

*

THE GRAMMAR OF RELATIVE ADJECTIVES AND COMPARISON

Renate Bartsch
FU Berlin and University of California, Los Angeles
Theo Vennemann genannt Nierfeld
University of California, Los Angeles

1. Earlier Approaches to the Problem of Relative Adjectives and Comparison.

Speakers of English know that the following sentences have something in common.

- (0) John is 5 feet tall.
- (1) John is tall.
- (1') Mary is tall.
- (2) John is taller than Mary.
- (3) John is as tall as Mary.
- (4) John is the tallest of Peter's sons.
- (5) John is short.
- (6) John is shorter than Mary.
- (7) John is as short as Mary.
- (8) John is the shortest of Peter's sons.

Contemporary syntacticians have tried to account for this knowledge. Most of them have assumed that sentences (1) and (1') are somehow involved in the derivation of sentences (2) - (4), viz. as part of their deep structures from which their more complex surface structures are derived by means of syntactic transformations. This procedure did not, of course, arise accidentally. It is suggested by the relative complexity of the surface structures of these sentences, and is thus a direct result of the preoccupation of contemporary syntacticians with surface-syntactic properties of languages.

The failure of most contemporary syntacticians to analyze and formulate the properties of relative adjectives and comparison properly is, of course, by no means novel. On the contrary, they perpetuate (or renew) a venerable tradition which started two and a half thousand years ago with Plato's *Theaetetus*, and may thus consider themselves in excellent company. Bertrand Russell (1945, p. 159) writes,

* This paper appears in Linguistische Berichte, 21. It is a short version of an article, "Relative Adjectives and Comparison", UCLA Papers in Syntax, 2, edited by Paul Schachter and George Bedell. The theory of grammar on which this article is based, as well as applications of this theory to other problems of grammatical analysis, is presented in our book, Semantic Structures: A Study in the Relation Between Semantics and Syntax, Frankfurt: Athenäum-Verlag 1972.

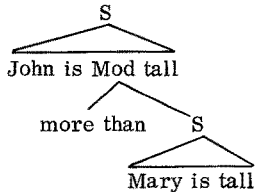
"There are, at this point, some puzzles of a very elementary character. We are told that since 6 is greater than 4 but less than 12, 6 is both great and small, which is a contradiction. Again, Socrates is now taller than Theaetetus, who is a youth not yet full grown; but in a few years Socrates will be shorter than Theaetetus. Therefore Socrates is both tall and short. The idea of a relational proposition seems to have puzzled Plato, as it did most of the great philosophers down to Hegel (inclusive)."

Logicians and semanticists emphasize that (2), (3), and (4) do not imply (1) and (1').

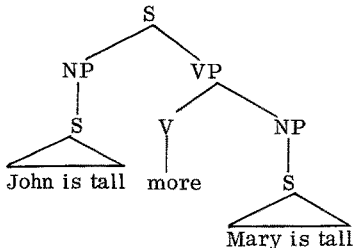
John may be a very short boy, which means that (1) is false; yet (2), (3), and (4) may be true. Nevertheless, even though there exists no implicational relation between sentences (2) - (4) on the one hand and (1) and (1') on the other, there clearly exists a semantic relationship which a theory of grammar must account for.

Transformational grammarians have proposed deep structures of the following kind in which (1) and/or (1') appear as constituents of (2)¹.

(A)



(B)



The syntactic properties in these approaches run counter to the semantics of sentences like (2): they present deep structures of comparative sentences which contain deep structures of the corresponding sentences with positives as constituents. This is incorrect, because the meanings of sentences with positives are not conceptual constituents of the meanings of the corresponding comparative sentences: one does not have to interpret the positive sentences in order to interpret the comparative sentences. In (B), the constituent $[[\text{John is tall}]_{\text{S}}]_{\text{NP}}$ suggests that terms are the arguments of the relation more. But as it stands we do not know what kind of nominalization of S is intended. It is even wrong because it suggests that the

¹ For details and references, cf. the complete article (Bartsch and Vennemann, 1972).

nominalization of the positive sentences John is tall and Mary is tall is intended. This would be something like John's being tall and Mary's being tall, or John's tallness and Mary's tallness where tallness is the property of being tall, even though sentence (2) does not imply that either John or Mary have the property of being tall. Note further that with analyses of type (A) and (B), sentences like Bill believes he is taller than he is are not representable, as observed by Ross and Perlmutter (1970).

Several semantic approaches to the problem of relative adjectives and comparison have been proposed. It would be too time-consuming to demonstrate in this lecture that those made by Reichenbach (1947), Seuren (1970), and Wierzbicka (1972) are inadequate. The only authors who have shown an understanding of the problem are Sapir (1944), Fillmore (1965), Wunderlich (1970), Bierwisch (1971) and Dik (1971). Of these, Sapir-Fillmore on the one hand, and Wunderlich on the other, represent contributions that are similar to our proposal. Sapir-Fillmore suggest that the comparative is not based on the positive semantically, despite the morphology. And Wunderlich introduces the concept of measure functions related to relative adjectives. Both proposals have several inadequacies.¹ Here, rather than pointing these out, we prefer to present our own analysis of relative adjectives and comparisons.

