ON GENERATIVE SEMANTICS[1]

George Lakoff

The University of Michigan

1. The Basic Theory

I would like to discuss some questions having to do with the theory of grammar. I assume that a grammar of a language is a system of rules that relates sounds in the language to their corresponding meanings, and that both phonetic and semantic representations are provided in some language-independent way. I assume that the notion 'possible surface structure' in a possible natural language is defined in terms of 'trees' or 'phrase-markers', with the root S, whose node-labels are taken from a finite set of node-labels: S, NP, V, ... The notion 'tree', or 'phrase-marker', is to be defined in one of the usual ways, in terms of predicates like precedes, dominates, and is labelled. Thus a grammar will define an infinite class of surface structures. In addition I assume that a grammar will contain a system of grammatical transformations mapping phrase-markers onto phrase-markers. Each transformation defines a class of well-formed pairs of successive phrase-markers, P_i and P_{i+1}. These transformations, or well-formedness constraints on successive phrase-markers, P_i and P_{i+1}, define an infinite class K of finite sequences of phrase-markers, each such sequence $P_1,...,P_n$ meeting the conditions:

(1) (i) P_n is a surface structure
     (ii) each pair P_i and P_{i+1} meet the well-formedness con-
           straints defined by some transformation
     (iii) there is no P_0 such that $P_0,P_1,....,P_n$ meets conditions
            (i) and (ii).

The members of K are called the syntactic structures generated by the grammar. I will assume that the grammar contains a lexicon, that is, a collection of lexical entries specifying phonological, semantic, and syntactic information. Thus, we assume

(2) a lexical transformation associated with a lexical item I maps
     a phrase-marker P containing a substructure Q which contains
     no lexical item into a phrase-marker P' formed by superimpos-
     ing I over Q.

That is, a lexical transformation is a well-formedness constraint on classes of successive phrase-markers P_i and P_{i+1}, where P_i is identical to P_{i+1} except that where P_i contains a subtree Q, P_{i+1} contains the lexical item in question. Various versions of this framework will differ as to where in the grammar lexical transformations apply, whether they apply in a block, etc.

In this sense, transformations, or well-formedness conditions on successive phrase-markers, may be said to perform a 'filtering function', in that they 'filter out' derivations containing successive phrase-marker pairs (P_i, P_{i+1}) which do not meet some well-formedness condition on such pairs. A system of transformations is essentially a filtering device which defines a class of well-formed sequences of phrase-markers by throwing out all of those sequences which contain pairs (P_i, P_{i+1}) which do not meet some such well-formedness condition, that is, are not related by some transformation. Since transformations define possible derivations only by constraining pairs of successive phrase-markers, I will refer to transformations as 'local derivational constraints'. A local derivational constraint can be defined as follows. Let 'P_i/C_i'

mean phrase-marker P_i meets tree condition C_i. A transformation, or local derivational constraint, is a conjunction of the form P_i/C_1 and P_{i+1}/C_2, as where C_1 and C_2 are tree conditions defining the class of input trees and class of output trees, respectively. It is assumed that:

$$C_1 = C_1' \text{ and } C_1''$$

$$C_2 = C_2' \text{ and } C_2''$$

$$C_1' = C_2'$$

$$C_1'' \neq C_2''$$

$$C_1' \text{ and } C_2' \text{ are both nonnull}$$

$$C_1'' \text{ and } C_2'' \text{ are not both null}$$

C_1' (which is identical to C_2') will be called the <u>structural description</u> (SD) of the transformation. C_1'' and C_2'' will be called the <u>structural correlates</u> (SC) of the transformation. The SD of the transformation defines the part of the tree condition which characterizes both P_i and P_{i+1}. The SC of the transformation defines the minimal difference between P_i and P_{i+1}. Thus, a pair (C_1, C_2) defines a local derivational constraint, or 'transformation'. A derivation will be well-formed only if for all i, $1 \leq i < n$, each pair of phrase-markers (P_i, P_{i+1}) is well-formed. Such a pair will, in general, be well-formed if it meets some local derivational constraint. There are two sorts of such constraints: optional and obligatory. To say that a local derivational constraint, or transformation, (C_1, C_2) is <u>optional</u> is to say:

(x) $(P_x/C_1 \supset (P_{x+1}/C_2 \supset (P_x, P_{x+1})$ is well-formed $)$ $)$

To say that (C_1, C_2) is <u>obligatory</u> is to say:

(x) $(P_x/C_1 \supset (P_{x+1}/C_2 \equiv (P_x, P_{x+1})$ is well-formed $)$ $)$

For a derivation to be well-formed it is necessary (but in general not sufficient) for each pair of successive phrase-markers to be well-formed.

In addition to transformations, or local derivational constraints, a grammar will contain certain 'global derivational constraints'. Rule orderings, for example, are given by global derivational constraints, since they specify where in a given derivation two local derivational constraints can hold relative to one another. Suppose (C_1, C_2) and (C_3, C_4) define local derivational constraints. To say that (C_1, C_2) is ordered before (C_3, C_4) is to state a global derivational constraint of the form:

(i) (j) $((P_i/C_1 \text{ and } P_{i+1}/C_2 \text{ and } P_j/C_3 \text{ and } P_{j+1}/C_4 \supset (i < j))$.

Another example of a global constraint is Ross' coordinate structure constraint which states that if some coordinate node A^1 dominates node A^2 at some point in the derivation P_i, then there can be no P_{i+1} such

that A^2 commands A^1 and A^1 does not dominate A^2, where "command" means "belong to a higher clause than." That is, A commands B if and only if the first S-node higher than A dominates B. This is a global derivational constraint of the form:

Let C_1 = A^i dominates conjunction $A^{j_1} \ldots A^{j_m}$

C_2 = A^i dominates $x^1 \, A^k \, x^2$

C_3 = A^k commands A^i

CSC: (y) ~($(P_y/C_1$ and $C_2)$ and $(P_{y+1}/C_3$ and $\sim C_2)$))

What the coordinate structure constraint does is to keep track of the derivational histories of pairs of nodes A^i and A^k. This is just what elementary transformations and their associated rules of derived constituent structure do: they define constraints on successive phrase-markers, keeping all but one or two nodes constant, and then tell what happens to those one or two other nodes in going from the first tree to the second. It seems reasonable on the basis of our present knowledge to limit individual derivational constraints, both local and global, to tracing the histories of at most two nodes—over a derivation in the case of global constraints, and over two successive trees in the case of local constraints. Other examples of global derivational constraints are Ross' other constraints on movement rules (Ross, [65]), the theory of exceptions (Lakoff, [40] and R. Lakoff, [44]), Postal's crossover principle (Postal, [57]), output conditions for pronominalization (Lakoff, [41]), etc. It should be clear that all theories of transformational grammar have included both local and global derivational constraints. The question arises as to what kinds of local and global derivational constraints exist in natural languages. I will suggest in section 2 below that there is a wider variety than had previously been envisioned. It is assumed that derivational constraints will be restricted to hold either at particular levels in a derivation (semantic representation, surface structure, shallow structure, and deep structure if such exists), or to range over entire derivations or parts of derivations occurring between levels. Constraints holding at particular levels. Constraints holding at particular levels define well-formedness conditions for those levels, and so are analogous to McCawley's node-acceptability conditions ([47]), which play the role of phrase-structure rules in a theory containing deep structures.

Given a syntactic structure $(P_1,\ldots,P_n)$ we define the semantic representation SR of a sentence as SR = $(P_1,$ PR, Top, F, ...), where PR is a conjunction of presuppositions, Top is an indication of the 'topic' of the sentence, and F is the indication of the focus of the sentence. We leave open the question of whether there are other elements of semantic representation that need to be accounted for. Perhaps some examples are in order. Let us start with presuppositions. Pedro regretted being Norwegian presupposes that Pedro is Norwegian. Sam's murderer reads Reader's Digest presupposes that Sam was murdered. In general, a sentence may be either true or false only if all its presuppositions are true. Since any propositions (at least nonperformative ones) may be presupposed, it is assumed that the elements of PR are of the same form as those of P_1, and that they are defined by the same well-formedness

conditions. The notation given above, with PR as a member of an ordered n-tuple, assumes that presuppositions are structurally independent of P_i. However, Morgan, in an important paper [52], has argued convincingly that there are cases where presuppositions must be linked to certain propositions embedded in P_i, and that such links are identical to, or share properties with, conjunctions. He has also shown that presuppositions may be attributed not merely to the speaker and addressee, but also to the subjects of certain predicates in P_i (e.g., verbs of saying, thinking, dreaming, etc.). For example, know is factive, and presupposes the truth of its complement. Everyone knows that I am a Martian presupposes that I am a Martian. Dream on the other hand is counterfactual, and presupposes the falsehood of its complement. I dreamt that I was a Martian presupposes that I'm not a Martian. Morgan has noticed sentences like: I dreamt that I was a Martian and that everyone knew that I was a Martian. If presuppositions were unstructured relative to P_i, this sentence would contain contradictory presuppositions, since dream presupposes that I'm not a Martian and know presupposes that I am a Martian. Morgan takes this to show that presuppositions cannot be unstructured relative to P_i; rather they must be associated with certain verbs in P_i. Since know is embedded in the complement of dream, the presupposition of know is only assumed to be true of the world of my dream. However, the presupposition of dream is assumed to be true of the world of the speaker. There is no contradiction since the presuppositions are true of different possible worlds. The notation given above, which represents the traditional position that presuppositions are unstructured with respect to nonpresupposed elements of meaning, would appear to be false on the basis of Morgan's argument. However, we will keep the above notation throughout the remainder of this work, since the consequences of Morgan's observations are not well-understood at present and since his examples are not directly relevant to the subsequent discussion.

The notion of 'topic' is an ancient one in the history of grammatical investigation. Grammarians have long recognized that sentences have special devices for indicating what is under discussion. Preposing of topics is common. For example, in John, Mary hates him, John is the topic, while in Mary, she hates John, Mary is the topic. Clearly, an adequate account of semantic representation must take account of this notion. As with presuppositions, it is usually assumed that topics are structurally independent of the other components of meaning, as our notation indicates. As we will discuss in section 3 below, this may well not be the case. However, throughout what follows, we will assume the traditional position in an attempt to minimize controversy.

'Focus' is another traditional notion in grammar. Halliday [24] describes the information focus as the constituent containing new, rather than assumed information. The information focus often has heavy stress. Thus, in JOHN washed the car yesterday, the speaker is assuming that the car was washed yesterday and telling the addressee that the person who did it was John. Again, it is usually assumed that the semantic content of the 'focus' is structurally independent of other components of meaning. Our notation reflects this traditional position, although, as in the case of topic and presupposition, nothing that we have to say depends crucially on the correctness of this position.

I will refer to the above theory of grammar as a "basic theory", simply for convenience and with no intention of suggesting that there is anything ontologically, psychologically, or conceptually 'basic'

about this theory. Most of the work in generative semantics since 1967 has assumed the framework of the basic theory. It should be noted that the basic theory permits a variety of options that were assumed to be unavailable in previous theories. For example, it is not assumed that lexical insertion transformations apply in a block, with no intervening nonlexical transformations. The option that lexical and nonlexical transformations may be interspersed is left open.

