ON SOME SEMANTIC RULES IN A GRAMMAR*

S. I. Harada

0. In this paper, we shall be concerned with the relation between syntax and semantics. In particular, we shall consider what kind of rules are necessary in the semantics of natural languages, and by doing so, we shall argue against the so-called "interpretive" theory of meaning, first formulated by Katz and Fodor (1963) and accepted, with revisions, by the "standard" and "extended standard" theories of generative-transformational grammar.[1] The theory of meaning which we shall find adequate will resemble the so-called "generative semantics," but we shall also demonstrate that an adequate semantic theory must incorporate the notion of "feature percolation rules," a type of rule which has not been recognized by generative semanticists.

1. The Interpretive Theory

The "interpretive" theory of meaning is based on the following two assumptions:[2]

(I) The semantic component of a generative grammar is an interpretive system that operates on phrase markers independently generated by the syntactic component to assign a semantic interpretation (or "reading") to each constituent of a sentence.

(II) The semantic component contains a system of rules that assign the reading of a constituent by combining the readings of its subconstituents.

---

* This is a revised and expanded version of a portion of my Master's Thesis, "Syntactic Aspects of Meaning," submitted to the Graduate School of Humanities, University of Tokyo, December 1971. Earlier versions of this paper were presented orally on several occasions, and I am indebted to those who attended these meetings and delivered me valuable comments on them. I wish to express my special gratitude to Professors Osamu Fujimura, Kinsuke Hasegawa, Kazuko Inoue, and to John V. Hinds. None of them are responsible for the errors that may remain, however.

Assumption (I) is a formulation of Chomsky's belief in Syntactic Structures that "[syntax] is autonomous and independent of meaning."[3] Put in a different way, a roughly equivalent but more formal assumption is:

(I') No syntactic rule may refer to any semantic information, and rules of semantic interpretation do not refer to applications of particular syntactic rules but to the resultant phrase markers.

It should be noted that this is an empirical hypothesis and thus must be examined in the face of empirical evidence. Below, we shall present evidence that assumption (I') (and hence (I) if it implies (I')) does not allow an adequate description of a certain phenomenon.

Note, incidentally, that although there are a number of different versions of the interpretive theory, the difference among them stems just from the difference in the more specific empirical hypotheses. Thus the version of the interpretive theory accepted by the standard theory differs from the version in the extended standard theory in that the former embodies the hypothesis

(III) The set of underlying phrase markers gives all the syntactic information necessary for assigning the semantic interpretation to a sentence.

while the latter incorporates a more specific, but in a sense somewhat weaker hypothesis, namely,

(IV) The set of underlying phrase markers gives all the syntactic information needed for the operation of combinatorial semantic rules (as hypothesized in assumption (II) above) ; other semantic rules may refer to the set of phrase markers of a much shallower level.

The choice of one hypothesis over the other, however, does not affect the basic assumptions (I) and (II). Since we shall be concerned with assumptions (I) and (II), the difference between (III) and (IV) will be of no concern to us.

2. Lexical Insertion and Combinatorial Semantic Rules

2. 0. The interpretive theory recognizes a type of semantic rules with a combinatorial function, as expressed in assumption (II). Such rules are called "projection rules." The empirical content of this notion, however, is rather obscure and it is doubtful whether assumption (II) goes beyond the

obviously correct but quite trivial statement that the meaning of a constituent is composed of the meanings of its subconstituents. This is surely true, but empirically vacuous, since it does not reveal the intricate interactions of syntactic structures and semantic information. And actual examples of projection rules (proposed in various works by Katz) seem just to confirm this evaluation. All that they do is simply to connect the readings of the immediate constituents or to insert the readings of subject, object, etc. into the designated positions of variables within the reading of the main verb. But there is a simpler way to effect the same result, and that is to generate semantic elements (in place of lexical items) in the terminal positions of underlying phrase markers. If one segments a single constituent from an underlying phrase marker, the result will be the reading of that constituent. Thus we can dispense with the notion "projection rules" altogether, by virtue of this automatic segmentation convention.

In order for the interpretivist to refute this alternative, he must show that underlying phrase markers must contain all the lexical items, even those with "complex" readings. As we shall show in the next section, however, the actual evidence seems to argue against the interpretivist assumption.

2. 1. Consider, by way of example, the following sentence:[4]

(1) John promised Nancy a necklace when he won the presidential election.

This sentence is ambiguous: in one sense, the time adverbial "when he won the presidential election" refers to the time at which John made a promise; and in the other sense, it refers to the time at which John is expected to give Nancy the necklace he promised. We shall refer to these interpretations as "the promise-time interpretation" and "the give-time interpretation," respectively. Within the interpretivist framework, this ambiguity will be accounted for by postulating the lexical entry (2) for the verb "promise":[5]

(2) (PROMISE x, y, (GIVE x, y, z, at t_{i+k}), at t_i)

alongside of the normal lexical reading (PROMISE s, y, w, at t_i), where PROMISE and GIVE are semantic elements and x, y, z, w, t_i, and t_{i+k} are categorized variables. Projection rules are supposed to insert the

derived readings of the subject, indirect object, and direct object into the positions of x, y, and z, respectively. The ambiguity of (1) will be accounted for by allowing the reading of the time adverbial to be inserted either into the position of t_i or into the position of t_{i+k} in (2).

No problem has arisen so far. But now consider:

(3) When he won the presidential election, John promised Nancy a necklace.

Although (3) clearly derives from the same underlying structure as (1) does through the rule of Adverb Preposing, it lacks the ambiguity found in (1); (3) has only the promise-time interpretation. How can we account for this phenomenon within the interpretivist framework?