2. Semantic Representations of (0) - (8).

In this section, we use f_T^M as an abbreviation for the measure function as applied to the dimension Height (Tallness). We will explain this abbreviation in Section 3. $N_{T, Y}$ represents the average of the heights of the objects in the reference set Y within which x is compared. The angular brackets are used to indicate presuppositions.

$$(0a) \quad f_T^M(x) = 5 \text{ feet}$$

$$(1a) \quad f_T^M(x) > N_{T, Y}$$

$$(2a) \quad f_T^M(x) > f_T^M(y)$$

$$(3a) \quad f_T^M(x) = f_T^M(y) \quad (\text{and } \langle f_T^M(y) > N_{T, Y} \rangle)$$

$$(4a) \quad \text{John} = (1x) \left[\text{Son}(x, \text{Pe}) \ \& \ (y) (\text{Son}(y, \text{Pe}) \ \& \ x \neq y \supset f_T^M(x) > f_T^M(y)) \right]$$

$$(5a) \quad f_T^M(x) < N_{T, Y}$$

$$(6a) \quad f_T^M(x) < f_T^M(y) \quad (\text{and } \langle f_T^M(y) < N_{T, Y} \rangle)$$

¹ The semantic approaches are analyzed in the complete article (Bartsch and Vennemann, 1972).

² Sentence (3) has a second interpretation with \geq rather than =. Similarly, (7) with \leq for =.

$$(7a) \quad f_T^M(x) = f_T^M(y) \text{ and } \langle f_T^M(y) < N_{T, Y} \rangle$$

$$(8a) \quad \text{John} = (1x) [\text{Son}(x, \text{Pe}) \ \& \ (y) (\text{Son}(y, \text{Pe}) \ \& \ x \neq y \supset f_T^M(x) < f_T^M(y))]]$$

3. The General Measure Function f^M .

The general measure function f^M is a 2-place operator where one argument is a dimension D , and the other, an object measured in this dimension, $f^M(x, D)$. We use two abbreviations:

$$\begin{aligned} f_D^M(x) &= \text{def } f^M(x, D) && \text{if } D \text{ is constant;} \\ f_x^M(D) &= \text{def } f^M(x, D) && \text{if } x \text{ is constant.} \end{aligned}$$

For example, D is constant in John is taller than Mary, and x is constant in This house is broader than high. The general measure function f^M maps pairs of objects x and dimensions D on equivalence classes of objects relative to D . The set of equivalence classes relative to D is linearly ordered.

In Section 2, D was exemplified with T (Tallness). Other dimensions for which both English and German have relative adjectives are: Breadth, Temperature, Speed; Intelligence, Beauty, Liveliness. Note that a dimension may or may not have a quantified scale associated with it. Of the above, the former three do, the latter do not (but attempts have been made to quantify intelligence). Many, more, most, and their antonyms few, fewer, fewest, mistaken by some linguists as quantifiers like all and some, are also relative adjectives, special in that their dimension ($D = P$, for Power, the number of elements in a set) applies to sets rather than individual objects. Much (more, most) and its antonym little (less, least) are relative adjectives referring to the dimension Quantity applicable to mass nouns.

4. On the Syntax of Relative Adjectives.

We now turn to the syntactic derivation of comparative sentences. The main types of rules we use are lexicalization rules, constituent structure rules, and serialization rules. By lexicalization we mean the choice of lexical items appropriate to convert semantic representations into phonetically interpretable structures. There are usually several ways to verbalize a given semantic representation. For example, the semantic representation

$$(2a) \quad f_T^M(x_J) > f_T^M(y)$$

may be expressed as in (2⁰) and (2).

$$(2^0) \quad \text{"Direct" verbalization:}$$

The height of John is greater than the height of Mary.

(2) "Better" verbalization:

John is taller than Mary.

Of these sentences, (2⁰) seems to verbalize (2a) more directly. As a consequence it is longer than (2). The difference between a "direct" verbalization and a "better" one is even more obvious in cases like

$$(1a) \quad f_T^M(x_J) > N_{T, Y}$$

(1⁰) John's height is greater than the average of the heights of boys.

and

(1) John is tall. [(1) used in a context which makes clear that John is compared to other boys.]

The function of syntax is not simply to verbalize semantic representation but to do it "well", i. e., to provide shortcuts which allow the speaker to express complex meanings in short and simple surficial structures. Thus, (2) is better than (2⁰) and (1) is better than (1⁰), where "better" here stands for "expressing the same semantic representation with fewer words and a simpler surface phrase structure".

From this point of view, the "invention" of the relative adjective (just like that of the relative verb) was a stroke of genius of the human linguistic mind. A dimension is inherently a nominal concept, and so is the value associated in it with an individual by the measure function f^M . The relations $>$, $<$, $=$ are the verbal elements in comparison sentences. The compared values are the arguments of the relations. The names of the objects compared are hidden in the argument of the measure function. Relative adjectives permit the speaker to present the second-degree arguments of the semantic representations as first-degree surface arguments. Consider the semantic and surface representations of the sentences (2a), (2), (1a), and (1): John is a second-degree argument in (2a) and (1a), but a first-degree argument in (2) and (1) (see Figure 1). Furthermore, the deletion of an average which can be inferred from the context allows semantic representations of the kind (1a) to assume surface structures like (1) which look like the most primitive kind of predications: $f(x)$, as in This rose is red, Mary is human (Figure 2). The gain in surface shortness and simplicity is optimal, while all semantic properties of the sentence remain completely recoverable.