As should be obvious from the above discussion, the basic theory does not assume any notion of 'directionality of mapping' from phonetics to semantics or semantics to phonetics.[2] Some writers on transformational grammar have, however, used locutions that might mislead readers into believing that they assume some notion of directionality. For example, Chomsky ([10]) remarks that "... properties of surface structure play a distinctive role in semantic interpretation." However, as Chomsky points out a number of times in that work, the notion of directionality in a derivation is meaningless, so that Chomsky's locution must be taken as having the same significance as "Semantic representation plays a distinctive role in determining properties of surface structure" and nothing more. Both statements would have exactly as much significance as the more neutral locution "Semantic representation and surface structure are related by a system of rules." The basic theory allows for a notion of transformational cycle in the sense of Aspects, so that a sequence of cyclical rules applies "from the bottom up", first to the lowermost S's, then to the next highest, etc. We assume that the cyclical transformations start applying with P_k and finish applying (to the highest S) at P_ℓ where k is less than ℓ. We will say in this case that the cycle applies 'upward toward the surface structure'(though, of course, we could just as well say that it applies 'downward toward the semantic representation', since directionality has no significance).

It should be noted that a transformational cycle defines an 'orientation' on a derivation, and readers should be cautioned from confusing the notion 'cyclical orientation of a derivation' with the notion 'directionality of a derivation'. The former is a real and quite important notion; the latter is meaningless. To say that a cycle is oriented 'upward-toward-the-surface' is the same as to say that it is oriented 'downward-toward-the-semantics', and such terminology makes no claim about where a derivation 'begins'. Most theories of transformational grammar that have been seriously entertained have assumed a cyclical orientation that is upward-toward-the-surface. However, it is possible to envision a theory with an upward-toward-the-semantics cyclic orientation. Moreover, it is possible to imagine theories with more than one cyclic orientation. Consider a sequence of phrase-markers $P_1,...,P_i, ...,P_n$. One could imagine a theory such that $P_1,...,P_i$ had an upward-toward-the-semantics cyclic orientation and $P_i,...,P_n$ had an upward-toward-the-surface orientation, or vice-versa.

The basic theory does not necessarily include a level of 'deep structure', and the question as to whether such a level exists is an empirical question in the basic theory. We assume that the notion of 'deep structure' is defined in the following way. (i) Assume that all lexical insertion rules apply in a block. (ii) Assume that all upwards-toward-the-surface cyclic rules follow all lexical insertion rules. We define the output of the last lexical insertion rule as 'deep structure'.

```
          all lexical              all upward-toward-surface
           insertion                      cyclic rules
                          ┌──────────┐      ┌──────────────────┐
P₁, . . . . . .│. . . . . .│ Pᵢ  . ·. . . . . . . . . . . . .│. . . Pₙ
                          ↑                                              ↑
                        deep                                        surface
                      structure                                    structure
```
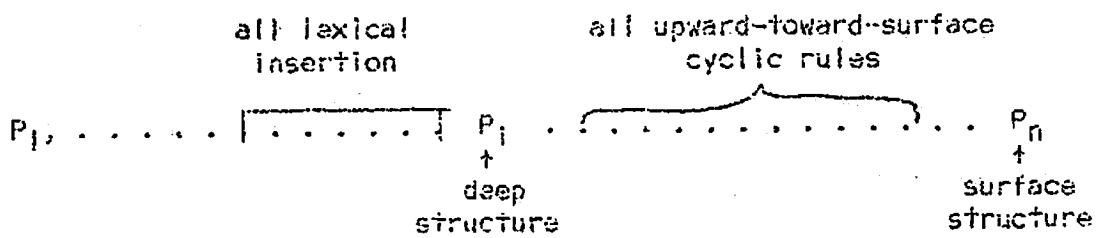
In the following section, I will discuss an example of a global deriva-
tional constraint and the evidence it provides for the question of whether
'deep structure' in this sense exists.

2. A Derivational Constraint Involving Quantifiers

Let us consider in detail an example of a global derivational constraint.[3]

(1) Many men read few books.
(2) Few books are read by many men.

As is obvious, (1) and (2) are not synonymous.

Sentences like (1) and (2) were brought up in a discussion by Partee
([55]) with regard to certain inadequacies of a proposal made in Lakoff ([40])
for the derivation of quantifiers from predicates of higher sentences. That
proposal was suggested by the observation that sentences like "Many men left"
are synonymous with those like the archaic "The men who left were many". It
was proposed that sentences like the former were derived from structures un-
derlying sentences like the latter, with "many" as an adjective, which is then
lowered. Under such a proposal, underlying structures like (3) and (4) would
be generated.[4]

(3)
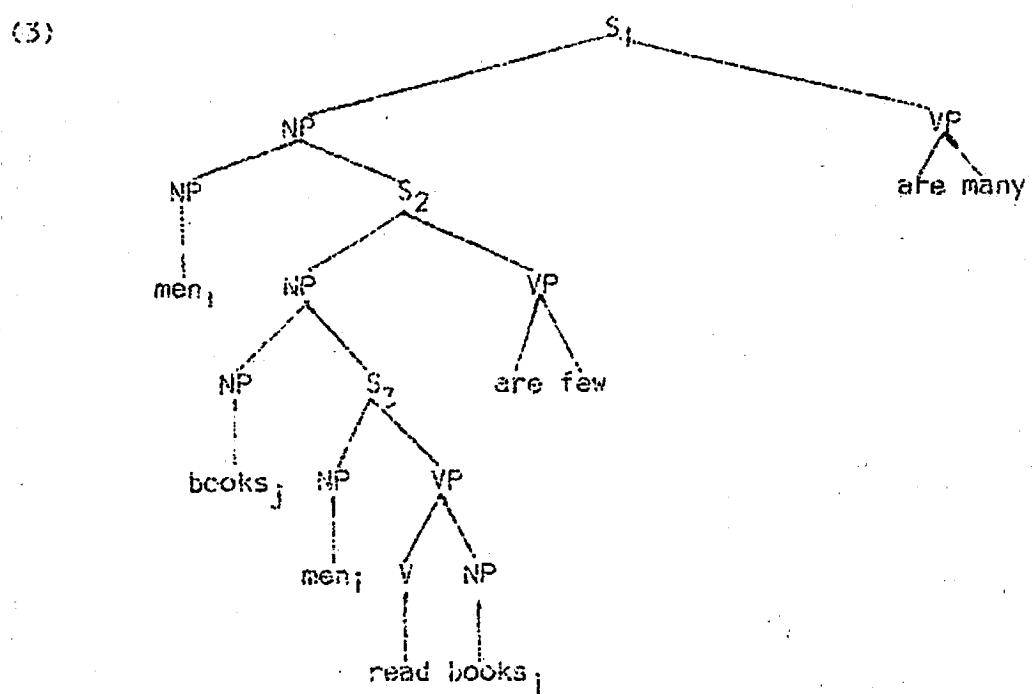
(3') Many are the men who read few books. There are many men who read few
      books.

(4)

S_1

NP — VP (are few)

NP (books$_j$) — S_2

NP — VP (are many)

NP (men$_i$) — S_3

NP (men$_i$) — VP

V (read) — NP (books$_j$)
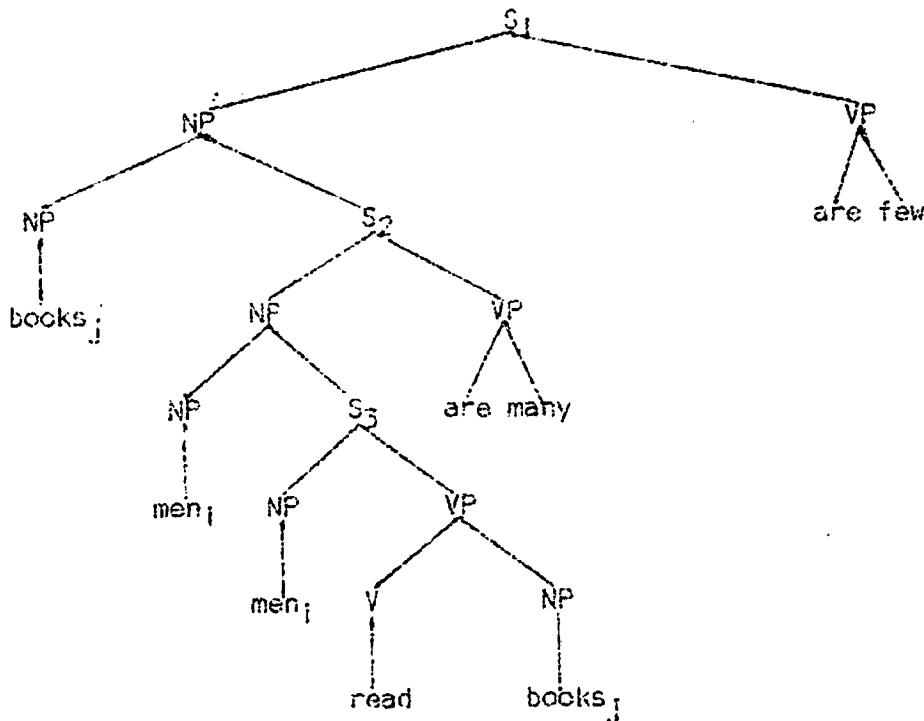
(4') Few are the books that many men read. There are few books that many men read.

In (3), a cyclical rule of quantifier-lowering will apply on the S_2-cycle, yielding men$_i$ read few books. The same cyclic rule will apply on the S_1-cycle, lowering many onto men$_i$ and yielding (1), many men read few books. In (4) let us suppose that the passive, a cyclical rule, applies on S_3, before any quantifier-lowering takes place: this will give us books$_j$ are read by men$_i$. On the S_2-cycle, many will be lowered onto men$_i$, yielding books$_j$ are read by many men. On the S_1-cycle, few will be lowered onto books$_j$ giving (2), few books are read by many men.

These derivations work as they should and account for the synonymy of (1) with (3') and (2) with (4'). However, if nothing more is said, this proposal will yield incorrect results. For example, if the passive applies to S_3 in (3), we will get books$_j$ are read by men$_i$, and then quantifier-lowering on the S_2- and S_1- cycles will yield (2), few books are read by many men. But if (2) were derived from (3) in this fashion, it should be synonymous with (3'), many are the men who read few books. This is false, and such a derivation must be blocked. Similarly, if the passive were not to apply to S_3 in (4), the application of quantifier-lowering on the S_2- and S_1-cycles would yield (1), many men read few books, again a mistake, since (1) does not have the meaning of (4'), few are the books that many men read.

Such a proposal would work in the first two cases, but would also predict the occurrence of two derivations that do not occur, at least for the majority of English speakers. If one inspects (1) - (4), one notices that the correct derivations have the property that the 'higher' quantifiers in (3) and (4) are the leftmost quantifiers in (1) and (2) respectively. Thus, we might propose a derivational constraint that would say something like: if one quantifier commands another in underlying structure (or rather, P_1), then that quantifier must be leftmost in surface structure. Such a constraint as it stands would be too strong. Consider cases like (5).

(5) The books that many men read are few (in number).

(5) would have an underlying structure like (4), where few is the higher quan-
tifier; however, few is to the right of many in surface structure. Thus cases
like (5) would have to be ruled out of any such derivational constraint. If
one inspects (2) and (5), one sees that they differ in the following way. In
(5), few commands many, but many, being in a relative clause, does not command
few: that is, few is higher in the tree than many in (5), just as it is in
the underlying structure of (4). In other words, (5) preserves the asymmetric
command-relationship between the quantifiers that occurs in (4). In (2), how-
ever, this is not the case. In (2), neither few nor many is in a subordinate
clause, and so each commands the other and the command-relationship is sym-
metric. Thus the asymmetry of the command-relationship in the underlying
structure, where few commands many but many does not command few, is lost in
(2). It is exactly these cases where the quantifier that was higher in under-
lying structure must be leftmost in surface structure. Where the asymmetric
command-relationship is lost it must be supplanted by a precede-relationship,
which is necessarily asymmetric.