One possibility that immediately suggests itself is to have a projection rule for time adverbials to refer to derived syntactic structures, as in:

(4) Given a sentence whose main verb has a 'complex' lexical reading i. e., of the form: $(\ldots(\ldots \text{at } t_n)\ldots \text{at } t_m)$,

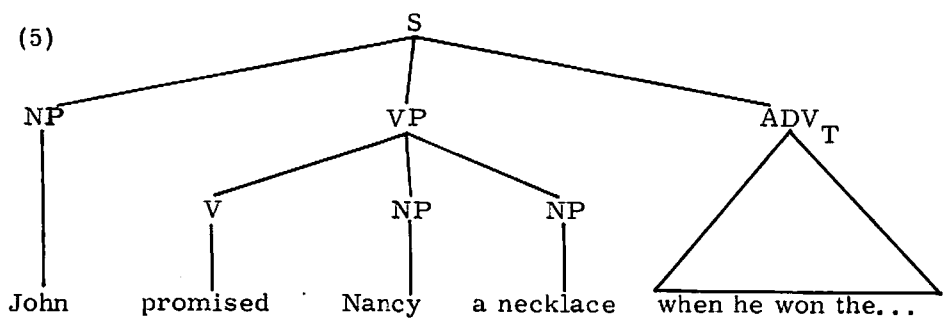
(i) insert the reading of ADV_T into the position t_m, not t_n, if the time adverbial ADV_T occurs at the beginning of the sentence at the level of surface structure;

(ii) the reading of ADV_T may be inserted into either position otherwise.

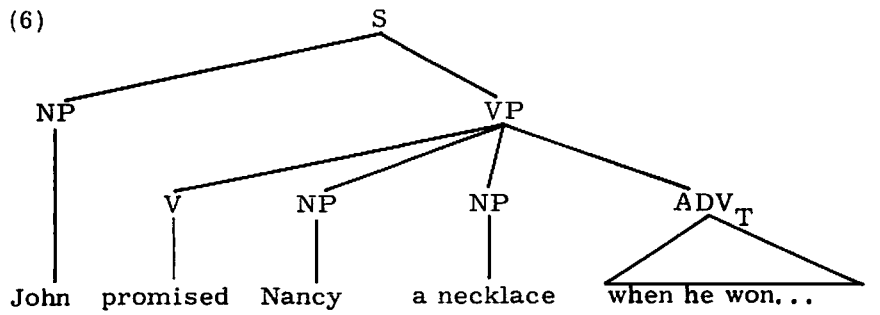
But such an account is obviously ad hoc; the ad-hocness of this approach can be best illustrated by the fact that other combinations of readings are determined solely in terms of underlying syntactic configurations.

A better account would be to set up two distinct underlying structures for the sentence (1), and to assign a unique interpretation to each underlying structure. The underlying structures would look something like the following:

(5)



(6)



The projection rule needed here would be something like this:

(7) Given a sentence whose main verb has a 'complex' lexical reading (. . . (. . . at t_n) . . . at t_m) and which contains a time adverbial ADV_T,

(i) insert the reading of ADV_T into the position of t_m if ADV_T is outside VP;

(ii) insert the reading of ADV_T into the position of t_n if ADV_T is inside VP.

With this rule, the underlying structure (5) will receive the promise-time interpretation, and the structure (6) the give-time interpretation. All that remains to be accounted for is the lack of ambiguity of (3). Since the only interpretation of (3) is the promise-time interpretation, it must be the case that the only possible source of (3) is the structure (5); the derivation of (3) from the structure (6) must somehow be blocked. But there is in fact an independently motivated mechanism that prevents the movement of the time adverbial to the sentence-initial position in structures like (6), namely, the "transportability convention" proposed by Keyser (1968). As he puts it (p. 368), this convention "permits a particular constituent to occupy any position in a derived tree so long as the sister relationships with all other nodes in the tree are maintained." Thus, the transformation of Adverb Preposing can front the time adverbial in (5) but not in (6), for in the latter case this would destroy the sister relationships that hold between the time adverbial and the other constituents in the VP.

Thus there appears, at first glance, to be a fair amount of plausibility to this account of the sentences (1) and (3); nevertheless it is still inadequate

in certain respects. Notice first the peculiarity of the projection rule (7). The formulation (7) would be strongly motivated if it were the case that the readings of any VP-internal constituents are to be associated with the "inner" reading [6] of the main verb. However, this is not the case. Compare the following sentence:

(8) John promised Nancy a necklace reluctantly.

The reading of (8) is such that John was reluctant in making the promise, not that John's promise was to give Nancy a necklace reluctantly. Thus, the reading of the manner adverbial "reluctantly" must be associated with the "outer" reading of "promise" and never with the "inner" reading of it. But in the framework under consideration, manner adverbials must originate inside VPs. [7] Thus, the formulation (7) of the relevant projection rule cannot be used as evidence for the postulation of distinct underlying structures for (1).

This leaves us with the facts about the transportation of time adverbials as the only support for the two distinct underlying structures (5) and (6). But here too, the interpretivist account fails to explain some additional facts. Observe first that manner adverbials, originating inside VPs, can be shifted to the VP-initial position, in conformity with the transportability convention:

(9) John reluctantly promised Nancy a necklace.

Compare this with the fact that the time adverbial in (6), though occurring inside the VP, cannot be fronted to the VP-initial position; although sentence (10) is ambiguous in the same way as (1), the corresponding sentence (11) is unambiguous and has only the promise-time interpretation:

(10) John promised Nancy a necklace later.

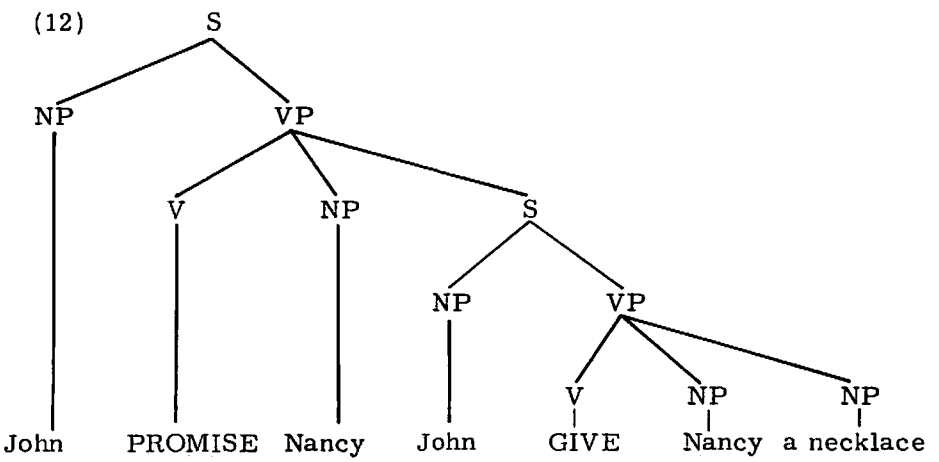
(11) John later promised Nancy a necklace.