The shortcut aspect of relative adjectives becomes most apparent from the fact that only the most frequently referred to dimensions have a corresponding lexical relative adjective. English has lexical units for "largeness", "liveliness", and "weight", but not for "being prepared for the contest" or "being more balanced racially". (Perhaps the deverbal adjective integrated will develop into a special relative adjective for the dimension of "racial balance" of a given set of people, as this dimension becomes more important socially and is

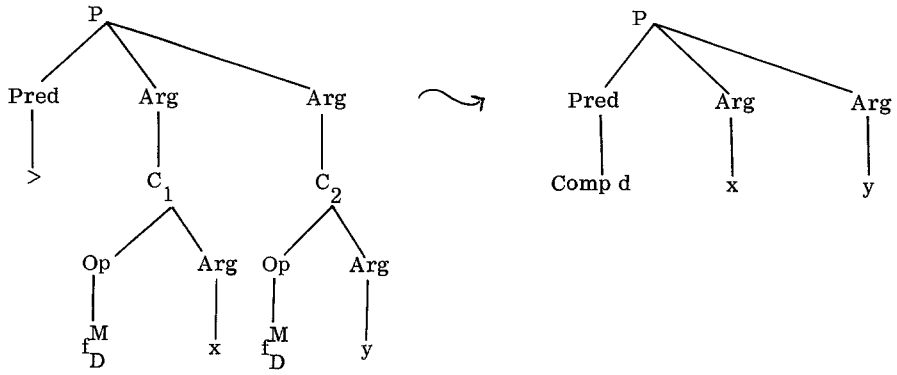


FIGURE 1

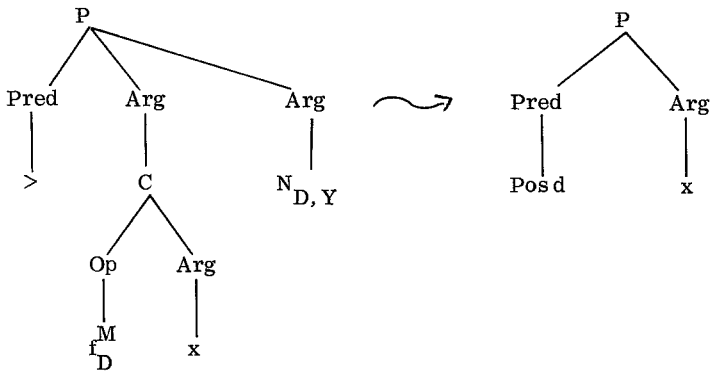


FIGURE 2

more frequently referred to.) Even certain simple geometrical dimensions do not have adjectives in English, e.g., the (maximal) diameter and circumference of a figure such as a rectangle or circle, while the length and width of appropriate figures do. There is no sentence *All circles are more circumferenced than diametered as there is the sentence All proper rectangles are longer (or more long) than wide. In such cases the nominal character of dimensions becomes apparent on the surface: The circumference is greater than the diameter for all circles, or All circles are greater in circumference than in diameter. The analogous situation seems to arise in some languages even with more common dimensions, e.g., in Igbo with Tallness: John ka Mary ogologo, "John is taller than Mary", literally "John (sur-)passes Mary (in) tallness", (examples due to Marianne Celce-Murcia, UCLA). Often a language neutralizes several semantically similar dimensions into one adjective English uses long for a spacial and a temporal dimension. German uses gross for the Tallness, Largeness, Bigness and Greatness dimensions separated in English. Certainly there are interesting universals of dimensional syncretism to be discovered here. It seems quite plausible to us that there may exist languages lacking relative adjectives altogether in which nouns (weight, duration) or verbs (weigh, last) are used instead. (English has heavy and long in addition to the nouns and verbs.) However, that a language renounces both the adjective and the verb shortcuts seems less likely. Note also that while many of the relative nouns lacking related relative adjectives are rather long, relative adjectives themselves tend to be short, with many of the most basic ones being monosyllabic in English: long, high, wide, broad, far, tall, large, big, great, smart (together with sharp and bright preferred to the unwieldy intelligent), late, good, bad, etc.

A relative adjective is represented in the lexicon with no other information than the dimension(s) to which it refers, and the information Adjective (in addition to its phonetic shape, and perhaps certain idiosyncratic properties it may have, such as suppletiveness). Let d be a basic ("unmarked") relative adjective with the phonetic shape \underline{d} and the dimension D . The lexical representation of d is $[\underline{d}, D, \text{Adjective}]$. The ("marked") antonym \tilde{d} of d is represented with a "mark", for which we may use $<$ itself: $[\tilde{d}, D, <, \text{Adjective}]$.

We will now give the typical lexicalizations for sentences containing positives, comparatives, superlatives, and as...as constructions. We use Y to designate the reference set. This reference set is needed for the superlative, where the "superlative" element has to be a member of it, and in the positive, where the average $N_{D, Y}$ is based on it; cf. the "different meanings" of big in This mouse is big and This elephant is big, which follows from the fact that $N_{\text{Big}, \{x:\text{mouse}(x)\}} \neq N_{\text{Big}, \{x:\text{elephant}(x)\}}$.

- (9) (a) $f_D^M(x) > N_{D, Y} \rightsquigarrow x(\text{Pos } d)$
 $f_D^M(x) < N_{D, Y} \rightsquigarrow x(\text{Pos } \tilde{d})$
- (b) $f_D^M(x) > f_D^M(y) \rightsquigarrow x(\text{Comp } d) y$
 $f_D^M(x) < f_D^M(y) \rightsquigarrow x(\text{Comp } \tilde{d}) y$
- (c) $(\perp x) [(y \in Y \ \& \ x \neq y) \supset f_D^M(x) > f_D^M(y) \ \& \ \langle x \in Y \rangle] \rightsquigarrow (x)(x(\text{sup } d) Y)$
 $(\perp x) [(y \in Y \ \& \ x \neq y) \supset f_D^M(x) < f_D^M(y) \ \& \ \langle x \in Y \rangle] \rightsquigarrow (x)(x(\text{sup } \tilde{d}) Y)$
- (d) as ... as constructions
 $f_D^M(x) = f_D^M(y) \rightsquigarrow x(\text{Equal } d) y$
 $f_D^M(x) = f_D^M(y) \ \& \ \langle f_D^M(y) < N_{D, Y} \rangle \rightsquigarrow x(\text{Equal } \tilde{d}) y$.