Such a derivational constraint may be stated as follows:

(6) Let C_1 = Q^1 commands Q^2

C_2 = Q^2 commands Q^1

C_3 = Q^1 precedes Q^2 '/' means 'meets condition'

Constraint I: $P_1/C_1 \supset (P_n/C_2 \supset P_n/C_3)$

Constraint I states that if two quantifiers Q^1 and Q^2 occur in underlying
structure P_1, such that P_1 meets condition C_1, then if the corresponding sur-
face structure P_n meets condition C_2, that surface structure P_n must also meet
condition C_3. In short, if an underlying asymmetric command-relationship
breaks down in surface structure, a precede-relationship takes over. Con-
straint I is a well-formedness constraint on derivations. Any derivation not
meeting it will be blocked. Thus, the derivations (3) → (1) and (4) → (2) will
be well-formed, but (3) → (2) and (4) → (1) will be blocked.

It is important to note that the fact that one of the two quantifiers is
in subject position in the sentences we have discussed so far is simply an
accident of the data we happened to have looked at. The difference in the
interpretation of quantifiers has nothing whatever to do with the fact that
in these examples one quantifier is inside the VP while the other is outside
the VP. Only the left-to-right order within the clause matters.

(7) John talked to few girls about many problems.

(8) John talked about many problems to few girls.

These sentences differ in interpretation just as do (1) and (2), that is the
leftmost quantifier is understood as the highest in each sentence, though both
quantifiers are inside the VP.

Although (1) and (2) are cases where the asymmetry of the underlying
command-relationship disappears in surface structure, it happens to be the
case in (1) and (2) that condition C_1, which holds in underlying structure
continues to hold in surface structure: that is, Q^1 continues to command Q^2.

We might ask if there exist any cases where this does not happen, that is, where Q^1 commands Q^2 in underlying structure, but Q^1 does not command Q^2 in surface structure. A natural place to look for such cases, and perhaps the only one, is in sentences containing complement constructions. Let us begin by considering sentences like (9).

(9) Sam claimed that John had dated _few_ girls.

(9) is open to both of the readings (10) and (11), though (10) is preferable.

(10) Sam claimed that the girls who John had dated were few (in number).

(11) The girls who Sam claimed that John had dated were few (in number).

(10) and (11) would have underlying structures like (10') and (11') respectively.

(10')
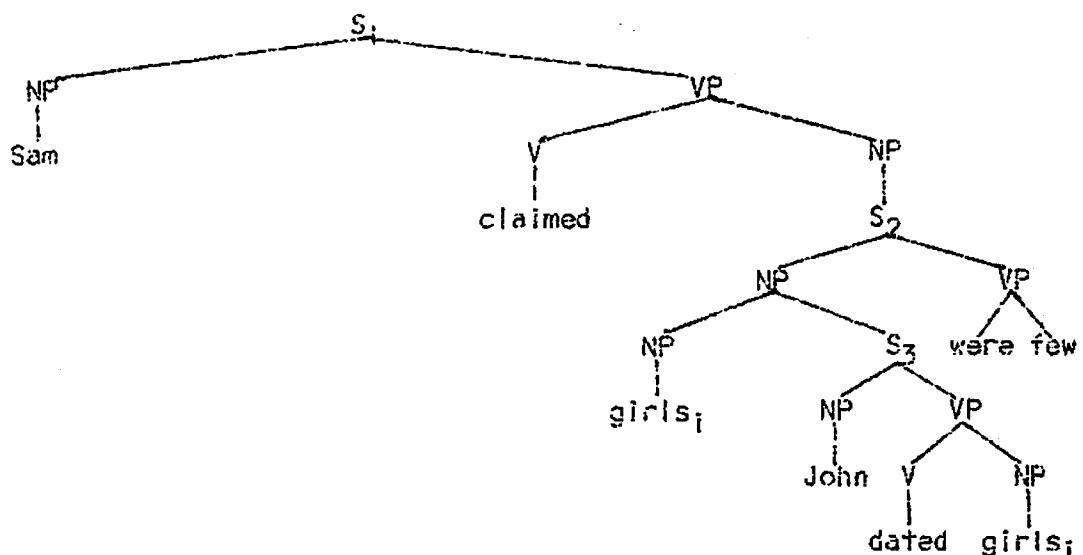
(11')
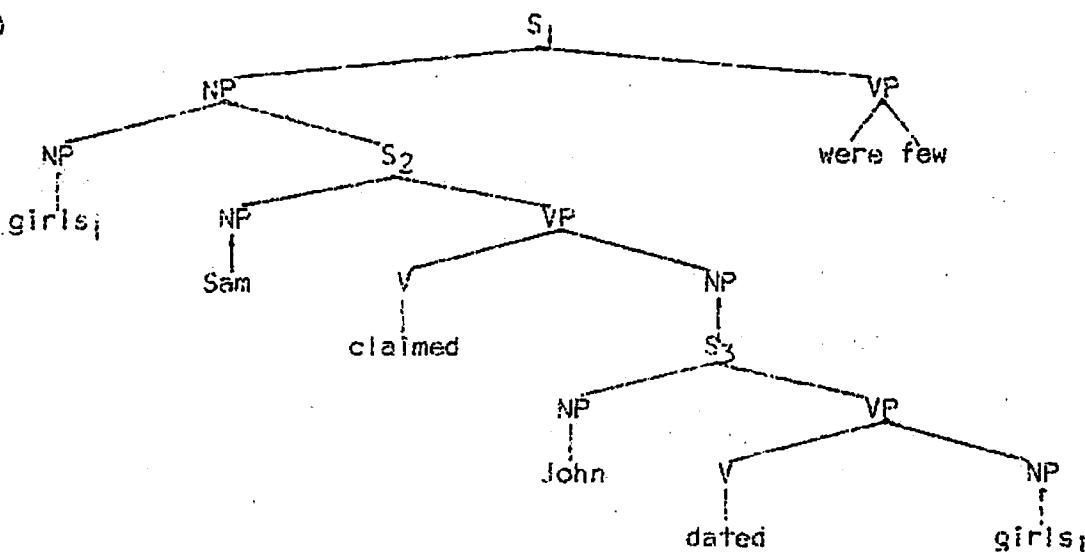
In each case quantifier-lowering will move _few_ down to $girls_i$. In (10),

girls$_i$ occurs in the S immediately below few; in (11), girls$_i$ occurs two sentences down from few. We are now in a position to test the conjecture that one quantifier may cease to command another in the course of a derivation. Consider (12), where few commands many.

(12)

```
                              S₁
              NP                              VP
        NP          S₂                      were few
      girls₁    NP       VP
             NP      S₃    were many
         reporters_j  NP       VP
               reporters_j  V         NP
                         claimed      S₄
                                  NP        VP
                                 John     V       NP
                                        dated   girls₁
```
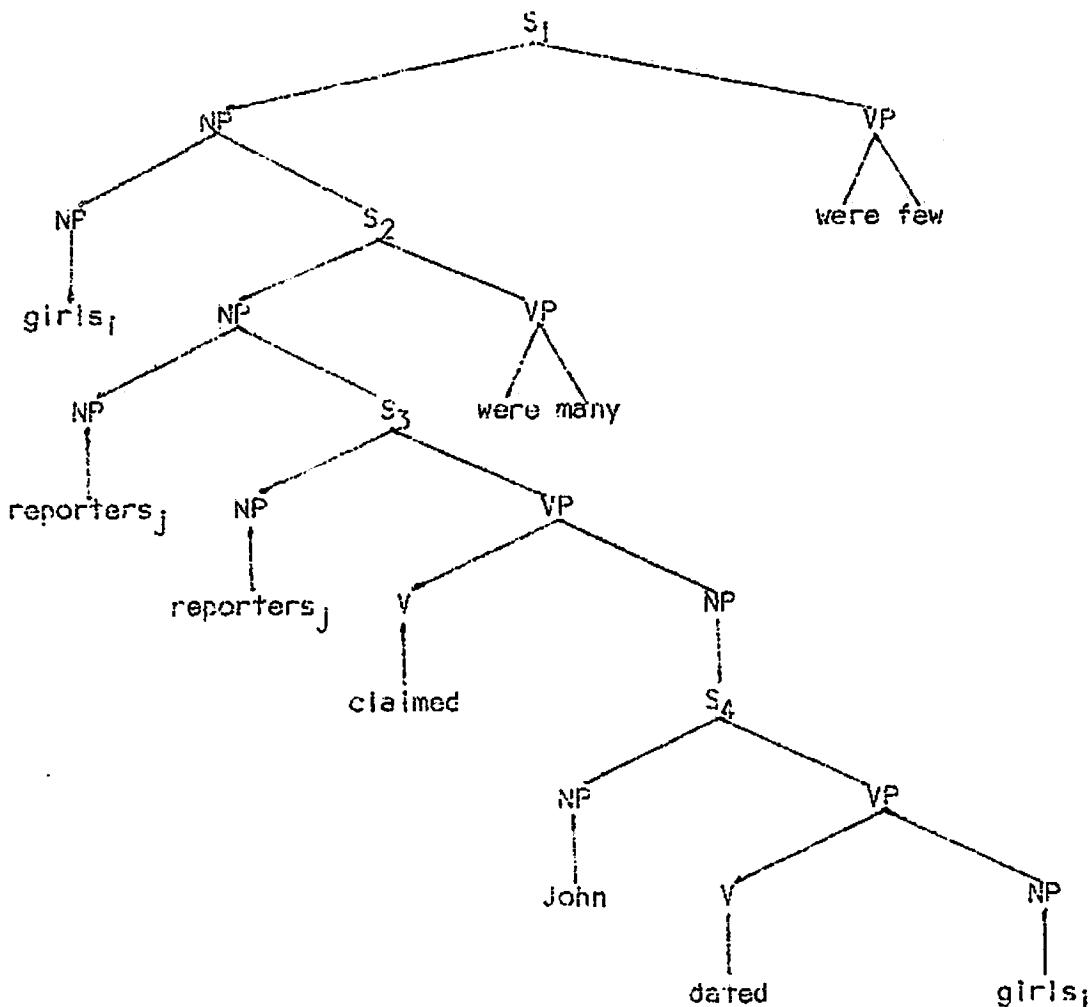
(12) would have the meaning of (12').

(12') Few were the girls who many reporters claimed John dated.

If we allow quantifier-lowering to apply freely to (12), many will be lowered to reporters$_j$ on the S$_2$-cycle, yielding many reporters claimed that John dated girls$_i$. The derived structure will now look just like (11), except that it will have the noun phrase many reporters instead of Sam. As in (11), quantifier-lowering will lower few onto girls$_i$ yielding (13).

(13) Many reporters claimed that John dated few girls.

In (13), few is in a subordinate clause and does not command many. Thus we have a case where few commands many in underlying structure, but not in surface structure. Note, however, that (13) does not have the meaning of (12'). It has the reading of (14),

Constraint 2. In (17) and (18), not commands many in underlying structure, just as in surface structure (16). In (19), many commands not in underlying structure, but many does not command not in surface structure (16). Thus we can generalize Constraint 2 in the following way. Let L stand for a 'logical predicate', either Q or NEG.

(20)   Let $C_i = L^1$ commands L^2 (L = Q or NEG)

Constraint 2': $(P_i/C_i) \supset (P_n/C_i)$

Conditions of this sort suggest that quantifiers and negatives may form a natural semantic class of predicates. This seems to be confirmed by the fact that Constraint 1 can be generalized in the same fashion, at least for certain dialects of English. Consider, for example, the following sentences discussed by Jackendoff ([32], [33]).

(21)   Not many arrows hit the target.

(22)   Many arrows didn't hit the target.

(23)   The target wasn't hit by many arrows.

Jackendoff reports that in his speech (23) is synonymous with (21), but not (22). I and many other speakers find that (23) has both readings, but that the (22) reading is 'weaker': that is, (23) is less acceptable with the (22) reading. However, there are a number of speakers whose dialect displays the facts reported on by Jackendoff, and in the remainder of this discussion we will be concerned with the facts of that dialect.

