Journal of Semantics 3: 1 - 77

COMPARING SEMANTIC THEORIES OF COMPARISON

ARNIM VON STECHOW

1. INTRODUCTION

> It is a wholesome plan, in thinking about semantics, to stock the mind with as many puzzles as possible, since these serve much the same purpose as is served by experiments in physical science.
>
> B. Russell, "On Denoting" (with "logic" for "semantics").

0. PRELIMINARIES

This is a critical and constructive review of recent semantic theories of the comparative.

It is critical because I want to show that none of the existing theories gives an adequate account of what I consider relevant data.

The article is constructive in so far as I try to show how deficiencies of existing theories can be overcome. At the end a picture emerges of what I believe to be a best theory of the comparative, given the actual state of research and the data discussed.

I will discuss proposals by B. Russell, P. Postal, E. Williams, P. Seuren, E. Klein, D. Lewis, M. Cresswell, and L. Hellan. Some other approaches are briefly mentioned (Bartsch-Vennemann, S. Wheeler). I try to evaluate the different accounts with respect to the question how certain linguistic phenomena are treated by them.

I am indebted to the following people for comments on an earlier version: Manfred Bierwisch, Max Cresswell, Cornelia Hamann, Irene Heim, David Lewis, Pieter Seuren, Wolfgang Sternefeld and Ede Zimmermann. None of these persons is responsible for my flaws (a necessary truth, I guess). In particular, I have to thank Max Cresswell and the editors of the Journal of Semantics for improving the quality of my English.

1. THE DATA

In particular, I will focus on the following data.

Russell's ambiguity (RA). How is the ambiguity of (i) explained?

(i) I thought your yacht was larger than it was

I am considering *ambiguous counterfactuals* (AC) of the kind (ii):

(ii) If Ede had smoked less (than he did), he would be healthier (than he is)

None of the existing theories treats this case adequately, though it could be treated adequately within a Russellian account. I will investigate the question in what way the existing theories will have to be revised in order to account for the nontrivial meaning of (ii). This will lead to the method of double indexing and it will have consequences for the analysis of (i).

A further phenomenon to be explained is the possible presence of *negative polarity items* (NPI) in the *than*-phrase:

(iii) Ede is cleverer than *anyone* of us

The same holds good for *as*-phrases:

(iv) Max is as well as *ever*

I will investigate what the different theories have to say about that. It will follow that Hellan's account can't explain this phenomenon. The accounts of Russell and Postal (and Williams) are not correct either in view of such data. On the other hand, Cresswell's, Seuren's, Lewis' and Klein's approaches give a nice explanation.

Closely connected with the previous phenomenon is the behaviour of *quantifiers and connectives* (Q & C) in comparative complements. There is a sense in which the following arguments are valid.

(v) Konstanz is nicer than Düsseldorf *or* Stuttgart
∴ Konstanz is nicer than Düsseldorf *and* Stuttgart

(vi) Ede is fatter than anyone of us [= someone of us]
∴ Ede is fatter than everyone of us

In a lot of cases we can replace the *than*-complement by something more informative. But we have to block unwarranted inferences (UI) like (vii):

(vii) Ede is fatter than Max
$\overline{\not\therefore \text{Ede is fatter than everyone}}$

In its actual form, the theories of Russell and Postal can't account for these phenomena. The approaches of Seuren, Lewis and Klein can explain arguments like (v) and (vi), but can't block (vii). Cresswell's theory is adequate, whereas Hellan has nothing to offer in view of such data.

If we embed *negative quantifiers* (NQ) into comparative complements we get statements that seem to be nonsensical in one reading:

(viii) *Ede is $\left\{\begin{matrix}\text{more}\\ \text{as}\end{matrix}\right\}$ intelligent $\left\{\begin{matrix}\text{than}\\ \text{as}\end{matrix}\right\}$ no one of us

(ix) *Ede is $\left\{\begin{matrix}\text{more}\\ \text{as}\end{matrix}\right\}$ intelligent $\left\{\begin{matrix}\text{than}\\ \text{as}\end{matrix}\right\}$ Bill is not

It is fairly obvious that the oddness of these is not due to syntactic factors. Hence an adequate semantic theory should explain it. It seems to me that only Russell's theory can explain this in a straightforward way, though Cresswell has something to say about this phenomenon, too.

The *possibility operator* ($\Diamond$) in comparative statements offers a nice touchstone for the adequacy of theories, as the following examples shows.

(x) A polar bear could be bigger than a grizzly bear could be

Only Seuren and Lewis can treat this example adequately. Russell's and Cresswell's approaches can be improved in order to cover this datum. Hellan's analysis is unimprovable in this respect.

An example which does not discriminate so much between different approaches but which is interesting for its own sake is the following.

(xi) More silly lectures have been given by more silly professors – than I expected. (Chomsky)

This is a case of a *multihead comparative.* There are people who believe that sentences like this don't have clear truth-conditions at all. I will say what the truth conditions of (xi) are and how we get them as a natural outcome of an adequate comparative semantics (and syntax!).

I will briefly go into the question whether degrees can be eliminated from an analysis of simple comparative statements like (xii):

(xii) John is taller than Mary

Klein assumes this. He thinks that *taller* is just a relation between John and Mary. Examples like the following make this claim doubtful:

(xiii) Ede is taller than he is broad

Furthermore such a theory can't express *differential readings* (DR) of comparative or equative statements, like the following ones.

(xiv) John is six inches taller than Mary

(xv) Ede is twice as fat as Angelika

The only theory that treats such cases adequately is Hellan's.

Iterated modality (IM) helps to decide the question whether we should prefer a scope solution from a double indexing solution. Take the following sentence.

(xvi) I thought Plato could have been more boring

I think that the only theory which can express all readings of this sentence is an improved version of Russell's account.

2. EVALUATION OF THEORIES

If we try to evaluate the different theories according to the data considered, we obtain the following score:

(xvii) *Descriptive adequacy*

	Russell Postal Williams	Seuren Lewis Klein	Cresswell	Hellan
RA	+	+	+	+
AC	(+)	–*	–*	–*
NPI	–*	+	+	–
Q&C	–*	+/–	+	–
UI	+	–	+	+
NQ	+	(–)	+	–
◇	–*	+	–*	–
DR	–*	–	–*	+
IM	(+)	–	–	–
Score	((5))	3½	5	3

The seals for the phenomena have been introduced before. The minuses with a star mark those gaps in a theory which can in principle be filled in. It is surprising to see that the oldest approach *viz.* Russell's is the one which can to my mind be accommodated to the facts in the best way.

It should be added that this table is not necessarily fair because it doesn't weight the different features. For instance, I didn't take into account syntactic merits of the different approaches. Klein's and Hellan's approach are syntactically much better motivated than the others.

3. PERSPECTIVE

In the last section I propose an analysis of the comparative, equative and positive which tries to overcome the weaknesses of the existing theories. In a way, this approach will combine the nice features of the existing theories. I don't claim much originality for my proposal: I hope it is just the right combination of long existing ideas. Therefore the last section is called "Synthesis".

I have tried to discuss the different theories in a maximally theory-neutral way. The only thing I presuppose is the ability to read expressions that look like formulas of predicate logic. Furthermore, some basic knowledge of possible world semantics is assumed. The discussion is informal throughout because I didn't want to bother you with unnecessary technicalities. But the paper is hopefully precise enough to enable the reader to translate the analyses into his favourite syntactic-semantic frame-work.

Though this article is basically a semantic one, I took care that my own approach is compatible with contemporary syntactic theories, especially with the so-called Governing-and-Binding theory of Chomsky. Section IX is even a purely syntactic one. The implications for the syntax are mostly obvious. I will say quite a number of things as to the structure of adjective phrases. Furthermore, I allude to the thesis that all sentential complements behave semantically as nominals and therefore have to be 'raised' in 'logical form'. I'll have, however, to work out the consequences of this claim on another occasion.

II. RUSSEL'S AMBIGUITY AND SCOPE SOLUTIONS

Bertrand Russell was perhaps the first person who had a semantical theory which could explain certain ambiguities found in comparative constructions. In his paper 'On Denoting' he writes:[1] "I have heard of a touchy owner of a yacht to whom a guest, on first seeing it remarked, 'I thought your yacht was larger than it is'; and the owner replied, 'No, my yacht is not larger than it is.' What the guest meant was, 'The size that I thought your yacht was is greater than the size your yacht is'; the meaning attributed to his is, 'I thought the size of your yacht was greater than the size of your yacht.' "

1. RUSSELL

So Russell obviously assumes that the sentence under discussion (1) has the analyses (2) and (3).

(1) I thought that your yacht was larger than it is

(2) The size x [I thought your yacht is x] is greater than the size x [your yacht is x]

(3) I thought the size x [your yacht is x] is greater than the size x [your yacht is x]

Thus, the ambiguity of (1) is explained by a scope difference of the nominal "the size x". In (2), it has wide scope, representing a consistent thought of mine. In (3), "the size x" has narrow scope, making my thought contradictory.[2]

This explanation is straight-forward and most modern theories proceed along these lines. Is Russell's solution the right one? Not entirely, I believe.

But before giving reasons for my view let me review some of the modern semantic theories dealing with Russell's ambiguity.

The standard method for explaining a non-lexical ambiguity is to assume a scope difference of quantifiers or operators. Every semantic approach to Russell's ambiguity I am aware of proceeds according to this idea (or allows us at least to proceed so). The only account which goes into a different direction is Horn [1981]. This is a pragmatic theory. In this section, I want to discuss semantic theories only. Therefore I will ignore Horn's approach in what follows.[3]

2. POSTAL

At first sight, the several approaches to Russell's ambiguity look very different. But, as I have said, they have a common link: They all work with scope-difference of operators. Let us show this.

Postal [1973] explains the ambiguity of (1) exactly as Russell did

(4) [= Postal's (37)]
Jim believes that Mary is older than she is

(5) a. MORE (x [Jim believes (Mary is old to x)], y [Mary is old to y])
b. Jim believes [MORE (x[Mary is old to x], y [Mary is old to y])]

In footnote 30, Postal gives an informal account of the intended readings of his 'formulas'. Thus (38a) [= our (5a)] can be very crudely and unnaturally realized as:

(i) The extent to which Jim believes Mary is old $\left\{\begin{matrix}\text{is more than}\\ \text{exceeds}\end{matrix}\right\}$ the extent to which she (Mary) is old.'

This is Russell's paraphrase.

3. WILLIAMS

A somewhat peculiar variant of this analysis is found in Williams [1977]. He analyses the two interesting readings of

(6) [= Williams' (111)]
John thinks Bill has more horses than he has

as (7)(a) and (b) [his (112)(a) and (b)]:[4]

(7a) a. John thinks [MORE$_i$ (Bill has x_i horses) than MORE$_i$ (he has x_i horses)]
b. MORE$_i$ (John thinks Bill has x_i horses) than MORE$_i$ (he has x_i horses)

On page 133, Williams gives a hint how (7)(a) and (b) have to be read. (7)(b), *e.g.* means the same as

(8) The amount x_i such that John thinks Bill has x_i horses is greater than the amount x_i such that Bill has x_i horses

Obviously, this is a Russellian solution. Let me briefly say why I think that Williams analysis is peculiar. He assumes that the 'S-structure' underlying (7)(a) is

(9) John thinks [(Bill has MORE horses)] than (He has MORE horses)]

(7)(a) is obtained from (9) by a twofold application of the rule of 'quantifier interpretation', which adjoins the 'quantifier' MORE to the sentence leaving a bound variable at the original place.

MORE means something like "the amount x". The relation "greater" found in (6) is expressed by *than*. My objection against this treatment is this.

I think it is pretty obvious that the relation 'greater than' is expressed by MORE or the comparative morpheme and not by *than*. This is seen from the contrast of the following two German sentences (not my German, but my grandmother's):

(10) (i) Wagner ist *so* bedeutend *wie* Mozart
(ii) Wagner ist bedeutend*er wie* Mozart

The contrast in meaning between the equative (10)(i) and the comparative (ii) can't be explained by means of a contrasting 'complementizer': we have *wie* in both cases. So it must have its source somewhere else.

4. BARTSCH & VENNEMANN

If worked out properly, then presumably also the analysis of Bartsch & Vennemann [1972] is a Russellian one. Bartsch & Vennemann analyse simple comparative statements like

(11) Ede is more intelligent than Sue

as

(12) f^M (Ede, Intelligence) > f^M(Sue, Intelligence)

f^M is a measure function that maps individuals plus dimensions into numbers. At first sight, it looks as if (12) were a Russellian paraphrase of (12) *viz.* something like

(13) Ede's intelligence exceeds Sue's intelligence

Consider now a case of Russell's ambiguity, *viz.* Bartsch & Vennemann's (107a), here represented as

(14) John is less intelligent than he believes

Bartsch & Vennemann analyze this as

(15) $f^M(\text{John, intelligence})(i_o) \;>\; f^M(\text{John, intelligence})(i_{\text{John},b})$

This is to be read as "The degree of intelligence John has in the real world i_o is greater than the degree of intelligence John has in the world he believes to be ($=i_{\text{John},b}$)." This analysis is not complete, because there is more than just one world compatible with John's beliefs.

It is exactly the task of the modal operator "John believes that" to create the evaluation worlds $i_{\text{John},b}$. Since this is not worked out in Bartsch & Vennemann's book, I will neglect their theory in what follows, despite the merits it has in other respects.

As I have said, I believe that Bartsch & Vennemann's account would be equivalent (in the important respects) with Russell's, when worked out properly. Therefore, everything I am going to say about Russell's solution presumably carries over to their theory. More or less the same could be said about Wunderlich [1973].

5. CRESSWELL

The analysis of Russell's ambiguity given in Cresswell [1976] is similar in spirit to the preceeding ones. The details are somewhat different, however. Cresswell would analyse Russell's (1) in the following two ways:

(16) a. λx[I thought your yacht is x-long]*er than* λx [it is x-long]
b. I thought (λx [your yacht is x-long] *er than* λx [it is x-long])

(16)(a) means something like (17):

(17) Every degree x such that I thought your yacht is x-long is greater than any degree x such that your yacht is x-long.[5]

Here, again the ambiguity of (1) is explained by a scope-difference of some operator, the λ-operator in this case. In the previous examples, we had a scope-difference of the definite article, *i.e.* Russell's ι-operator. So Cresswell's (16)(a) corresponds to Russell's (2), whereas (16)(b) corresponds to (3). Notice, incidentally, that despite its superficial similarity to the previous analyses, Cresswell's analysis is essentially weaker in 'expressive power': it can't treat certain constructions which Russell's analysis can treat. I shall come back to this point. Variants of Cresswell's approach are Kaiser [1979] and Hamann, Nerbonne & Pietsch [1980]. Since these papers don't offer any new aspects for the problems discussed in this paper, they won't be mentioned anymore.

6. SEUREN

Seuren [1973] doesn't discuss Russell's ambiguity, but his system can account for it. The ambiguity of (17)(a) would roughly be described by (17)(b) and (c).

(17) a. I thought you worked harder than you did
b. $(\exists e)$ [I thought you worked hard to e & ~(you worked hard to e)]
c. I thought $(\exists e)$ [you worked hard to e & ~(you worked hard to e)]

(17)(b) means something like (18):

(18) There is an extent e such that I thought you worked to (at least) e but you didn't work to (at least) e (*i.e.*, you worked less).

This is an interesting alternative to Russell: We have only one operator binding the variables in the main and the subordinate clause. Furthermore, there is some syntactic evidence that comparatives contain a negative element (in the *than*-clause).[6] It can be shown, however, that Seuren's analysis is weaker in expressive power than Russell's, a feature it shares with Cresswell's account. Notice again, that the ambiguity is explained via scope-difference of some operator. (17)(a) is the consistent, (17)(b) the inconsistent thought.

7. LEWIS

Lewis [1972] assumes that a sentence like

(19) The water is warm

denotes a set of (temperature) delineations,[7] where a delineation is the boundary temperature between warm and cool things. (The idea is, of course, that the actual boundary may change from occasion to occasion.) Let us denote this and similar sets in the following way:

(20) |The water is warm|

A sentence like

(21) The tomato soup is warmer than the mineral water

would then have to following truth conditions:

(22) |The tomato soup is warm| $\supsetneq$ |The mineralwater is warm|)

If we express this by means of 'extent variables', which we used when we discussed Seuren's theory, it comes to something like the following:

(23) ($\forall$e) (The mineral water is warm to e $\Rightarrow$ The tomato soup is warm to e) & ($\exists$e) [The tomato soup is warm to e & $\sim$ (The mineral water is warm to e)]

This paraphrase shows that Seuren's theory is virtually identical with Lewis'.[8]

It is clear, then, that Russell's ambiguity can be analysed in the same way in both theories. So (24)(a) can be analyzed as (24)(b) and (c):

(24) a. I hoped the soup be warmer than it is
b. |I hoped the soup would be warm| $\supsetneq$ |The soup is warm|
c. I hoped (|The soup is warm| $\supsetneq$ |The soup is warm|)

My contradictory hope in (c) is here represented as the proposition that a particular set is a proper subset of itself. Using quantifiers and extent variables, (c) would be represented as (c′):

(24) c′. I hoped [($\forall$e) (The soup is warm to e $\Rightarrow$ The soup is warm to e) & ($\exists$e) (The soup is warm to e & $\sim$ (The soup is warm to e))]

A similar 'translation' can be given for (24)(b). This shows that Lewis' approach, too, can explain Russell's ambiguity by means of scope difference.

8. KLEIN

As to the aspects relevant to this discussion Klein's analysis of the comparative is virtually the same as Lewis' or Seuren's (cf. Klein [1980]). Klein's approach has, however, other interesting features, to which we will return in section XI.

9. HELLAN

Hellan treats Russell's ambiguity explicitly (cf. Hellan [1981, p. 205ff]). He analyses, among others, the sentence

(25) John thinks Mary is more intelligent than she is

The consistent thought of John's is represented as (25)(a), whereas (25)(b) is the inconsistent thought

(25) a. $(\exists d_1,d_2,d_3)$ (John thinks Mary is d_1-intelligent & Mary is d_2-intelligent & $d_1=d_2+d_3$&$d_3>0$)
b. John thinks $(\exists d_1,d_2,d_3)$ (Mary is d_1-intelligent & Mary is d_2-intelligent & $d_1=d_2+d_3$&$d_3>0$)

In section III, I will discuss how Hellan obtains these analyses from the surface.

III. AMBIGUOUS COUNTERFACTUALS AND "RAISING"

The criticism developed in this section can be summarized in this way: No one lowers or raises the right thing. The right solution would be to lower or raise the entire comparative complement, *i.e.* the *than*-clause. My own approach will do this. But no author realizes that $\overline{\mathrm{S}}$-complements behave semantically exactly like nominals and hence have scope. Everyone seeks the scope somewhere else.