The transportability convention incorrectly predicts that (11) is derivable from an underlying structure analogous to (6), as was the case with (9). But in fact the give-time interpretation is totally absent in (11). Thus, the transportability convention cannot be used as evidence for the interpretivist analysis.

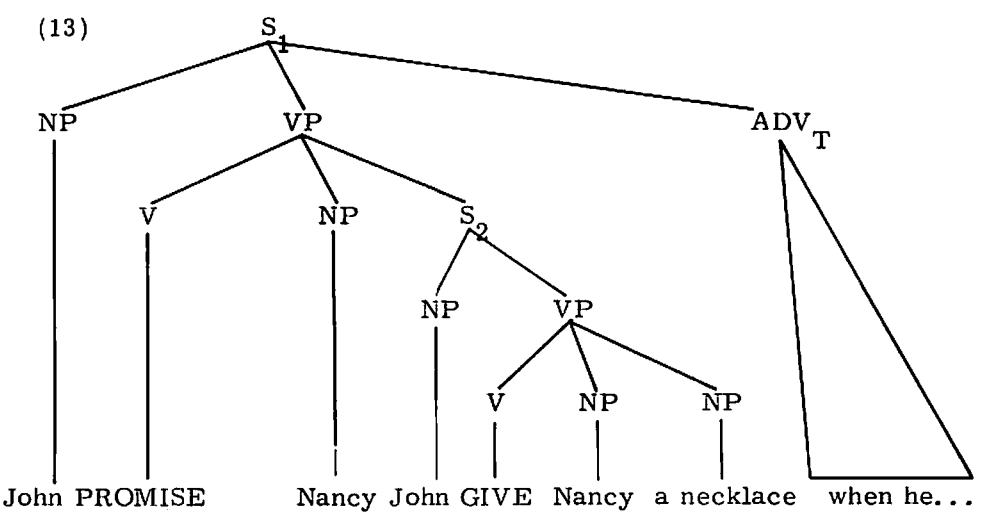
We have shown that none of the proposals formulable within the

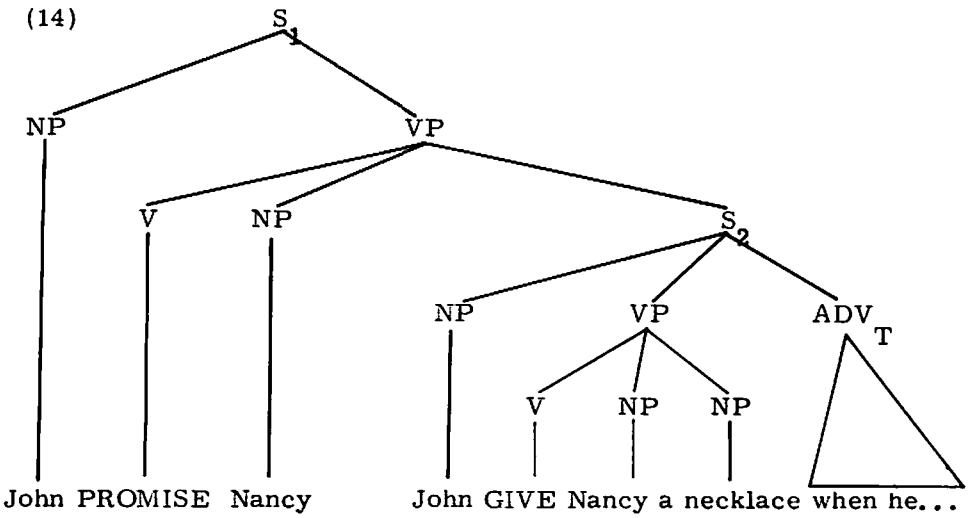
interpretivist framework succeeds in offering an adequate account of the verb "promise" with a complex reading. We shall now go on to show that these facts can be adequately accounted for when one abandons this framework.

Suppose we derive the sentence "John promised Nancy a necklace" from an underlying structure with an embedded sentence of the form (12).



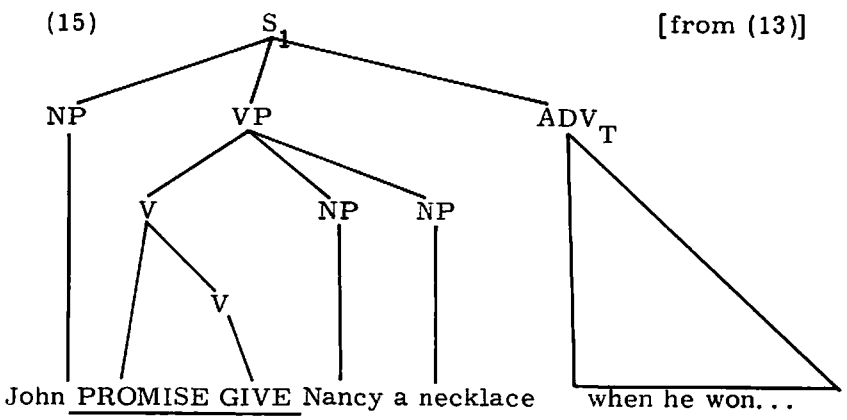
Then for sentence (1), we have two possible underlying structures, (13) and (14), differing only in the S node with which the time adverbial is associated:



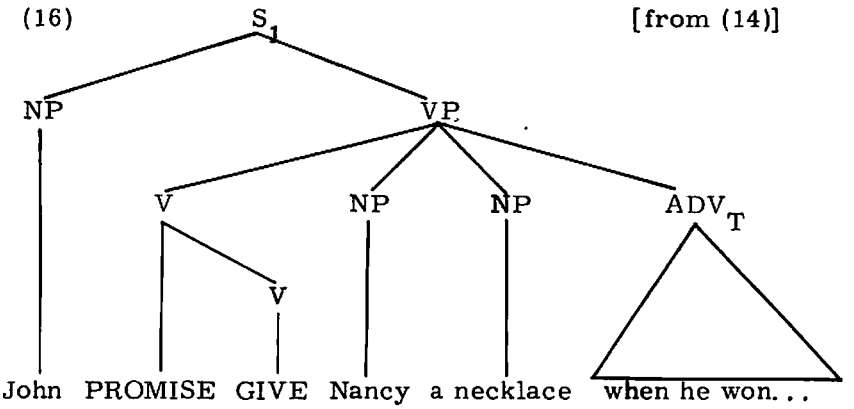


We can now account for both syntactic and semantic properties of the examples discussed so far.