These lexicalization rules require a number of clarifying remarks. First, we note that no other aspects leading to the lexicalization of these semantic representations have been dealt with than those leading to the introduction of the relative adjectives themselves. Other rules are needed to lexicalize the constituents Pos (Positive) [usually lexicalized as zero], Comp (Comparative), Sup (Superlative), and Equal, and to introduce further morphological markers required for a given language. We will show in some detail what these rules are for English.

Second, we must point out that (9) is a simplified representation which we have introduced in the interest of expository clarity. A false impression one may get from (9) is that the constituents of the forms left and right of the arrow are necessarily arranged in a linear sequence. It is a priori the case that the logical forms to the left of the arrow are logically devoid of any spacial or temporal sequentiality. The fact that they are ordered on paper is based on the desire to save additional notational apparatus. To make explicit that the linear arrangement of logical forms is accidental (i. e., has only pragmatic but no logical reasons), we must introduce a notation which makes explicit what the relationship of each argument to its relation is. Thus, we label the two arguments occurring in the definition of $>$ in some arbitrary fashion, e. g., by means of numbers 1, 2 (but an asterisk and a heart would serve the same purpose). Let 1 mark that place in the relation $>$ which in the conventional notation $a > b$ is occupied by a ; let 2 mark the place of b . We write $\overset{1, 2}{>}$ to indicate that we use this particular notation for the relation "greater (more) than". (Note that this procedure is not circular because the relationships of the two arguments to the relation $>$ follows directly from the meaning of the relation.) Now all possible linear

arrangements of the constituents $a, >, b$ of $a > b$ are equivalent:

$$(10) \quad \left(\begin{array}{c} 1, 2 \ 1 \ 2 \\ \succ \\ a \ b \end{array} \right), \quad \left(\begin{array}{c} 1, 2 \ 2 \ 1 \\ \succ \\ b \ a \end{array} \right), \quad \left(\begin{array}{c} 1 \ 1, 2 \ 2 \\ a \succ b \end{array} \right) \\ \left(\begin{array}{c} 1 \ 2 \ 1, 2 \\ a \ b \succ \end{array} \right), \quad \left(\begin{array}{c} 2 \ 1, 2 \ 1 \\ b \succ a \end{array} \right), \quad \left(\begin{array}{c} 2 \ 1 \ 1, 2 \\ b \ a \succ \end{array} \right).$$

In fact, vertical arrangements would likewise be equivalent. We enclose constituents whose order is irrelevant in braces, with an arbitrary arrangement of the constituents

$$(11) \quad \left\{ \begin{array}{c} 1, 2 \ 1 \ 2 \\ \succ \\ a \ b \end{array} \right\}$$

Thus, we write for $f_D^M(x) > f_D^M(y)$:

$$(12) \quad \left\{ \begin{array}{c} 1 \\ \succ \\ f_D^M(x) \end{array} \right\} \left(\begin{array}{c} 2 \\ f_D^M(y) \end{array} \right) \left. \right\} .$$

More explicitly, since f_D^M and its argument are likewise unordered, we write:

$$(13) \quad \left\{ \begin{array}{c} 1, 2 \\ \succ \\ \left\{ f_D^M \ x \right\} \left\{ f_D^M \ y \right\} \end{array} \right\} .$$

Note that the operation f_D^M and its two arguments can likewise be represented as an unordered set, e.g., $\left\{ f_D^M \ x \ y \right\}$; but we use the abbreviatory notation, as less complex.

We suggest that the output of the lexicalization rules (9) is likewise unordered. Thus, we intend our rules in (9) to be interpreted on the analogy of (9b'), which is a formal representation of (9b) in the model just outlined:

$$(9b') \quad \left\{ \begin{array}{c} 1 \\ \succ \\ \left\{ f_D^M \ x \right\} \left\{ f_D^M \ y \right\} \end{array} \right\} \rightsquigarrow \left\{ \left\{ \text{Comp } d \right\} \begin{array}{c} 1 \ 2 \\ x \ y \end{array} \right\} \\ \left\{ \begin{array}{c} 1, 2 \\ \succ \\ \left\{ f_D^M \ x \right\} \left\{ f_D^M \ y \right\} \end{array} \right\} \rightsquigarrow \left\{ \left\{ \text{Comp } \tilde{d} \right\} \begin{array}{c} 1 \ 2 \\ x \ y \end{array} \right\} .$$

Our basis for this suggestion is that different languages serialize (linearize) the constituents in comparison sentences differently, and that some languages serialize the same constituents differently in different contexts. Thus, English serializes $\left\{ \left\{ \text{Comp } d \right\} \begin{array}{c} 1 \ 2 \\ x \ y \end{array} \right\}$ as $\left[\left\{ \text{Comp } d \right\} \begin{array}{c} 1 \ 2 \\ x \ y \end{array} \right]$, Hindi as $\left[\begin{array}{c} 1 \ 2 \\ x \ y \end{array} \left\{ \text{Comp } d \right\} \right]$, Sanskrit in any order; German serializes $\left\{ \text{Comp } d \right\}$ as $\left[d \ \text{Comp} \right]$, Spanish as $\left[\text{Comp } d \right]$, English as $\left[d \ \text{Comp} \right]$ if Comp is lexicalized as -er, otherwise as $\left[\text{Comp } d \right]$. On the other hand, serialization depends on lexicalization: Whether a relational concept is lexicalized as a relative adjective, or a verb, or a noun, can have incisive consequences for the serial arrangement of the whole sentence.