But let us come back to our theories. Is there a way to choose among the different proposals? One criterion would be, of course, the adequacy of the analyses in terms of truth-conditions. If one theory gives a correct account of the intuitive truth-conditions of our sentences whereas the other does not, then the former is better than the latter.

At the present stage, however, this criterion doesn't tell us much. As far as the truth-conditions, are concerned, any of the above analyses looks reasonable – at least, as a first approximation.

What we have to do is to look more closely at the 'theory of interpretation', which tells us how the 'logical forms' discussed so far are obtained from the 'surface'. I will try to show that the following sentence represents a serious difficulty for theories that try to explain Russell's ambiguity by means of scope-difference, *i.e.* that work with 'raising' in a sense to be explained.

(26) If Mary had smoked less (than she did), she would be healthier (than she is)

This sentence is ambiguous.[9] There is a reading where both the antecedent and the consequent of the counterfactual are consistent and there are readings where this is not so. Hence we would expect an explanation of this ambiguity along the lines of the preceding section, *i.e.*, via difference of scope.

It is surprising to see that this is not always possible. The theories of Seuren, Klein, Lewis and Cresswell can't represent the informative reading at all. The theories of Postal and Williams can't represent the ambiguity in the way we would expect. Hellan's proposal is the only one which can deal with (26). Yet, as we will see later, his theory is semantically inadequate.

In order to represent counterfactuals, I will use the counterfactual operator $\Box\!\!\rightarrow$ with the Stalnaker - Lewis - semantics:

(SL) $S_1 \Box\!\!\rightarrow S_2$ is true at a world w iff either (1) there are no possible S_1-worlds or (2). Some S_1-world where S_2 holds is closer to w than is any S_1-world where S_2 does not hold.[10]

An S-world is a world in which S is true.

Let us now have a look at the theory of Seuren (Lewis and Klein).

It is easy to see that this approach can't account for the informative reading of (26) at all. The reason is that a comparative construction contains only one existential quantifier binding both the variable in the main- and in the subordinate clause:

(27) $(\exists e)$ [Mary smoked e-little & $\sim$(Mary smoked e-little)] $\Box\!\!\rightarrow$ $(\exists e)$ [Mary is healthy to e & $\sim$(Mary is healthy to e)]

This is the trivial reading of (26). The only thing we can do is to give wide scope over the counterfactual operator to one or both quantifiers. In each case we get an absurd reading:

(28) a. $(\exists e_1)(\exists e_2)$[(Mary smoked e_1-little & $\sim$(Mary smoked e_1-little)) $\Box\!\!\rightarrow$ (Mary is healthy to e_2 & $\sim$(Mary is healty to e_2))]
b. $(\exists e_1)$[(Mary smoked e_1-little & $\sim$(Mary smoked e_1-little)) $\Box\!\!\rightarrow$ $(\exists e_2)$(Mary is healthy to e_2 & $\sim$(Mary is healthy to e_2))]
c. Like (b) with $(\exists e_1)$ having narrow and $(\exists e_2)$ having wide scope

The reason for the stranding of our attempts is that, in this theory, either both the main-clause and the subordinate of a comparative are transparent or both clauses are opaque, *i.e.* under the counterfactual operator. What we want is that the main-clause of the comparatives be opaque whereas the subordinates be transparent. But we can't express this. So these theories are too weak in expressive power.

An objection of the same kind must be raised against Cresswell. He can have only one reading, namely the 'fully-opaque' one, *i.e.*, the reading which corresponds to Seuren's (27).

(29) $(\lambda x[\text{Mary smoked x-little}]$*er than* $\lambda x[\text{Mary smoked x-little}])$
$\Box\!\!\rightarrow (\lambda x[\text{Mary is x-healthy}]$*er than* $\lambda x[\text{Mary is x-healthy}])$

I have said that Postal's and Williams' theories can't represent the ambiguity of (26) in the way we would expect. What would we expect?

Well, the trivial reading of (26) should be something like (30) in Postal's and something like (31) in Williams' approach.

(30) MORE (x[Mary smoked x-little], y[she smoked y-little]) $\Box\!\!\rightarrow$ MORE (x[she is healthy to x], y[she is healthy to y])

(31) [MORE$_i$ (Mary smoked x_i-little) than MORE$_i$ (she smoked x_i-little)] $\Box\!\!\rightarrow$ [MORE$_i$ (she is healthy to x_i) than MORE$_i$ (she is healthy to x_i)]

(30) and (31) are structurally very similar to Creswell's (29). And for the same reason Cresswell can't have the informative reading of (26), Postal and Williams can't have it either – provided we take (30) and (31) as possible logical forms of (26). We can't 'raise' the MORE in (30) and (31) anymore. Raising would give us no well-formed expressions.

On the other hand, there *is* a way to represent the nontrivial reading of (26) in a notation of the style used by Postal and Williams. In order to see this, let us reparaphrase (30) and (31) in a notation which uses Russell's ι-operator. We thus get:

(32) ιx(Mary smoked to x) $<$ ιy(she moked to y) $\Box\!\!\rightarrow$
ιx(she is healthy to x) $>$ ιy(she is healthy to y)

Now the ι-terms are nominals and nominals have scope. We can use this fact in order to represent the informative reading of (26). It is (33).

(33) ι(she smoked to y) $\lambda y[\iota$(she is healthy to y) $\lambda z[\iota x$(Mary smoked to x) $<$ y) $\Box\!\!\rightarrow$ (ιx(she is healty to x) $>$ z)]]

In the system of Montague's PTQ, the ι-terms would be of the NP-type whereas the λ-abstraction would be over individual variables.

(33) is, indeed, the correct representation of the informative reading of (26). But (33) presupposes a very different system from Postal's or Williams': We would have to 'lower' (or to 'raise') the whole *than*-phrase of the comparative construction and not only the *more*-bit, a possibility (and even necessity) never considered by either Postal or Williams. This new system would also suggest a different analysis of Russell's ambiguity in

epistemic contexts. Remember that Postal and Williams represent the consistent thought of John which is reported by (34)(a) as (34)(b). Now, our discussion suggests that (34)(c) would be a safer analysis.

(34) a. John thinks Bill has more horses than he has
b. ιx (John thinks Bill has x-many horses)> ιx(Bill has x-many horses)
c. ιx(Bill has x-many horses) λy [ιx(John thinks Bill has x-many horses)> y]

(34)(c), however, is so different from (34)(b) that almost all the syntactic considerations - which I did not discuss - leading Postal and Williams to their analysis would be inconclusive, if this theory were correct.

And notice something else. If we allow clauses to have scope, we must forbid that the main-clause of the comparative construction ever gets wide scope in counterfactual contexts. Consider again (26), here repeated as (35).

(35) If Mary had smoked less (than she did), she would be healthier (than she is)

Raising the main-clauses in 'logical form' would yield (36):

(36) ιx[Mary smoked x-much] λz [ιy[Mary is healthy to y] λw[(z < ιx[Mary smoked x-much]) $\Box\!\!\rightarrow$ (w > ιy [Mary is healthy to y])]]

But this means something very different from (35), namely (37):

(37) If Mary had smoked more than she did, she would be less healthy than she is

If we allow 'non-parallel' raising of the clause, *e.g.* raising of the main-clause of the antecedent and raising of the subordinate of the consequent, we can even increase the number of absurd interpretations.

Let me conclude this discussion by saying that (26) presents a serious difficulty for Postal's and Williams' theory.

Now, Postal actually doesn't treat (26) in the way sketched here. He does something different, which is - to my mind - self-defeating for his whole approach. I will come to this in the next section.

It is interesting to see that Postal's (and Williams') method of 'lowering' or 'raising' the comparative morpheme, *i.e.* MORE, gives the right result for an analysis of Russell's ambiguity but fails in view of our ambiguous counterfactual. This shows to my mind that these authors lower or raise the wrong thing.

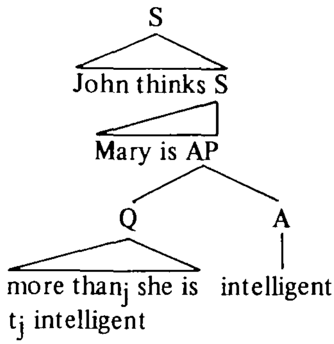
More or less the same point can be made with respect to Hellan's theory. He raises the comparative morpheme plus the *than*-complement. Again this yields the right result for Russell's ambiguity but is not entirely correct for our (26). Thus, Hellan too, seems to raise the wrong thing. Let me show this in some detail.

Hellan doesn't discuss ambiguous counterfactuals, but it is clear from his treatment of Russell's ambiguity, how he has to analyse them. So, let us return to his (25), here repeated as (38).

(38) John thinks Mary is more intelligent than she is

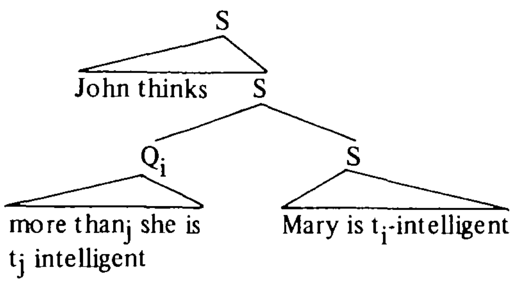
Hellan's analysis is influenced by Bresnan [1975], who assumes that the comparative morpheme *er* forms a quantifier phrase together with the *than*-clause. The underlying structure of (38) is something like

(39)

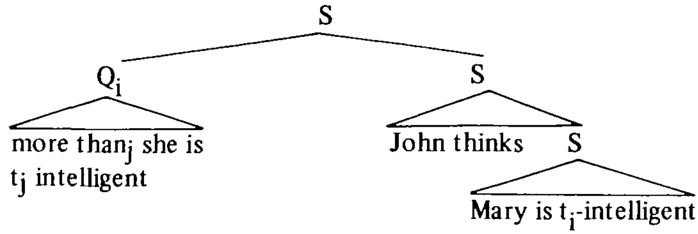


In order to obtain a 'logical form', we have to 'raise' the quantifier Q, *i.e.* (in Bresnan's terms) we Chomsky-adjoin it to an S leaving a coindexed trace. Thus 'raising' will give us two logical forms namely (40)(a) and (b).

(40) a.



(40) b.



Obviously, (40)(a) will serve as 'logical form' for the representation of John's inconsistent belief, whereas (40)(b) represents his consistent belief, *i.e.* (25)(b) and (a) respectively, here represented as (41)(a) and (b).

(41) a. $(\exists d_1,d_2,d_3)$(John thinks Mary is d_1-intelligent & Mary is d_2-intelligent & $d_1 = d_2+d_3$ & $d_3 > 0$)
b. John thinks $(\exists d_1,d_2,d_3)$(Mary is d_1-intelligent & Mary is d_2-intelligent & $d_1 = d_2+d_3$ & $d_3 > 0$)

We get these readings from (40)(a) and (b) by means of the following assumptions.

1. The *than*-clause expresses a property of degrees. Thus *than$_j$ she is t_j intelligent* is that property P which is true of any degree d *iff* she is intelligent to degree d.

2. 'Quantifier raising' makes a λ-abstract out of the sentence to which the raised quantifier phrase is Chomsky-adjoined. Thus we can read the highest embedded S in (40)(b) as λd_i [*John thinks Mary is d_i intelligent*]. This is, of course, the property true of any degree d whatsoever iff John thinks that Mary is intelligent to degree d.

3. We assume the following meaning rule for *er*.

(42) *Meaning rule for the comparative*
|*er*| takes a property of degrees and gives us a nominal (in the sense of Montague) that applies to properties of degrees again. Let P, Q be any properties of degrees. Then |*more*|(P)(Q) is true *iff* $(\exists d_1,d_2,d_3)[P(d_2) \& Q(d_1) \& d_1 = d_2+d_3 \& d_3 > 0]$

It follows from these assumptions that (40)(a) and (b) can be read as (43)(a) and (b) respectively.

(43) a. *more* (λd_j[she is d_j-intelligent]) (λd_i[John thinks Mary is d_i-intelligent])
b. John thinks *more* (λd_j[she is d_j-intelligent]) (λd_i[Mary is d_i-intelligent])

If we evaluate these two readings according to our meaning rule (42), we obtain (41)(a) and (b).

I have grossly trivialized Hellan's original approach in order to bring out what I consider to be its essential features. Notice that Hellan's approach is, in a way, perfectly standard. It is entirely reasonable to assume that *than*-clauses express properties.

Furthermore, the *more* is of the same logical type as determiners, which also take a property and give you a nominal, *i.e.* a quantifier phrase. And, within a Chomskyan frame-work, quantifier phrases always have to be 'raised' in order to obtain a logical form. Nevertheless, even this approach is not correct, as we can see from the fact that it makes the wrong predictions for our (26). At least, it seems to me that this is so.

Hellan can have four logical forms for (26), of which I will represent only the reasonable candidates, namely the reading which represents the trivial conditional and the one which represents the informative conditional:

(44) a. *If* $[Q_i$ more than$_j$ she has smoked t_j little$]({}_S$ Mary has smoked t_i little) *then* $[Q_n$ more than$_m$ she is t_m healthy$]({}_S$ she is t_n healthy)

b. $[Q_i$ more than$_j$ she has smoked t_j little$]({}_S[Q_n$ more than$_m$ she is t_m healthy$]({}_S$ *If* Mary has smoked t_i little, *then* she is t_n healthy))

These representations neglect thesubjunctive.

It is easily checked that (44)(a) gives us, indeed, the trivial reading, *i.e.* the counterfactual with inconsistent antecedent and consequent. If we evaluate it according to Hellan's assumptions, we obtain the following reading:

(45) $(\exists d_1, d_2, d_3, d_1 = d_2 + d_3, d_3 > 0$, she smoked d_2-little)
$(\exists d'_1, d'_2, d'_3, d'_1 = d'_2 + d'_3, d'_3 > 0$, she is d'_2-healthy)
[Mary smoked d_1-little $\Box\!\!\rightarrow$ Mary is d'_1-healthy]

Let me say now, why I think that this reading is not entirely the intended one. Suppose we have four worlds were the following facts obtain. (Worlds with larger numbers are meant to be further away from the real world w_0).

(46)

	Cigarettes smoked by Mary	Mary's degree of health
w_0:	20	1
w_2:	15	1
w_3:	10	1
w_4:	5	2

It seems to me, that the counterfactual (26) should be intuitively wrong in w_0. Smoking less doesn't always make healthier. Only smoking considerably less would help. Hellan, however, predicts, that (26) is true in w_0, because if we take 5 cigarettes for d_1 and the health degree 2 for d_1', then (47) is true in w_0:

(47) Mary smoked 5-cigarettes $\Box\!\!\rightarrow$ Mary is healthy to degree 2

Therefore, (45) is true in w_0 as well.

On the other hand, the analysis (33) would predict that (26) would be wrong in w_0. To my mind, the latter prediction is the correct one. This shows that Hellan's truth conditions are too weak.

Therefore, his account isn't correct either. So perhaps an accommodation of the Russellian analysis along the lines discussed in connection with (33) is the right way out of our dilemma. But it will be shown, that this analysis is defective on other grounds. Before showing this, let me investigate the question, whether there is a way to repair the other theories discussed so far.

IV. POSTAL AND DOUBLE INDEXING

An affirmative answer to this question is suggested, if we analyse Postal's actual treatment of our critical sentence (26). As I have said, Postal would not analyse the trivial reading of (26) as (30) - though this would be in the spirit of his theory. He chooses a different formalization. Let us consider his own example:[11]

(48) If Bob had been taller than he was, he would have made the team

The two readings of this sentence are represented as follows.

(49) a. IF(MORE x [WOULD(Bob was tall to x)w]y[Bob was tall to y]) (WOULD(Bob made the team)w)
b. IF(WOULD(MORE x[Bob was tall to x]y[Bob was tall to y])w)(WOULD(Bob made the team)w)

It requires some good will to interpret Postal's somewhat helpless attempts to build a semantics for counterfactuals in view of the fact that the serious proposals existed quite a while before Postal's article was written.[12]

Postal explains his notation in the following way. WOULD is a two-place predicate, requiring a sentence and a world-variable as arguments. "WOULD(S_a) w" means 'S_a holds in w', where w is a hypothetical world distinct from the real world (cf. p. 392). In footnote 48, Postal explains that the w's in (49) had better be bound by a universal quantifier. Furthermore, it is clear that the embedded term y[Bob was tall to y] is not evaluated with respect to the hypothetical worlds. It rather has to be evaluated with respect to the real world, as Postal's comment on (49)(a) and (b) (his (77a) and (b)) shows:

> "(77a) [=(49a)] thus represents a reading on which what is compared is Bob's actual height and a hypothetical height reached only in w. (77a) [=(49a)] then says that, had that hypothetical height in w exceeded his actual height, Bob would have made the team in w." (p. 392)

Postal goes on and argues that the antecedent of (49a) is consistent whereas the antecedent of (49b) is not.

Let us use the variable w_0 in order to represent the actual world. What Postal has in mind with (49)(a) can then be restated in the following way:

(50) $(\forall w)$(w is hypothetical & $w \neq w_0$ & ιx[Bob was tall to x in w]> ιy[Bob was tall to y in w_0]$\Rightarrow$ Bob made the team in w)

It is rather obvious how we have to restate (49)(b).

Details aside, (50) strikes us as a reasonable account of the consistent reading of (48). The essential point of this analysis is that the sentences not in the scope of a WOULD-predicate are evaluated with respect to the actual world. Their inherent world-variable escapes the quantifying force of IF. I will make this remark clearer below. But, then, Postal's notation misses something. It is not always the case that a sentence not in the scope of a WOULD-operator is to be evaluated with respect to the actual world. To see this, reconsider Postal's treatment of Russell's ambiguity in belief-contexts.

Remember that Postal analyses the consistent belief of Bob's reported by (51)(a) as (51)(b).

(51) a. Bob believes that Mary is older than she is
b. MORE x[Bob believes (Mary is old to x)]y[she is old to y]

Now, if we assume that sentences not in the scope of the WOULD-predicate are evaluated with respect to the actual world then the embedded sentence "Mary is old to x" escapes the belief-operator and (51)(b) expresses a plain contradiction.

In order to make this argument more perspicuous, let us represent the idea that a sentence has to be evaluated with respect to the real world by means of an 'actuality-operator':

(A) ACTUALLY (S) is true in a world w iff S is true in w_0, where w_0 is the real world

Now, if it were correct that sentences outside the scope of a WOULD-predicate always have to be evaluated with respect to the real world, we could represent (51)(b) equivalently as (52):

(52) MORE (x [ACTUALLY (Bob believes (ACTUALLY (Mary is old to x)))], y [ACTUALLY (she is old to y)])

Given Hintikka's semantics for believing, (52) is true iff Mary is older than herself.