Consider first the syntactic derivations from the underlying structures (13) and (14). The structures in (13) and (14) first undergo Equi NP Deletion, and the repeated occurrences of "John" and "Nancy" are deleted. Then those structures undergo the rule of Predicate Raising (cf. McCawley 1971) to yield the derived structures shown in (15) and (16).



(16) [from (14)]



The syntactic derivations of sentence (1) will be completed by the insertion of the lexical item "promise" into the position of "PROMISE + GIVE." The results will be precisely the structures (5) and (6). These structures could not receive a full justification as "syntactic" underlying structures, since it was found necessary to state certain restrictions on Adverb Preposing in terms of structural differences such as those found in (5) and (6). Our analysis has, however, enabled us to derive these structures from well-motivated sources (13) and (14), and has thus given the structures (5) and (6) a rather solid status in the syntactic derivations of (1).

When we proceed to the account of adverb preposing phenomena, we immediately find that Keyser's transportability convention must be reformulated as a global derivational constraint. Though a full-scale investigation goes far beyond the scope of this paper, we can at least propose the following tentative formulation:

(17) Global Constraint on Adverb Preposing:

If in an underlying structure there is an adverbial ADV_i and an element A such that A asymmetrically commands ADV_i, then ADV_i cannot precede A in surface structure.

This will then account for the nontransportability of "later" in structures like (6), since if "later" were shifted to the VP-initial position, it would precede the elements "promise" and "Nancy" which belonged to a higher clause in the underlying structure.[8] In the case of manner adverbials like "reluctantly," however, there is no difficulty in preposing. In (9), for exam-

ple, "reluctantly" originates in the VP of S_1 (but not in the VP of S_2) and therefore the preposing of this adverb to the position immediately before "promise" will not be blocked by the constraint (17).

Assuming, then, that our proposed account is preferable to any of the interpretivist accounts,[9] let us now consider the theoretical implications of our account. In the first place, it is obvious that the interpretivist assumption of autonomy of syntax (cf. (I) and (I') above) cannot be seriously maintained, since this assumption leads us to the theorem

> (V) No semantic element may occur in the underlying phrase marker without being lexicalized.

with which our proposed underlying structures (13) and (14) are in direct conflict. The issue is not simply about the point of a derivation at which lexicalization takes place; it concerns rather with the question of what prelexical structures look like. As Katz (1971:322) points out, the only alternative to the deep lexicalization hypothesis available in the interpretivist framework is something like this:

> (VI) Certain major categories may be left unlexicalized until some nonlexical transformations have applied; rules of semantic interpretation refer to the stage of derivation at which such categories are lexicalized.

As we have argued, however, this solution does not work in the "promise" examples; for, in this case, we must set up underlying structures in which two separate items must be assumed for what forms a single lexical item in surface structure. Therefore, if our account is correct, we will have a strong empirical evidence against the interpretivist assumption of autonomy of syntax.

Note incidentally that exactly the same conclusion was reached by Harada and Saito (1971), where evidence was presented that reflexive pronouns must be derived by a transformation and that this transformation must be stated with reference to the semantic property of coreference, an account in direct conflict with the interpretivist assumption (I'). Coupled with the present findings, it seems beyond any doubt that the interpretivist assumption of autonomy of syntax is empirically incorrect.

Another important implication of our proposal is that no specific "projection rule" is necessary now that structures like (12) are assumed for underlying structures containing "promise." The reading of a constituent will be obtained, quite automatically, from such "semantic" underlying structures by means of a fairly general convention, as we have pointed out above. If, as some linguists (Katz (1971) in particular) have argued, syntactic labels such as S, NP, V, etc. are of no use in semantic interpretation, all we have to do is simply to supplement this general convention by a subsidiary convention that deletes all syntactic labels but leaves the bracketings intact. The result will be exactly the same as the semantic interpretation given by the interpretivist account.[10)]

3. On Feature Percolation Rules

3. 0. The theory of meaning behind our proposed analysis of the "promise" examples bears a close resemblance to the so-called "generative semantics"[11)] in that the underlying syntactic representation is itself regarded as the semantic representation. In both theories, the base component of a grammar generates a set of phrase markers with semantic elements in the terminal positions: there are no 'projection rules', and a uniform set of transformational rules interrelate such underlying pharas markers with surface phrase markers. But we shall argue in this section that we must have a set of rules which have not been recognized within generative semantics, namely, what we shall call "feature percolation rules." We shall conclude that any adequate semantic theory must incorporate an equivalent of such rules.

3. 1 As an illustrative example, let us consider the following fact noted by Kamio (1971). As is widely known, there is a restriction on the main verb of an embedded sentence occurring after an element of the class (18):

(18) Imperative, try to, want to, I will, encourage NP to...

This is usually considered as a "Verb-Verb selectional restriction" holding between a matrix verb of the class (18) and a complement verb.[12)] That is, verbs of the class (18) require the complement main verb to have the semantic property of "controllability"[13)]; non-controllable verbs such as

"be tall," "know the answer," etc. are excluded from these contexts. Usually, the presence of a negative element in the complement clause does not affect this selectional restriction:

(19) a. John tried to kiss Nancy. (controllable)
b. John tried not to kiss Nancy.

(20) a. *John tried to be tall. (non-controllable)
b. *John tried not to be tall.