We will now show with a sample derivation how rules (9') [= (9) reinterpreted along the lines of (9b')] work, and what other lexicalization and serialization rules may be needed in a given language. We use an English sentence, (2) John is taller than Mary. Its semantic representation is (14a), where x_J and x_M are indices for the two individuals which are lexicalized as John and Mary, respectively (rather than as, say, Smith's son and Miller's daughter). This lexicalization, with all the syntactic and semantic properties of the lexical items involved, has taken place in (14b). One must recall the notational difference between d and \underline{d} in order to interpret lexical representations properly:

$$(14) \quad (a) \quad \left\{ \begin{array}{c} 1 \\ \supseteq \\ \left\{ \begin{array}{c} f_T^M \\ x_J \end{array} \right\} \end{array} \right\} \left\{ \begin{array}{c} 2 \\ \supseteq \\ \left\{ \begin{array}{c} f_T^M \\ x_M \end{array} \right\} \end{array} \right\}$$

$$(b) \quad \left\{ \begin{array}{c} 1 \\ \supseteq \\ \left\{ \begin{array}{c} f_T^M \\ \text{John} \end{array} \right\} \end{array} \right\} \left\{ \begin{array}{c} 2 \\ \supseteq \\ \left\{ \begin{array}{c} f_T^M \\ \text{Mary} \end{array} \right\} \end{array} \right\}$$

Application of (9b') yields (15):

$$(15) \quad \left\{ \left\{ \text{Comp tall} \right\} \begin{array}{c} 1 \\ \text{John} \end{array} \begin{array}{c} 2 \\ \text{Mary} \end{array} \right\}$$

Lexicalization of Comp follows the well-known rule according to which a certain class of common monosyllabic and disyllabic words (with certain phonological conditions imposed on disyllables) take -er, while all others take more: taller, redder (or more red), happier, commoner (or more common) vs. more serene, more beautiful, more jaune, more dead. Let us designate the class of adjectives taking -er be ER; i.e., let us consider all such adjectives as marked ER in the lexicon. The rule can then be formulated as (16), where both -er and more are, of course, marked lexically as containing the feature Comp.

$$(16) \quad (a) \quad \text{Comp} \rightsquigarrow \text{-er/ER}$$

$$(b) \quad \text{Comp} \rightsquigarrow \text{more} .$$

Equivalently, we can write (17):

$$(17) \quad \text{Comp} \rightsquigarrow \left\{ \begin{array}{c} \text{-er/ER} \\ \text{more} \end{array} \right\} ,$$

but we do not intend any order of application to be involved here. The theory of grammar provides conventions which prohibit the general case from applying if a special mark requires the application of a restricted rule.

The need for at least two more lexicalization rules of a morphological nature is obvious. The first introduces than as part of the second argument of a comparative:

$$(18) \quad \left\{ \left\{ \begin{matrix} 2 \\ \text{Comp X} \end{matrix} \right\} \text{Y} \right\} \rightsquigarrow \left\{ \left\{ \text{than} \begin{matrix} 2 \\ \end{matrix} \right\} \left\{ \text{Comp X} \right\} \text{Y} \right\}$$

Note that the information Comp is still present even after the lexicalization of the constituent Comp, because it has to be part of the lexical representation of both -er and more in order for rules (16a, b) to find them in the lexicon as suitable lexicalizations of Comp. Note further that than is introduced in such a way that it builds a new constituent together with the old constituent 2.

The second morphological rule is that which introduces the copula in verbless subject-predicate sentences. We formulate only that part of the rule which applies to adjectives as predicates, because only that part is needed in our fragment. We do not formalize certain restrictions related to the modality of the sentence, or to such information as "predicative use of the adjective". Note that different from than, be is not adjoined to any one of the constituents of the sentence. The reason for this is that we want this rule to be neutral to the modality of the sentence, e.g., to the information "declarative", "interrogative", "imperative":

$$(19) \quad \emptyset \rightsquigarrow \text{be} / \left\{ \text{---} \begin{matrix} 1 \\ \end{matrix} \right\} \left\{ \text{Adjective X} \right\} \text{Y} \left. \right\} .$$

The remaining rules are familiar enough. The verb of an English sentence takes the number and person of the first argument:

$$(20) \quad (a) \quad \emptyset \longrightarrow \left[\alpha \text{ Number} \right] / \left\{ \left[\frac{\text{---}}{\text{Verb}} \right] \left\{ \alpha \text{ Number} \right\} \right\} \\ (b) \quad \emptyset \longrightarrow \left[\beta \text{ Person} \right] / \left\{ \left[\frac{\text{---}}{\text{Verb}} \right] \left\{ \beta \text{ Person} \right\} \right\} .$$

These are feature spreading rules, or agreement transformations. The verb (be is characterized in the lexicon as a verb) likewise adopts properties from the modality of the sentence. We have not given any part of the modality here. In a complete grammar, the situation must be presented in such a way that the verb can adopt the features equivalent to Present and Indicative. The lexicalization rules for the verbal categories can now operate. If the verb is regular, they add -s, given the verbal categories of our sentences. Since be is a special case, no -s is added. Instead, the lexicalization rule for verbal categories refers us back to the lexical item be, where a paradigm search finds is as a match, whereupon be is replaced with is.

The next rules are examples of constituent structure rules:

- (21) (a) $\{\{-er\ Adjective\}\{than\ \{^2\}\}X\}$
 $\longrightarrow \{\{\{-er\ Adjective\}\{than\ \{^2\}\}\} X\}$
- (b) $\{\{be\ Adjective\}\ \{^1\}\} \longrightarrow \{\{be\ \{Adjective\}\}\ \{^1\}\}$

Rule (21b) must be restricted to declarative and imperative sentences. Note that it orders itself intrinsically after (21a), and that it is formulated in a sufficiently general way to accommodate positives, adverbially expanded adjectives, and other kinds of adjective constructions as well, because they are all categorically relabeled as Adjectives. In other words, in a more comprehensive grammar of this kind, we would have to systematically label constituent braces with category labels. For example, rule (21b) presupposes that the total expression $\{\{-er\ Adjective\}\{than\ \{^2\}\}\}$ is labeled as an Adjective.