Hintikka's semantics for believing is something like this:

(H) "x believes S" is true in w_0 iff $(\forall w)$(If w is compatible with everything x believes in w_0, then S is true in w)

The reader has to bear in mind that the whole discussion is very sloppy. In particular, I am not distinguishing properly between meta-language and object language. But this is common practice and, hopefully, no confusion will arise from that.

According to (H), (52) is true in the actual world w_0 *iff*

(i) The x: $(\forall w)$(If w is compatible with Bob's beliefs in w_0, then ACTUALLY (Mary is old to x) is true in w) is greater in w_0 than y (ACTUALLY (she is old to y))

This is equivalent to (ii):

(ii) The x: $(\forall w)$(If w is compatible with Bob's beliefs in w, then Mary is old to x in w_0) > the y (she is old to y in w_0)

In view of the fact that the universal quantifier doesn't bind the relevant variable w_0, this is equivalent to

(iii) The x (Mary is old to x in w_0) > the y (she is old to y in w_0)
Q.E.D.

This consideration shows that in belief-contexts, sentences not in the scope of the WOULD-operator don't escape the quantifying force of the belief-operator: (51)(b) - not (52) - is the correct formalization of Bob's consistent belief.

On the other hand, in Postal's (49)(a), the sentence not in the scope of a WOULD-operator is supposed to escape the quantifying force of Postal's IF. But then we have to apply the ACTUALLY-operator to it, because it makes this escape possible. To see this, let us try to reconstruct Postal's ideas underlying the notation (49). Using the usual procedure of possible world semantics, we want to drop the world-variable in the WOULD-sentences of (49). (An alternative would be to have a world variable in each sentence, i.e. also in sentences not in the scope of a WOULD-predicate.) Let us represent, therefore, Postal's (49)(a) rather as (53):

(53) IF (MORE x[WOULD (Bob was tall to x)] y[Bob was tall to y]) (WOULD (Bob made the team))

We further assume the following meaning rules - fair reconstructions of Postal's comments on (49), I believe:

(54) a. "WOULD (S)" is true in a world w *iff* w is hypothetical & $w \neq w_0$ & S is true in w
b. "IF S_1, S_2" is true in a world w iff $(\forall w')$(If S_1 is true in w', then S_2 is true in w'

In view of these rules, (53) is true in the real world w_0 *iff* $(\forall w'($ [(w' is hypothetical & $w' \neq w_0$ & (Bob's tallness in $w' >$ Bob's tallness in w')) $\Rightarrow$ (w' is hypothetical & $w' \neq w_0$ & Bob made the team in w')].

This shows that the world-variable of the *than*-phrase doesn't escape the quantifying force of the IF.

Thus, (53) can't represent what Postal has in mind with (49)(a). We get the right reading, if we apply the ACTUALLY-operator to the *than*-phrase:

(55) IF (MORE x [WOULD (Bob was tall to x)] y [ACTUALLY (Bob was tall to y)]) (WOULD (Bob made the team))

I leave it to the reader to check that (55) is true in the real world w_0 iff (50) is the case, and (50) was the reading Postal intended with (49)(a).

If we reconstruct Postal's ideas in this way - a fair reconstruction, again, in my view - then his treatment is fatal for his previous analysis of Russell's ambiguity. Once we have the ACTUALLY-operator, we can explain the ambiguity by means of it alone without any recourse to a scope difference of the ι-operator, the MORE-predicate, and the like.

The ambiguity arising with belief-sentences is represented by the contrast of (56)(a) and (b):

(56) a. Bob believes (MORE x[Mary is old to x] y[ACTUALLY (she is old to y)])
b. Bob believes (MORE x[Mary is old to x] y[she is old to y])

The consistent belief (a) is distinguished from the inconsistent belief (b) simply by the presence of the ACTUALLY-operator.

In a language like German, a verb under an intensional operator *not* in the scope of the ACTUALLY-operator can be in the subjunctive mood whereas a verb in the scope of an ACTUALLY-operator is in the indicative.

(57) Robert meint, daß Marie älter sei (subj.) als sie ist (ind.)

An unembedded verb is in the indicative, of course. We therefore roughly have the following correspondence between intensional operators and moods:

(58) Unembedded verbs and verbs in the scope of an ACTUALLY-operator are in the indicative. Verbs under intensional operators which are *not* in the scope of an ACTUALLY-operator can be in the subjunctive

This consideration shows that representations like (56)(a) and (b) are not without linguistic motivation.

This explanation of the ambiguity in our examples carries over in a straightforward way to counterfactuals. Since I don't want to defend Postal's WOULD-operator, which I believe to be misconceived for a number of reasons, I return to the Stalnaker-Lewis-account of counterfactuals, which was sketched on p. 16. Postal's two readings (49)(a) and (b) are then represented in a much simpler way as (59)(a) and (b).

(59) a. MORE (x[Bob was tall to x], y[ACTUALLY (Bob was tall to y)]) $\Box\!\!\rightarrow$ Bob made the team
b. MORE (x[Bob was tall to x], y[Bob was tall to y]) $\Box\!\!\rightarrow$ Bob made the team

(a) is the reading with a consistent antecedent. According to the rule (SL) on p. 16, (a) is true in w_0, *iff* some world in which Bob's tallness was greater than his tallness in w_0 and in which he made the team is closer to w_0 than any world in which Bob's tallness was greater than his tallness in w_0 but in which he didn't make the team.

And notice that our principle (58) predicts exactly the mood distribution of (59)(a)'s German realisation:

(60) Wenn Robert größer gewesen wäre (subj.) als er war (ind.), wäre (subj.) er in die Mannschaft aufgenommen worden

According to (58), (59)(b) is realized as (61):

(61) Wenn Robert größer gewesen wäre (subj.) als er wäre (subj.), wäre (subj.) er in die Mannschaft aufgenommen worden

Perhaps , the logical form underlying (61), *viz.* (59)(b) explains the extreme oddness of (61): The antecedent of (59)(b) is contradictory. Personally, I am inclined to say that (61) is not a grammatical sentence at all, but it seems hard to rule it out on purely syntactic grounds.

Let us summarize this discussion. Postal's treatment of counterfactuals seems to require the introduction of an ACTUALLY-operator. Once we have gone so far, we don't need the scope-device anymore for explaining Russell's ambiguity. A genuine scope solution *without* the use of the ACUTALLY-operator was sketched in the previous section (cf. p. 18). This solution, however, does not seem compatible with the syntactic considerations leading Postal or Williams to their proposals. A scope solution has to block, furthermore, unwanted readings (cf. (36)). Hellan's solution is certainly better than these. But it isn't correct either, because it doesn't yield the right truth conditions. On the other hand, I have shown that all the ambiguities discussed so far can be explained by the presence or absence of the ACTUALLY-operator, *i.e.* by a minimal difference in logical form. I have further given some evidence that this operator is reflected in the indicative-subjunctive distinction. It seems, then, that we have found an explanation of Russell's ambiguity which is simpler and better linguistically motivated than explanations which work with scope-differences.

V. ACCOMMODATING DOUBLE INDEXING INTO THE DIFFERENT THEORIES

Once we use the ACTUALLY operator, the relation between surface and logical form becomes rather simple, regardless whether we represent comparative-constructions in the style of Postal, Seuren, Cresswell or Hellan. We can assume that the surface-structure of sentence (62) is something like (63):[13]

(62) Mary is taller than Bill is.

(63) Mary is $[_{AP}$ tall-er $[_{\bar{S}}$than$_i$ Bill is i-tall$]]$

Let us assume that the *than*-phrase represents a property of degrees. In Seuren's approach, the *than*-phrase represents the two properties:[14]

(64) $\lambda d_i[(\text{ACTUALLY})\ (\text{Bill is tall to at least } d_i)]$

(64) represents two properties, because the ACTUALLY-operator may be present or not in English. (In German, it is obligatory if the verb is in the indicative, otherwise we don't have it.) I am assuming that a property is a function from individuals into propositions.[15] In the above case, it is a function from degrees into propositions. (64) is that property which is true of a degree d in a world w iff Bill is tall to at least d in w (or w_0, if the ACTUALLY-operator is present in (64)).

In Seuren's approach, the adjective phrase of (63) will then express the properties

(65) $\lambda x\ (\exists d)\ [x \text{ is tall to at least } d\ \&\ \sim(64)(d)]$

If we evaluate this, we get the properties (66):

(66) $\lambda x\ (\exists d\ [x \text{ is tall to at least } d\ \&\ \sim[(\text{ACTUALLY})$
$(\text{Bill is tall to at least } d)])$

So it is very easy to write a compositional semantics which combines the comparative AP in (63) with semantics in Seuren's style. The same holds for the other approaches.

In Cresswell's theory, the *than*-clause $[_{\bar{S}}$ than$_i$ Bill is i-tall] would represent the following properties:

(67) $\lambda d_i[(\text{ACTUALLY})\ (\text{Bill is tall to degree } d_i)]$

And the whole AP in (63) would express the following properties:

(68) $\lambda x(\forall d_1,d_2)\ [x \text{ is tall to } d_1\ \&\ ((67)(d_2) \Rightarrow d_1 > d_2]$

Finally, in Postal's theory, the AP should express the properties (69):

(69) $\lambda x[\iota d\ (x \text{ is tall to } d) > \iota d\,(67)(d)]$

Thus, each of these approaches is compatible with the ACTUALLY-operator and it is very easy to write a compositional semantics for comparative adjective phrases.

In section III, I have said that Russell's and Postal's theory can live also without the ACTUALLY-operator. It doesn't increase their expressive power. And, for Hellan's approach, the introduction of the ACUTALLY-operator does not make sense at all. So, *if* we adopt the operator, considerations of conceptual economy speak in favour of the approach of either Seuren or Cresswell. I will, however, ignore metatheoretical arguments of this kind. I rather want to evaluate the different approaches in the light of another observation, *viz.* that *than*-phrases are environments which trigger so-called negative polarity items. Can each of the above theories, in its original or in its modified form, account for this phenomenon?

VI. NEGATIVE POLARITY

One of the arguments Seuren gives for his analysis is that the *than*-phrase of a comparative AP admits for negative polarity items but not for positive polarity items.[16] He concludes from this observation that the *than*-phrase must contain a (hidden) negative element, because negative polarity items require that. Negative polarity items are, *any, ever, much, be all that, lift a finger, cease to VP* and others. Positive polarity items are, for instance, *already, rather, just as well, still.* Consider the following contrasts:

(70) *Any of my friends could ever solve these problems faster than Ede

(71) Ede could solve these problems faster than any of my friends could ever do

(70) and (71) illustrate the behaviour of the negative polarity items *any* and *ever.*

(72) a. You have already got less support than he has
b. *He has got more support than you already have

(72) illustrates the behaviour of the positive polarity item *already.*[17] Now, Seuren obviously assumes that a negative polarity item must be in the immediate scope of a negation, whereas a positive polarity item is precluded from such an environment. This view goes back to Klima and justifies, indeed, Seuren's analysis (*i.e.* Lewis' or Kaplan's analysis, if we want to be historically just). Consider *e.g.* the examples (72). Possible readings are (73):

(73) a. ($\exists$d) [You have *already* got at least d-less support & $\sim$(he has got least d-less support)]
b. ($\exists$d) [He has got at least d-much support & $\sim$(you have *already* got at least d-much support)]

The explanation of (73)(b)'s ungrammaticality is straightforward in Seuren's theory. In (73)(b), the logical form of (72)(b), *already* occurs in the immediate scope of a negation. Since *already* is a positive polarity item, this is not admissible and the sentence is bad. Similarly, we can explain the ungrammaticality of (71)(a). Here, the negative polarity item *any* is not in the scope of a negation.

If an account along these lines were the only possible one, then Seuren's theory would be preferable to the competing ones we have discussed. Now, recently, Bill Ladusaw has proposed a theory of negative polarity which generalizes the idea that negative polarity items can occur only in the scope of a negation. The generalization is this: Negative polarity items occur only in the scope of a downward-entailing operator. It is not possible to give the definition of a downward-entailing operator without going into technical details. I can't do this here and I therefore want to illustrate this notion by means of examples.

The first example is the following:

(74) If O is a sentential operator, then O is downward-entailing, if O(S) entails O(S′), for any S, S′ such that S′ is at least as informative as S (*i.e.* S′ entails S)

Clearly, negation is a downward-entailing operator. So the theory which says that negative polarity items have to occur in the scope of a negation is subsumed under Ladusaw's theory which says that they can only occur in the scope of a downward-entailing operator. Let us consider a second example:

(75) A determiner Det is downward-entailing for the CN-position, if any sentence of form (i) entails any sentence of form (ii)
(i) Det CN VP
(ii) Det CN′ VP,
where CN′ is at least as informative as CN.[18]

I am assuming that CN-phrases express properties. And a property P′ is at least as informative as a property P iff for any x, the proposition P′(x) entails the proposition P(x).

Definition (75) enables Ladusaw to predict the grammaticality of the following sentence:

(76) Every boy who has *ever* met Tania is delighted

The negative polarity item *ever* occurs within the CN-phrase, and *every* is downward-entailing with respect to the CN-position. This is seen from the following sentence pair:

(76) a. Every boy who likes potato-chips or Coca-Cola is welcome
b. Every boy who likes Coca-Cola is welcome

Clearly the CN-phrase *boy who likes Coca-Cola* is more informative than *boy who likes potato-chips or Coca-Cola* and, therefore, *a fortiori* at least as informative as the latter. Furthermore, (a) entails (b). This illustrates the fact that *every* is downward-entailing for the CN-argument.

Notice that the old theory, which requires that negative polarity items be in the scope of a negation, could explain the grammatically of (76) only by very artificial moves. It would have to analyse *every* perhaps as something like "$\sim\exists\sim$". But Ladusaw discusses cases, where such analyses do not seem feasible. For instance, *few* is downward-entailing for the CN-position and *regret* is downward-entailing for the complement position. A theory that is not forced to introduce negations in logical form in order to explain these cases is certainly preferable. And it is more general. It applies also for cases where there is no negation in logical form. Comparative constructions are possible candidates for an application of Ladusaw's theory. The theory can, indeed, explain that negative polarity items can appear in *than*-phrases, even if we adopt the approach of Cresswell, Postal or Hellan, where the *than*-phrase contains no negative element. It seems, however, that Seuren and Cresswell can do better than Postal or Hellan. Furthermore, it will turn out that Cresswell can do better than Seuren. In the rest of this section, I will establish these claims.

Remember that the syntax of the comparative AP-s was something like this:

(77) $[_{AP}$ Aer $[_{\overline{S}}$ than S$]]$

Let us define the following.

(78) Aer in (77) is downward-entailing with respect to the $\overline{S}$-argument iff any sentence of form (i) entails any sentence of form (ii):
(i) NP is Aer than $\overline{S}$
(ii) NP is Aer than $\overline{S}'$
where $\overline{S}'$ is at least as informative as $\overline{S}$.

It should be clear, by now, that any theory which says that *Aer* is downward-entailing for the $\overline{S}$-argument is good, whereas any theory which doesn't say so is bad.[19] Only theories of the former kind can predict the occurrence of negative polarity items in *than*-phrases. Let us check the different approaches with respect to property (78). We consider the following sentence-pair:

(79) a. Ede is fatter than Otto or Max is
b. Ede is fatter than Otto is

Remember that, in the theories discussed, the *than*-phrase, *i.e.*, our $\overline{S}$, expresses a property of degrees. The *than*-phrase of (79)(a) expresses the property (80)(a) and the *than*-phrase of (79)(b) expresses (80)(b).

(80) a. λd[Otto or Max is fat to d]
b. λd[Otto is fat to d]

Clearly (b) is more informative than (a): any degree satisfying (b) also satisfies (a). Therefore, in any theory where *fatter* is downward-entailing with respect to the *than*-phrase, the proposition expressed by (79)(a) should entail the proposition expressed by (79)(b). Let us check the different approaches with respect to this criterion.

In Cresswell's theory, (79)(a) and (b) express the propositions (81)(a) and (b) respectively.

(81) a. $(\forall d_1, d_2)$[Ede is fat to degree d_1 & Max or Otto is fat to degree $d_2 \Rightarrow d_1 > d_2$]
b. $(\forall d_1, d_2)$[Ede is fat to degree d_1 & Otto is fat to degree $d_2 \Rightarrow d_1 > d_2$]

Clearly (a) entails (b). Therefore, *fatter* is downward-entailing with respect to the *than*-phrase.

In Russell's approach, (79)(a) and (b) express (82)(a) and (b).

(82) a. ιd[Ede is fat to degree d] > ιd[Max or Otto is fat to degree d]
b. ιd[Ede is fat to degree d] > ιd[Otto is fat to degree d]

Also in this case, (a) entails (b). This is so, because (a) can only be true, if Max and Otto have exactly the same degree of fatness. Otherwise the definite description ιd[Max or Otto is fat to degree d] wouldn't be true of exactly one degree.

Hellan has nothing to say to the problem we are discussing. He would analyze (79)(a) and (b) as (83)(a) and (b) respectively. And the latter two propositions are logically independent from each other.

(83) a. $(\exists d_1, d_2, d_3,\ d_1 = d_2 + d_3$ & $d_3 > 0$ & Max or Otto is d_2-fat) [Ede is d_1-fat]
b. $(\exists d_1, d_2, d_3, d_1 = d_2 + d_3$ & $d_3 > 0$ Otto is d_2-fat) [Ede is d_1-fat]

Thus Hellan is ruled out, and Russell and Postal survive our test only if we stick very closely to our criterion that the comparative adjective is downward entailing with respect to its complement.

There is, however, the problem that we also want the inference from (79)(a) to (b) if Otto and Max are not fat to exactly the same degree. Seuren (and Lewis) can explain this, but Russell and Postal can't. Therefore, the first two theories are better in this respect.[20]

There is, however, another problem with Seuren's approach: it will license the following inference:

(84) $\underline{\text{Ede is fatter than Otto}}$
∴ Ede is fatter than everyone

This is seen at once if we write down Seuren's formalization of this argument:

(85) $\underline{(\exists d)\ [\text{Ede is} \geqslant d\text{-fat} \ \&\ \sim(\text{Otto is} \geqslant d\text{-fat})]}$
$\therefore (\exists d)\ [\text{Ede is} \geqslant d\text{-fat} \ \&\ \sim(\text{Everyone is} \geqslant d\text{-fat})]$

I have used, here, the notation "$\geqslant$" for "at least to degree".

On the other hand, in Russell's, Postal's and Cresswell's theory, the argument (84) is only valid, if *everyone* means 'everyone else' and everyone else has the same degree of fatness. Intuitively, this is correct.