Assuming the framework discussed above, (21) and (22) would have underlying structures basically like (24) and (25).

(24) $[_S$ not $[_S$ arrows$_i$ $[_S$ arrows$_i$ hit the target $]$ were many $]]$

(24') The arrows that hit the target were not many.

(25) $[_S$ arrows$_i$ $[_S$ not $[_S$ arrows$_i$ hit the target $]]$ were many $]$

(25') The arrows that didn't hit the target were many.

If Constraint 1 is generalized to include 'logical predicates', both negatives and quantifiers, then the facts of (21) - (23) will automatically be handled by the new Constraint 1', given the underlying structures of (24) and (25) and the rule of quantifier-lowering. Constraint 1' would be stated as (26).

(26)   Let $C_1 = L^1$ commands L^2

         $C_2 = L^2$ commands L^1

         $C_3 = L^1$ precedes L^2 (L = Q or NEG)

   Constraint 1': $P_i/C_i \supset (P_n/C_2 \supset P_n/C_3)$

Any derivation not meeting this condition will be ill-formed.[5]

(21) and (22) work as expected. Take (21): not (L^1) commands many (L^2) in underlying structure (24), many commands not in surface structure (21), and not precedes many in surface structure. So (24) → (21) meets Constraint 1'. Take (22): many (L^1) commands not (L^2) in underlying structure (25) and in surface structure, not commands many in surface structure, and many precedes not in surface structure. So (25) → (22) meets Constraint 1'. Now consider (23), which is the interesting case in this dialect. If one allows the passive transformation to apply to the innermost S of (24) and (25), and then allows quantifier-lowering to apply, both (24) and (25) will yield (23). First consider the derivation (24) → (23). Not (L^1) commands many (L^2) in underlying structure (24) and in surface structure, many commands not in surface structure, many commands not in surface structure (21), and not precedes many in surface structure. Thus, the derivation (24) → (23) meets Constraint 1'. Now consider the derivation (25) → (23). Many (L^1) commands not (L^2) in underlying structure and in surface structure, not commands many in surface structure (23), but many (L^1) does not precede not (L^2) in surface structure (23). Therefore, the derivation (25) → (23) does not meet Constraint 1', and so the derivation is blocked. This accounts for the fact that (23) is not synonymous to (21) and that there is no passive corresponding to (25) in this dialect.

It should also be noted that that part of Constraint 1' which says that L^2 must command L^1 in surface structure (P_n/C_2) if the precede-relationship (P_n/C_3) is to come into play, is necessary for the cases discussed. Consider, for example, sentence (25'), The arrows that didn't hit the target were many. In (25'), many (L^1) commands not (L^2) in P_1, and many also commands not in P_n. Thus, the if part of the conditional statement of Constraint 1' is not met, and the fact that not (L^2) precedes many (L^1) in surface structure (that is, that (P_n/C_3) does not hold) does not matter: since the if-condition is not met, the constraint holds and the sentence is grammatical with the reading of (25).

We have assumed so far that Constraints 1' and 2' mention the surface structure. But this is just an illusion which results from considering only simple sentences. Suppose, for example, that we consider complex sentences where deletion has taken place. Consider (27) and (28).

(27) Jane isn't liked by many men and Sally isn't liked by many men either.

(28) Jane isn't liked by many men and Sally isn't either.

Note that the sentence fragment Sally isn't either does not contain many in surface structure, but it receives the same interpretation as the full Sally isn't liked by many men either, and does not have the reading of There are many men who Sally isn't liked by. Constraint 1', as it is presently stated will not do the job, in that it mentions surface structure P_n rather than some earlier stage of the derivation prior to the deletion of liked by many men.

This raises a general problem about constraints like 1' and 2'. Since they only mention underlying structures P_1 and surface structures P_n, they leave open the possibility that such constraints might be violated at some intermediate stage of the derivation. My guess is that this will never be the case, and if so, then it should be possible to place much stronger constraints on derivations than 1' and 2' by requiring that all intermediate stages of a derivation P_i meet the constraint, not just the surface structure P_n. Using quantifiers, we can state such a stronger constraint as follows

(29) Let $C_1 = L^1$ commands L^2

$C_2 = L^2$ commands L^1

$C_3 = L_1$ precedes L_2

Constraint I": $P_i/C_1 \supset ((i)(P_i/C_2 \supset P_i/C_3))$

(29) will now automatically handle cases like (28), since it will hold at all points of the derivation up to the point where the deletion rule applies; after that point, it will hold vacuously. The reason is that many will not appear in any phrase-marker after the deletion takes place and so C_2 will not hold in such phrase-markers; and where C_2 does not hold, then C_3 need not hold.

However, this is still insufficient, since Constraint I" still requires that if C_2 holds in surface structure, then C_3 must hold in surface structure, as well as in earlier stages of a derivation. But there are late rules which make gross changes in derived structure and produce surface structures in which the constraints do not hold. Compare (30) and (31).

(30) Sarah Weinstein isn't fond of many boys.

(31) Fond of many boys, Sarah Weinstein isn't.

The rule of Y-movement as discussed in Postal, [57], will produce (31) from the structure underlying (30). (Needless to say, (31) is not grammatical in all American dialects. We are considering only those in which (31) is well-formed.) Note that (30) works exactly according to Constraint I": the reading in which many commands not in P_1 is blocked since many does not precede not in surface structure. But (31), where the surface order of not and many is reversed, shows the same range of blocked and permitted readings as (30). The outputs of Y-movement do not meet Constraint I", though earlier stages of the derivation do. Thus, it appears that Constraint I" has some cutoff point prior to the application of Y-movement. That is, there is in each derivation some 'shallow structure' P_a defined in some fixed way such that

(32) Constraint I"': $P_i/C_1 \supset ((i)(P_i/C_2 \supset P_i/C_3))$, where $i \leq a \leq n$

This raises the interesting question of exactly how the 'shallow structure' P_a is to be defined. One possibility is that P_a is the output of the cyclical rules. However, there aren't enough facts known at present to settle the issue for certain. Still, we can draw certain conclusions. Passive, a cyclic rule, must be capable of applying before P_a, since Constraint "' must apply to the output of passive. Y-movement must apply after P_a.

Let us now consider the constraints we have been discussing with respect to the process of lexical insertion. Let us constrain the basic theory so that some notion of 'deep structure' can be defined along the lines discussed above. Let it be required that all lexical insertion rules apply in a block and that all upward-toward-the-surface cyclic rules apply after lexical insertion. Since passive is a cyclic rule and since passive must be able to apply before P_a is reached, it follows that if there is such a 'deep structure' all lexical insertion must occur before P_a is reached. Thus, it is an empirical question as to whether such a notion of 'deep structure' is correct. If there exist lexical items that must be inserted after P_a, then such a notion of 'deep structure' would be untenable since there would exist upward-

toward-the-surface cyclic rules (e.g., passive) which could apply <u>before</u> some case of lexical insertion.

all lexical insertion   cyclic rules

$$P_1, \ldots \quad \boxed{\ldots \ldots} \quad P_i \quad \vdots \ldots \ldots \vdots \quad P_a \quad \ldots \ldots \ldots P_n$$

deep structure   passive   shallow structure cutoff point   Y-movement
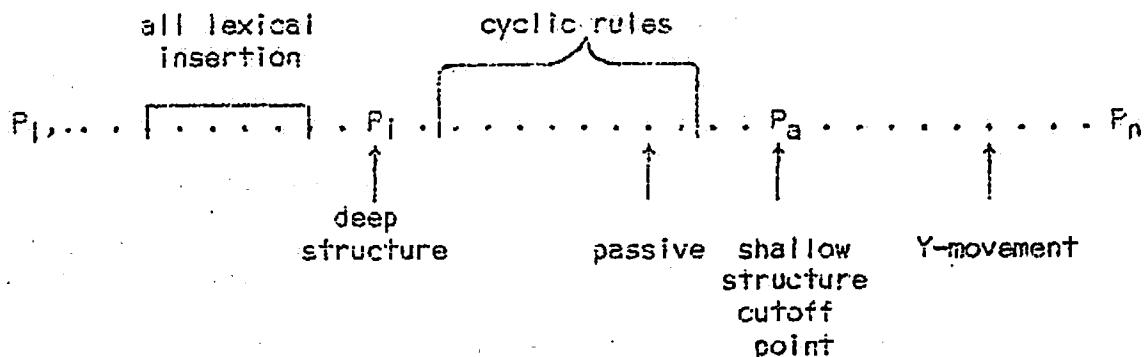
Empirical question: Does there exist a lexical item that must be inserted between P_a and P_n? If so, then 'deep structure' does not exist.

Let us pursue the question somewhat further. Consider the sentences

(32) a. I persuaded Bill to date many girls.
      b. I persuaded many girls to date Bill.

(33) a. I persuaded Bill not to date many girls.
      b. I persuaded many girls not to date Bill.

(34) a. I didn't persuade Bill to date many girls.
      b. I didn't persuade many girls to date Bill.

If we consider the meanings of these sentences, it should be clear that these cases work according to the two derivational constraints that we have stated thus far.

The difference in the occurrence of <u>not</u> is crucial in these examples. In (33), <u>not</u> in semantic representation would occur inside the complement of <u>persuade</u>, while in the semantic representation of (34) <u>not</u> will occur in the sentence above <u>persuade</u>. That is, in (33) <u>persuade</u> commands <u>not</u> in SR, while in (34), <u>not</u> commands <u>persuade</u>. This difference in the occurrence of <u>not</u> accounts for the fact that (33a) is unambiguous, while (32a) and (34a) are ambiguous. (32a) can mean either (35) or (36)

(35) There were many girls that I persuaded Bill to date.

(36) I persuaded Bill that the number of girls he dates should be large.

(34a) can mean either (37) or (38)

(37) There weren't many girls that I persuaded Bill to date.

(38) It is not the case that I persuaded Bill that the number of girls he dates should be large.

But (33a) can mean only (39)

(39) I persuaded Bill that the number of girls he dates should not be large.

The reason (39) is unambiguous is that since not precedes many in derived structure, not must command many in underlying structure (by Constraint I"). Since not originates inside the complement of persuade, and not must command many, many must also originate inside the complement of persuade. In (32a) and (34a), many may originate either inside the complement of persuade or from a sentence above persuade, which accounts for the ambiguity.

Now compare (33b) and (34b). In (33b) many both precedes and commands not in derived structure; therefore, many must command not in underlying structure. (33b) only has the reading of (40)

(40)   There were many girls that I persuaded not to date Bill.

In (34b), on the other hand, not precedes many in derived structure and so must command many in underlying structure. So (34b) can mean (41) but not (40).

(41)   There weren't many girls that I persuaded to date Bill.

Let us now consider the lexical item dissuade.

(42)   a.   I dissuaded Bill from dating many girls.
       b.   I dissuaded many girls from dating Bill.

In (42) the word not does not appear. The only overt negative element is the prefix dis-. Thus, the postlexical structure of (42) would not have the negative element inside the object complement of -suade, but in the same clause, as in (43).

(43)   a.   I NEG-suaded Bill from dating many girls.
       b.   I NEG-suaded many girls from dating Bill.

Moreover, the negative element would precede rather than follow the object of dissuade . In terms of precede and command relations, the postlexical structure of (42) would look like (34) rather than (33). Suppose P_3 were postlexical, that is, suppose that the command relationship between not and many in semantic representation were predictable from the precede relationship at some point in the derivation after the insertion of all lexical items. We would then predict that since NEG precedes many in (42), NEG must command many in the underlying structure of (42). That is, we would predict that the sentences of (42) would have the meanings of (34). But this is false. (42a) and (42b) have the meanings of (33a) and (33b).

