of a scale can have the effect of universal quantification: Fauconnier related a set of propositions \( \{p_1, p_2, p_3 \ldots p_n\} \) to a pragmatic scale such that the scale imposes an ordering on the propositions: if \( p_1 \succeq p_2 \) and \( p_1 \) is true, then \( p_2 \) is true too; based on the pragmatic ordering, then if the proposition that is ordered highest is true, all proposition in the set are true. In more
general terms and independently of pragmatically supplied scales, we can define entailment along a scale as follows.

Assume the elements of a set $X$ are all ordered relative to a scale $S$ such that for all $x, y \in X$, either $x \geq_S y$ or $y \geq_S x$. Assume further that some predicate $\phi$ is true of an element $a$, $a \in X$, where $a$ is ordered above all other elements in $X$; that is, for all $x \in X$: $a \geq_S x$. If the truth of $\phi(a)$ entails that $\phi$ is true for all elements in $X$, then we want to call this entailment along the scale. Note that instead of one element $a$, we could also have a subset $A$ of $X$ ($\rightarrow A \subseteq X$) such that for all elements $a \in A$, each one is ordered above all other elements in $X$. If the truth of $\phi$ applied to the elements in $A$ entails that $\phi$ is true for all elements in $X$, this is entailment along the scale. Obviously, $\phi$ has to be of a specific kind to enable such an entailment; furthermore, we need a scale $S$ to impose an ordering on $X$. We will see now that a construction like the more-than-comparative satisfies those needs.

Since we are concerned with NPI-licensing, entailment along the scale shall apply to the set $P$ restricted by the NPI-function: ultimately, we want a universal statement about $P$. Crucially, the NPI-function’s work in a scale context is to zoom in on a subset of $P$: if all elements of $P$ can be ordered relative to the scale, then $\delta(P)$ is the subset of $P$-elements that correspond to a point which is highest relative to all other servants. Take example (43) to see why this allows satisfaction of the licensing condition.

(43) The gardener is taller than any servant is.

In (43), we have a clausal comparative which is verified (or falsified) by a comparison on the height scale $S_{\text{TALL}>\text{SHORT}}$. That is, $[\ldots \exists x[[\delta(P)](x)] \ldots] \psi(t)$ is true if ‘the gardener is taller than $\exists x[[\delta(\text{SERVANT})](x)]$’. Given that, the NPI-licensing condition is satisfied if $\delta$ zooms in on the set of servants who are tallest relative to all other servants. This is because if the gardener is taller than the tallest servant(s), then he is taller than all servants. Hence, (43) has the following representation with $\delta(\text{SERVANT})$ denoting a proper subset of all servants, the set of tallest servants.

(44) The gardener is taller than any servant is K-tall.

a. $\delta(\text{SERVANT}) = \text{TALLEST-SERVANT}$

b. The gardener is taller than $\exists x[[\delta(\text{SERVANT})](x) \land x \text{ is K-tall}]$

Note that assuming the NPI-function, we can work with Schwarzschild & Wilkinson’s Interval analysis and we can assume existential quantification without getting in danger of predicting the wrong
meaning: existential quantification will pick just one element out of \(\delta(\text{SERVANT})\) and the gardener has to be taller than that one, but since \(\delta(\text{SERVANT})\) is the set \(\text{TALLEST-SERVANT}\), by transitivity we get that the gardener is taller than all servants. Given that, we ensure not only licensing but also the right universal interpretation. (44) is therefore logically equivalent to (45).

(45) \(\forall x[\text{SERVANT}(x)] \rightarrow \text{the gardener is taller than } x\) is K-tall

Can we say some more about how we determine the set \(\delta(P)\) in a specific context? For (44), in order to pass the licensing condition, \(\delta(\text{SERVANT})\) must denote the set of tallest servants. Furthermore, in order to verify the entire comparative, the gardener must be taller than some element of this set. But what is the set of tallest servants? We can spell out the following condition which must be true for all elements of \(\delta(\text{SERVANT})\):

(46) \(\forall x[[\delta(\text{SERVANT})](x) \rightarrow \neg\exists y[\text{SERVANT}(y) \land y \neq x \land y \text{ is taller than } x]]\)

\(\delta(\text{SERVANT})\) denotes the set of tallest servants if there is no element in the entire set of servants that is taller than a member of \(\delta(\text{SERVANT})\). Note that all members of \(\delta(\text{SERVANT})\) are also members of the set SERVANT, so if there is more than one element in the set \(\delta(\text{SERVANT})\), those elements have to be of equal height.