That this is so, is clear for the first three authors, because in their analysis, the *than*-clause is represented as a definite description. In order to facilitate the understanding of my claim for Cresswell, I write down his formalization of (84). It is this:

(86) $\underline{(\forall d_1, d_2)\ [\text{Ede is } d_1\text{-fat} \ \&\ \text{Otto is } d_2\text{-fat} \Rightarrow d_1 > d_2]}$
$\therefore (\forall d_1, d_2)\ [\text{Ede is } d_1\text{-fat} \ \&\ \text{everyone is } d_2\text{-fat} \Rightarrow d_1 > d_2]$

Thus, it seems that Cresswell is the winner of the game we played in this section. Most theories predicted the occurrence of negative polarity items, but Russell and Postal couldn't account for cetain intuitively valid inferences *viz.* (79). On the other hand, Seuren's and Lewis' theory licenses too many inferences, *viz.* (84). Cresswell did exactly right.

In the example (84) we embedded a quantifier in the *than*-clause. Quantifiers in *than*-clauses have some interesting properties. These we will study in the next section.

VII. QUANTIFIERS AND CONNECTIVES IN COMPARATIVE COMPLEMENTS

It has been noticed by several people that one reading of (87)(a) entails (87)(b).[21]

(87) a. Ede is wiser than Pericles *or* Socrates
b. Ede is wiser than Pericles *and* he is wiser than Socrates

Thus, the *or* seems to have the force of an *and* in the scope of a comparative adjective. The existential quantifier shows a similar behaviour. Let us assume with Ladusaw that the existential quantifier *someone* is realised as *anyone* in the scope of a downward-entailing operator. Consider now the following pair of sentences:

(88) a. Ede is fatter than anyone of us
b. Each one of us is less fat than Ede

Here, the existential quantifier in the *than*-phrase seems to have the force of a universal quantifier. In this section, I will be concerned with the question what the different comparative theories have to say about this phenomenon.

I am assuming here that comparative adjectivals of the form [Aer than NP] are interpreted exactly in the same way as [Aer than NP is].

Consider now first Seuren's theory. Here, (87)(a) and (b) express the following propositions:

(89) a. $(\exists d)$(Ede is wise to at least d & $\sim$ (Pericles *or* Socrates is wise to at least d))
b. $(\exists d)$(Ede is wise to at least d & $\sim$ (Pericles is wise to at least d)) *and* $(\exists d)$(Ede is wise to at least d & $\sim$ (Socrates is wise to at least d))

(89)(a) entails (89)(b) and vice versa. This is so because of de Morgan's law, which can be applied to (a) and gives us the following proposition as an intermediate result between (a) and (b):

(89) c. $(\exists d)$(Ede is wise to at least d & $\sim$ (Pericles is wise to at least d) & $\sim$ (Socrates is wise to at least d))

(89)(b) directly follows from this.

The equivalence of the sentence pair (88) is explained in an analogous way. (88)(a) expresses the proposition (90)(a):

(90) a. $(\exists d)$(Ede is fat to at least d & $\sim$ $(\exists x)$(x is one of us & x is fat to at least d))

By the usual logical laws, this is equivalent to (90)(b):

(90) b. $(\exists d)$(Ede is fat to at least d & $(\forall x)$(x is one of us $\Rightarrow$ x is not fat to at least d (*i.e.*, x is less fat than d)))

Now, this is exactly what (88)(b) means.

If an account predicts that a noun phrase of the form [NP *or* NP] is 'converted' into a noun phrase of the form [NP*and* NP] in the scope of a comparative adjective, then it will obviously also predict that an existential quantifier is converted into a universal one. Therefore, I will henceforth consider the behaviour of the connectives *or* and *and* under the scope of comparatives only. If they are treated adequately, then $\exists$ and $\forall$ are treated adequately as well.

To summarize the preceeding discussion: Seuren's theory seems to be able to account for the phenomenon initially mentioned.

Cresswell's theory can do equally well. His account predicts the following truth-conditions for our sentences (87):

(91) a. $(\forall d_1,d_2)$[Ede is d_1-wise & (Pericles or Socrates is d_2-wise) $\Rightarrow$ $d_1 > d_2$]
b. $(\forall d_1,d_2)$[Ede is d_1-wise & Pericles is d_2-wise $\Rightarrow$ $d_1 > d_2$] & $(\forall d_1,d_2)$[Ede is d_1-wise & Socrates is d_2-wise $\Rightarrow$ $d_1 > d_2$]

A bit of reflection shows that (91)(a) does indeed entail (91)(b). (The premises of (b) are stronger).

Postal's account will predict the following truth-conditions for (87)(a) and (b):

(92) a. ιd[Ede is d-wise] $>$ ιd[Pericles or Socrates is d-wise]
b. (ιd[Ede is d-wise] $>$ ιd[Pericles is d-wise])
&(ιd[Ede is d-wise] & ιd[Socrates is d-wise])

Now, (92)(a), in fact, entails (92)(b). But the truth-conditions of (a) are very weak. (a) is falsified if Pericles and Socrates don't have exactly the same IQ. This is not what (87)(a) means. So this account doesn't seem right.

Exactly the same objection can be raised against Hellan.

If we have to choose among Cresswell and Seuren, it seems that Cresswell's theory is preferable, because the following sentence is a problem for Seuren:

(93) Ede is fatter than Niko *and* Senta

Seuren predicts that (93) means, among others, (94).

(94) Ede is fatter than Niko *or* he is fatter than Senta

I think, this is not correct. We certainly want to say that (93) entails (94), but it is not equivalent with it. The reading of (93) which is (94) is this:

(95) $(\exists d)$ (Ede is $\geqslant$ d-fat & $\sim$(Niko and Senta $\geqslant$ d-fat))

Thus, Seuren's (and Lewis' and Klein's) account faces the problem of barring this interpretation. I have no idea by which principles this could be done.

We can, of course, also have the reasonable reading (96) for (93):

(96) Ede is fatter than Niko *and* he is fatter than Senta

We get this by "quantifying in" the NP *Niko and Senta,* as can be seen from the following representation:

(97) Niko and Senta $\lambda x[(\exists d)$(Ede is $\geqslant$ d-fat & $\sim$(x is $\geqslant$ d-fat))]

Cresswell has no difficulties with these examples. He predicts the following two readings for (93):

(98) a. $(\forall d_1,d_2)$[Ede is d_1-fat & Niko and Senta are d_2-fat $\Rightarrow d_1 > d_2$]
b. Niko and Senta $\lambda x[(\forall d_1,d_2)$[Ede is d_1-fat & x is d_2-fat $\Rightarrow d_1 > d_2$]

Both readings exist. So, this is another point for Cresswell.

Another problematic point is the behaviour of *neither-nor* (or *no one*) in the scope of a comparative adjective.

Consider the following sentences.

(99) a. Irene is prettier than neither Ede nor Senta
b. Irene is prettier than no one of us

These sentences are extremely odd and semantical theories of the comparative have to explain this fact.

The theory of Russell gives a straightforward explanation of these data. For him, (99)(a) and (b) don't express propositions at all, as can be seen from his analysis of (99)(a):

(100) ιd[Irene is d-pretty] > ιd[neither Ede nor Senta is d-pretty]

This analysis uses the term ιd[*neither Ede nor Senta is d-pretty*], which, for abvious reasons, doesn't denote.

Seuren would have to say that (99)(a) expresses a tautology, *viz.* (101):

(101) $(\exists d)$[Irene is $\geq$d-pretty & $\sim\sim$(Ede or Senta is $\geq$d-pretty)]

One might object against this solution that (99) is rubbish and should therefore not express a tautology, *i.e.* something very precious to the philosopher or mathematician. On the other hand, it takes some time to figure out that (101) is in fact a tautology. In ordinary life, we calculate in real time. Perhaps we are not quick enough to perceive the logical structure of (99), *viz.* (101). We stop at a certain point of conceptual complexity, hence the oddness of (99).

The same kind of explanation is perhaps valid for Cresswell's account. He would predict that (99) expresses a logical falsehood, namely (102):[22]

(102) $(\forall d_1,d_2)$[Irene is d_1-pretty & $\sim$(Ede or Senta is d_2-pretty) $\Rightarrow d_1 > d_2$]

This time it is even more difficult to find out what proposition (102) is. Each time, I am faced with (102), I have to draw a picture in order to convince myself that (102) is indeed a logical falsehood. Therefore, to my mind, Cresswell's theory gives quite a good explication of the oddness of (99).

Notice that we have to prevent the wide-scope reading for the NP *neither Ede nor Senta,* if we adopt the approach of Seuren or Cresswell. The wide-scope reading would yield the following propositions, (103)(a) in Seuren's, (103)(b) in Cresswell's system:

(103) a. Neither Ede nor Senta $\lambda x(\exists d)$[Irene is pretty to at least d & $\sim$(x is pretty to at least d)]
b. Neither Ede nor Senta $\lambda x(\forall d_1,d_2)$[Irene is d_1-pretty & x is d_2-pretty $\Rightarrow d_1 > d_2$]

(103)(a) and (b) mean that neither Ede nor Senta does have the property of being less pretty than Irene. This is a perfectly sensible proposition, but none that is intuitively expressed by (99)(a). Similarly, (99)(b) - here repeated as (104) - does not express the proposition (105).

(104) Irene is prettier than no one of us

(105) No one of us $\lambda x(\exists d)$(Irene is pretty to at least d & $\sim$(x is pretty to at least d)

(105) means that no one of us is less pretty than Irene, or that everyone of us is at least as pretty as Irene. Again, this is not what (104) means.

Notice that the problem of distinguishing the behaviour of these NPs is not tied to Cresswell's or Seuren's approach. It arises with the other theories as well. So this is a general problem which does not decide between the different theories.[23]

Let me try to draw a conclusion from the discussion of this paragraph. Cresswell's and Seuren's theory can say something about the behaviour of *or* and the existential quantifier in the scope of a comparative adjectives. Seuren, however, makes the wrong predictions about *and* and the universal quantifier in that environment. The other theories can't deal with these facts at all.

Concerning a negative quantifier like *no one* or *neither-nor*, the theories treating the *than*-phrase as a definite term seem to do best. But also Seuren and Cresswell can say something about the relevant phenomenon. If we take all this together, than Cresswell's theory seems to be the winner of this paragraph. This is not so astonishing, because this section is closely related to the preceding one, where also Cresswell was the favourite.

VIII. THE POSSIBILITY OPERATOR IN COMPARATIVE CONSTRUCTIONS

In a way, modal operators are quantifiers. They quantify over invisible world variables. In the preceding section we found that definite comparative analyses (Russell and Postal) could not adequately treat embedded existential quantifiers. This result carries over in an obvious way to an embedded possibility operator $\Diamond$. This operator has the usual semantics, that is, we assume the rule (107):

(107) $\Diamond$ S is true in a world w iff there is an (accessible) world w' such that S is true in w'

One might think that it is clear from our comments on Russell's ambiguity how $\Diamond$ is treated. But this is not so. The *believe*-operator is, in a way, a universal quantifier and behaves quite differently from the possibility operator, an existential quantifier. Consider the following sentence.

(108) A polar bear could be bigger than a grizzly bear could be[24]

It is no problem to analyse this sentence by means of Seuren's and Lewis' theory. On the other hand, it is not possible to analyse (108) within the other approaches. *Prima facie,* this is a strong point for Seuren and Lewis. I will, however, show that the other theories can be repaired in such a way that (108) is no longer an insurmountable difficulty. This will have a positive effect concerning the examples of the last two sections. Most of the difficulties encountered there can be overcome.

Let's turn to Seuren and Lewis first. In their system, (108) would be analysed as (109):

(109) $(\exists d)[\Diamond(\text{A polar bear is} \geqslant \text{d-big}) \mathbin{\&} {\sim}\Diamond(\text{A grizzly bear is} \geqslant \text{d-big})]$

In order to get this we have to assume that the auxiliary *could* expresses $\Diamond$, a standard assumption.

Thus, (109) means

(110) $(\exists d)[\exists w(\text{A polar bear is} \geqslant \text{d-big in w}) \mathbin{\&} {\sim}(\exists w)(\text{A grizzly bear is} \geqslant \text{d-big})]$

exactly as we would expect.

Cresswell's theory yields the wrong result, *viz.* (111).

(111) $(\forall d_1,d_2)[\Diamond(\text{A polar bear is } d_1\text{-big}) \mathbin{\&} \Diamond(\text{A grizzly bear is } d_2\text{-big}) \Rightarrow d_1 > d_2]$

These truth-conditions are too strong. We only have to find a polar bear which is smaller than a grizzly bear in order to falsify this proposition. This is very easy. Indeed, (111) is trivially false, whereas I don't know whether (108) is true. We could try to give wide scope to one or two of the modal operators in (111). I leave it to the reader to check that such a move wouldn't help.

Hellan's approach fails for similar reasons as Cresswell's. He would analyse (108) as (112). I leave it to the reader to check that these truth-conditions are far too weak:

(112) $(\exists d_1,d_2,d_3)[\Diamond(\text{A polar bear is } d_1\text{-big}) \mathbin{\&} \Diamond(\text{A grizzly bear is } d_2\text{-big}) \mathbin{\&} d_1 = d_2+d_3 \mathbin{\&} d_3 > 0]$

The straightforward representation of (108) in Russell's and Postal's theory fails entirely, because the definite terms obviously don't denote anything:

(113) $\iota d\Diamond$(A polar bear is d-big) > $\iota d\Diamond$(A grizzly bear is d-big)

Again, I leave it to the reader to convince himself that giving wide scope to both model operators could not save the situation.

It seems then that our example (108) gives strong support to Seuren's and Lewis' account. Alas, this appearance is not cogent. I think that we should not be content with a failure of Russell's analysis. The reason is, that we need a definite analysis of comparative statements on quite independent grounds. Consider the following sentence.

(114) The size a polar bear can have exceeds the size a grizzly bear can have

This is a perfectly understandable English sentence which we have to treat semantically anyway. Now, in (114), the modal *can* obviously must be within the scope of the definite article. A bit of reflection will show then, that (114) is a problem for any of our theories. This problem, however, is not caused by the semantics of the comparative morpheme. It arises together with the theory of descriptions. Once we have understood that, it is pretty obvious that formalization (113) is too naive. If we use the definite article, we often leave certain parameters inexplicit. *The book* means 'the book at such and such place', *the money* means 'the totality of the money in my wallet'[25] etc., and, obviously, the intended sense of (114) is something like (115):

(115) The *maximal* size a polar bear can have exceeds the *maximal* size a grizzly bear can have

In Russell's theory, we have to formalize this as something like (116):

(116) the maximal d $\Diamond$(A polar bear is d-big) > the maximal d $\Diamond$(A grizzly bear is d-big)

Take "the maximal d... d..." as short for "$\iota d[\mathrm{Max}(\lambda d[\ldots d\ldots])(d)]$". *Max* is an operator modifying properties of degrees, having the following semantics:

(117) Max (P) is true of d *iff* P(d) and $\sim(\exists d')[P(d') \;\&\; d' > d]$

Given that degrees are linearly ordered, Max(P) is true of exactly one degree; Therefore, the definite description $\iota d[\mathrm{Max}(P)(d)]$ will make sense, if P is the predicate $\lambda d\Diamond$(A polar bear is d-big).

Thus, the more explicit spelling of (116) is (118):

(118) $\iota d[\mathrm{Max}(\lambda d \Diamond(\text{A polar bear is d-big}))(d)] > \iota d[\mathrm{Max}(\lambda d \Diamond(\text{A grizzly bear is d-big}))(d)]$

Thus, Russell's theory is saved, and the objections against his theory can be met on quite general grounds: Once we have a pragmatic theory which tells us with respect to which implicit parameters definite terms are evaluated, we obtain the correct readings. And it is clear that we need such a pragmatical theory anyway.

This salvation procedure crucially hinges on the fact that Russell's analysis of the comparative is a definite one.

If we look at Cresswell's analysis, we find that it is not a definite one. But, if we want to save him from the objection raised in this section, we obviously have to adopt the same method. A correct analysis of (108) within an improved theory of Cresswell would be the following:

(119) $(\forall d_1,d_2)((d_1 = \text{the maximal } d[\Diamond\text{A polar bear is d-big}]\ \&\ d_2 = \text{the maximal } d[\Diamond\text{A grizzly bear is d-big}]) \Rightarrow d_1 > d_2)$

This accommodation is not so obvious as in the previous cases, because, as said, Cresswell's analysis is not of the definite type. We can eliminate the definite article from Cresswell's account in view of the fact that "the maximal d" and "a maximal d" mean, in fact, the same thing. In other words, a better accommodation than (119) would, perhaps, be (120):

(120) $(\forall d_1,d_2)([\mathrm{Max}(\lambda d \Diamond(\text{A polar bear is d-big}))(d_1)\ \&\ \mathrm{Max}(\lambda d \Diamond(\text{A grizzly bear is d-big}))(d_2)] \Rightarrow d_1 > d_2)$

Here, *Max* is the operator we have described in (117). The difference between Cresswell's and the other theories is that the presence of the *Max*-operators is a natural outcome of the pragmatics of the definite article for Russell whereas it has to be stipulated in Cresswell's system, *i.e.*, it has to be built into the semantics of the comparative. This seems to be a slight weakness of Cresswell's account. Hellan's theory can't be saved at all.

The introduction of the *Max*-operator has a positive effect for all definite theories in two respects.

First. It obviously has the impact that the comparative complement, *i.e.* the *than*-phrase is the right environment for negative polarity, because the *Max*-operator is downward-entailing. I leave it to the reader to convince himself of the correctness of this assertion. Thus, the objections raised against Russell in section VI can be met now.

Second, quantifiers appearing in the *than*-clause can be treated adequately now by Russell. Consider first the article case where a negative existential is in the scope of the *than*-phrase.

(121) *Ede drinks more than no one of us

Russell's analysis would predict the following reading:

(122) the maximal d[Ede drinks d-much] > the maximal d[no one of us drinks d-much]

As in (101), the second definite term doesn't denote, because although there certainly are degrees d, such that no one of us drinks d-much, there is no maximal degree of this kind, at least, if we assume that there is no upper limit of degrees, a reasonable assumption. On the other hand, it is no problem to embed existential or universal quantifiers within the *than*-phrase:

(123) Ede drinks more than anyone of us

(124) Ede drinks more than everyone of us

In Russell's system, these will be analyzed as (125)(a) and (b), respectively:

(125) a. The maximal d[Ede drinks d-much] > the maximal d [someone of us drinks d-much]
b. The maximal d[Ede drinks d-much] > the maximal d [everyone of us drinks d-much]

(125)(a) means, indeed, (126):

(126) For everyone of us, it is the case that Ede drinks more than him

This is exactly the reading we were after in section VII. (125)(b) is correct, too. It means that Ede drinks more than everyone of us, in case we all drink the same amount.