Summary of Majority Dialect[6]

Persuade --	32a	means	35, 36
Persuade not --	33a	means	39
	33b	means	40
Not persuade --	34a	means	37, 38
	34b	means	41
Dissuade --	42a	means	39
	42b	means	40

Dissuade means persuade - NP - not rather than not - persuade - NP, and the constraints on the occurrence of quantifiers in derived structure reflect this meaning, and must be stated prelexically. The lexical item dissuade must be

inserted at a point in the derivations of (35) and (34) after Constraint I'''
has ceased to operate. Now recall that Constraint I''' must operate on the out-
put of the passive transformation. Consider (44).

(44) a. Many men weren't dissuaded from dating many girls.
      b. Not many men were dissuaded from dating many girls.
      c. I didn't dissuade many men from dating many girls.

(44) shows both the characteristics of (34) and (21) - (23). In (44), Con-
straint I''' must operate both after the passive transformation and before the
insertion of dissuade. Thus we have cases where an upward-toward-the-surface
cyclic rule must apply before the insertion of some lexical item. This shows
that any conception of 'deep structure' in which all lexical insertion takes
place before any upward-toward-the-surface cyclic rules apply is empirically
incorrect. It also shows that the passive transformation may apply to a verb
before the overt lexical representation of the verb is inserted, which means
that prelexical structures must look pretty much like postlexical structures.
In the case of dissuade, one might be tempted to try to avoid such a conclu-
sion with the suggestion that dissuade is derived by a relatively late trans-
formation from a structure containing the actual lexical item persuade. Under
such a proposal, dissuade would not be introduced by a rule of lexical inser-
tion, but rather by a rule which changes one actual lexical item to another.
Such a solution cannot be made general, however, since lexical items like
prohibit, prevent, keep, forbid, etc., which do not form pairs like persuade-
dissuade, work just like dissuade with respect to the properties we have dis-
cussed.

A particularly tempting escape route for those wishing to maintain a level
of 'deep structure' might be the claim that the lexical item dissuade is in-
serted precyclically, that dissuade requires a not in its complement sentence,
and that this not is deleted after shallow structure. Thus, the not would be
present at the time that the constraints shut off, and all of the above facts
would be accounted for. This proposal has some initial plausibility since
similar verbs in other languages often have a negative element that appears
overtly in its complement sentence. For example, in Latin we have "Dissuasī
Marcō nē īret" (I dissuaded Marcus from going), where nē, the morphological
alternant of nōn in this environment, occurs in the complement sentence.

Let us suppose for the moment that such a solution were possible. This
would mean that the complement sentence of dissuade would contain a not at the
level of shallow structure, but not at the level of surface structure. Now
consider the following sentences:

(45) I dissuaded Mary from marrying no one.

(46) *I persuaded Mary not to marry no one.

(47) *Mary didn't marry no one.

(48) I didn't persuade Mary to marry no one.

(46) and (47) are ungrammatical in standard English, and in all dialects if the
two negatives are both considered as underlying logical negatives (e.g., if
(47) has the reading It is not the case that Mary married no one). As is well-
known, this prohibition applies only for negatives in the same clause (cf. ex-
ample (48), where the negatives are in different clauses). The question arises

as to where in the grammar the No-double-negative (NDN) constraint is stated.
(a) If 'deep structure' exists, it could be stated there; (b) it could be
stated at shallow structure; (c) it could hold at all levels between deep
structure and shallow structure; or (d) it could hold only at surface struc-
ture.

Now consider (45). Under the above proposal for post-shallow-structure
deletion of not in dissuade-sentences, not would still be present at the
level of shallow structure. (In fact, it would be present at all points be-
tween deep structure and shallow structure, that is, all points in the deri-
vation where constraints 1 and 2 hold.) At the level of shallow structure,
(45) would have the form:

(45)   I dissuaded Mary from not marrying no one.

Thus under the above proposal, the No-double-negative (NDN) constraint could
not hold at the level of shallow structure, or at any previous point in the
derivation back to deep structure; if it did, (45) would be ruled out. Thus,
under the proposal for post-shallow-structure not-deletion, the NDN-constraint
could only be a constraint on surface structure, not on shallow structure.
Let us now take up the question of whether this is possible.

The following sentences are in accord with the NDN-constraint whether it
holds at shallow or surface structure, since these sentences have essentially
the same representation at both levels.

(49)   Max said that Sheila Weinstein was spurned by no one.

(50)   Max didn't say that Sheila Weinstein was spurned by no one.

(51)   *Max said that Sheila Weinstein wasn't spurned by no one.

Note that in the appropriate dialects, Y-movement can apply to sentences of
this form moving the participial phrase of the embedded clause.

(52)   Max said that Sheila Weinstein wasn't spurned by Harry.

(53)   Spurned by Harry, Max said that Sheila Weinstein wasn't.

In (53), Harry has been moved from a position where it was in the same clause
as n't to a position where it is in a higher clause. Now consider once more:

(51)   *Max said that Sheila Weinstein wasn't spurned by no one.

If the NDN-constraint holds for surface, not shallow, structures, then the
application of Y-movement to (51) would move no one out of the same clause as
n't and should make the NDN-constraint nonapplicable at surface structure.
Thus, (54) should be grammatical.

(54)   *Spurned by no one, Max said that Sheila Weinstein wasn't.

But (54) is just as bad as (51)--the NDN-constraint applies to both. As we
have seen, the NDN-constraint cannot apply to the surface structure of (54),
since no one is not in the same clause as n't, having been moved away by Y-
movement. Thus, in order to rule out (54), the NDN-constraint must apply be-
fore Y-movement; that is, it must apply at the level of shallow structure.

But this contradicts the post-shallow-structure not-deletion proposal, since under that proposal, (45), which is grammatical, would contain two negatives in the same clause at the level of shallow structure (cf. (45')). Thus, the post-shallow-structure not-deletion proposal is incorrect.

A similar proposal might say that prior to shallow structure, from replaced not with verbs like dissuade and prevent, while to replaced not with a verb like forbid, and that from and to 'acted like negatives' (whatever that might mean) with respect to constraint 1 and 2 at shallow structure and above. However, it is clear from sentences like (45) that from does not act like a negative with respect to the NON-constraint at shallow structure. Such a proposal would thus require from both to act like a negative and not to act like a negative, a contradiction. Neither of these routes provides an escape from the conclusion that dissuade must be inserted after shallow structure.

Another route by which one might attempt to avoid this conclusion would be by claiming that derivational constraints did not operate on the internal structure of lexical items, and that therefore dissuade could be inserted before the passive, preserving the notion of 'deep structure'. It would then not interact with the constraints. Such a claim would be false. Dissuade does interact with the constraints. As the summary of meaning-correspondences given above shows, dissuade acts just like persuade not, not like persuade. In particular, (42a) is unambiguous, just like (33a). It only has a reading with many originating inside the complement, as the constraints predict. If dissuade were impervious to the constraints, then we would expect it to act like persuade: in particular, we would expect (42a) to be ambiguous, just like (32a), where many may be interpreted as originating either inside or outside of the complement. But as we have seen, (42a) does not have the outside reading. John dissuaded Bill from dating many girls cannot mean Many were the girls who John dissuaded Bill from dating.

It should be noted that this argument does not depend on the details of Constraint 1"' being exactly correct. It would be surprising if further modifications did not have to be made. Nor does this argument depend on any prior assumption that semantic representation must be taken to be phrase-markers, though the discussion was taken up in that context. It only depends on the facts that persuade-not and not-persuade obey the general constraints on quantifiers and negatives, and that dissuade acts like persuade-not. Thus, in any version of transformational grammar there will have to be stated a general principle relating semantic representations of sentences containing quantifiers and negatives to the left-to-right order of those corresponding quantifiers and negatives in 'derived structure'. If the general principle is to be stated, the notion 'derived structure' will have to be defined as following the passive rule, but preceding the insertion of dissuade. Thus, in no non-ad-hoc transformational grammar which states this general principle will all lexical insertion precede all cyclical rules.

Let us sum up the argument.

(i)   Suppose "deep structure" is defined as a stage in a derivation which follows the application of all lexical insertion rules and precedes the application of any upward-toward-the-surface cyclic rules.

(ii) Evidence was given for derivational constraints 1 and 2, which relate semantic command-relationships to precede- and command- relationships at some level of 'derived structure'.

(iii) But the passive transformation must precede the insertion of lexical items such as dissuade, prohibit, prevent, keep, etc.

(iv) Since passive is an upward-toward-the-surface cyclic rule, (iii)

shows that the concept of 'deep structure'given in (i) cannot be maintained.

3. Further examples

The constraints discussed in Chapter 3 extend to more cases than just quantifiers and negatives. For example, the same constraints seem to define at least in part the limits of conjunction reduction.

(1) John claimed that he robbed the bank and he claimed that he shot Sam.

(2) John claimed that he robbed the bank and claimed that he shot Sam.

(3) John claimed that he robbed the bank and shot Sam.

In (1) and (2), it is understood that two claims have been made. In (3), it is understood that only one claim has been made. This suggest that (1) and (2) are derived from an underlying structure like (4), while (3) is derived from something like (5).

(4)

(5)
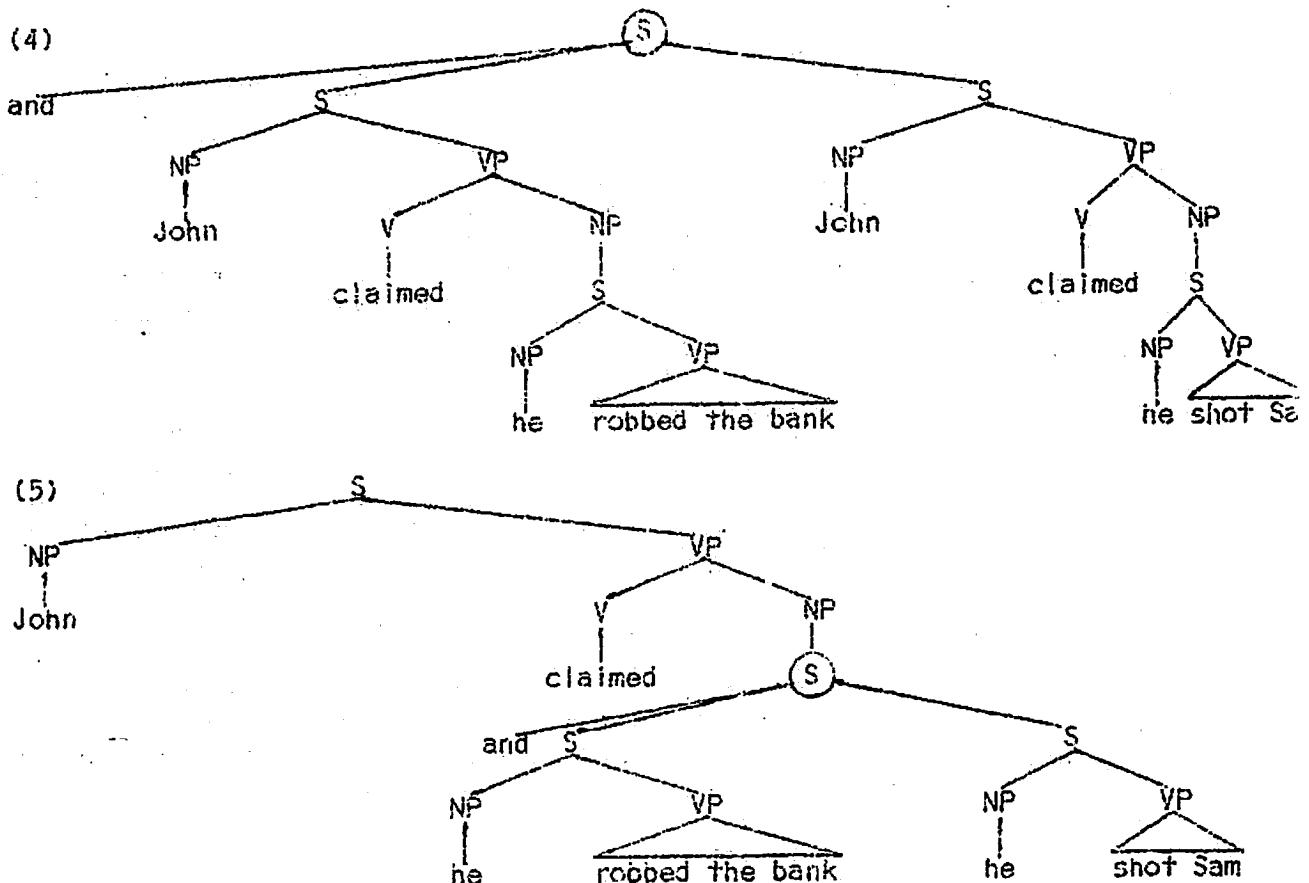
In (4), there are two instances of claim, while in (5) there is only one. The following question now arises: If conjunction reduction applies to (4) so that it yields (2), what is to keep it from applying further to yield (5)? If it applies freely to yield (5) from (4), we would get the incorrect result that (3) should be synonymous with (1).