There is, however, a class of predicates which behave differently in the complement clause depending on the presence or absence of a negative element:

(21) a. take courage, calm oneself, take it easy, keep cook, cheer oneself up, . . .
b. worry about NP, feel nervous, lose one's patience, fret about NP, . . .

Without negatives, predicates of the class (21) can, and those of the class (21) cannot, occur in the context of (18).

(22) a. John tried to take courage.
b. *John tried to feel nervous.

When the predicate is negated, however, the restriction is reversed, and it is now the predicates of the class (21) that are excluded from the context under consideration:

(23) a. *John tried not to take courage.
b. John tried not to feel nervous.

The only framework in which facts like this can be accounted for is the one proposed by Kuno (1970); we shall now try to describe the phenomenon along his lines.[14]

As a preliminary to the account we shall propose, we argue that selectional restrictions should be reformulated as semantic restrictions imposed by some constituent on relevant phrase nodes (not necessarily lexical nodes), most usually sister constituents.[15] Thus, the selectional restriction discussed above with respect to the items in (18) should be restated as a restriction involving a verb of the class (18) and the complement sentence which occurs as a sister constituent of that verb, not as a verb-verb selectional restriction (that is, not as a selectional restriction holding

between two lexical categories). The verb try (or the semantic element TRY; the choice is immaterial), for example, will be given the selectional feature [+____S[+controllable]] in our framework. In general, the semantic properties of the main verb are, at the same time, those of the entire sentence, and when a sentence is embedded in a noun phrase (as a noun phrase complement), then the semantic properties of the sentence are those of that noun phrase, and so on. Thus, for instance, the sentence "John kissed Nancy" denotes a physical action just as the verb "kiss" denotes a physical action, and so on. We must therefore recognize rules of the following sort, which "percolate" the semantic properties of lower categories up to higher categories:[16)]

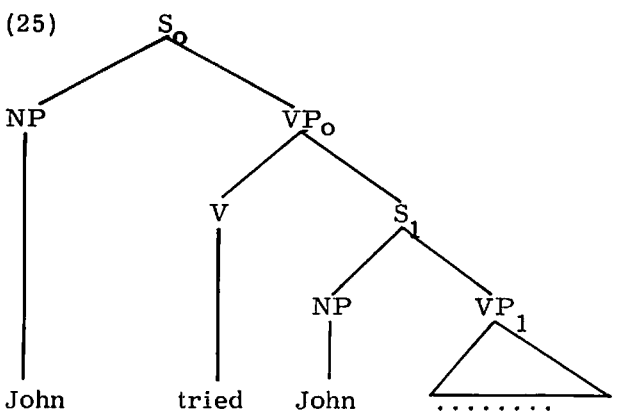
(24) a. $V[\alpha F] + \ldots \longrightarrow VP[\alpha F]$

b. $NP + VP[\alpha F] \longrightarrow S[\alpha F]$

c. $it + S[\alpha F] \longrightarrow NP[\alpha F]$

By rules of this sort, we can formulate the filtering processes in the following manner. Sentences (19) and (20) are derived from the following underlying structure:

(25)



The predicate that occupies the position of dots in this diagram may contain either the semantic feature [+controllable] or [-controllable]. Such features are copied onto the node VP_1 by rule (24), and then to S_1 by rule (24). If the predicate is [+controllable], then S_1 = [+controllable] and the entire sentence is filtered out as involving a violation of the restriction. In the case

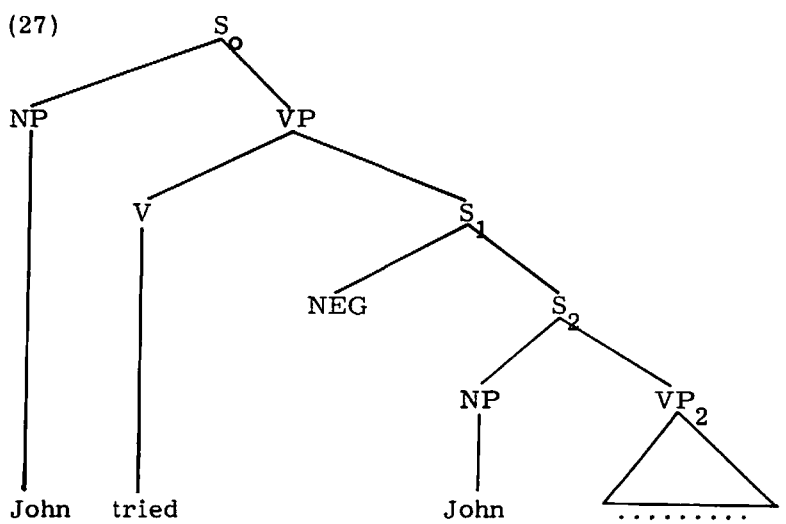
of sentences (19) and (20), the underlying structure contains an additional element, NEG (negative), perhaps as the leftmost daughter of S_1.[17] We take it, however, that NEG is a sort of "transparent" element, and its presence does not affect the application of rules like (24), unless otherwise specified. The selection processes for these sentences will be, therefore, essentially the same as in the case of sentences (19) and (20).

Now consider sentences in (22) and (23). The predicates in (21) all appear to belong to a natural semantic class, and we shall assume that they are associated with the features [+psychological, +polar] in the lexicon. We may distinguish between (21) and (21) by means of a new feature, say, [± GMC] (for "good mental condition"), assigning the plus value to the former class and the minus value to the latter. Since [+GMC] predicates pass and [-GMC] predicates fail the selectional restriction of <u>try</u>, we can assume that there is a semantic redundancy rule[18] like the following:

(26) [α GMC] ⟶ [α controllable]

Now the sentences in (22) occur in the same underlying structure as (25), and the selection processes are the same as in the more normal cases. In the case of sentences in (23), however, there is a need for an additional mechanism. They originate in the underlying structure of the following form:

(27)