On the other hand, Lewis' and Seuren's system could not treat (124) adequately. Cresswell's system was adequate with respect to these phenomena, and nothing has changed with the introduction of the *Max*-operator. So Lewis, Seuren and Hellan seem to be the losers of this game.

There is, however, a negative effect for Russell's theory, which should be mentioned. The particular behaviour of the quantifiers is restricted to the *than*-phrase, whereas Russell predicts that we would find it also with respect to the subject term. The same holds with respect to negative polarity. This, however, is not borne out by the facts. According to Russell, there should be a reading of (127)(a) which means (127)(b) and

(128) should, in one reading, express no proposition at all. Clearly, this is inadequate.

(127) a. Someone of us is taller than Ede
b. Everyone of us is taller than Ede

(128) No one of us is taller than Ede

Cresswell's analysis is inadequate for exactly the same reason.

The outcome of this section is that all the theories considered so far are inadequate with respect to our data. It will be the task of the last section to offer an alternative that can meet all the difficulties discovered.

IX. ARE COMPARATIVE COMPLEMENTS ALWAYS REDUCED CLAUSES?

So far we have assumed that comparative complements always go back to reduced clauses. In other words, we assumed that the comparative adjectival in (129)(a) is interpreted exactly as (129)(b):

(129) a. Aer than NP
b. Aer than NP is

This suggests that the *than*-NP is always interpreted as the subject of a missing *than*-S. This is, however, not true without further qualifications. Hellan [1981, p. 56] observes that:

(130) John thinks that Mary is taller than herself

could not mean the same as:

(131) John thinks that Mary is taller than she is

We would, however, predict this meaning, if we assumed that the underlying structure of (130) were something like (132):

(132) *John thinks that Mary_1 is taller than herself_1 is tall

Any of our systems would predict that (132) may report a consistent thought of John's. Take, for convenience, Seuren's representation:

(133) $(\exists d)(\text{John thinks that Mary}_1 \text{ is} \geqslant \text{d-tall} \;\&\; {\sim}\text{Actually [herself}_i \text{ is} \geqslant \text{d-tall]})$

One might object that this reading is blocked on independent grounds. Chomsky's binding theory (cf. Chomsky [1981]) says that the reflexive *herself* must have its antecedent within the minimal governing category, here the *than*-clause. Since there is no possible antecedent for *herself* in the *than*-clause, the sentence is ungrammatical. But this would leave unexplained that there is a possible interpretation of (130) *viz.* the case where John holds a contradictory thought about Mary. The only way out of this difficulty seems to be the assumption that the construction (129)(a) does not go back to (129)(b) but rather to (134):

(134) Aer $[_{PP}$ *than* NP$]$

A structure of this kind is advocated also in Chomsky [1977]. Thus, the surface structure of (130) would be (135):

(135) John thinks that Mary$_i$ is $[_{AP}$ taller $[_{PP}$ than herself$_i]]$

Here, *Mary* can be an antecedent of *herself*, because the two are in the same minimal governing category *viz.* the *that*-clause.

This leaves open the question how (135) is interpreted exactly. I will assume that we proceed exactly as we did before. The PP *than herself*$_i$ is interpreted exactly as if it were the reduced clause *than herself*$_i$ *is*. This will give us the right reading, *viz.* (133) without the ACTUALLY-operator. We have to say, of course, that the ACTUALLY-operator may be present in logical form only, if the *than* NP goes back to a reduced surface clause. This stipulation will bar the unwanted reading (133).

Notice that such a treatment is consistent with Chomsky's binding theory, because the binding theory applies to the surface structure *i.e.* to (135) and not to the logical form, *i.e.* (133).

Thus we obtain the following tentative conclusion. The NP *herself* in *taller than herself* is a subject on the level of logical form but not the subject of a reduced clause on the level of surface structure. Furthermore, it seems to me that an example like (130) supports Chomsky's view that the binding conditions apply on the surface and not on logical form.

X. MULTIHEAD COMPARATIVES

In this section I will be concerned with the semantics of comparative constructions where the comparative complement is associated with more than one 'head'. These examples typically occur with plural nouns or mass nouns. I don't think that these cases represent genuine semantic problems. They can essentially be treated with the methods we already have at our disposal. But it took me quite a while to realize this, because the examples are conceptually rather complicated and are neglected in the literature. Furthermore, we can perhaps say something about the question of why such cases are restricted to plural or mass nouns. Chomsky [1981, p. 81] consider cases like the following:

(136) a. [more silly lectures] have been given by [more boring professors] - than I would have expected
b.*[more silly lectures] have been given by [more boring professors] - than I met yesterday

Chomsky says that the first sentence is a case of construal whereas the second one is a case of unwarranted application of Move-α.

(The *than*-phrase is naturally associated with the second bracketed NP. But, then, subjacency is violated. I will not go into the question of the correctness of the latter explanation.)

The construal process Chomsky assumes for (136)(a) associates the *than*-clause with the two bracketed NPs. Hence, the input for the semantic rules is not (136)(a) but rather (137).

(137) more silly lectures [than I would have expected] have been given by more boring professors [than I would have expected]

As it stands, (137) is not yet sufficient for the semantic rules to apply. The bracketed $\overline{S}$'s lack an internal sentential complement. Each $\overline{S}$ must be reconstructed as a sentence of the form (138):

(138) than I would have expected that silly lectures would be given by boring professors

Thus, the entire structure determining the interpretation of (136)(a) is (139):

(139) More silly lectures [than I would have expected that silly lectures would be given by boring professors] have been given by more boring professors [than I would have expected that silly lectures would be given by boring professors]

If we compare (139) with (136)(a) we see that greater explicitness is not necessarily connected with better understanding.

It would be interesting to investigate the syntactic processes by which (136)(a) and (139) are related. I will say a few speculative words about some aspects of this problem. But let us ask first the following question: Which proposition does (139) express?

It is advisable to consider first a simpler example of the same kind, before we answer the question.

Consider the following sentence.

(140) More dogs ate more rats than cats ate mice

I take it that (140) has the following truth-conditions (in its most easily available reading):

(141) The number of dogs that ate rats is greater than the number of cats that ate mice & the number of rats that were eaten by dogs is greater than the number of mice that were eaten by cats

The best way to convince oneself that this is indeed the right reading of (140) is to check that other formulations found occasionally are not correct. I have heard once the following truth-condition for (140):

(142) The number of those <x,y> such that x is a dog and y is a rat and x ate y is greater than the number of those <z,w> such that z is a cat and w is a mouse and z ate w

But this is too weak. Suppose three dogs ate the same rat (they share it) and one cat ate two mice. Then (142) would be satisfied because we would have more rat-eating dogs than mouse-eating cats. But, intuitively, (140) is false in this situation. This is so because the number of mice that were eaten by cats (=2) is larger than the number of rats that were eaten by dogs (=1), *i.e.* the second conjunct of (141) is not met.

Thus, I take it that (141) is essentially correct. The question we have to answer is now: How can we obtain the information (141) from the surface (140)?

Let us assume, as Chomsky does in his comments on (136)(a), a rule of construal that associates the *than*-phrase in (140) with both *more* NPs:

(143) [More dogs]^i ate [more rats]^i - [than cats ate mice]^i

I have indicated this association by means of the 'association-index' i, an *ad hoc*-invention for the purposes of the present discussion. This notation is to be interpreted as if the *than*-phrase belonged to both NPs. So let us put it there, for convenience:

(144) More dogs [than cats ate mice] ate more rats [than cats ate mice]

The next step of construal involves considerations of parallelism: the subject *dogs* of the matrix-clause is correlated with the subject of the *than*-clause *cats* and the direct object of the matrix clause *rats* is correlated with the direct object of the *than*-clause mice. We can express this correlation in the following way:

(145) More dogs^{SU} [than cats^{SU} ate mice]$_{DO}$
ate more rats^{DO} [than cats ate mice^{DO}]

Semantically, this correlation will have the following effect: the first *than*-clause will express property (146) and the second one will express property (147):

(146) λn[n cats ate mice]

(147) λm[cats ate m mice]

We are now assuming a semantics for comparative NPs which parallels the semantics of comparative APs. Let us assume, for convenience, a meaning rule in the syle of Lewis-Seuren.

(148) $[_{NP}$ more Ns $\overline{S}_{than}]$ is that higher-order property which is
true of a plural property P[26] *iff* $(\exists x)(\exists n)$
[x is a set of Ns & x has at least n-many members
& P(x) & ~(For at least n: $\overline{S}_{than}$ is true of n)]

Let us assume now that the *ate* in (148) is a two-place plural predicate. Furthermore, we assume that we have a quantifier rule that raises the two NPs in (148) leaving a λ-bound variable at the original place. Thus we get (149) from (145):

(149) [More dogs^{SU} than cats^{SU} ate mice] λx([more rats^{DO} than cats ate mice^{DO}] λy(x ate y))

If we evaluate this according to our assumptions, (146) and (147), we get the following proposition:

(150) $(\exists x)(\exists n)$[x is a set of dogs & x has at least n-many members $(\exists y)(\exists m)$[y is a set of rats & y has at least m-many members & x ate y & ~(For at least m: cats ate m mice)] & ~(For at least n: n cats ate mice)]

I claim that (150) is exactly what we wanted to have. In order to see this, let us restructure (150) a bit. This will give us the more readable formula (151):

(151) $(\exists xy)(\exists nm)$[x is a set of dogs & y is a set of rats & x has at least n members & y has at least m members &x ate y & ~(For at least n: n cats ate mice) & ~(For at least m: cats ate m mice)]

Notice first (151) does not imply that cats ate mice. This seems to be correct to me because

(140) More dogs ate more rats than cats ate mice

is true, if, for instance, three rats were eaten by dogs but no mouse was eaten at all.

In order to see that (151) is equivalent with our intuitive account of (140)'s truth-conditions, *i.e.*, with (141), let us leave out from (151) the last conjunct, *i.e.* the condition ~(For at least m: cats are m mice). This will give us (152):

(152) $(\exists xy)(\exists mn)$[x is a set of dogs & y is a set of rats & x has at least n members & y has at least m members & x ate y & ~(For at least n: n cats ate mice)]

This means that the number of dogs that ate rats is greater than the number of cats that ate mice. This is the first half of condition (141). In order to obtain the second half of (141), we leave away the condition ~(For at least m: cats ate m mice).

This will give us the second conjunct of (141) by analogous reasoning, as before.

Let us return now to Chomsky's example (136)(a). We have asked what its truth conditions are. We have gained a certain security from the treatment of the simpler example (140) and can therfore answer the question in a rather straightforward way. The truth-conditions of (136)(a) are the following ones:

(153) The number of silly lectures given by boring professors is greater than the number of silly lectures such that I expected them to be given by boring professors
and
the number of boring professors who gave silly lectures is greater than the number of boring professors such that I expected them to give boring lectures.

We get (153) from (139) if we associate the subject of (139) with the subject of the reconstructed *than*-clause and *by*-NP of the matrix with the *by*-NP of the *than*-clause, as is indicated in (154):

(154) *More[silly lectures$]_i^{SU}$ than I would have expected that $[_{S_1}$ [Silly lectures$]_i^{SU}$ would be given by boring professors] have been given by more [boring professors$]^{BY}$ than I would have expected that $[_{S_2}$ silly lectures would be given by [boring professors$]^{BY}$]

We have to make sure that the two *than's* express the following two properties, respectively:

(155) a. λn[I expected that n silly lectures would be given by boring professors]
b. λm[I expected that silly lectures would be given by m boring professors]

Using the same procedure as before, this treatment will give us exactly the truth-conditions (153) for (154).

Let us come back to the question asked initially. Why is the phenomenon of multiple head comparatives restricted to plural and mass nouns? Before answering the question, let me give two more examples: (156) is a case where the heads of the 'extraposed' *than*-clause are mass nouns, (157) is a mixed case.

(156) [Less land$]^i$ produces [more corn$]^i$ [than ever before$]^i$

(157) No airline saves you more money in more ways than Delta.[27]

Let us compare (156) with (158) in order to answer our question, *i.e.* why is the former better than the latter?

(158) ? [A greater man$]^i$ would be [a better man$]^i$ - [than Otto$]^i$

As before, the superscript indicates that the *than*-clause has to be associated with both bracketed NPs.

Let us consider first (156). This sentence expresses the following proposition.

(159) The amount of land which produces corn is smaller than the amount of land which produced corn ever before
and

the amount of corn produced by land is larger than the amount of corn produced by land ever before.

We get this reading by familiar reasoning, if we associate the reconstructed *than*-clause with both head - NP and if we 'link' the subject of the *than*-clause to the subject of the main clause and the object op the *than*-clause to the object of the main clause, as indicated in (160).

(160) [Less land]^{SU} [than [land]^{SU} ever before produced corn] produce [more corn]^{OBJ} [than land ever before produced [corn]^{OBJ}]

If we try to apply this method to (158) we see at once, where the difference is with respect to (156). Before associating the *than*-clause with the two head-NPs, we have to ask how we have to complete it. Let us repeat (158) as (161) for convenience:

(161) [A greater man]^{i} would be [a better man]^{i} - [than Otto]^{i}

We have two possibilities to fill up the *than*-clause:

(162) a. than Otto is a great man
b. than Otto is a good man

If we choose the first alternative, we get a complement that fits the first NP but not the second, and if we choose the second alternative, we have it the other way round. We would, of course, first associate the *than*-clause in (161) with the two different NPs and only then complete it in different ways. But the low acceptability of (161) shows that our grammar doesn't work that way. It seems, then, that the restrictions for the reconstruction of a *than*-phrase are rather syntactic than semantic.

As said initially, the treatment of multihead comparative constructions offers no new insight into the semantics of comparison. But they are a nice example of an application of the theory. Rather simple principles enable us to analyse sentences which are at first sight so complex that some people even doubt that they are semantically well-formed.[28]

XI. CAN DEGREES BE ELIMINATED?

All the analyses we have considered so far rely on the notion of degree. We have not been concerned with the question, what degrees are. (For this question, cf. Cresswell [1976] and Klein [1980].) Suffice it to say, whatever they are, they are highly abstract entities. Consider now the following two sentences.

(163) a. John is taller than Mary
b. John's height exceeds Mary's height

We have the feeling that the former is somehow more concrete than the second. It states a relation between John and Mary, two concrete objects, whereas the second states a relation between John's height and Mary's height, *i.e.* between two degrees (abstract objects).

Now, details aside, the semantic theories we have been considering in this article, treat (164)(a) and (b) exactly alike, *viz.* as propositions involving degrees.

It seems desirable to develop an analysis that does not involve the notion of degree for sentences like (163)(a). This is one of the motivations for E. Klein's theory of comparison (Klein [1980]).

In this section I will show two things.

First, in its actual form, Klein's analysis is incorrect, because it makes certain wrong predictions.

Secondly, degrees cannot be eliminated in the general case. Therefore Klein's theory is necessarily incomplete.

Now, what is Klein's theory? His main idea (which goes back to Wheeler [1972]) is that an adjective like *tall* has to be interpreted relative to a *comparison class,* called *context.* If c is a context, *tall* devides c into the definitely tall things (if there are any), the definitely not-tall things (if there are any) and those things which are neither definitely tall nor definitely not-tall. Thus a sentence like:

(164) *Mary is tall*

is true with respect to c, if Mary is among the definitely tall things in c with respect to c, false if she is among the definitely not-tall things in c with respect to c, and truth-valueless otherwise. The sentence

(163) a. *John is taller than Mary*

is true with respect to c iff (165) holds:

(165) $(\exists c' \subseteq c)$(*John is tall* is true w.r.t. c' & $\sim$(*Mary is tall* is true w.r.t. c'))

It can be shown that, in Klein's theory, (165) is equivalent with (166):

(166) *John is tall* is true w.r.t. {John, Mary} & $\sim$(*Mary is tall* is true w.r.t. {John, Mary})

(166) grasps the idea that (163)(a) is about John and Mary and that in order to verify this sentence we must compare the two.

Let us make this more precise. I will restrict the discussion to an extensional version of an adjective semantics incorporating these ideas. It is obvious how it carries over to an intensional frame-work.

Let a *proposition* be a partial function from the set of comparison classes into the set of truth-values $\{0,1\}$. The proposition p is *true* w.r.t. the comparison class c iff p(c) = 1, and p is false w.r.t. c iff p(c) = 0.

Let us denote the denotation of an expression α by $|\alpha|$. Then, in Klein's semantics, the following is assumed:

(167) |*Mary is tall*|(c) = 1 iff Mary is definitely tall w.r.t. c and |*Mary is tall*|(c) = 0 *iff* Mary is definitely not tall w.r.t. c

Now, Klein presumably identifies the definitely not tall objects with the short objects (though he is not entirely explicit about this). Hoepelman [1982: 11] remarks that this has the consequence that the following two sentences are treated alike, given Klein's interpretation rules for the connectives (Klein [1980: 11]).

(168) a. Mary is tall and short
b. Mary is neither tall nor short

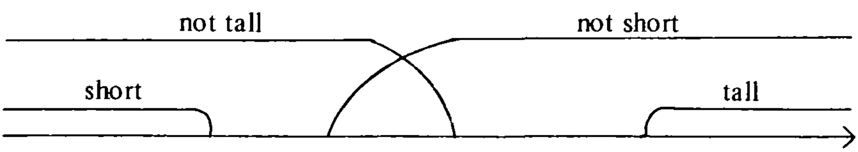
But, as Hoepelman correctly notices, there is an intuitively given difference: (a) is always false whereas (b) is contingent.

This criticism assumes the same meaning for *not tall* and *short*. Indeed, if *short* where the same as *not tall,* then (168)(b) would mean the same as (168)(c):

(168) c. Mary is not tall and she is not not-tall

This can never be true. We could try to save Klein's adjective semantics against Hoepelman's criticism if we assume the following partition for any comparison class:

(169)



The arrow symbolizes an ordering of the members of c according to increasing tallness.

Such a move would help Klein against the particular criticism of Hoepelman. But this move seems rather artificial. For instance, we have to stipulate that the extensions *not tall* and *not short* can overlap, because otherwise the sentence

(170) Mary is not tall and she is not short

can't be true. It is not clear then, what the general semantics of negation would be within that system.

And an accommodation of Klein's theory along these lines would not save it from other objections. It seems to me that the following sentences could never be treated within Klein's frame-work:

(171) a. John is six inches taller than Mary
b. Ede is twice as fat as Angelika
c. Ede is more tall than broad

I have no idea how (171)(a) and (b) can be incorporated in Klein's system. I am inclined to think that this is not possible. I have tried hard, but without success. In order to convince the reader of the plausibility of my conjecture, I should say a bit more about Klein's comparative rule, but I can't go into this for reasons of restricted space.