Observe that in (4), and commands claim, but claim does not command and. In (5), the reverse is true: claim commands and, but not vice versa. To say

that (5) cannot be derived from (4) is to say that the asymmetric command relation between and and claim cannot be reversed in the course of a derivation. But that is the same as allowing constraint 2 of the previous section to hold for and and claim.

This constraint will hold for other predicates like claim as well, though this is not obvious and to discuss it here would take us far afield.

It should also be noted that, although constraint 2 holds for and and verbs like claim, constraint 1 does not. For example, consider

(6) a. John claimed something outrageous and something quite reasonable.
    b. John claimed something outrageous and he claimed something quite reasonable.

(6-b')

(6-a')
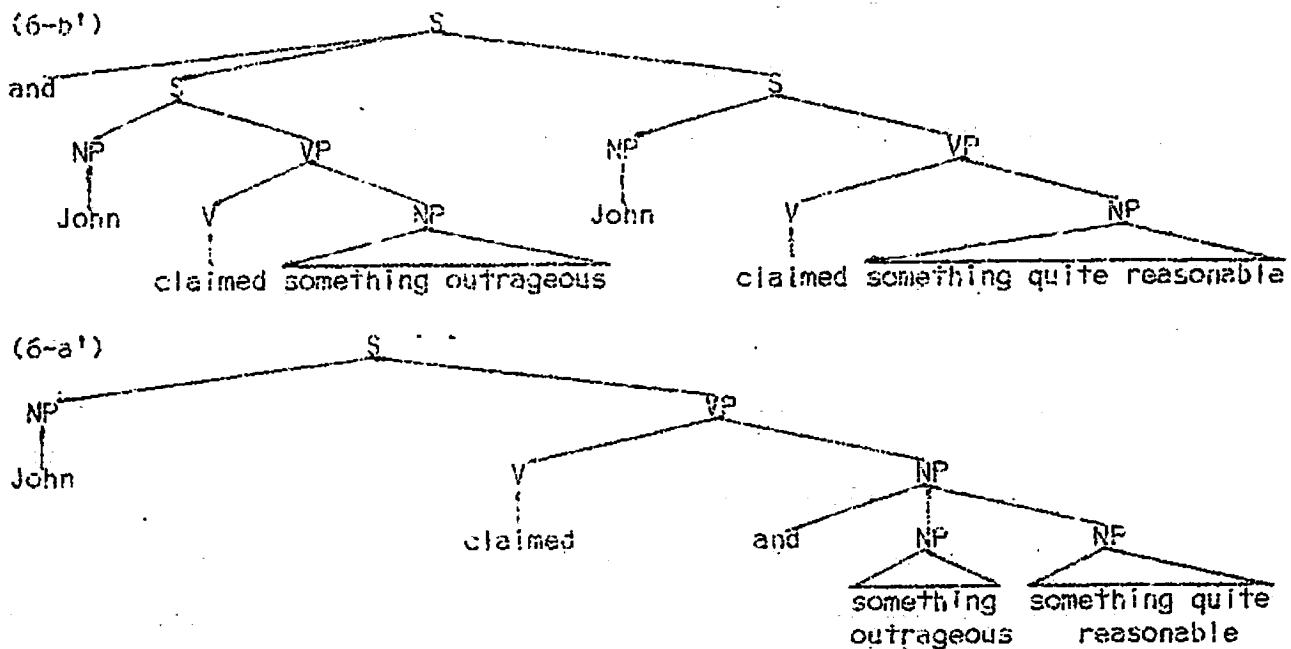
(6a) involves two claims, and so is synonymous with (6b). Let us assume then that (6a) is derived from (6b) by conjunction reduction. Prior to conjunction reduction, as in (6b'), and commands the two occurrences of claim. After conjunction reduction, claim precedes and. If claim and and obeyed constraint 1, this would be impossible. Exceptions to constraints, like this one, are not rare and are discussed below.

The fact that claim and and obey constraint 2, but not constraint 1, can be made the basis for an explanation of a rather remarkable minimal pair.

(7) a. John claimed that he robbed the bank and that Sam shot him.
    b. John claimed that he robbed the bank and Sam shot him.

In (7a), John is making two claims, while in (7b), he is making one. How can we account for this fact? Let us assume, as is usual, that the complementizer that is Chomsky-adjoined to the S of an object complement, as in (8a). Other alternatives are given in (8b and c); since they will yield exactly the same result, there is no need to choose among (8a, b, and c) for the sake of this argument. (8a) seems at present to be the least problematic alternative.

(8) a.   NP
          |
          S
         / \
      That   (S)
             / \
            /___\

b.   NP
      |
     (S)
     / \
  That /___\

c.   NP
       / \
    That   (S)
          / \
         /____\
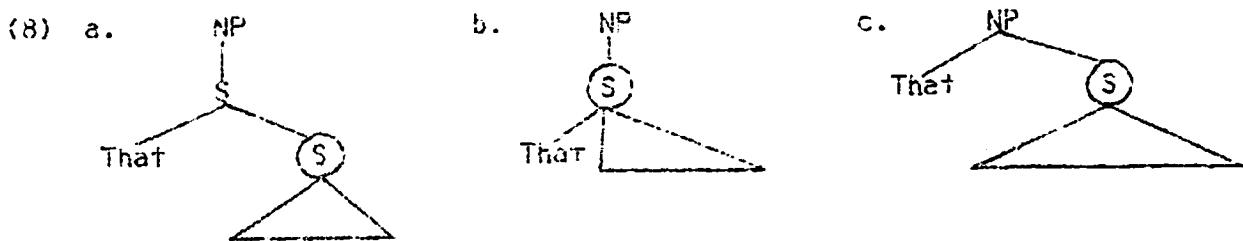
In order to explain the sentences in question, we need to show that the transformation inserting the complementizer _that_ inserts only one occurrence of that complementizer per noun phrase complement, even if the complement S is a conjunction. This can be shown readily, if one looks at sentences of the form:

(9) NP₁ and NP₂ are both correct.

'Both' can occur in sentences of this form if there are exactly two NP's conjoined in the subject. Thus, there are no sentences of the form (10) or (11),

(10) *NP₁ are both correct.

(11) *NP₁ and NP₂ and NP₃ and NP₄ are both correct.

Let NP₁ and NP₂ in (9) each contain a pair of conjoined sentences as its complement.

(12)   NP₁
                 |
                 S
               / | \
           and   S   S
              /____\ /____\
         Sam robbed Bill shot him
         the bank

                 NP₂
                  |
                  S
                / | \
            and   S   S
               /____\ /____\
          Sally got   Her mother spanked
          pregnant   her
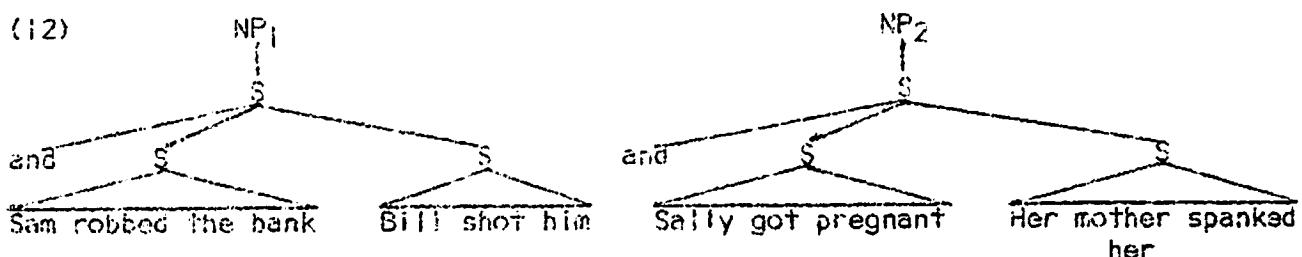
Observe the following facts:

(13) *That Sam robbed the bank and Bill shot him are both correct.

(14) That Sam robbed the bank and Bill shot him and that Sally got pregnant and her mother spanked her are both correct.

(15) *That Sam robbed the bank and that Bill shot him and that Sally got pregnant and that her mother spanked her are both correct.

These sentences indicate that the rule of complementizer placement may introduce at most one occurrence of _that_ for each noun phrase complement. (13) provides evidence that _that_ is not subject to conjunction reduction. Since the sentence (16) That Sam robbed the bank and that Bill shot him are both correct. is grammatical, and since (13) would result if conjunction reduction applied to _that_ it appears that conjunction reduction may not apply to _that_. Let us now return to:

(7) a. John claimed that he robbed the bank and that Sam shot him.
     b. John claimed that he robbed the bank and Sam shot him.

Since (7a) contains two occurrences of _that_, it must contain two noun phrase

complements, and since (7b) contains only one occurrence of <u>that</u>, it cannot contain two noun phrase complements.