Suppose the predicate that occupies the position of dots in (27) is "take courage," with the feature [+GMC]. By (24a), VP_2 receives the specification [+GMC] and hence [+controllable]. The same set of features is assigned to S_2 by rule (24b). However, since in this case the entire sentence (S_0) must be filtered out, S_1 must have the specification [-controllable). As it now stands, there is no way to prevent the assignment of the feature [+GMC] and hence [+controllable] to the node S_1. What we need is a rule of the form:

(28) NEG + S[+GMC] ⟶ S[-GMC]

Given this rule, we can now assign the feature [-GMC] to S_1, and S_1 will automatically receive the specification [-controllable] due to the semantic redundancy rule (26). Likewise, in the case of [-GMC] predicates, what we need is a rule like:

(29) NEG + S[-GMC] ⟶ S[+GMC]

Combined into a single schema, rules (28) and (29) form:

(30) NEG + S[α GMC] ⟶ S[~α GMC]

This is just another instance of what Kuno (1970) called a "feature-changing" semantic rule. Since Kamio's fact cannot be accounted for without such a rule, the account proposed above serves as additional evidence for incorporating such rules into semantic theory.

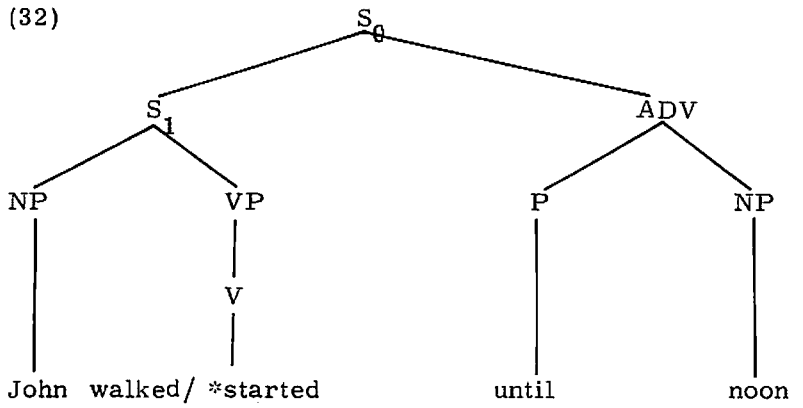
3.2 Other evidence for the same conclusion can be adduced ad nauseam. Consider for instance the rather well-known restriction involving until-adverbials. The restriction is that an until adverbial can co-occur only with a main verb which expresses a durative action or state:

(31) a. John walked until noon. (durative)

b. *John started until noon. (non-durative)

Assuming that sentences like these are derived from underlying structures like the following.[19]

(32)



the relevant restriction can be stated as a selectional feature [+S [+durative] ____ . . . .] imposed by until on the main clause. When the main verb has the feature [+durative], S_1 = [+durative] by rules (24a, b), and the entire sentence S_0 passes the selectional restriction in question. When the main verb has the feature [-durative], S_1 = [-durative], and the entire sentence fails this selectional restriction.

When the verb is negated, however, the situation is a bit more complicated:

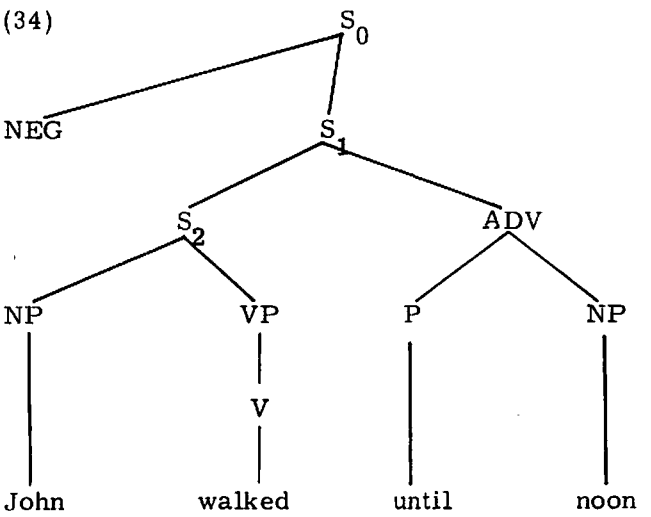
(33) a. John didn't walk until noon.

b. John didn't start until noon.

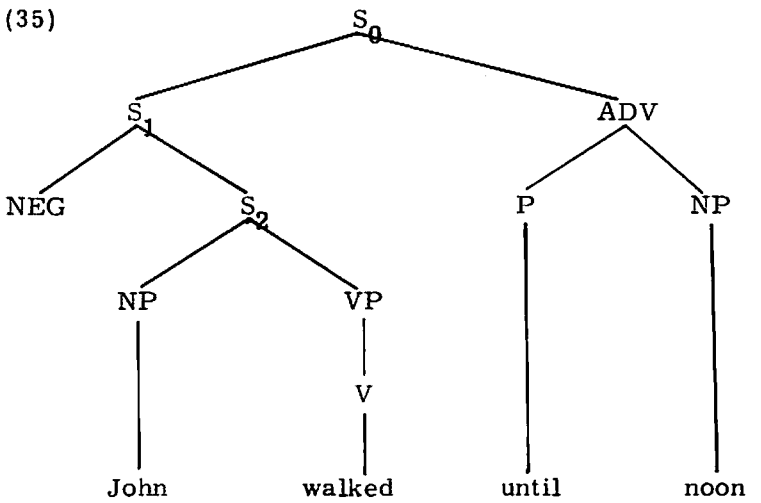
It is obvious that a sentence like (33a), with a durative main verb, is ambiguous and that a sentence like (33b), with a non-durative main verb, is now grammatical (but unambiguous).