Concerning (171)(c), this example shows very clearly that degrees can't always be replaced by concrete objects. In order to find out whether (171) is true, we don't compare Ede with himself. We compare Ede's height with his breadth. It is reasonable to assume that this amounts to a comparison of two degrees. An other theory would be to say, that we compare two aspects of Ede. Ede *qua* tall and Ede *qua* broad. No one has, however, proposed a theory of the latter kind. And anyway, such a theory would involve abstract entities, height and breadth, exactly as before.

The main conclusion of this section is that degrees cannot be eliminated in the general case. Thus Klein cannot offer a unified analysis of comparative constructions. I don't know how strong this criticism is.[29] Perhaps a more concrete analysis in Klein's style (without the defect mentioned) should be upheld as long as possible. Only when this is impossible we will switch over to degrees. The second conclusion is that we have to respect the data of (171) for an adequate analysis of the comparative.

Notice that the only theory which has treated cases lik (171)(a) and (b) explicitly, is, to my knowledge Hellan [1981]. It is not easy to treat these examples without more or less drastical changes in Lewis', Seuren's or Cresswell's system. On the other hand, it is precisely cases like (171) which motivate Hellan's analysis, which looks so complicated at first sight. Hellan would analyse (171)(a) and (b) as (172)(a) and (172)(b), respectively:

(172) a. $(\exists d_1,d_2,d_3)$(John is d_1-tall & Mary is d_2-tall & $d_1=d_2+d_3$ & $d_3 = 6$ inches)
b. $(\exists d_1,d_2,d_3)$(Ede is d_1-fat & Angelika is d_2-fat & $d_1=d_2 \cdot d_3$ & $d_3 = 2$)

The positive and equative arise simply as special cases: the 'differential'-variable d_3 is specified by the context. It is a degree > 0 in the case of the comparative and it is the number 1 in the case of the equative. Thus (173)(a) and (b) are analysed as (174)(a) and (b), respectively, in his system:

(173) a. John is taller than Mary
b. John is as tall as Mary

(174) a. $(\exists d_1,d_2,d_3)$ (John is d_1-tall & Mary is d_2-tall & $d_1 = d_2+d_3$ & $d_3 > 0$)
b. $(\exists d_1,d_2,d_3)$ (John is d_1-tall & Mary is d_2-tall & $d_1 = d_2 \cdot d_3$ & $d_3 = 1$)

Thus, this extra complication is well motivated after all, and it must be incorporated in an adequate theory of the comparative.

X. ITERATED MODALITIES AND SCOPE OF *THAN*-CLAUSES

In section VIII, we found a way of justifying a theory that regards the *than*-complement of a comparative adjective as a definite description. Since the time has come to risk a proposal which brings the loose ends of the previous sections together, I want discuss the question of whether we should prefer such an approach or one working with double indexing.

There is a classical argument, in Cresswell [1973: 149f], that can help us decide this question: In certain cases, definite descriptions have a greater expressive power than double indexing. Cresswell discusses the (admittedly somewhat peculiar) sentence:

(175) The major allowably presides necessarily

If we treat *the mayor* as a Russellian description we can have a reading which is roughly represented as (176):

(176) Allowably (the mayor λx[necessarily (x presides)])

This means something like (177):

(177) In every permitted world w, the one who is mayor in w, presides in every world w′, which is accessible from w

The point is that the nominal *the mayor* is within the scope of the modal operator *allowably* but outside the scope of the model *necessarily.* If we treat *the mayor* by means of double indexing, that is, as a name where *the mayor* (w) denotes the mayor in w and *the actual mayor* (w) denotes the mayor in w_0, where w_0 is the real world, we can't express this, because then, the mayor is either in the scope of both modal operators or it is so to speak, not in the scope of any modal operator.

We can now try to carry this observation over to comparative constructions. Consider the following sentence.

(178) I thought Plato could have been more boring

I think there is a reading of (178) which can be rendered as something like (179):

(179) For every world of my thoughts w, there was a world w′ accessible from w, such that the degree of Plato's boringness in w′ exceeded the degree of Plato's boringness in w

We can represent this reading[30] easily in a theory which treats the *than*-complement of a comparative as a definite description. But we can't represent it in one of our double indexing approaches. A Russellian account of (179) would be (180):

(180) I thought (the d [Plato is d-boring] $\lambda d'[\Diamond$(the d″[Plato is d″-boring] $> d')])$

The only reasonable candidate for a representation of (179) in a double indexing theory, say Seuren's, would be (181):

(181) I thought $\Diamond(\exists d)$(Plato is $\geqslant$ d-boring & $\sim$ACTUALLY(Plato is $\geqslant$ d-boring)]

But this is not the reading (179). (181) rather means something like (182):

(182) In the worlds of my thoughts it was possible that Plato was more boring than he really was

One could try to express (179) by means of a more complicated paraphrase, say (183):

(183) I thought $\Diamond$[($\exists$d)(Plato is $\geqslant$ d-boring & ~(ACTUALLY(I thought [Plato is $\geqslant$ d-boring]))]

But such a move wouldn't help. Some reflection will show that (183) is inconsistent.

The question is, of course, whether the reading (179) exists for (178). I think it does. If (178) is not the right kind of example then it shows at least what kind of example is needed in order to show that the Russellian approach enjoys a generality the double-indexing theories can't reach.

On the other hand, we know that we can't take over Russell's solution in its existing shape. It will be the task of the last section to develop a proposal which is adequate with respect to all facts observed in this article.

XI. SYNTHESIS

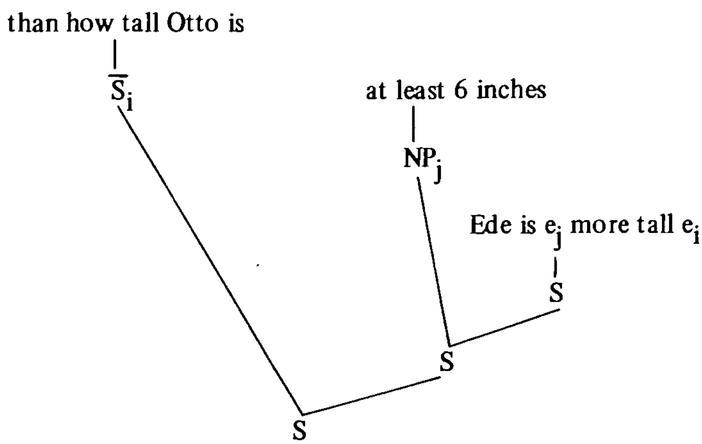
1. *The comparative*

Let us start with the comparative. We use Montague's strategy and generalize to the worst case. Here is one.

(184) Ede is at least 6 inches taller than Otto is

We get the right interpretation for this if we assume that the 'logical form' of (184) is something like:

(185)



Before I explain this picture let me give an idea of the interpretation I have in mind for (185). It is this:

(186) the d[Otto is d-tall] $\lambda d_i(\exists d_j, d_j \geqslant 6$ inches) [Ede is $d_j + d_i$-tall])

The wide scope of the description doesn't matter in this case. Therefore, this amounts to:

(187) $(\exists d_j, d_j \geqslant 6$ inches) (Ede is d_j + the d [Otto is d-tall] -tall]

This seems an adequate account of the truth-conditions of (184). Let me explain next how we get to (187) from the surface. I will give an informal account which is supposed to be precise enough for everyone to translate into his favourite grammatical frame-work. The surface of (185) is:

(188) Ede is $[_{NP}$ at least 6 inches] $\left\{\begin{matrix}\text{more}\\ \text{er}\end{matrix}\right\}$ tall $[_{\bar{S}}$ than how tall Otto is]

The modifier *more* or *-er* is the comparative morpheme.

The analysis of the *than*-complement is inspired by Chomsky [1977]. I am assuming that the 'deep-structure' of the *than*-clause is:

(189) than Otto is how tall

By *wh*-movement we get (190)(a) or, perhaps, (b):

(190) a. than [how tall]$_i$ Otto is e_i
b. than how$_i$ Otto is e_i tall

The *than*-complement might also be the clause (191)(a) or (b), possibilities discussed by Chomsky but immaterial for my purposes.

(191) a. than what$_i$ Otto is e_i
b. $[_{PP}$ than Otto]

The only thing that matters here is the following assumption:

A1. The *than*-complement determines a property of degrees.

The most plausible surface structure for illustrating this assumption is (190)(b). We may think of the semantics of *wh*-movement simply as property abstraction. Thus, (190)(b) determines the property λd[Otto is d-tall]. We can form this in a very sloppy way as a rule:

R1. *Property abstraction*
how$_i$ [.....i-A...] expresses the property of degrees λd[...d-A...].

The idea is now that any *than-* or *as*-complement is, semantically, a nominal. Since nominals have scope, they must in general be raised in the way indicated in (185). This holds for sentential complements generally. It is interesting to note that only very few people have noticed this, one of these being Cresswell [1973: 165ff]. But, as far as I can see, this far-reaching idea has not entered into the linquistic literature.

A proper way of nominalizing the property of degrees expressed by a comparative complement is to make a definite description out of it. In the case we are discussing here, the application of the usual definite article will yield the right result. But from the discussion in section VIII we know that we have to 'maximize' the property of degrees to get the right results. Thus the nominalization of the *than-* or *as*-clause can be described in the following way:

R2. *Nominalization of comparative complements.*

Suppose, the "S" in "*than* S" or "*as* S" determines the property of degrees P. Then the whole phrase determines the nominal-meaning:

the (*Max*(P))

Let us repeat, for convenience the definition of the functions, *Max* and *the*.

(192) Let P_1, P_2 be say first-order properties. *the*$(P_1)(P_2)$ is the proposition true in a world w if $(\exists x)(\forall y)[[w \in P_1(x) \Leftrightarrow x = y]$ & $w \in P_2(x)]$

I am assuming an ontology like Cresswell [1973], i.e., I think of propositions as sets of possible worlds and of properties as functions mapping individuals into propositions. And I am assuming, of course, the usual semantic vocabulary.

Thus, *the* is just Russell's definite description-operator.

(193) Let P be any property of degrees. Than *Max*(P) is that property which is true of any degree d in a world w iff $P(d)$ & $\sim(\exists d')$ $[w \in P(d')$ & $d' > d]$

This is just a repetition of (117). If we take our rules together, then the meaning of our $\overline{S}$ *than how tall Otto is* can be roughly described in the following way.

(194) Let P be any property of degrees and let w be any world. Then ‖than how tall Otto is‖ (P) is true in w iff the maximal degree such that Otto is d-tall has property P in w

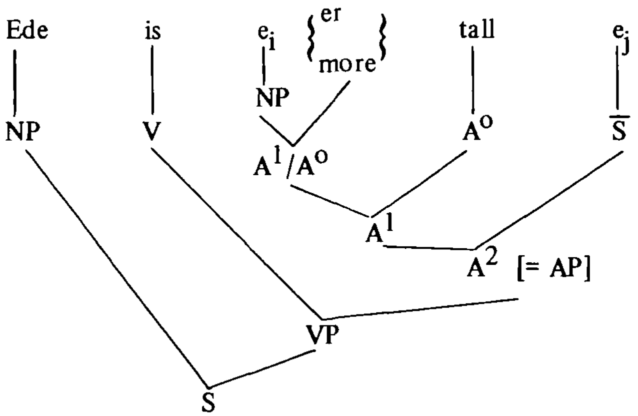
(If we spell that out properly according to (192) and (193), the metalinguistic locution gets much longer, of course.)

There is a second nominal in (185), *viz. at least 6 inches.* If we disregard its internal structure, its meaning can be described in the following way:

(195) ‖*at least 6 inches*‖ is true in world w of a property of degrees P iff $(\exists d)(d \geqslant 6 \text{ inches} \;\&\; w \in P(d))$

What we need in order to interpret (185) completely is to have an idea what the comparative morpheme *-er* (which we represented as *more*) means. I think of the internal structure of the deepest S in (185) as something like the following:

(196)



As Cresswell [1976] and Hellan [1981], I am assuming, that *tall* denotes a two-place relation connecting individuals with degrees, *i.e.* we have the following rule:

R3: Let x be an individual and d a degree, then we have for any world w: $w \in$ ‖*tall*‖ (x,d) *iff* d is a degree of tallness & x has d in w.

For a precise analysis of the metalinguistic locution used in the definiens, *vide* Cresswell [1976: 267].

Before spelling out the meaning rule for *more* it is useful to reflect on the logical type of its meaning. We can read it off from (196) in the following way. In Cresswell's ontology, we would have the following correspondence between the syntactic categories assumed and types:

$$A^o \longrightarrow \langle 0,1,1\rangle$$
$$A^1 \longrightarrow \langle\langle 0,1\rangle,1\rangle$$

Hence, $A^1/A^o \longrightarrow \langle\langle\langle 0,1\rangle,1\rangle,\langle 0,1,1\rangle\rangle$

Since *more* takes a degree and gives us a A^1/A^o, it must be of type $\langle\langle\langle\langle 0,1\rangle,1\rangle,\langle 0,1,1\rangle\rangle,1\rangle$. In what follows, I will denote this type by $\langle\overline{A^1/A^o},1\rangle$. Similarly for other complex types. I don't think of $\rightarrow$ as a function but as a relation. One category determines a lot of types, in general.

Now, the meaning of *more* (*-er*) is roughly this:

R4: *Comparative*

Let d_1,d_2 be any degrees, A^o an appropriate relation of type $\langle 0,1,1\rangle$, x any individual and w any world. Then

$$w \in \| \begin{Bmatrix} \textit{more} \\ \textit{er} \end{Bmatrix} \| (d_1)(A^o)(d_2)(x) \textit{ iff } w \in A^o(x,d_1+d_2)$$

As is seen from a glance at the definiens, this meaning rule is, in fact, very simple. The complicated type of the function $\|$*more*$\|$ is due to the particular descriptive frame-work I am assuming. In Montague's PTQ, *more* would even be of type $\langle\langle s,e\rangle\langle\langle s,\langle\langle s,e\rangle,\langle\langle s,e\rangle,t\rangle\rangle\rangle,\langle\langle s,e\rangle,\langle\langle s,e\rangle, t\rangle\rangle\rangle\rangle$, I guess.

In order to evaluate (185) we need the generally accepted assumption:

A2. Raised nominals are evaluated according to the familiar quantifying in rule (cf. Montague [1974]).

It is convenient to repeat (185) in a way that reflects the logical structure which it has according to our assumptions:

(197) the maximal d.s.t. Otto is d-tall λd_2[at least 6 inches λd_1 [Ede is d_1+d_2-tall]]

According to our rules, this proposition is true in a world w, iff (198) holds good.

(198) $(\exists d_1)[d_1 \geqslant 6$ inches & Ede is d_1+d_2-tall in w & d_2 = the max. d, s.t. Otto is d-tall in w]

If we look at the meaning rule (R4), we find that $\|$*more*$\|$ is a four place relation which has two degrees among its arguments. In the example we have discussed, these arguments where specified linguistically. This is, however, often not the case. Take the sentence:

(197) Ede is taller than Otto

Here, the 'differential' degree d_1 is not expressed. It has therefore to be supplied by the context. Hence, the 'logical form' of (197) is something like:

(198) For some degree $d_1 > 0$, Ede is d_1+d_2-tall, where d_2 is the d s.t. Otto is d-tall

In (198), the nominal "For some degree $d_1 > 0$" has to be supplied by the context. There are cases, where also the second degree d_2 has to be reconstructed from the context, for instance, if we say:

(199) Plato is more boring

The present analysis follows the strategy that we make explicit those parameters in the logical analysis, which can be made explicit in a natural way linguistically. As (184) shows, both degree positions can be explicit. Thus, our rule (R4) is not artificial, it seems to me.

2. *The equative*

We can analyse the equative in a perfectly parallel way. The meaning rule for *as* is this:

R5. $w \in [\![as]\!](d_1)(A^o)(d_2)(x)$ *iff* $w \in A^o(x, d_1 \cdot d_2)$

Here, d_1 is a real number and the other parameters are as in (R4).
Let us apply this.

(200) Aristotle is at most half as boring as Plato

The logical form of (200) is roughly:

(201) the maximal d s.t. Plato is d-boring λd_1 [at most half (times) λn [Aristotle is $n \cdot$ d-boring]]

In order to see that this is right, we only need an appropriate meaning rule for *at most half*:

(202) $w \in [\![$*at most half(times)*$]\!](P)$ iff $(\exists n)(n \leq \frac{1}{2}$ & $w \in P(n))$

Here, w is any world and P is a property of real numbers.

I leave it to the reader to check that this semantics has exactly the desired effects.

One of the nice features of this analysis of the equative is that we don't have to worry (as most authors have, *e.g.* Cresswell), whether sentence (203)(a) means (203)(b) or (c).

(203) a. Tristan is as heavy as Randi
b. Tristan is exactly as heavy as Randi
c. Tristan is at least as heavy as Randi

(203)(a) can have both meanings. In one case the context supplies the quotient ‖*(exactly) one (times)*‖ and the second case it supplies ‖*at least one (times)*‖. Since ordinary language hardly ever makes the quotient explicit, equatives are vague with respect to the two readings. It would be wrong to build the 'exactly' or 'at least' into the meaning of *as*, I believe. In a way, *as* means simply multiplication and *more* means addition.

It should be obvious from remarks at the end of section IX that I am indebted to Lars Hellan for this insight.

3. *The positive*

We have to speak about the positive because the adequacy of an adjective semantics can be judged only if it covers in a natural way both the positive and the comparative adjective.

Recall that *fat* is a relation between things and degrees (of fatness). The simplest way of representing

(204) Ede is fat

is to leave the degree variable open:

(205) Ede is d-fat

We would say than that it is the task of the context to specify the degree of fatness which Ede really has. In the same way, the context had to specify the d_1-variable in the case of the comparative (cf. section 1). This approach, however, would not be satisfactory. The meaning of (204) can't be satisfied by just any degree. Everyone is fat in some sense (cf. for this, Lewis [1970: 65]). So (204) would be trivially true. Since this is obviously not the case, (204) means something as 'Ede is positively fat'. The operator 'positively', call it *pos,* is invisible, which made E. Klein think that it doesn't exist. I will come back to this.

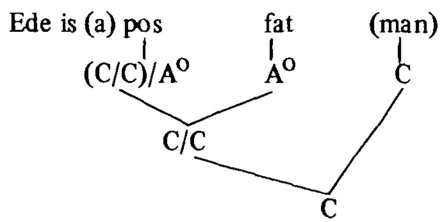
Now, what should the semantics of *pos* be? From the literature we know that it is heavily context-dependent (cf., *e.g.* Cresswell [1976] or Kaiser [1979] or Aristotle [Categories]). Take Aristotle's examples.