Now we must sketch the serialization process leading from the still unordered though lexicalized and constituent-braced string to the surface representation (2). Some of the processes outlined here are of greater generality than comparison sentences require. We omit the information which is drawn from the modality to determine whether or not a rule applies. We use brackets [] to indicate ordered strings. Note, however, that an ordered string X can be referred to as either [X] or {X}. The notation {X} simply abstracts away from the order. Thus, the serial order imposed by the following rules is entirely determined by their structural descriptions and the inputs they operate on, but independent of their order of application.

$$(22) \quad \{-er\ Adjective\} \Longrightarrow [Adjective\ -er]$$

$$(23) \quad \{than\ \{^2\}\} \Longrightarrow [than\ \{^2\}]$$

$$(24) \quad \{\{-er\ Adjective\}\{than\ \{^2\}\}\} \\ \Longrightarrow [\{-er\ Adjective\}\{than\ \{^2\}\}]$$

$$(25) \quad (a) \quad \{be\ \{Adjective\}\} \Longrightarrow [be\ \{Adjective\}]$$

$$(b) \quad \{\{^1\}\ \{be\ \{Adjective\}\}\} \Longrightarrow [\{^1\}\ \{be\ \{Adjective\}\}]$$

This group of rules formulates special applications of the general English word order principles that in declarative clauses the verb precedes its complement, and the subject its verb.

After all rules have been applied, we need to apply a convention which erases all non-phonetic information from the string. We indicate this by replacing constituent symbols

by their underlined forms. (Recall the relation between d and \underline{d} defined above.)¹

We summarize the derivation of sentence (2) in Table 1. It must be kept in mind that the particular order showing in this derivation is in part arbitrary. Rules of grammar are not themselves ordered. They apply whenever their structural description is met. This implies that a rule cannot apply as long as its structural description is not met, which may impose a certain amount of intrinsic order on the application of the rules. This order of application need not, however, be stated in the grammar, because it is completely a consequence of the structure of the rules and their input.

Finally, we sketch the rules needed for the other comparison constituents resulting from the application of (9). (9a) requires no further rules. Since Pos is not lexicalized by any rule of English, it will finally disappear from the string through the application of the non-phonetic feature-deletion convention.

The further derivation of superlative terms, cf. (9c), is difficult to describe in a general way, because so much depends on the way in which Y is defined and lexicalized: John is a) the tallest of Peter's sons, b) the tallest son of Peter's. Assume that Y is defined as $\{y: \text{Son}(y, \text{Peter})\}$. Now $(x)(x(\text{Sup } d)\{y: \text{Son}(y, \text{Peter})\})$ can be lexicalized as $\{\text{the}(\text{Sup } d)\{y: \text{Son}(y, \text{Peter})\}\}$. The set can then be lexicalized either as a genitive of (Sup d), or as a head noun with genitive for (Sup d). The former yields a) the tallest of Peter's sons (equivalently: the tallest of the sons of Peter), the latter, b) the tallest son of Peter. (Sup d), or more properly, $\{\text{Sup } d\}$, is lexicalized and serialized by rules which are analogous to (16) and (22). In order to get from here to sentence (4), John is the tallest of Peter's sons, we have to introduce the semantic representation of the superlative term the tallest of Peter's sons, cf. (9c), into an equality relation $\text{John} = \boxed{\quad}$ (or $\{\overset{1}{\underline{1, 2}} \text{John } \overset{2}{\boxed{\quad}}\}$), which is lexicalized and serialized as John is $\boxed{\quad}$.

The as ... as constructions of (9d) present no new problems. One way of lexicalizing the right-hand side of (9d) is by rule (26):

$$(26) \quad \left\{ \left\{ \text{Equal } d \right\} \overset{1}{x} \overset{2}{y} \right\} \rightsquigarrow \left\{ \left\{ \text{as } d \right\} \overset{1}{x} \left\{ \text{as } \left\{ \overset{2}{y} \right\} \right\} \right\} .$$

The serialization rule for as is straightforward. All other rules needed are available from our study of the comparative sentence (2).

¹ Note that the erasure of all non-phonetic information at the end of a derivation may well be an oversimplification. For example, it may well be the case that the argument markers 1 and 2 are superfluous after the application of the serialization rules (24) and (25) in English, while in a language relying on morphology rather than position (e.g. Sanskrit) the argument markers become superfluous after the application of the case marking rules (which are lexicalization rules in our system).