At the same time, crucially notice that \(\delta\) is free to give back any kind of subset of \(P\). There is no requirement to zoom in on a specific subset corresponding to a specific point or interval on the scale, nor is there one to choose a subset at all (instead of \(P\)). That we get the subset of the tallest servants is simply a consequence of the fact that only such restriction will survive the licensing condition.

Finally, notice that the comparative is not the only scale context of the right kind to enable entailment along the scale. Prepositions like English above or beyond seem to do the same job. In the following example, beyond introduces the scale \(S_{\text{HIGH-Low}}\) such that \(\delta\) zooms in on the domain of THING to those things that have the highest price relative to what I could possibly afford; if the item in question is more costly than the highest price I could pay, then it is more costly than everything I could possibly afford.\(^{18}\)

\(^{18}\)Thanks to Markus Hiller, Koichi Nishitani and Scot Zola for bringing this example to my attention.
It was priced beyond anything I could have paid.

Altogether, we see that we can explain the universal interpretation of existential *any* and *ever/jemals* in the comparative as the consequence of existential quantification over a highest point on the scale. Restricting the domain to this highest point is required in order to satisfy the licensing condition. Licensing does not require a downward entailing function. Thus, we don’t need to prove that the meaning of the comparative contains one. Rather, licensing demands entailment of a universal statement which precisely forces quantification over a highest scale point and causes a universal interpretation. Therefore, reviewing previous comparative examples, we can paraphrase their meanings as follows.\(^{19}\)

\[ (48) \text{Babajaga is meaner than } \text{any other witch is.} \]
\[ \text{Babajaga is meaner than some witch who is meaner than the rest of the witches } \Leftrightarrow \text{Babajaga is meaner than all of the other witches} \]
\[ \text{(where the witches can have different degrees of malice).} \]

\[ (49) \text{Today it is hotter than it } \text{ever was before.} \]
\[ \text{This day is hotter than some days in the past which are hotter than all other days in the past } \Leftrightarrow \text{This day is hotter than all days in the past} \]
\[ \text{(where the days in the past can have different temperatures).} \]

But wait. Now, we are saying that the ‘superlative behavior’ of the NPIs in the context of a scale is due to restricting the domain of quantification to a *highest* point. NPIs if analyzed as ‘scale endpoints’ commonly have been associated with the *lowest* point (for reference, recall section 3.1).

In the analysis outlined above, the correspondence to a highest point follows from the fact that the sentential context introduces an explicit scale: the truth of the comparative is determined relative to a scale provided by the degree quantifier and the scalar predicate. Crucially,

\(^{19}\) Look back and compare this proposal with the idea that maximality is built into the semantics of the comparative: here, the equivalent of maximality is built into the semantics of the NPI which, in the comparative, zooms in on the ‘maximal’ elements of the set.

Pushing this idea further and considering Schwarzschild & Wilkinson’s (1999: 29) observation that the modal *can* might have a negative polar variant, notice the following parallel with *any* and *ever/jemals*: if NPI-*can* contains a function \(\delta\) that restricts the domain of existential quantification over possible worlds, then the ‘maximality’-readings observed by Stechow and Rullmann are a simple consequence of the requirement to entail a universal statement:

\[ (i) \text{Obelix is stronger than a soldier can be.} \]

In (i), the NPI-function of the modal would zoom in on the subset of possible worlds that contain the strongest soldiers: if Obelix is stronger than one of those strongest soldiers, it is entailed that for all possible worlds that contain some K-strong soldier, Obelix is stronger than that.
for this scale, what is high and what is low is defined by the ordering relation ‘⪰’. For example, in (44), we had the height scale with TALL ≻ SHORT; that is, if \( a ≻ b \), then \( a \) is taller than \( b \). Notice that in order to determine what is high and what is low, we necessarily need to know the ordering of a scale. To observe that idiomatic NPIs denote smallest units like a red cent, a jot doesn’t allow us to conclude that we are dealing with a low point; just remember that also the smallest unit is a superlative itself which corresponds to the highest point on the scale that orders SMALL ≻ BIG. In general, we can say that identification of a point (or an interval) on a scale relative to other points (intervals) is only possible if we can identify the scale, which means that we can identify the ordering. That is, talking about high and low is undefined unless we relate this to an explicit ordering.