(206) a. This corn is big
b. This mountain is small
c. In the summer, there are only few people in Athens, but there are a lot of people in the village

According to Aristotle, (a) means that this corn is big with respect to other corns, (b) means that this mountain is small with respect to other mountains, and (c) means that the number of the people dwelling in Athens in summer time is small with respect to the number of people who stay in Athens during the winter, but the other way round for the village.

We can incorporate this insight into the semantics of the positive operator by stipulating that the degree determined by *pos* depends on a comparison class (*vide* Cresswell [1976] for this).

The comparison class can be given explicitly (= linguistically) or implicitly (*i.e.*, by the context). These considerations suggest that the 'logical form' of (204) is something like (207)

(207)


I have used the label C for 'comparison class' (or 'common nouns'). Semantically, C's are properties, *i.e.* functions of type ⟨0,1⟩. Hence *pos* takes an adjective and makes a C-modifier out of it. We are now ready for a rough semantics of *pos*.

R6. *Positive*
Let A^O be any adjective meaning, C be any appropriate property, x be any appropriate individual and w be any world. Then w ϵ $[\![pos]\!]$ $(A^O)(C)(x)$ iff $(\exists d)$ [d is an A^O-degree & d > average $[A^O,C]$ & x has d in w & w ϵ C(x)].

I am not interested in analysing the notion "average $[A^O,C]$". If A^O = $[\![tall]\!]$ and C = $[\![man]\!]$, then this can be read as "the average tallness for men". It is an empirical question how the notion "average" has to be analysed; it must depend on the parameters A^O and C, or perhaps even on w. Whatever the truth may be, I think the meaning of the positive must be something like the function described in (R6).

Let me make a brief remark about the difference between predicative and attributive adjectives. I think the semantic difference consists in the fact that the morphology indicates that in the first case the comparison class is implicit whereas it is explicit in the second case. Thus, (208)(a) is analysed as (208)(b), where C is a property variable the context has to specify.

(208) a. Ede is $[_{AP}[_{\substack{A\\ +pred}}$ fat$]]$

b. Ede is $[[_{C/C}$ pos$[_{A^o}$ fat$]]C]$

It is the impact of the feature [+ predicative] which requires the presence of the property variable C in logical form. The attributive adjective is analysed as (209)(a)-(b):

(209) a. Ede is a $[_{N}[_{\substack{A\\ +attr}}$ fat$]$ man$]$

b. Ede is a $[[_{C/C}$ pos $[_{A^o}$ fat$]]$man$]$

We can think of a positive adjective as modifying always the next common noun, simple or complex. This assumption would force us to assume an *and* in 'logical form' between appositive adjectives:

(210) Ede is a fat young scholar

The logical form (211)(a) determines the reading (212)(a) whereas (211)(b) represents (212)(b)

(211) a. Ede is a pos-fat (pos-young(scholar))

b. Ede is a [pos-fat and pos-young] scholar[31]

(212) a. Ede is fat as young scholars go and he is a scholar

b. Ede is fat as scholars go and young as scholars go and he is a scholar

The positive operator specifies in a way the degree variable of the adjective. Therefore we can't have it if the degree is overtly specified, as in (213):

(213) Ede is 6 feet tall

(213) will roughly have the logical form

(214) tall (Ede, 6 feet)

This creates a certain problem for the interpretation of

(215) Ede is a 6 feet tall scholar

"6 feet tall' is simply a property and not a relation between individuals and properties like "pos tall". There are a number of possibilities for solving the problem. The simplest one is to work with property conjunction, i.e. we make out of the two properties "6 feet tall" and "scholar" the complex property of "being a scholar and 6 feet tall". We need this rule anyway, if we want interpret relative clauses, e.g.

(216) Ede is a scholar who is 6 feet tall

Thus, (215) requires no new principles of interpretation.

Let me conclude this section with a remark concerning the relation between positive and comparative. Ewan Klein [1980: 2] says that we should derive the meaning of *Aer* from the meaning of *A*, where *A* is the positive adjective. This must be so by the principle of compositionality. Obviously we have not followed Klein's advice.

I find Klein's argument not convincing for the following reasons. I have assumed that the positive adjective may be regarded as

A-stem + Positive morpheme

whereas the comparative is analysed as

A-stem + Comparative morpheme

Since the positive is the unmarked form, the positive morpheme is not visible. Obviously, my analysis conforms the compositionality principle. Klein would admit this but he thinks that this is merely lip-service. He stipulates that a positive adjective should express a one-place property and that the comparative has to be shaped out of that. I have argued elsewhere, at length, that adjectives can't be one-place properties (cf. v. Stechow [1982]). I think Klein's flaw is this: if certain parameters are left implicit in the language we should not conclude from this fact that they are logically dispensable. The situation of having to introduce extra parameters for the semantic analysis of language is not uncommon. Indeed, this is what almost always happens. Take, for instance, the modals *must* or *can*. Syntactically these take naked infinitives as complements and make VPs out of it. Think for convenience of them as sentential operators, i.e. as one-place functors. It is a common place that, semantically, these operators have to be two-place operators. |*must*|(p) means something like "p follows from q", where q is a proposition to be specified by the context ('the modal background', cf. Kratzer [1978]). Similarly, |*can*|(p) means that p is compatible with the background proposition q. It would be hopeless to attempt to think of |*must*| and |*can*| as one-place functions simply because they are syntactically one-place connectives.

If there were a straightforward connection between the positive and the comparative then it would be just the opposite of what Klein assumes: *tall* means "taller than x", and the context tells you what x is.

4. THE CROSS-CATEGORY GENERALIZATION OF THE COMPARATIVE AND POSITIVE

In this section I want to show that my treatment of the comparative and positive generalizes across the relevant categories: adjectives, mass nouns, plural nouns and certain adverbs. I will show in detail only how the analysis carries over to plural and mass nouns. It is obvious how adverbs have to be treated. I am relying heavily on ideas found in Cresswell [1976] in order to formulate the relevant generalizations. Let me start with comparative plural (and mass) nouns:

(217) At least 6 more toads than frogs croak

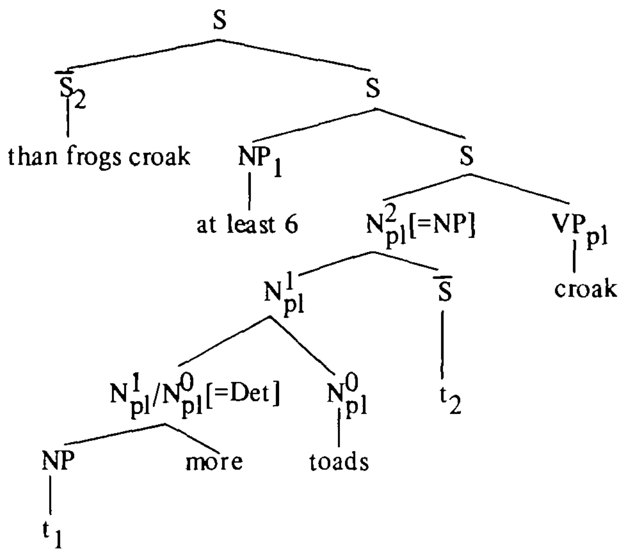
(218) Ede owns at most 3 ounces more gold than Kurt

The truth-conditions of (217) can be read off from the following formula.

(219) the maximal n s.t. $\exists X$[X is an n-membered set of frogs who croak]λd_2 [At least 6 λd_1[$\exists Y$[Y is a d_1+d_2-membered set of toads who croak]]]

In order to see how we get these truth-conditions, it is useful to get an idea of how we analyse the nominal *at least 6 more toads*. Cresswell's idea is that semantically, plural nouns (and mass nouns) are similar to adjectives. ‖*toads*‖ is a relation connecting a set X of individuals and a cardinality n('set-degree') iff X is a set of toads and X has the cardinality n. It is pretty obvious now that we can apply the meaning of ‖*more*‖ to ‖*toads*‖, two cardinalities d_1 and d_2 and a set X. The only thing we have to make sure of is that *at least 6 more toads* is a nominal. This is achieved by analysing *at least 6 more* as a complex determiner. Before stating the relevant meaning rules, it is good to have an idea of what the logical structure of our nominal is. We illustrate this by drawing a tree which is, in all relevant respects, parallel to (196), i.e. the tree motivating our rules (R3) and (R4). This observation will allow an immediate understanding of the following meaning rules:

(220)



We have more or less the same category-type correspondence as in the case of (196), i.e.:

N^0_{pl} — $<0,1,1>$
N^1_{pl} — $<<0,1>,1>$
N^1_{pl}/N^0_{pl} — $<\overline{N^1_{pl}}, \overline{N^0_{pl}}>$
more — $<\overline{N^1_{pl}/N^0_{pl}},1>$

Let us indicate the meaning rules now. I will begin with the plural rule (cf. Cresswell [1976; Def. 3.6]). A plural noun like *toads* has the following structure: *toad-pl,* where *pl* is the plural morpheme. *pl* is interpreted as an operator of type $<<0,1,1>,<0,1>>$. Its meaning is this:

R7. Let P be a first-order property, X be a set of individuals, n be a number, and let w be any world. Then $w \in [\![pl]\!](P)(X,n)$ iff $(\forall x \in X)$ [$w \in P(x)$ & the cardinality of X is n].

This rule presupposes that 'plural-degrees' are numbers with their usual ordering. Thus, '+' and '·' have their ordinary meaning in this case. According to this rule, ‖*toads*‖, i.e. ‖*pl*‖(‖*toad*‖) applies to <X,n> iff X is an n-membered set of toads.

The meaning rule of our determiner forming *more* is exactly parallel to adverb forming *more* (R4). It is this:

R8. $w \in$ ‖*more*‖ $(d_1)(N^0_{pl})(d_2)(VP_{pl}) \Leftrightarrow \exists X[w \in N^0_{pl}(X, d_1 + d_2)$ & $w \in VP_{pl}(X)]$.

This rule presupposes that N^0_{pl} is the meaning of a plural noun, i.e. something like ‖*toads*‖. VP_{pl} must be the meaning of a plural-VP. The difference between singular and plural-VPs is simply that the latter apply to sets of individuals. Thus, ‖*croak*$_{pl}$‖ is true of a set of individuals X iff each member of X croaks. I don't want to spell this out here. For a detailed account, vide v. Stechow [1980].

According to our rules, the NP [t_1 *more toads* t_2] in (220) means $\lambda VP[\exists X(X$ is a d_1+d_2-membered set of toads which VP)].

I leave it to the reader to check that the entire (220) does indeed mean (219).

Constructions involving mass nouns are analysed in exactly the same way. The only difference is that mass nouns are inherently relational. There is no need to apply an operator like ‖*pl*‖ on them. Thus the meaning of *gold* is something like this.

R8. ‖*gold*‖ is a two-place relation which holds good of a mass x and an amount *a* iff x is gold and *A* is the amount of the totality of x.

It should be obvious that this semantics (Cresswell's) enables us to analyse (218) in a way which is exactly parallel to our analysis of (217). I don't want to spell this out here.

Everything we have said carries over to the analysis of positive plural and mass nouns. Since the details are messy I will content myself with a few remarks. Consider the following data.

(221) a. How much gold [$_{VP}$ is in Ede's pocket]?
b. Much gold VP
c. At least 6 ounces (of) gold VP
d. A large amount of red gold VP
e. Gold VP

(221)(a) suggests that we should analyse the NP *how much gold* as (222).

(222) d-much gold

with the following semantics for the determiner *d-much*:

R9. $w \in \|d\text{-}much\|(N_{mass})(VP)$ iff $\exists X[w \in N_{mass}(X,d)\ \&\ w \in VP(X)]$.

The logical form of (221)(c) is then roughly (223):

(223) $(\exists d, d \geqslant 6$ ounces) [d-much gold VP]

The determiner *d-much* is determined by an appropriate rule from an amount degree *d* and the mass adjective *much* whose meaning is:

R10. $w \in \|much\|(X,a)$ iff *a* is the amount of the totality of X in w.

This presupposes that X is a mass and *a* is an amount. Since *much* is an ordinary adjective stem, we can apply the positive operator to it. This will give us *pos much*. Since this adjective occurs only in determiner position, we have to describe its semantics analogously to the determiner *d-much*. Thus we arrive at the following meaning rule:

R11. $w \in \|pos\ much\|(N_{mass})(C)(VP)$ *iff* the maximal d [d is an N_{mass}-degree & $\exists X(w \in N_{mass}(X,d)\ \&\ w \in VP(X))] >$ average $[N_{mass},C]$

According to this rule, *Much gold C is in Ede's pocket* is true iff the maximal *d* such that *d* is a degree of gold in Ede's pocket which exceeds the average of gold with respect to the comparison class C. If we assume an analogous rule for *pos little* with "<" instead of ">", then we obtain the desired result that *Much gold is in Ede's pocket* is incompatible with *Little gold is in Ede's pocket.*

There is a great deal more to be said about this. For instance, what is the nature of the comparison property C, which is practically never specified linguistically? The only thing I have assumed is that it is a property of masses. Clearly, we want to know more about the relevant contextual parameters which determine C. I have nothing to say about this. But, details aside, the analysis of positive mass nouns seems correct. And it carries over, of course, to positive plural nouns. Here we have to work with the determiners *how many, n-many* and *pos many.*

It should be obvious that we can analyse positive and comparative adverbs in the same spirit:

(224) a. Tristan yells loudly
b. How loudly does Tristan yell?
c. Tristan yells three times as loudly as Otto

I don't want to work out the details for these cases. The general strategy is as before. The logical form of (224)(c) is roughly this.

(225) [as how loudly Otto yells]$_2$ [three times]$_1$ Tristan yells $d_1 \cdot d_2$-loudly

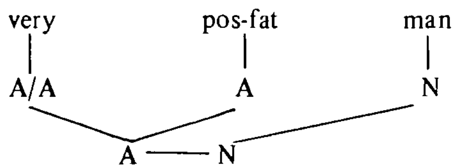
I leave it as homework for the reader to develop an appropriate semantics for the adverb forming suffix-*ly*. If we have it we get the right semantics, given the assumption that *loud* is an ordinary adjective. Thus our semantics of the comparative and equative generalizes over all relevant categories.

5. DEGREE MODIFIERS

Let me add some remarks about words like *very* or *fairly*. Wheeler [1972] has proposed that the property *very fat man* gives us the individuals which are fat with respect to fat men. And a *fairly fat man* is a fat man but not a very fat man. This idea has been taken up by Klein [1980]. I don't know whether Wheeler's view is correct, because this semantics seems too precise to me. But, if you want to adopt it, it is easy to build in. If you want another semantics, you can have it, too. The following is an accommodation of Wheeler's and Klein's account.

The complex noun *very fat man* has the following logical structures:

(226)



Here A is short for N/N, where N determines the type $\langle 0,1\rangle$. Hence, *very* is of type $\langle N/N, N/N\rangle$. And it has the following meaning:

R12. $w \in \|$*very*$\|(A)(N)(x)$ *iff* $w \in A(A(N))(x)$.

This presupposes that A is a function of type N/N, i.e. something like *pos-fat*, that N is a function of type $\langle 0,1\rangle$, x is an individual, and w is a world. According to our rules, (226) will express the property which is true of an x in a world w iff $(\exists d)$[d is a degree of fatness & d > average [fat, $\|$*pos-fat man*$\|$] & x is d-fat in w & $w \in \|$*pos-fat man*$\|(x)$].

If we evaluate this further we find that this holds good iff x is a man in w who is fatter in w than the average fat man. Thus (R12) is no more than a translation of Wheeler's idea into this frame-work.

The meaning rule for *fairly* is now obvious.

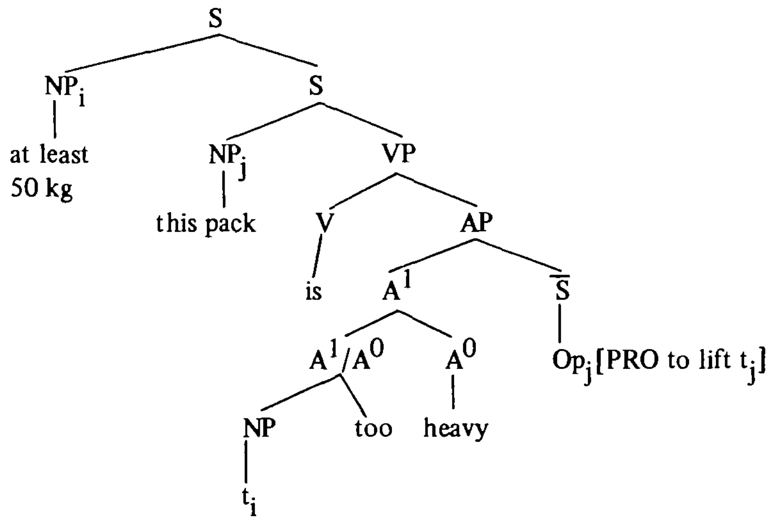
6. *TOO*: A COUNTERFACTUAL COMPARATIVE MORPHEME

The word *too* has the most complicated semantics of the comparative morphemes, as far as I know. Hence it is worth describing semantically. Consider this sentence:

(227) This pack is at least fifty kilos too heavy to lift

The 'logical form' of this is:

(228)



The idea for an analysis of this structure is that it should express the following proposition:

(229) If one could lift this pack, then it would be at least 50 kg less heavy than it actually is

Let us assume that the embedded $\overline{S}$ in (228) expresses a proposition, *viz.* 'One lifts it$_j$', where 'it$_j$' means 'this pack'.[32] According to what we have said in previous sections, the $\overline{S}$ should semantically be a nominal. But in this particular case we can ignore this. So let us assume that the $\overline{S}$ determines the type of propositions 0. Hence *too* is of type $\langle\langle\langle\langle 0,1\rangle, 0\rangle, A^0\rangle,1\rangle$. The difference with the comparative morpheme is simply that the $\overline{S}$-complement of a *too*-adjective doesn't denote a property but a proposition. In the case of (228) it is not obvious why this should be so, but the following example illustrates this point.

(230) The weather is too good to stay at home

Clearly, the embedded infinitival does not express a property which applies to the weather.

Instead of giving a full account of the meaning of *too,* I will simply write down a very sloppy meaning rule which hopefully makes the idea clear:

R13. $\|too\|\ (d_1)(A^0)(p)(x)$ = the max.d $[x$ is d-$A^0]\ \lambda d_2\ [p \,\Box\!\!\rightarrow A^0\,(x,d_2-d_1)]$

Our semantic rules predict the following truth-conditions for (228):

(231) the d [this pack is d-heavy] λd_2 [at least 50 kg λd_1 [One lifts this pack $\Box\!\!\rightarrow$ it is d_2-d_1-heavy]]

A little reflection will show that these truth-conditions are adequate.

7. REVIEWING THE RELEVANT DATA

In this section I will go again through the data considered initially. The claim is that my proposal can deal adequately with all of them.

(RA) *Russell's ambiguity*
I thought your yacht was larger than it was.