TABLE I

$\left\{ \begin{matrix} 1 & 2 \\ \geq \{ f_T^M x_J \} \{ f_T^M x_M \} \end{matrix} \right\}$	(13)
$\left\{ \begin{matrix} 1 & 2 \\ \geq \{ f_T^M \text{John} \} \{ f_T^M \text{Mary} \} \end{matrix} \right\}$	cf. (14)
$\left\{ \left\{ \text{Comp tall} \right\} \begin{matrix} 1 & 2 \\ \text{John} & \text{Mary} \end{matrix} \right\}$	(9b'), cf. (15)
$\left\{ \left\{ -\text{er tall} \right\} \begin{matrix} 1 & 2 \\ \text{John} & \text{Mary} \end{matrix} \right\}$	(16a)
$\left\{ \left\{ -\text{er tall} \right\} \begin{matrix} 1 \\ \text{John} \end{matrix} \left\{ \text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right\} \right\}$	(18)
$\left\{ \left\{ -\text{er tall} \right\} \text{be} \left\{ \begin{matrix} 1 \\ \text{John} \end{matrix} \right\} \left\{ \text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right\} \right\}$	(19)
$\left\{ \left\{ -\text{er tall} \right\} \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \begin{matrix} 1 \\ \text{John} \end{matrix} \left\{ \text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right\} \right\}$	(20)
$\left\{ \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \begin{matrix} 1 \\ \text{John} \end{matrix} \left\{ \left\{ -\text{er tall} \right\} \left\{ \text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right\} \right\} \right\}$	(21a)
$\left\{ \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \left\{ \left\{ -\text{er tall} \right\} \left\{ \text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right\} \right\} \right\} \begin{matrix} 1 \\ \text{John} \end{matrix} \right\}$	(21b)
$\left\{ \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \left\{ \left[\text{tall} \text{ -er} \right] \left\{ \text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right\} \right\} \right\} \begin{matrix} 1 \\ \text{John} \end{matrix} \right\}$	(22)
$\left\{ \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \left\{ \left[\text{tall} \text{ -er} \right] \left[\text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right] \right\} \right\} \begin{matrix} 1 \\ \text{John} \end{matrix} \right\}$	(23)
$\left\{ \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \left\{ \left[\left[\text{tall} \text{ -er} \right] \left[\text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right] \right] \right\} \right\} \begin{matrix} 1 \\ \text{John} \end{matrix} \right\}$	(24)
$\left\{ \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \left\{ \left[\left[\text{tall} \text{ -er} \right] \left[\text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right] \right] \right] \right\} \begin{matrix} 1 \\ \text{John} \end{matrix} \right\}$	(25a)
$\left[\begin{matrix} 1 \\ \text{John} \end{matrix} \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \end{matrix} \right] \left[\left[\text{tall} \text{ -er} \right] \left[\text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right] \right] \right]$	(25b)
$\left[\begin{matrix} 1 \\ \text{John} \end{matrix} \left[\begin{matrix} \text{be} \\ 3\text{rd Person} \\ \text{Singular} \\ \text{Present} \\ \text{Indicative} \end{matrix} \right] \left[\left[\text{tall} \text{ -er} \right] \left[\text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right] \right] \right]$	(unnumbered rule)
$\left[\begin{matrix} 1 \\ \text{John} \end{matrix} \left[\text{is} \left[\left[\text{tall} \text{ -er} \right] \left[\text{than} \left\{ \begin{matrix} 2 \\ \text{Mary} \end{matrix} \right\} \right] \right] \right] \right]$	(unnumbered rule)

John is taller than Mary

Although we have not illustrated this, it is obvious that our rules work in both directions, from semantic structure to surficial structure and conversely. Our fragment of a grammar is thus neutral as to production or interpretation.

This discussion of lexicalization and serialization has been somewhat technical and tedious. We feel that its explicitness may be useful, because it represents the first fairly complete description of a derivation of certain language-specific surface representations from language-independent (universal) semantic structures in a grammar without ordered deep structures, without movement transformations, and without extrinsic rule order (natural generative grammar).

5. Further Evidence and Special Applications.

Within our framework, we can account for a number of phenomena relating to relative adjectives and comparison which have puzzled earlier analysts. For reasons of time limitations, we can here only sketch our solutions.

5.1 Sentences containing positives typically compare a measurement value of an individual to an average or norm presupposed as part of the cultural background of speakers of the language. Sentences containing comparatives typically compare a measurement value of one individual to that of another which is not part of the listener's cultural background, but has to be generated in his mind for the specific purpose of the comparison. We therefore consider the comparative as semantically "marked", the positive as semantically "unmarked". It is this semantic asymmetry which is reflected in the morphological "marked" : "unmarked" contrast of the comparative vs. the positive, such as in English, tall + er : tall + \emptyset .

5.2 Just as two individuals can be compared in one dimension, two dimensions can be compared for one individual.

(27) This house is wider (or: more wide) than high.

$$(27a) \quad f_x^M(D_1) > f_x^M(D_2) \quad .$$

5.3 A measurement value $f^M(x, D)$ can be quantified in two different scales, with units $1M_1$ and $1M_2$. We express this by representing the measurement value as a function over scalar units.

(28) Mary is heavier in American pounds than in Continental pounds.

$$(28a) \quad (f^M(x, D))(1M_1) > (f^M(x, D))(1M_2) \quad .$$

5.4 One and the same object may have different measurement values in the same dimension at different times.

(29) The U. S. was larger in 1900 than in 1800.

$$(29a) \quad (f^M(x, D))(t_1) > (f^M(x, D))(t_2) \quad .$$

5.5 Just as in the case of time, a measurement value can also be a function over places.

(30) Explorer 5 is lighter on the moon than on earth.

(30a) $(f^M(x, D))(p_1) < (f^M(x, D))(p_2)$.

5.6 A measurement value may further vary according to different propositional attitudes of different individuals. We represent this with the concept of "possible worlds" developed in model theory (cf. Hintikka, 1969).

(31) John believes he is taller than he is.

(31a) $(f^M(x, D))(i_{x, \text{believe}}) > (f^M(x, D))(i_0)$.

5.7 Again, the measurement value of an individual in a dimension can be understood as a function over circumstances under which the Measurement is taken.¹

(32) Mary is livelier with her lovers than with her parents.

(33) Mary is prettier in a nightgown than in a raincoat.

(32, 33a) $(f^M(x, D))(j_1) > (f^M(x, D))(j_2)$.

6. Conclusion.

"Comparative constructions in any language have proven themselves resistant to satisfactory analysis " (Jacobs and Rosenbaum, 1970, viii). We agree with this assessment of the research situation and conclude that linguists in the past have not fully appreciated the role of semantics in determining syntactic structures. We consider the semantic properties (in production) and the physical properties (in reception) of a sentence as "given"; the role of syntax is determined by the exigencies of semantic structure on one side and by the constraints arising from the physical nature and requisite efficiency of communication on the

¹ We would like to point out that for sentences like those in Sections 5.4 - 5.7, alternative interpretations are possible in which the indices t, p, i, j are indexical to the individuals, or in the case of positives, also to the average or norm. For example, sentence (29) could also be interpreted as in (29a').