Reversing perspective, we can say that we can construct a scale as soon as we have an ordering relation ‘⪰’. Now recall what we said about licensing by entailment along the scale: take some predicate \( \phi \) which is true of \( a \), where \( a \) is an element of a set \( X \) and \( a \) is ordered above all other elements in \( X \). If in this situation \( \phi \) is necessarily true for all other elements in \( X \) as well, then we have entailment along the scale. Notice that to verify such entailment, we rely on an ordering, because we rely on the axiom that \( a \) is ordered above all other elements in \( X \). What does this mean in terms of licensing by entailment along the scale? It means that, independent of the fact that the NPI-function \( \delta \) can choose any subset of \( P \) whatsoever, verification of entailment relies on an ordering of the elements in \( P \), and thus relies on ‘⪰’. Now, why is this important?

Once we realize the significance of the ordering relation, we can explain the following opposition.

(50) a. *My semantics professor will ever be richer than I am.
   b. I am poorer than my semantics professor will ever be.

For both (50a) and (50b), ever-licensing would require the restriction of the domain of quantification to time intervals where my professor has his poorest moments: if my professor is still richer than I am on his poorest day in the future, then he will always be richer than I am. By the same token, if I am poorer than my professor will be on his poorest day(s), then I am poorer than he will be at all times in the future. If for both sentences, we can find a determination of \( \delta(T) \) that ensures the required entailment, namely that for all times in the future, my professor will be richer than I am, then why is just (50b) good, but (50a) bad?

Here is a reason why: First, we have just noticed that entailment along the scale relies on the ordering relation ‘⪰’. In the comparative, ‘⪰’
is an essential part of the degree quantifier; that is, -er, and likewise more and less are degree quantifiers that operate by the means of \( \geq \). Now, in (50b), the NPI ever is in the semantic scope of the degree quantifier -er, but in (50a), it is not.\(^{20}\)

Second, recall that at the end of the last section, we defined the notion of a trigger: given that NPIs are licensed in assertions that entail a particular universal statement, a trigger is a quantifier(phrase) which is likewise contained in the assertion and which by substitution or pronominalization can cancel the universal entailment. In comparative constructions in general, we can identify the degree quantifier as a trigger. Take as an example (51).

(51) She is more hysterical than any woman I have met before.

\begin{itemize}
  \item a. \( \delta(\text{WOMAN I HAVE MET BEFORE}) = \text{MOST HYSTERICAL WOMAN I HAVE MET BEFORE} \)
  \item b. She is more hysterical than \( \exists x [\text{MOST HYSTERICAL WOMEN I HAVE MET BEFORE}(x) \land x \text{ is K-hysterical}] \)
  \item c. She is less hysterical than \( \exists x [\text{MOST HYSTERICAL WOMEN I HAVE MET BEFORE}(x) \land x \text{ is K-hysterical}] \)
  \item \( \Rightarrow \text{less} \Rightarrow \forall x [\text{WOMEN I HAVE MET BEFORE}(x)] \rightarrow \text{she is less hysterical than } x \text{ is K-hysterical} \)
\end{itemize}

Substituting the degree quantifier more by a quantifier of the same type such as less cancels the entailment about the entire set of woman I have met before, given that \( \delta(\text{WOMAN I HAVE MET BEFORE}) \) is the set containing the most hysterical woman.

Putting the pieces together, (50a) might be bad not because the licensing condition as such is not satisfied, but because satisfaction of the licensing condition is restricted by locality concerning the structural relation of a trigger and the NPI. In the comparative, the degree quantifier is a trigger which furthermore operates on the ordering relation \( \geq \), where \( \geq \) is necessary to verify entailment along the scale. Thus, in the comparative, the trigger of NPI-licensing brings in an operation that is crucial for the type of licensing at stake. Acknowledging this, let us formulate the following locality constraint on NPI-licensing if it involves entailment along a scale:

\(^{20}\) Recall that we are assuming Schwarzschild & Wilkinson’s Interval analysis: the subordinated clause (containing the NPI in (50b)), as well as the differential, are arguments of the degree quantifier; the saturated function which denotes an interval is itself an argument that has to satisfy the main clause; cf. Schwarzschild & Wilkinson (1999: 22), Schwarzschild & Wilkinson (2002: 27).
(52) **Locality**-condition on NPI-licensing in a scale context:
If NPI-licensing is **by entailment along a scale**, then the NPI’s meaning must be in the semantic scope of a trigger for NPI-licensing which operates on ⪰.