(7a')

(7b')
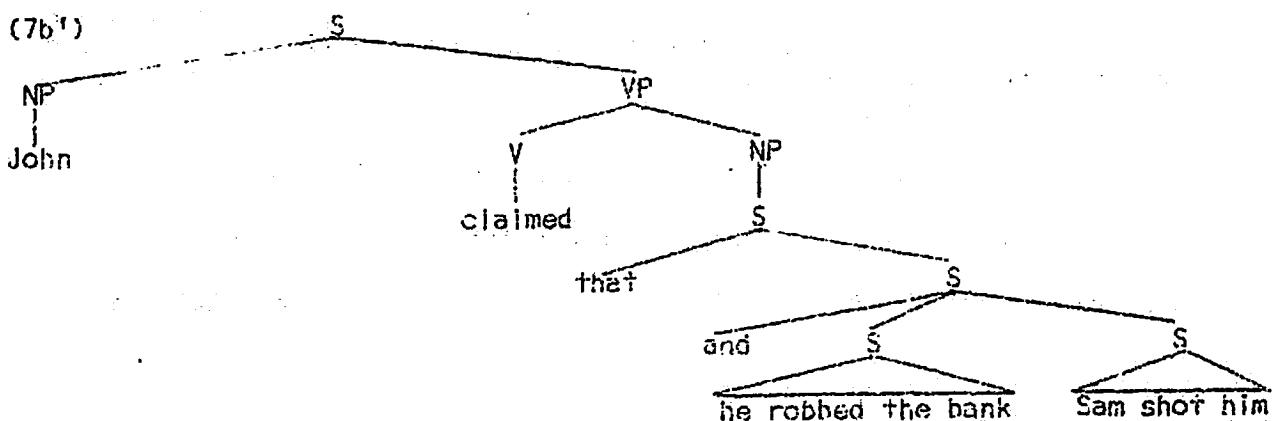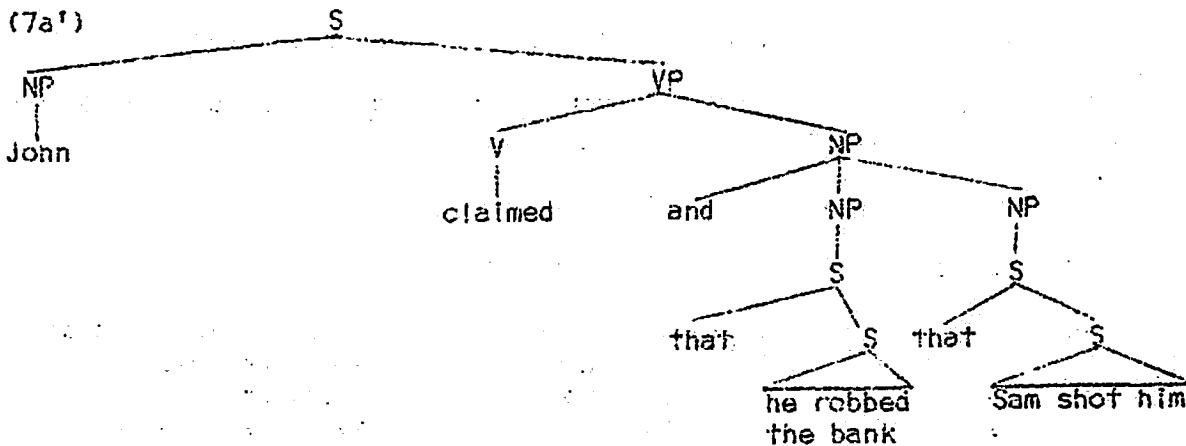
The essential feature of the analysis in (7a') is that (7a) is represented as containing a noun phrase conjunction, not a sentence conjunction. Thus, (7a) and (7b) would differ in structure in that the former would contain an NP-conjunction, whereas the latter would contain an S-conjunction. Given this analysis, the difference in meaning between (7a) and (7b) is an automatic consequence of the fact that <u>claim</u> and <u>and</u> obey constraint 2, but not constraint 1. Since (7a') does not contain an embedded S-conjunction, the only possible source would be that of (4), where <u>and</u> commands <u>claim</u> and two claims are indicated. Since constraint 1 does not apply to <u>claim</u> and <u>and</u>, such a derivation is possible. (7b') on the other hand, contains an embedded conjunction, and so has two conceivable sources: (4) and (5). But since <u>and</u> commands <u>claim</u> in (4) while <u>claim</u> commands <u>and</u> (but not vice versa) in (7b'), such a derivation is ruled out by the fact that constraint 2 holds for <u>and</u> and <u>claim</u>. Thus, the only possible derivation for (7b') would be from (5), which indicates only one claim. Thus, the difference in meaning between (7a) and (7b) is explained by the fact that <u>and</u> and <u>claim</u> obey constraint 2, but not constraint 1.

Constraint 2 works for <u>or</u> as well as for <u>and</u>, as the following examples show.

(17) (Either) John claimed that he robbed the bank or he claimed that he shot Sam.

(18) John either claimed that he robbed the bank or claimed that he shot Sam.

(19) John claimed that he either robbed the bank or shot Sam.

These sentences parallel (1) - (3).[7] Thus, it should be clear that constraint 2 applies to or.

Constraint 1 holds for or, though not for and. Consider the following examples, pointed out by R. Lakoff ([46]).

(20) Either you may answer the question, or not.

(21) You may either answer the question or not.

(20) and (21) are not synonymous, (20) says that there are two possibilities: Either it is the case that you are permitted to answer the question or it is not the case that you are permitted to answer the question. (20) exhausts the range of possibilities. (21), on the other hand, says that you are permitted the choice of answering or not answering. Lakoff points out that the difference in meaning can be accounted for, given Ross' analysis of modals as verbs that take complements. And, the difference between (20) and (21) is paralled by the difference between (22) and (23), where there is an overt verb with a sentential complement.

(22) Either you are permitted to answer the question, or not.

(23) You are permitted either to answer the question or not.

She proposes that (20) and (21) differ in structure as do (24) and (25).

(24)

25

(25)

```
                    S
          NP                 VP
          |             V         NP
         you           |          |
                      may         S
                        or     S          S
                            NP    VP    NEG    S
                            |                NP    VP
                           you  answer the    |
                                  question   you  answer the question
```
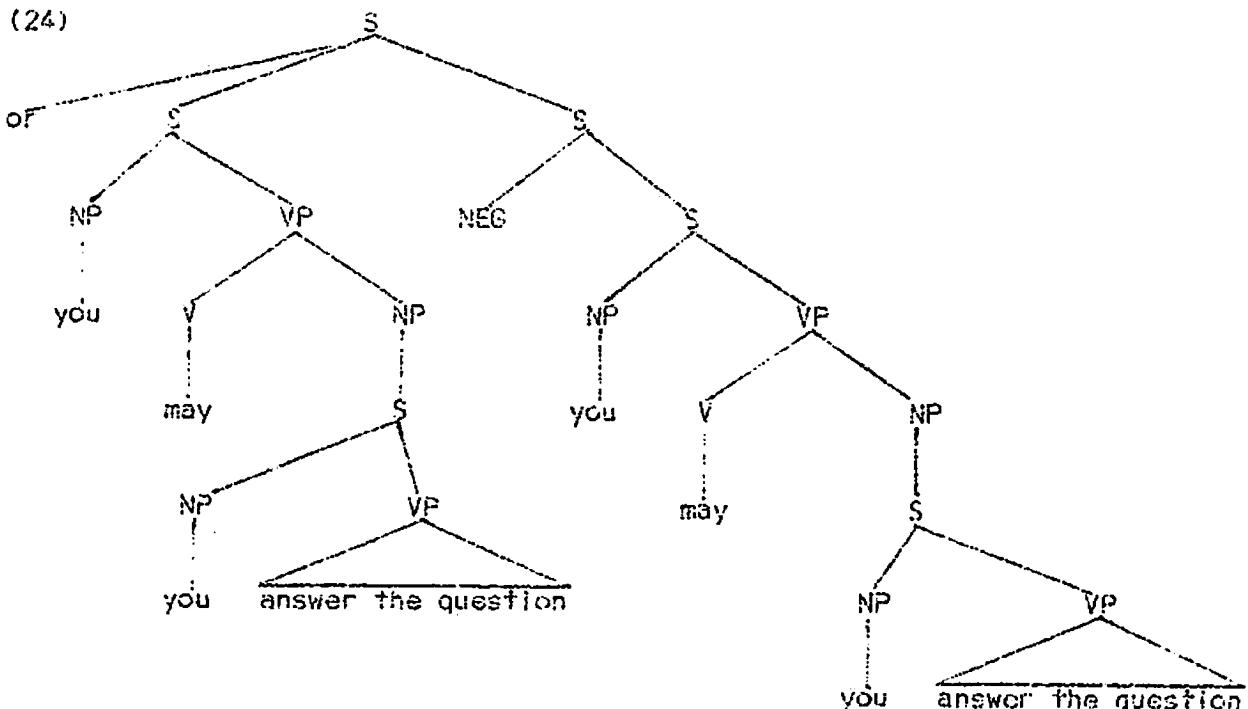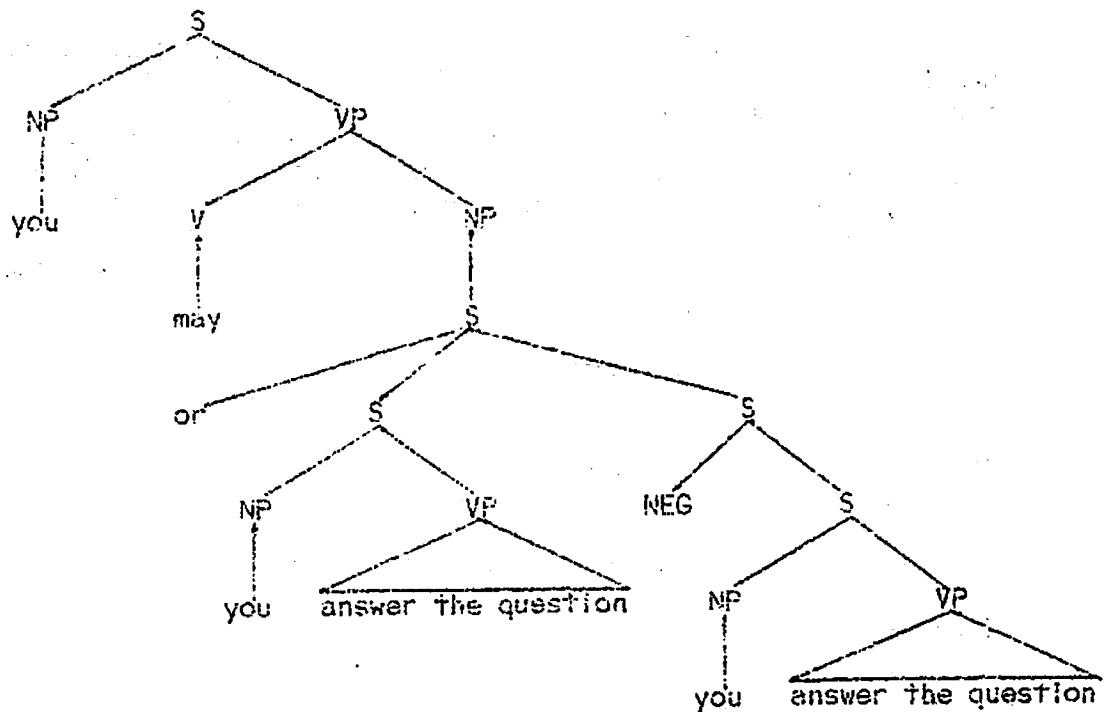
In (24), <u>or</u> commands <u>may</u> and not vice versa. In (25), <u>may</u> commands <u>or</u> and not
vice versa.