(29a') $f^M_{\text{Large}}((x)(x = \text{U.S.})(t_1)) > f^M_{\text{Large}}((x)(x = \text{U.S.})(t_2))$.

In this case, the terms left and right of the comparative sign have the same denotations in the two interpretations. The same is not true of the following example: The president of the U.S. was older in 1965 than in 1971. In the absence of factual knowledge this sentence could be interpreted as a single individual's having grown younger in time:

$(f^M_A(x))(t_1) > (f^M_A(x))(t_2)$.

However, the intended interpretation is, of course, "Of the two different individuals who held the office in the years 1965 and 1971, respectively, the first was older in 1965 than the second was in 1971":

$(f^M_A((x)(P(x))(t_1)))(t_1) > (f^M_A((x)(P(x))(t_2)))(t_2)$

Determining the conditions under which such intensionally different interpretations are extensionally equivalent goes beyond the scope of the present paper.

other. Our analysis of comparison sentences starts with an analysis of their (language-independent) semantic properties. We explain their seemingly paradoxical overt properties by making explicit some of the applicable constraints imposed by communication.

Specifically, we propose that semantic representations are logical forms which are directly mappable from the level of logical syntax onto models of states-of-affairs. The reason for positing this kind of semantic representation is twofold: 1. the conviction that semantics is universal, in the sense that only the basic words, the morphology and the word order of sentences such as (0-8) are specifically English, but their meanings, which represent exclusively the structure and the results of conceptual operations, are independent of any particular language; 2. the inferability principle, i. e., the postulate that all valid inferences, and only these, must be possible on the semantic level by no other means than logical inference rules (and the language-specific meaning postulates for lexical items): e. g. while the surface structures of (2), (21), and (1) below do not reveal that (2) does not entail (1) while (2') does, the semantic structures must reveal this according to this principle, and do so in our notation:

(34) (2) John is taller than Mary. $\not\rightarrow$ (1) John is tall.

(35) (2') John is even taller than Mary. \rightarrow (1) John is tall.

(34a) $f_D^M(x) > f_D^M(y) \not\rightarrow f_D^M(x) > N_{D, Y}$
 $a > b \not\rightarrow a > c$

(35a) $f_D^M(x) > f_D^M(y) \quad \& \quad \langle f_D^M(y) > N_{D, Y} \rangle \rightarrow f_D^M(x) > N_{D, Y}$
 $a > b \quad \& \quad b > c \rightarrow a > c .$

We consider comparing (measuring) as a universal capacity of the human mind, and represent it semantically by the general measure function which is a two-place operation, with individuals as first arguments and dimensions as second arguments. While positives and comparatives are, in our system, both based on comparison, they are defined independently of each other. It follows that they are not based conceptually one upon the other. We thus eliminate the basic defect of the syntactically oriented analyses. We characterize and exemplify the syntactic rules relating these semantic representations in terms of their overt manifestations. The rules lexicalizing comparisons by means of relative adjectives are of a type that has not been discussed in the literature before. We argue that the semantic representations we provide are not inherently linearly ordered (a consequence of both the universality and the inferability principles), and we have found no reason to postulate extrinsic order for our syntactic rules.

It seems significant to us that earlier approaches, in addition to leaving doubts as to their semantic adequacy, have invariably left - or created - residual problems of even greater magnitude than they solved. By contrast, all of these problems are resolved in our theory in a uniform way without any ad hoc apparatus. It seems to us that our approach reveals these problems to be pseudo-problems created by models of language which treat syntax as the generative component of a grammar. We therefore feel encouraged to propose that our mode of analysis, which is based on the universality and inferability principles for semantic representations, be extended to other syntactic problems and to the foundation of language theory in general.

References

- Bartsch, R. and T. Vennemann (1972), "Relative Adjectives and Comparison", in P. Schachter and G. Bedell, eds., UCLA Papers in Syntax, 2. Univ. of Calif., Los Angeles.
- Bierwisch, M. (1971), "On Classifying Semantic Features", in D. D. Steinberg and L. A. Jakobovits, eds., Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology, Cambridge: University Press, 410-435.
- Dik, S. C. (1971), "Elsje is twee jaar ouder dan twee jaar geleden", MS, Univ. of Amsterdam.
- Fillmore, C. J. (1965), "Entailment Rules in a Semantic Theory", Pola Reports, 10 60-82 (Columbus: The Ohio State University).
- Hintikka, J. (1969), "Semantics for Propositional Attitudes", in J. W. Davis and D. J. Hockney, eds., Studies in Philosophical Logic, Dordrecht: D. Reidel, 21-45.
- Jacobs, R. A. and P. S. Rosenbaum (Eds.) (1970), Readings in English Transformational Grammar, Waltham, Mass.: Ginn.
- Reichenbach, H. (1947), Elements of Symbolic Logic, paperback edition, 1966, New York: Free Press.
- Ross, J. R. and D. M. Perlmutter (1970), "A Non-source for Comparatives", Linguistic Inquiry, 1, 127-128.
- Russell, B. (1945), A History of Western Philosophy (Clarion paperback), New York: Simon and Schuster.
- Sapir, E. (1944), "Grading: A Study in Semantics", Philosophy of Science, 11, 93-116.
- Seuren, P. (1970), "The Comparative", MS, Magdalen College, Oxford.
- Wierzbicka, A. (1972), "The Deep or Semantic Structure of the Comparative", Linguistische Berichte, 16, 39-45.
- Wunderlich, D. (1970), "Vergleichssätze", MS, Freie Universität, Berlin.