Given (52), (50a) is bad because NPI-licensing would require that the entailment of a universal statement is entailment along a scale, but at the same time, *ever* is not in the scope of the degree quantifier *-er* which is a trigger and furthermore operates on ⪰.

Notice a further, more general consequence of (52): it gives a reason for why, in the *more-than*-comparative, NPI-licensing always involves a set δ(P) that contains only elements that correspond to a highest scale point. Here is why: if the NPI has to be in the scope of the degree quantifier, then licensing is always relative to the scale on which the degree quantifier imposes an ordering. That means, licensing is always relative to the scale the comparison is about; and since we want entailment of a statement about entire P and the comparison which has to be true is a *more-than*-comparison, therefore δ(P) is always a set that corresponds to a highest point on that scale.\(^{21}\)

Adding the condition on locality to what we had in (34), let us close this section by summarizing the crucial points of the analysis in (53) below.

(53) Proposed meaning of the NPIs *any* and *ever/jemals*:

a. The NPIs *any, ever, jemals* are functions that restrict the domain of existential quantification over a set: δ[P[δ(P)]](x)  
   [δ is of type ⟨⟨e, t⟩, {e, t}⟩; λP[δ(P)](x) is of type ⟨⟨e, t⟩, t⟩, an open proposition that undergoes existential closure]

b. δ(P) ⊆ P

c. **Licensing condition**: ‘NPI P’ is licensed in an assertion ψ of type t only if  
   \[\ldots \exists x[\delta(P)](x)\ldots]_ψ(t) \Rightarrow \forall x P(x) \rightarrow [\ldots x \ldots]_ψ(t)

d. Definition of a **trigger** for NPI-licensing:  
   A trigger is a quantifier Q or a quantifier phrase QP, contained in an NPI-assertion such that, **under maintenance of the set δ(P), substitution** of Q by an alternative quantifier of the same type or **pronominalization** of QP **does not necessarily maintain the required entailment of a wide scope universal statement.**

\(^{21}\) By the same token, NPI-licensing in *less-than*-comparatives will always involve a set δ(P) which corresponds to a *lowest* point of the scale introduced by the scalar predicate and the degree quantifier.
e. **Locality**-condition on NPI-licensing in a scale context: If NPI-licensing is **by entailment along a scale**, then the NPI’s meaning must be in the semantic scope of a trigger for NPI-licensing which operates on $\geq$.

- The indefinite NPIs neither denote only scale endpoints nor are they only licensed in the semantic scope of a downward entailing function. However, both the context of a downward entailing function and the context of a scale are licensing environments.
- In the scope of a downward entailing function, we get existential quantification over the widest domain; in the context of a scale, we get existential quantification over a scale endpoint. Both solutions accomplish satisfaction of the licensing condition which demands entailment of a wide scope universal statement.
- The clausal comparative doesn’t need to be downward entailing in order to license NPIs.

4 WHAT ABOUT FREE CHOICE ANY?

Given the suggested semantics of NPIs, can we say some more about the relation of NPI-\textit{any} and Free Choice (FC) \textit{any}? Here are two possibilities for further investigation:

A. Following Dayal (1998: 471ff): FC-\textit{any} has a different semantics, that is, it is a universal quantifier, but it is related to NPI-\textit{any}, since both \textit{any}$_{NPI}$ \textit{P} and \textit{any}$_{FC}$ \textit{P} intend to yield universal statements about \textit{P}; the former by embedding \textit{any} \textit{P} into an appropriate assertion \(\psi\) that entails universal quantification, the latter by introducing direct universal quantification.

B. FC-\textit{any} is also \(\lambda P[\delta(P)](x)\) such that there is only one \textit{any}, but the distinction is between \textit{ever}/\textit{jemals} on the one hand and \textit{any} on the other: \textit{ever} and \textit{jemals} are constrained by the Locality condition on licensing in a scale context, but \textit{any} is not. Otherwise, the same licensing condition holds for both \textit{any} and \textit{ever}/\textit{jemals}.