The explanations are quite straightforward. The ambiguity of a sentence like (33a) results from the two possibilities in the choice of the S node with which NEG is associated in the underlying structure (31). (33a) may derive from either of the following underlying structures: [20]

(34)



(35)



Roughly, the underlying structure (34) corresponds to the reading of (33a) which presupposes that John had been walking in the morning and asserts that he stopped walking before noon, and the structure (35) corresponds to the reading that presupposes that John had not been walking at all in the morning and asserts that he began walking at noon. The selectional facts may need no explicit comment, except for the fact that in (35) NEG functions as a "transparent" element, and that the feature [+durative] on walk is

carried over to the node S_1.[21]

Let us now consider the sentence (33b). As in the case of (33a), we can conceive of two putative underlying structures in principle, namely those of (33a) with "started" substituted for "walked". Introspection of the meaning of (33b) reveals that it is only the underlying structure corresponding to (35) that is permitted. The selectional properties of until must thus be accounted for to reflect this intuition. Suppose we have the following feature-changing rule:

(36) NEG + S [-durative] ⟶ S [+durative]

Then the putative underlying structure of (33b) which parallels (34) is filtered out but that which parallels (35) is passed as well-formed. By rules (24a, b), S_2 receives the specification [-durative], but by (36) S_1 is specified as [+durative] and thus passes the selectional restriction of until.

The correctness of an intuition that the grammatical source for (33b) is like (35) and not like (34) can be demonstrated by the fact that the following sentences are grammatical (and, in the case of (37a), unambiguous with the sense derivable from (34) as its only meaning):

(37) a. Until noon, John didn't walk.

b. Until noon, John didn't start.

The global constraint on adverb movement that we proposed above prevents the until-adverbial from being shifted to the sentence-initial position in structures like (34).[22]

3. 3 We have thus arrived at two new feature-changing rules that seem indispensable to the description of two phenomena, one recently discovered and the other well-known. Notice, however, that these rules share certain properties in common with the rules in (24); they both belong to the same format schematized in (38).

(38) $C_1 [\alpha_{i_1} F_{i_1}] + C_2 [\alpha_{i_2} F_{i_2}] + \ldots + C_n [\alpha_{i_n} F_{i_n}] \rightarrow C_0 [\alpha_{i_0} F_{i_0}]$

where C_1, C_2, ..., C_n are the immediate constituents of C_0. Let us call rules of this sort "feature percolation rules." In the simplest case, if C_k is the "head" of C_0, then $\alpha_{i_0} = \alpha_{i_k}$ and $F_{i_0} = F_{i_k}$ (for any i), and any

other constituents C_j ($j \neq k$, of course) are irrelevant. Feature-changing rules are exceptional in that while $F_{i_0} = F_{i_k}$, $\alpha_{i_0} \neq \alpha_{i_k}$. Some of the feature-changing rules proposed by Kuno require further relaxations of the condition on (38), but a full exploration of this matter lies beyond the scope of this paper.[23]

A comment seems necessary on the ways feature-percolation rules operate. It is clear that non-feature-changing rules are exceptions to these general rules. This means that given two feature-percolation rules which are identical except that one is a feature-changing rule while the other is not, the feature-changing rule must apply before the non-feature-changing rule and when it applies the latter rule cannot apply further with respect to the feature specified in the rules under consideration. Technically, these rules must be disjunctively ordered, and the feature-changing rule must precede the non-feature-changing rule. We can express this matter by extending the notation proposed in Chomsky and Halle (1968). Suppose we have a feature-percolation rule like the following:

(39) NEG + S [<-durative>] ⟶ S [<+durative>]

This is an abbreviation of two rules, (36) and the more general, non-feature-changing rule. Let us stipulate, in cases like this, that the non-feature-changing counterpart is a rule like this:

(40) NEG + S [αF] ⟶ S [αF]

where αF = <-durative>. In fact, (40) is not a rule but rather a rule schema, abbreviating a set of rules which blindly assign the features to the higher S node (including the feature <+durative>).

4. Summary and Conclusions

We have shown in this paper that (i) neither of the interpretivist assumptions (1) and (2) is supported by empirical evidence, and (ii) the adequate semantic theory must contain feature percolation rules, while it need not (or, for the sake of explanatory adequacy, should not) contain projection rules. The difference between necessary and unnecessary semantic rules is that the former set of rules operates on semantic "features," while the latter is supposed to operate on semantic elements. Our present findings imply that the rules that necessarily operate on semantic elements are

just transformations.

We thus propose to distinguish between semantic features and semantic elements, as it is necessary to distinguish between phonological features and phonological segments. Although we cannot give a fuller justification for this distinction here, it seems quite fruitful to pursue the consequences of this distinction.

NOTES

1) For these terms, see Chomsky (1972).

2) These are explicitly stated and discussed at length in Katz (1970). See now also Katz (1972).

3) Chomsky (1957 : 17)

4) Examples of this sort have been called to my attention by Masaru Kajita.

5) See Katz (1971) and Katz (1972) for Katz' own formulation of lexical entries.

6) When a lexical entry has an embedding, i. e. of the form (X(Y)Z), we shall refer to the part (XZ) as the "outer" reading and to the part (Y) as the "inner" reading of the entry.

7) See, e. g., Chomsky (1965) for justification.

8) Although this analysis predicts that sentences like

(a) *John promised Nancy later a necklace.

are grammatical, they are not acceptable. But this seems to be due to an independently motivated surface structures constraint that prevents a non-complex NP from being separated from the main verb by an adverbial element. See Keyser (1968) for some discussion.

9) The Japanese counterpart of "promise, " namely "yakusoku-suru, " offers a similar, but not identical piece of evidence against (18). The Japanese predicate "yakusoku-suru, " unlike "promise, " never has the ambiguity of (5); thus the following, intended as a translation of (5), is not ambiguous:

(a) John wa daitooryoo-senkyo ni katta toki ni Nancy ni nekkuresu o yakusoku-sita.

(a) only has the first interpretation of (5). But consider sentences like

(b) John wa Nancy to go-zi ni yakusoku-sita.
5 o'clock at

'John has an appointment with Nancy at 5 o'clock.'

(This example was pointed out to me by Osamu Fujimura.) The predicate "yakusoku-suru," in the normal sense, cannot co-occur with NP-to-phrases:

(c) John wa Nancy ni/*to tegami o dasu to yakusoku-sita.
letter send

"John promised (*with Nancy) that he would send her a letter."

Actually, however, (b) has the sense of "John promised Nancy to see her at 5 o'clock," and indeed the NP-to-phrases can co-occur with the predicate awu (see). Thus, here is a good reason for deriving structures like (b) from underlying structures with an embedded sentence whose main verb is awu.