The two readings are analysed as (232)(a) and (b):

(232) a. I thought (the max. d[your yacht is d-large] λd_2 [some positive degree λd_1 [your yacht is d_1+d_2-large]])
b. the max. d [your yacht is d-large] λd_2[I thought (Some positive degree λd_1 [your yacht is d_1+d_2-large])]

(232)(a) represents the inconsistent, (b) the consistent thought.

(AC) *Ambiguous counterfactuals*
If Ede had smoked less than he did, he would be healthier than he is

I will represent the nontrivial reading only. It is clear, that we can have the uninformative readings, too. In order to improve readability I leave out the nominal *some positive degree,* which specifies the open 'differential variable' d_1, and leave that variable free.

(233) the max. d [Ede has smoked d-much] λd_2 [the max. d [he is d-healthy] $\lambda d'_2$ [Ede has smoked d_1+d_2-little $\Box\!\!\rightarrow$ he is $d'_1 + d'_2$-healthy]]

Notice that our raising analysis prevents the unwarranted reading (36), which we discussed in section III. This is so because, in our account, the main clause is not a description, which could have wide scope.

(NPI) *Negative polarity*
a. Ede is cleverer than *anyone* of us
b. Max is as well as *ever*

We have to show that the comparative complement is an NPI-context, i.e. where $\overline{S}$ is a comparative complement, we can, sometimes, substitute a more informative $\overline{S}'$. We show this for a special case. The following argument is valid:

(234) Ede is fatter than Randi or Otto is
∴ Ede is fatter than Randi is

Here, $\overline{S}$ is *Randi or Otto is d-fat* and $\overline{S}'$ is *Randi is d-fat.* Clearly, $\overline{S}'$ is more informative than $\overline{S}$. The relevant formalization of (234) in our theory is this:

(235) the(Max(λd[Randi or Otto is d-fat]))λd_2 [$\exists$d,d > 0) λd_1 [Ede is d_1+d_2-fat]]
∴ the (Max(λd[Randi is d-fat])) λd_2 [($\exists$d,d > 0) λd_1 [Ede is d_1+d_2-fat]]

Recall that the semantics of the operators *the* and *Max* has been introduced in (192) and (193). Clearly, this argument is valid. Notice that it is essential that the 'differential variable' is bound by the appropriate existential quantifier in order to make the argument work.

It is obvious that the explanation of the validity of the argument:

(Q&C) *Quantifier and connective embedding*
Ede is fatter than anyone [= someone] of us
∴ Ede is fatter than everyone of us

proceeds along exactly the same lines.

The relevant formalization is the one where *anyone* is within the scope of the *Max*-operator but *everyone* is not, *i.e.* (236):

(236) the (Max (λd[($\exists$x, x is one of us) x is d-fat])) λd2[($\exists$d, d > 0) λd_1 [Ede is d_1+d_2-fat]]
∴ Everyone λx[the (Max(λd[x is one of us & x is d-fat])) λd_2 [($\exists$d, d > 0) λd_1 [Ede is d_1+d_2-fat]]

Recall that the presence of the *Max*-operator is essential for the soundness of the last two arguments. Since the *Max*-operator is present in any 'comparative' complement, i.e. also in complements of equatives, we predict exactly the same effects for equatives. This seems correct. We know that we find negative polarity items in *as*-clauses. Moreover, the quantifiers behave similarly, as the validity of the following arguments shows:

(237) Ede is as fat as {Randi or Otto / anyone of us}
∴ Ede is as fat as {Randi / everyone of us}

Let us see next how my analysis blocks unwarranted inferences like the following:

(UI) *Blocking unwarranted inferences*
Ede is fatter than Otto
∴ Ede is fatter than everyone

There are two formalizations of this argument, *viz.* (238) and (239).

(238) the max. d[Otto is d-fat]λd_2[$\exists$d,d > 0) λd_1[Ede is d_1+d_2-fat]]
∴ the max.d[everyone is d-fat]λd_2[($\exists$d,d > 0)λd1[Ede is d_1+d_2-fat]]

(239) the max.d[Otto is d-fat]λd_2[$\exists$d,d > 0)λd_1[Ede is d_1+d_2-fat]]
∴ Everyone λx[the max.d[x is d-fat]λd_2[($\exists$d,d >0) λd_1[Ede is d_1+d_2-fat]]]

Clearly, (239) is invalid. (238) could be true if *everyone* did mean "every one except Ede" and everyone except Ede would be fat to the same degree. If Ede is fatter than Otto he must then be fatter than everyone except himself. I think these predictions are correct.

We now explain the oddness of comparative constructions with embedded negative quantifiers:

(NQ) Ede is $\begin{Bmatrix}\text{more}\\ \text{as}\end{Bmatrix}$ intelligent $\begin{Bmatrix}\text{than}\\ \text{as}\end{Bmatrix}$ no one of us

There are two formalizations of this. (240) does not express a proposition since the description does not denote. (241), however, says that (NQ) expresses contingent propositions. I have no idea how this reading is to be barred. Yet, it should be clear from the discussion in section VII that any theory has a difficulty here.

(240) the max.d [No one of us is d-intelligent] λd_2 [$\begin{Bmatrix}(\exists d, d > 0)\\ (\exists d, d = 1)\end{Bmatrix}$

λd_1 [Ede is $\begin{Bmatrix}d_1 + d_2\\ d_1 \cdot d_2\end{Bmatrix}$ -intelligent]]

(241) No one of us λx[...x...]

Here, ...x... is (240) with x instead of *no one of us.*

The next datum on our list is:

(◇) A polar bear could be bigger than a grizzly bear could be

A satisfactory analysis is:

(242) the max.d [$\Diamond\exists x$[x is a grizzly bear & x is d-big]] λd_2[($\exists d, d > 0$) λd_1 [$\Diamond\exists y$[y is a polar bear & y is $d_1 + d_2$-big]]]

So our theory is adequate also with respect to this case. Let us repeat now the examples for the *differential readings*:

(DR) a. John is at least six inches taller than Mary
b. Ede is at most twice as fat as Angelika

These are analysed as (243) and (244), respectively.

(243) the max.d [Mary is d-tall] λd_2[$\exists d, d \geqslant 6$ inches) λd_1 [Ede is $d_1 + d_2$-tall]]

(244) the max.d [Angelika is d-tall] λd_2[$\exists d, d \leqslant 2$) λd_1 [Angelika is $d_1 \cdot d_2$-fat]]

Iterated modality is the last example on our check list:

(IM) I thought Plato could have been more boring

The interesting sense, which we were discussing in section X, is represented by the following analysis:

(245) I thought (the max.d [Plato is d-boring] λd_2 [$(\exists d, d > 0) \lambda d_1$ [$\Diamond$(Plato is $d_1 + d_2$-boring)]]

Thus we can be happy in this respect as well.

If we look back at our evaluation tabel (xv) of section I, we find that our proposal gets 9 points out of 9 possible points.

On these assumptions the following statement is a truth:

(246) My proposal is almost twice as good as Russell's or Cresswell's.

University of Konstanz
Linguistics Department
P.O. Box 5560
7750 Konstanz, BRD.

NOTES

1. Cf. Russell [1905: 489]. I am indebted to Horn [1981], who on his turn, is indebted to Abott, for pointing out that Russell brought up this problem some seventy years before Postal [1974] made it fashionable in linguistics.
2. There can hardly be any doubt that (2) and (3) are the correct 'protoformalizations' of Russell's analyses, because the paper where Russell's remark occurs is in fact his theory of descriptions, *i.e.* the birth of the idea that nominals have scope. What Russell calls a "size" is a special case of what is more commonly called a "degree of comparison" or just a "degree". In this paper I have nothing to say about the ontological status of degrees. I assume that degrees of the same kind form a scale so that we can say things like $d_1 < d_2$. I also assume that in many cases degrees have a metric so that we can speak of $d_1 + d_2$. (In the case of lengths for instance it is quite obvious that we can form a length which is the sum of two given lengths). Predicates like $<$ or $+$ are of course relative to the kind of degree in question, although this relativity is not usually indicated. It is not clear when and how we can perform operations on different kinds of degrees. For instance if d_1 is a degree of length and d_2 a degree of cleverness it is not clear that there is *any* $<$ for which $d_1 < d_2$ makes sense.
3. In an earlier version of this paper, I had included an extensive discussion of Horn's proposal (Cf. Stechow [1982]). The usual restrictions in space forced me to sacrifice that section.
4. Williams' formulas are slightly more complicated. I have simplified them by performing some λ-conversions.
5. I am grossly simplifying. For the details of Cresswell's semantics for the comperative operator *er than*, see Cresswell [1976: 268].
6. This idea goes back to at least Jespersen and is made explicit in Ross [1968: 294]. Cf. Seuren, p. 530ff.
7. Cf. Lewis [1972: 64].
8. Lewis' theory is a bit nicer, because delineations are implicit parameters, like *e.g.* the world-and-time-parameter. I think, this is an attractive view.

9. At least, this is claimed in the literature. Cf. Postal, 392f. In German, where we distinguish between the indicative and the subjunctive of the past, it is not possible to have a counterfactual of this sort with an inconsistent antecedent or consequent.

(i) Wenn Maria weniger geraucht hätte (subj.) als sie geraucht hat (ind.), wäre (subj.) sie gesünder als sie ist (ind.)

Counterfactuals with the subjunctive both in the main-clause and in the *als*-clause of the comparative construction are impossible:

(ii) *Wenn Maria weniger geraucht hätte (subj.) als sie geraucht hätte (subj.) wäre (subj.) sie gesünder als sie wäre (subj.)

(ii) is felt as incomplete. We would expect further embedded *wenn*-clauses:

(iii) Wenn Maria weniger geraucht hätte, als sie geraucht *hätte, wenn* sie eine ordentliche Arbeit gehabt hätte, wäre sie ...

This cross-linguistic comparison makes it very hard for me to believe in the ambiguity of (26). For the purposes of the present discussion, however, my disbelief is immaterial. I am interested in the 'consistent' reading only. And this is the one we have in German, too.

10. Cf. Lewis [1973a: 556].
11. Postal, p. 392f. Our (48) is Postal's (76)(a) and our (49) is his (77).
12. Cf., for instance, Stalnaker [1968].
13. Cf. Klein [1980].
14. The possibility of treating these cases by means of double indexing is mentioned in Lewis [1973b].
15. Cf. Cresswell [1973], for this use of properties.
16. Cf. Seuren, p. 533 ff.
17. (72) is Seuren's (24). (71) goes back to Ross [1968, p. 294], where a similar example is discussed. Notice that one might dispute the claim that *any* is a negative polarity item in view of (71) alone. I am not going into this, here.
18. It should be obvious that I am not using the symbols CN,VP etc. in the sense of nonterminals. CN denotes an arbitrary but fixed common noun phrase. Similarly for the other symbols.
19. It will be seen later that this claim is too strong. *Aer* is only downward entailing for some $\bar{S}$-arguments, as the next section will show. This suggests that we have to weaken Ladusaw's criterion.
20. In this connection it should be mentioned that it is very doubtful whether the ι-operator should be regarded as genuitively downward-entailing with respect to its domain.

If we assume that the definite article is represented by the ι-operator, then the definite article should be downward-entailing with respect to the CN-position. This, however, is not the case, intuitively. Thus, the definite article should not be downward-entailing with respect to the CN-positions. But the CN-position of the definite term corresponds exactly to the $\bar{S}$ in Postal's analysis. *i.e.* we have the correspondence $\iota x\mathrm{CN} \approx \iota x\bar{S}$. But then the ι-operator should not be downward-entailing with respect to the $\bar{S}$, This would rule out Russell's and Postal's theory on independent grounds.
21. *Vide, e.g.*, Cresswell [1976].
22. In the original account of Cresswell, which I have been simplifying throughout,

(99) doesn't express a proposition at all. That's the reason why I gave Cresswell a "+" in the evaluation in I.2.

23. Similarly, (i) doesn't have the reading (ii):

(i) *John is taller than Bill is not

(ii) John is not taller than Bill is

Clearly, these are examples which are related to our discussion. (I have raised this point with Cresswell, but he claims that the wide scope readings are possible. Since my own theory doesn't block these readings either it would be good if he were right).

24. A sentence of this kind was brought to my attention by David Lewis, though in a rather different connection.

25. Richard Sharvey once gave a talk in Wellington, where he built in a totality-condition for definite articles modifying mass nouns. I am indebted to him for this idea.

26. *I.e.*, a property that applies to sets of individuals. Vide Stechow [1980].

27. This example is taken from Chomsky [1981: 83]. Another nice example is: "Never in the history of human conflict has so much been owed by so many to so few." (Churchill).

28. So, for example, B. Partee, if I remember well a discussion we had in the Trianon at Nijmegen.

29. What worries me more with Klein's approach is that he thinks that the notion of the comparative is essentially derived from the fact that the predicates in natural language are vague. This idea is inspired by Kamp [1973]. I believe that this idea is misguided. We have comparative notions in precise languages. Furthermore, intuitively, comparative notions like *taller* are always applicable without any problem whereas this is not so for *tall* and *short.* So I see no reason to explain the clearer relation by two more obscure concepts.

30. The existence of this reading can be made plausible by the following story: "Before I had read Plato I thought him as boring as Spinoza (whom I find very boring). But when I was reading him I found him only as boring as Heidegger (whom I find less boring, than Spinoza). At this time, Plato had in every world w of my thoughts the same degree of boringness as Heidegger had in w, but for every such world w there was an accessible world w′, where Plato was as boring as Spinoza was in w. Since for any world of my thoughts w, Spinozas boredom in w exceeds Heidegger boredom in w, we can describe this situation as (179), whose linguistic expression is (178). It is pretty obvious that Plato's "objective" boringness (if there is such a thing), *i.e.* his boringness in the actual world, plays no role in this comparison. Perhaps the example is not well chosen. But the structure of the argument should be clear.

31. I am assuming a cross-category generalization of *and* in the sense of v. Stechow [1974] or Partee-Rooth [1983].

32. The structure of the purposive in (228) is the same as Chomsky [1981, p. 205] assumes.

REFERENCES

Aristoteles, *Aristotelis Categoriae et Liber de Interpretatione*. Ed. L. Minio-Paluello. Oxford University Press, Oxford, 1963.

Bartsch, R. & Vennemann, Th., 1972: *Semantic Structures*. Athenäum, Frankfurt/M.

Bresnan, J., 1975: Comparative deletion and constraints on transformations. *Linguistic Analysis* 1,1: 25-74.

Chomsky, N., 1977: On wh-movement. In: P. Culicover, T. Wasow & A. Akmajian (eds.), *Formal Syntax*. Academic Press, New York.

Chomsky, N., 1981: *Lectures on Government and Binding*. Foris, Dordrecht.

Cresswell, M.J., 1973: *Logics and Languages*. Methuen, London.

Cresswell, M.J., 1976: The semantics of degree. In: B. Partee (ed.), *Montague Grammar*. Academic Press, New York. pp. 261-292.

Gazdar, G., 1979: A Phrase Structure syntax for comparative clauses. In: T. Hoekstra, H. v.d. Hulst & M. Moortgat (eds.), *Lexical Grammar*. Foris, Dordrecht.

Hamann, C., Nerbonne J. & Pietsch, R., 1980: On the semantics of comparison. *Linguistische Berichte* 67: 1-22.

Hellan, L., 1981: *Towards an Integrated Analysis of Comparatives*. Narr. Tübingen.

Hoepelman, J., 1982: Adjectives and nouns: a new calculus. In: R. Bäuerle *et al.* (eds.) *Meaning, Use and Interpretation of Language*. De Gruyter, Berlin. pp. 190-220.

Horn, L., 1981: A pragmatic approach to certain ambiguities. *Linguistics and Philosophy* 4: 321-358.

Kaiser, G., 1979: Hoch und gut - Überlegungen zur Semantik polarer Adjektive. *Linguistische Berichte* 59: 1-26.

Kamp, H., 1978: Two theories of adjectives. In: E. Keenan (ed.) *Formal Semantics of Natural Language*. Cambridge University Press, Cambridge. pp. 123-155.

Klein, E., 1980: A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4: 1-45.

Kratzer, A., 1978: *Die Semantik der Rede, Kontexttheorie, Modale, Konditionale*. Scriptor, Kronberg/Ts.

Ladusaw, W.A., 1979: *Polarity Sensitivity as Inherent Scope Relations*. PhD diss. Austin, Texas.

Lewis, D.K., 1970: General Semantics. *Synthese* 22: 18-67.

Lewis, D.K., 1973a: Causation. *Journal of Philosophy* 70: 556-567.

Lewis, D.K., 1973b: Counterfactuals and comparative possibility. *Journal of Philosophical Logic* 2: 436-437.

Montague, R., 1974a: *Formal Philosophy*. R.H. Thomason (ed.). Yale University Press, New Haven - London.

Montague, R., 1974b: The proper treatment of quantification in ordinary English (= PTQ). In: Montague 1974a, pp. 247-270.

Montague, R., 1974c: English as a formal language (= EFL). In: Montague 1974a, pp. 188-221.

Partee, B. & Rooth, M., 1982: Generalized conjunction and type ambiguity. In: R. Bäuerle, Ch. Schwarze & A. von Stechow (eds.), *Meaning, Use, and Interpretation of Language*. De Gruyter, Berlin.

Postal, P.M., 1974: On certain ambiguities. *Linguistic Inquiry* 5: 367-425.

Ross, J.R., 1968: A proposed rule of tree pruning. In: D.A. Reibel & S. Schane (eds.), *Modern Studies in English*. Prentice Hall, Englewood Cliffs. pp. 288-299.

Russell, B., 1905: On denoting. *Mind* 14: 479-493.

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

Stalnaker, R., 1968: A theory of conditionals. In: N. Rescher (ed.), *Studies in Logical Theory.* Blackwell, Oxford.

Von Stechow, A., 1974: ε-λ-kontextfreie Sprachen. Ein Beitrag zu einer natürlichen formalen Semantik. *Linguistische Berichte* 34: 1-33.

Von Stechow, A., 1980: Modification of Noun Phrases. A challenge for compositonal semantics. *Theoretical Linguistics* 7,1/2: 57-109.

Von Stechow, A., 1981: Topic, focus, and local relevance. In: W. Klein & W. Levelt (eds.), *Crossing the Boundaries in Linguistics.* Reidel, Dordrecht. pp. 95-130.

Wheeler, S.C., 1972: Attributives and their modifiers. *Nous* 6: 310-334.

Williams, E., 1977: Discourse and logical form. *Linguistic Inquiry* 8, 1: 101-139.

Wunderlich, D., 1973: Vergleichsätze. In: F. Kiefer & N. Ruwet (eds.), *Generative Grammar in Europe.* Reidel, Dordrecht. pp. 629-672.