Option (B) might be more attractive, because it would lead to a uniform account of \textit{any}, but it might also be more dangerous: we have to watch out for over-generalization.

But before we address the difficulties, let us first see some examples of how option (B) would work. In general: if FC-\textit{any} is also \(\lambda P[\delta(P)](x)\), then it is also an indefinite which undergoes (at least by
default) existential closure. FC-\textit{any} is always interpreted universally, but at the same time, Free Choice contexts are never downward entailing functions. Following this, if nothing else intervenes, universality must be the result of existential quantification over a scale endpoint plus entailment along the scale. Now, if \textit{any}-licensing is not restricted by the Locality condition, we get the following consequence: as noticed above, if the NPI has to be in the scope of a trigger that operates on \(\geq\), trivially, the licensing will always be relative to the scale on which the trigger imposes an ordering. But if Locality does not hold, then licensing becomes available that is relative to all kinds of scales/orderings; where those scales could either be still introduced by the sentential context, or maybe even pragmatically. Here are some cases:

(54) a. Any professor is richer than I am.
   b. \(\exists x[\delta(\text{PROFESSOR})](x) \land x\) is richer than I am K-rich

Example (54) is the counterpart to the earlier example (50a), ‘My semantics professor will ever be richer than I am’, whose badness we explained by reference to locality. Now, (54a) is good, and this could be due to the fact that this is FC-\textit{any} and FC-\textit{any} is something else. But if we believe in just one \textit{any}, we could assume the representation in (54b) and say: \(\delta(\text{PROFESSOR})\) must denote the set of professors who are the poorest relative to all other professors; this is because if the poorest professor is still richer than I am, then all professors are richer than I am. Thus, we accomplish entailment of the corresponding wide scope universal statement.

(55) a. \(\exists x[\text{POOREST-PROFESSOR}(x) \land x\) is richer than I am K-rich]
   b. \(\forall x[\text{PROFESSOR}(x)] \rightarrow x\) is richer than I am K-rich

However, \textit{any} is not in the semantic scope of any trigger that operates on \(\geq\). The only adequate trigger in (54) is -er such that \textit{any} should be embedded in the \textit{than}-clause in order to satisfy locality. Therefore, \textit{any} is licensed because \(\delta(\text{PROFESSOR})\) can find a domain restriction which guarantees universal entailment, but, since this is entailment along a scale, \textit{any} crucially violates the Locality condition on licensing in such a scale context.

Similar to (54), we can have something like (56), where the assertion contains a scalar predicate but no degree quantifier (just the positive adjective).\footnote{Thanks to Roger Schwarzschild for bringing examples of this type to my attention.}
a. Any of his assignments is damn hard.
b. \( \exists x [ [\delta(\text{ASSIGNMENT})](x) \land x \text{ is damn } K\text{-hard}] \)

When missing the degree quantifier, the assertion doesn’t contain a trigger at all. Nevertheless, we can find a scale in order to license \textit{any}, that is the scale \textit{EASY} \succ \textit{HARD}, or if you like also \textit{HARD} \succ \textit{EASY}: \( \delta \) has to restrict the set of assignments to those that are the most easy relative to the rest. If even the easiest is hard, then all assignments are hard. Relative to the scale \textit{EASY} \succ \textit{HARD}, \( \delta(\text{ASSIGNMENT}) \) corresponds to a highest point, relative to the scale \textit{HARD} \succ \textit{EASY}, it corresponds to a lowest point. In either case, universal entailment is granted, but based on a violation of Locality.