Although (b) may refer to indefinitely many different situations, they need not be represented differently in underlying structure, for this is simply a matter of the motivation for John's seeing Nancy (i.e., whether to dine with her, or to discuss something with her, or to commit adultery, or the like). The crucial fact is that (b) cannot be paraphrased by means of any predicate with an NP-to-phrase, e.g.

(d) John wa Nancy to (*go-zi ni) onazi syumi o motu to
same pastime

yakusoku-sita.

'John promised Nancy to have the same pastime as she (*at 5 o'clock).

10) The matters are not as simple as this, however. The actual lexical readings proposed by Katz and others often contain quite complex internal structures, and it is not wholly clear whether one can account for the same facts in the "lexical decomposition" framework advocated here. For instance, Katz (1967: 172) proposes the following lexical reading for the verb "sell" (and, with minor adjustments, for "buy" as well):

((Condition (Possession of Y) of X at t_i) → (Condition (Possession of Y) of Z at t_{i+k})) & ((Condition (Possession of sum of money W) of Z at t_i) → (Condition (Possession of sum of money W) of X at t_{i+k}))

But it seems to me that this lexical reading confuses several different kinds of semantic information. In the current terminology, it misleadingly confuses the "assertive" part of the reading and the "presuppositional" part of it. Thus, in the case of the lexical reading of "sell, " the part "(Condition (Possession of Y) of X at t_i)" forms the presupposition, while the part "(Condition (Possession of Y) of Z at t_{i+k})" forms the assertion. This is witnessed by the fact that we cannot say "John didn't sell the encyclopedia to Mary " when John never was the owner of the encyclopedia, while we can normally say the same sentence when John was the owner of the encyclopedia but the transition of the possessorship did not take place; the former situation is an instance of "infelicity, " arising from the falsity of the presupposition of the utterance. But the presuppositions of a sentence do not have to be represented within the semantic representation of the sentence; in fact, to do so is utterly impossible. Rather, such relations as "presupposition of, " "implication of, " "equivalent to" etc. are best stated 'transderivationally, ' to exploit the recent terminology. Thus, what we need is presumably a transderivational constraint that states, e. g. , that the semantic representation of "John possessed the encyclopedia before t_1" is the "presupposition of" the semantic representation of "John sold the encyclopedia to Mary at t_1, " etc. With the hope that such extension of semantic analysis is generally compatible with the lexical decomposition hypothesis, we have drawn the conclusion reached in the main text. See Harada (in preparation) for a fuller explanation.

11) See Lakoff (1971) in particular.

12) In order to state the restriction more generally, the imperative marker must be analyzed as an underlying predicate. See Harada (1971) for some discussion.

13) Kuno calls this "self-controllability. "

14) Kamio (1971) presents a lucid discussion of how this phenomenon cannot be adequately treated in the classic framework of selectional restrictions as proposed in Chomsky (1965) or in Katz (1966).

15) Notice, incidentally, that if there is no VP node in underlying structure, we can strengthen the generalization, dropping the qualification "most usually. "

16) The formal nature of such rules will be discussed at length at the end of this section.

17) We assume here that NEG is introduced by the rule:

$S \rightarrow NEG\ S$,

but the actual form of the rule is not crucial to the present discussion.

18) Semantic redundancy rules are considered here as a sort of convention, applying at any stage in the derivation when they are applicable.

19) Again, details of this structure are immaterial to the present discussion. It is perfectly all right, in the present context, to introduce the until-adverbial by a rule like $S \rightarrow NP\ VP\ ADV$, as in the case of when-clauses discussed at length in the previous section. However, in order to state the relevant restriction in the most general way, the underlying structure must be such that until is an underlying predicate, with "noon" as the object and "John walked" as the subject. The selectional restriction of until will then be nothing but an instance of normal subject-verb-object selection.

20) Again, the structural details are immaterial. If the until-adverbial is introduced by a rule like $S \rightarrow NP\ VP\ ADV$ and NEG by a rule like $S \rightarrow (NEG)\ NP\ VP$, the relevant structural difference is to be obtained by simply allowing VP to have its own NEG. In this analysis, (34) would be of the form [NEG NP VP ADV] and (35) of the form $[\ NP\ [_{VP} NEG\ V\ ]_{VP} ADV\ ]$ The argument to follow would apply, mutatis mutandis, even if this analysis were accepted.

21) See the discussion at the end of this section.

22) See McCawley (1971) for more examples of co-occurrency phenomena.

23) Kuno (1970) proposes, e. g., a rule of the following sort:

[+destructive] + NP [-open class] $\rightarrow$ VP [-repeatable]

If this is the correct formulation, we have to considerably relax the contitions on feature-percolation rules.

References

N. Chomsky (1957), Syntactic Structures, Mouton and Co., The Hague.

_____ (1965), Aspects of the Theory of Syntax, The M. I. T. Press, Cambridge, Mass.

_____ (1972), Studies on Semantics in Generative Grammar, Mouton and Co., The Hague.

S. I. Harada (1971), "Where Do Vocatives Come From?," English Linguistics 5, 2-43.

_____ and S. Saito (1971), "A Non-Source for Reflexives," Linguistic Inquiry II, 546-557.

A. Kamio (1971), "On the Nature of a Certain Type of Selectional Restriction," Studies from Meiji Gakuin University 174, 159-186.

J. J. Katz (1970), "Interpretive Semantics vs. Generative Semantics," Foundations of Language 6, 220-259.

_____ (1971), "Generative Semantics Is Interpretive Semantics," Linguistic Inquiry II, 313-331.

_____ (1972), Semantic Theory, Haper & Row, New York.

_____ and J. A. Fodor (1963), "The Structure of a Semantic Theory," Language 39, 170-210.

J. Keyser (1968), "Review of Sven Jacobson, Adverbial Positions in English," Language 44, 357-374.

S. Kuno (1970), "Feature-Changing Rules in Semantics," Mathematical Linguistics and Automatic Translation Report No. NSF-24, Harvard University, pp. 69-89.

J. D. McCawley (1971), "Prelexical Syntax," Monograph Series on Languages and Linguistics No. 24, 19-33.