(57) non-negative episodic statement (for discussion, compare Dayal 1998: 452ff):
Babajaga talked to any student who came up to her.
\( \exists x [ [\delta(\text{STUDENT})](x) \land x \text{ came up to Babajaga} \land \text{Babajaga talked to } x] \)

(57) gives us an example which is crucial for the question of whether we have to allow scales that are not supplied by the direct sentential context but rather pragmatically. Once more, we can account for \textit{any} as a function that zooms the domain of existential quantification towards a scale endpoint: that is, \( \delta(\text{STUDENT})](x) \) in (57) could be about the set of students for which it is most unlikely that Babajaga talked to her/him such that if it is true that Babajaga talked to such a student, it is ‘pragmatically’ entailed that she talked to all students who came up to her (‘pragmatic’ entailment cf. Fauconnier 1975a: 363). Notice that in (57), we are not only dealing with a scale that is pragmatically supplied, but also with licensing based on ‘scalar implicatures’ rather than logical entailment. As pointed out by Haspelmath (1997: 112), Fauconnier’s ‘pragmatic’ entailment differs to logical entailment in the following way: logically, it is possible that Babajaga talked to some jerk she would normally not consider worth talking to, but at the same time, she missed talking to some student she normally wouldn’t ignore. This differs from all that we have seen so far. For example, recall the \textit{taller}-comparative: here, entailment along the scale is logical, because you cannot be taller than the tallest person without being also taller than some person who is shorter than the tallest.

Given this liberation with respect to entailment, we can formulate our concerns about hypothesis B more explicitly: by saying that \textit{any} can be licensed by all kinds of pragmatically supplied scales, and furthermore by allowing ‘pragmatic’ entailment, doesn’t this open the door for an
undesirable over-generalization? We know that FC-\textit{any} is also context sensitive. Even if we add up FC-contexts and NPI-contexts, we don’t get back to 100%. What about those context where neither FC-\textit{any} nor NPI-\textit{any} is licit? Crucially, what about mass terms?

   
   a. *Any sugar will be enough for me.
   b. *You can dissolve that substance in any water.
   c. If any sugar spills on the floor, ants will invade the kitchen.
   d. Kim did not dissolve that substance in any water.
   e. *A little synthetic oil lubricates better than any margarine.

(∗ for quantity reading)

In fact, (58) provides us with a revealing data sample. Recall that the proposal of the NPI-function cuts the pie of licensing contexts into two pieces: downward entailing functions on the one hand, in which \( ρ \) gives back entire \( P \); and scale contexts on the other hand, where \( ρ \) zooms in on a scale endpoint such that we get universality by entailment along the scale. Now, (58c) and (58d) are examples of downward entailing contexts in which entailment of a universal statement is due to narrow scope existential quantification under the scope of the downward entailing function. Crucially, \( ρ \) doesn’t need any scale in order to determine the domain but rather gives back \( P \) (here, \( ρ(\text{SUGAR}) = \text{SUGAR}, \, ρ(\text{WATER}) = \text{WATER} \)). But in both (58a/b) and (58e), licensing should be accomplished by entailment along a scale. Under a quantity-reading, (58e) is as ungrammatical as (a) and (b). Why that?

Suggestion: in general, independent of pragmatically supplied scales and orderings, mass terms cannot be ordered relative to a scale, and there is no relative endpoint. This is because there is no set that contains elements that \textit{could} be counted or \textit{ordered}. Only if we construct a set that allows ordering along the scale does licensing by zooming in on the domain of an endpoint become available. We can make (58a/b) and (58e) grammatical by ordering either ‘kinds’ or ‘amounts’.

(59) a. You can dissolve that substance in any kind of liquid.
   b. A little bit of synthetic oil lubricates better than any amount of margarine.

So far, so good; more work has to be done to show that other contexts that reject both FC-\textit{any} and NPI-\textit{any} can be accounted for. But not only this, there is also the opposite question: can all contexts that license FC-\textit{any} be explained by relying on entailment along a scale? A potential problem arises in \textit{generic} contexts:
a. Any graduate student wants to finish his dissertation.
b. Any linguist is a little bit crazy.

As proposed by Kadmon & Landman (1993: 357f), a uniform account of any can be based on the assumption that indefinites in general get generic readings, dependent on context. If any is indefinite, universal readings due to binding by a generic operator are likewise expected, once we abandon the idea that licensing requires necessarily a downward entailing context.

Saying that FC-any is $\lambda P[\delta(P)](x)$ also allows, at least potentially, binding of $x$ by a generic operator. However: Is the NPI-licensing condition still satisfied? We want entailment of a wide scope universal statement. Generic statements are always statements about the entire class/set $P$, but crucially, they allow exceptions.23

An alternative would be to argue that all apparent ‘generic’ statements are in fact, just as the prior examples, cases in which $\delta$ zooms in on the domain of a scale endpoint. That is, also in (60a) and (b), the universal interpretation would be due to entailment along a scale. Then, (60a) would existentially quantify over the set of graduate students who are most enthusiastic about their second qualifying paper such that if one of those graduate students wants to finish, all graduate students want to finish; and in parallel, (60b) could be paraphrased as ‘even the most normal linguist is a little bit crazy’.

To analyze statements like those in (60) also as cases of entailment along a scale might make sense when we look for a consistent explanation for the distribution of any/ever-modification by almost. We know that in general, FC-any can be modified by almost (just as other universal quantifiers can). Crucially, such modification is also possible in the comparative (for both any and ever), but it is rejected in downward entailing contexts like negation. Given the proposed split of licensing by either a downward entailing function ($\delta(P) = P$) or by entailment along a scale ($\delta(P) \subset P$), it would be obvious to say that almost exclusively modifies a scale endpoint, expressing that something reaches almost the upper extreme, but not quite.

a. Almost anybody likes to goof around sometimes.
b. The gardener is sexier than almost anyone I met so far.
c. Today it is colder than it almost ever was.
d. *Today, the butler doesn’t have almost any energy.

23 This might be a heretical thought, but: if any is allowed to be licensed by scalar implicatures, doesn’t this mean that it must accept exceptions in any case? Precisely because this is the definition of ‘pragmatic’ entailment: to allow for exceptions that couldn’t arise in the case of logical entailment?
Conclusion at this point: It might be possible to unify FC-\textit{any} and NPI-\textit{any} by analyzing both as involving a function $\lambda P[\delta(P)](x)$ which is licensed in existentially quantified statements that entail a wide scope universal statement. The difference would then be one between \textit{any} on the one hand and \textit{ever/jemals} on the other: if licensing happens by entailment along a scale, then

- \textit{ever} and \textit{jemals} have to be in the semantic scope of an NPI-licensing trigger which operates on an ordering relation $\succeq$.
- \textit{any} allows licensing by violating the Locality condition on licensing in a scale context. As a consequence, \textit{any} allows licensing relative to all kinds of scales, either provided by the sentential context or pragmatically.

5 CONCLUSION

This paper has shown that we can explain the occurrence of negative polarity items (NPIs) in the subordinated clause of a clausal comparative without reliance on a proof that the meaning of the comparative marker is a downward entailing function.

We have seen several reasons why it might be undesirable to insist on such a proof. Those reasons concerned both the semantic analysis of the comparative in general and the analysis of the meaning and licensing condition of the NPIs as such.

On the count of the comparative, besides the difficulties in identifying the comparative marker as a downward entailing function, achieving the goal comes together with the burden of explaining why most quantifier phrases (but not all) have to leave the $\text{than}$-clause. Such a QR-analysis is syntactically questionable and semantically (at least until we have an explanation) unmotivated.

On the count of the NPI’s meaning, acknowledging the evidence for the relevance of a downward entailing function as licenser, we still might want to know more about \textit{why} this is the case.

This paper proposed that the actual licensing condition of NPIs is the requirement to be contained in a particularly strong statement. The actual semantic representation containing existential quantification has to entail a wide scope universal statement. The NPIs \textit{any} and \textit{ever/jemals} are functions that can help to accomplish entailment by restricting the domain of existential quantification. The idea leads to dividing the licensing environments into two kinds: contexts of a downward entailing function on the one hand, and contexts of a scale with existential quantification over an endpoint on the other.
Licensing in the comparative can be explained by identifying it as a scale context such that the universal interpretation of *any* and *ever/jemals* is caused by entailment along a scale.

Finally, we have seen that a uniform account of *any* (collapsing FC and NPI) might be possible if we distinguish *any* and *ever/jemals* on the basis of their obedience to locality: whereas *ever/jemals* has to obey a Locality condition on licensing in a scale context, *any* is looser and accepts violation of locality.

**Acknowledgements**

I would like to thank Veneeta Dayal, Karina Wilkinson and in particular Roger Schwarzschild for the many intense and fruitful discussions which have made this paper possible. Thank you also to the participants of the Seminar in Semantics on Comparatives and Quantification at Rutgers, Fall 2000. Last but not least, a special thank you to the reviewers for their useful and fair, non-aggressive comments. This paper was originally written in 2001.

**REFERENCES**