However, in derived structure, this asymmetry of command can be neutralized

(26)

```
                    S
          NP                 VP
          |             VP           ?
         you        either   VP    or    not
                           V      VP
                           |
                          may  answer the question
```

(27)

```
                    S
          NP                 VP
          |             V         VP
         you            |       VP        ?
                       may   either VP   or   not
                                 answer the question
```
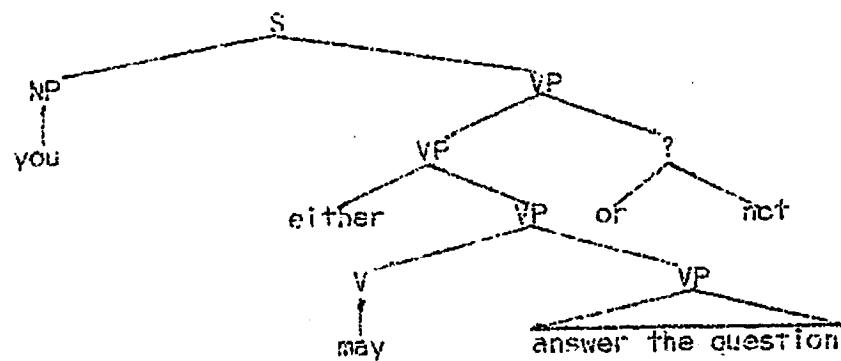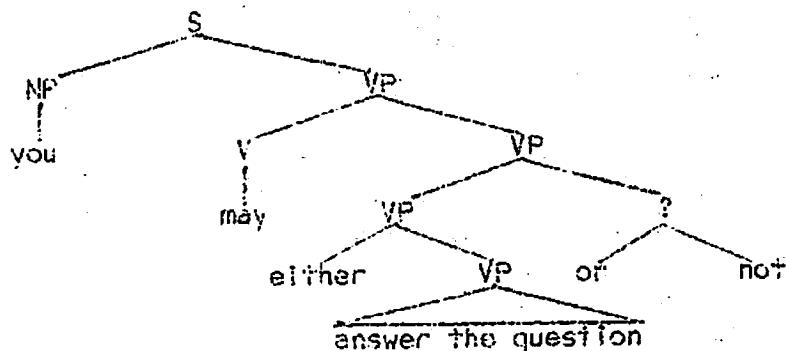
In (26) and (27), the S's have been pruned, and so <u>may</u> and <u>either</u> command each
other in both cases. The assymetric command relation of (24) and (25) is neu-
tralized. However, (26) <u>you (either) may answer the question or not</u> has the

meaning of (24), while (27) you may either answer the question or not has the meaning of (25). The generalization is that if either precedes may in derived structure, it must command may in underlying structure, and conversely. Thus, we have exactly the situation of constraint I in the previous section. And under Ross' analysis of modals the fact that the constraints work for modals follows from the fact that the constraints work for the corresponding verbs taking complements (e.g., permit).

Adverbs which are understood as predicates that take complements show the same property. Compare (28) and (29)

(28) a. It isn't obvious that John is a communist.
      b. It is obvious that John isn't a communist.

(29) a. John isn't obviously a communist.
      b. John obviously isn't a communist.

(28a and b) have underlying structures like:

(30) a.
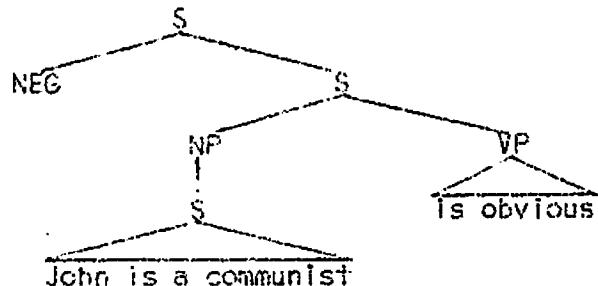
b.
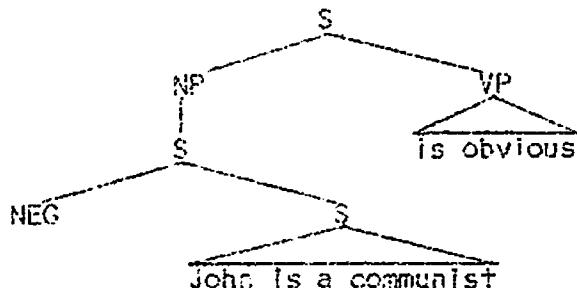
If we assume that obviously is derived from obvious by a rule of adverb lowering, and if it is assumed that obvious is one of those predicates taking a complement for which the constraints of section 2 hold, then it follows that (29a) should have the meaning of (28a), and that (29b) should have the meaning of (28b). Since not precedes obvious in (29a), it must command obvious in underlying structure (30a). Since obvious precedes not in (29b), it must command not in underlying structure, as in (30b).

The word only also obeys the constraints, though this follows automatically from the meaning of only. Only Bill means Bill and no one other than Bill. Since the latter expression contains a quantifier, we would expect the constraints of the previous section to hold. They do.

(31) a. John didn't hit Bill and no one else.
      b. Bill and no one else wasn't hit by John.

(32) a. John didn't hit only Bill.
      b. Only Bill wasn't hit by John.

The (a) sentences contain the reading It wasn't the case that there was no one other than Bill that John hit. The (b) sentences contain the reading There wasn't anyone other than Bill that John didn't hit. This is exactly what the constraints predict.

It should be noted again that the difference between subject and non-subject position in the clause has nothing to do with these constraints. Sentences like

(33) I talked to few girls about only these problems.

(34) I talked about only these problems to few girls.

show the predicted difference in meaning even though both few and only are in the VP in both examples.

It should be clear from these examples that at least some global derivational constraints do not serve just to limit the scope of application of a single rule, but rather can limit the applicability of a whole class of rules -- in this case, quantifier-lowering, conjunction-reduction, and adverb-lowering, together with rules like passive that interact with them. This result is similar to Ross' findings [65] that certain constraints hold for all movement rules of a certain form, not just for individual rules. Similar results were found by Postal [57] in his investigation of the crossover principle.

It has been known for some time that global derivational constraints have exceptions, as well as being subject to dialectal and idiolectal variation. Consider, for example, Ross' constraints on movement transformations. The coordinate structure constraint, if violated at any point in a derivation, yields ill-formed sentences, e.g., *Someone and John left, but I don't know who and John left. However, if the coordinate node is later deleted by some transformation, the sentence may be acceptable, e.g., Someone and John left, but I don't know who. Thus, the coordinate structure constraint applies throughout derivations, but with the above exception which takes precedence over the constraint. Ross' constraints are also subject to dialectal and idiolectal variation. For the majority of English speakers sentences like John didn't see the man who had stolen anything, and John didn't believe the claim that anyone left are ill-formed as predicted by Ross' complex NP constraint. However, for a great many speakers the latter sentence is grammatical and for some speakers even the former sentence is grammatical. So, it is clear that the global derivational constraints discovered by Ross are subject to such variation.

It should not be surprising that the global derivational constraints discussed above also have a range of exceptions and are subject to dialectal and idiolectal variation. For example, constraint 2 does not hold for the rule of not-transportation, the existence of which has been demonstrated by R. Lakoff ([45]). Thus, when L^i is a nonlogical predicate and L^j is a negative, the constraint does not hold. Similarly, constraint 1 does not hold when L^i is a negative and L^j is an auxiliary verb. Thus, John cannot go can mean it is not the case that John can go. Like Ross' constraints, constraints 1 and 2 admit of a great deal of dialectal and idiolectal variation. There are a great many people (more than one-third of the people I've asked) for whom constraint 1 does not hold for quantifiers and negatives. For such people, Few books were read by many men is ambiguous, as is The target wasn't hit by many arrows. Other sorts of differences also show up in the constraints. For example, for some people constraints 1 and 2 mention surface structure, not just shallow structure. Individuals with such constraints will find that I dissuaded Bill

from dating many girls can mean There are many girls that I dissuaded Bill from dating. Guy Carden [5] has pointed out that some speakers differ as to whether a constraint can hold at one level as opposed to holding throughout the grammar. McCawley [50] has shown that for some speakers the no-double-negative constraint holds only late in the grammar (at shallow structure). These speakers can get sentences like: John doesn't like Brahms and Bill doesn't like Brahms, but not Sam---he loves Brahms. (Before deletion, this would have the structure underlying *Sam doesn't not love Brahms.) However, many people find such sentences impossible. Carden has pointed out that this could be accounted for if the no-double-negative constraint held throughout the grammar down to shallow structure for such speakers. With other speakers, the same constraint seems to hold over other segments of the derivation (cf. Carden, [5]).

Such facts seem to show that constraints 1 and 2 are the norm from which individuals may vary. It is not clear at present how such variations from the norm can best be described, and it would seem that the basic theory will eventually have to be revised to account for such variations on basic constraints.

As the above discussion, as well as those of Postal ([57]) and Ross ([65]) have shown, global derivational constraints are pervasive in grammar. For example, interactions between transformational rules and presuppositions are handleable in a natural way using derivational constraints. Consider Kim Burt's observation (cf. Lakoff, [43]) that future will can optionally delete if it is presupposed that the speaker is sure that the event will happen. Suppose the rule of will-deletion is given by the local derivational constraint (C_1, C_2). Suppose tree-condition C_3 describes the presupposition in question. Then Burt's observation can be stated in the form:

(35) $(P_i/C_i \ \& \ P_{i+1}/C_2) \supset PR/C_3$

Presuppositions of coreferentiality can be treated in the same way. For example, consider the shallow structure constraint that states that a pronoun cannot both precede and command its antecedent. Suppose C_1 states that two NPs are coreferential, C_2 states that the pronoun precedes the antecedent, and C_3 states that the pronoun commands the antecedent. The constraint would then be of the form:

(36) $PR/C_1 \supset P_a/\sim(C_2 \ \& \ C_3)$ where P_a is a shallow structure

Another phenomenon that can be handled naturally in the basic theory is Halliday's ([24],[25]) notion of 'focus'. Halliday ([24]) describes focus in the following terms:

...the information unit, realized as the tone group, represents the speaker's organization of the discourse into message units: the information focus, realized as the location of the tonic, represents his organization of the components of each such unit such that at least one such component, that which is focal, is presented as not being derivable from the preceding discourse. If the information focus is unmarked (focus on the final lexical item), the nonfocal components are unspecified with regard to presupposition, so that the focal is merely cumulative in the message (hence the native speaker's characterization of it as 'emphatic'). If the information focus is marked (focus elsewhere than on the final lexical item), the speaker is treating the non-focal components as presupposed.

                                          --Halliday, [24], page 8

Halliday's account of focus has been adopted by Chomsky [10]. Assume for the moment that Halliday's account of focus as involving the location of stress on surface structure constituents were correct. Then the content of the sentence would be divided into a presupposed part and a part which is 'new' or focussed upon. Recall that derivational constraints enable one to trace the history of nodes throughout a derivation. This is necessary if one is to pick out which parts of P_i correspond to which surface structure constituents. Given such a notion, the correspondence between PR and FOC, which are part of the semantic representation, and the corresponding surface constituents can be stated by a global derivational constraint. Thus, the Halliday-Chomsky notion of 'focus' can be approached naturally within the basic theory. What is needed to make such representations precise is a precise definition of the notion 'semantic content corresponding to derived structure constituents'.

Of course, the Halliday-Chomsky account of focus is not quite correct. For example, Halliday and Chomsky assume that the constituent bearing main stress in the surface structure is the focus, and therefore that the lexical items in that constituent provide new rather than presupposed information. This is not in general the case. Consider (37).

(37) The TALL girl left.

Here the main stress is on TALL, which should be the focus according to Halliday and Chomsky, and should therefore be new, not given, information. However, in (37), TALL is understood as modifying the noun in the same way as the restrictive relative clause who was tall. Since restrictive relative clauses are presupposed, it follows that in (37), it is presupposed, not asserted, that the girl being spoken of was tall. Thus, the meaning of the lexical item TALL cannot be new information. Another possible candidate for focus might be the whole NP the tall girl. But none of the lexical content of this NP is new information, since it is presupposed that the individual under discussion exists, it is presupposed that that individual is a girl and it is presupposed that she is tall. None of this is new. In (37), it is presupposed that some girl left, and it is presupposed that some girl is tall. The new information is that the girl who was presupposed to have left is coreferential with the girl who was presupposed to be tall. The semantic content of the focus is an assertion of coreferentiality. In this very typical example of focus, the lexical semantic content of the surface structure constituent bearing main stress has nothing whatever to do with the semantic content of the focus.

So far, we have assumed that the Halliday-Chomsky account of focus in terms of surface structure constituents is basically correct. But this too is obviously mistaken. Consider (38).

(38) a. John looked up a girl who he had once met in Chicago.
      b. John looked a girl up who he had once met in Chicago.

(38) a'

```
b'                          S
        NP                  VP                        S
        |                                      ┌──────────────────
        John          V      NP    up      who he had once met in Chicago
                      |
                  looked a  girl
```
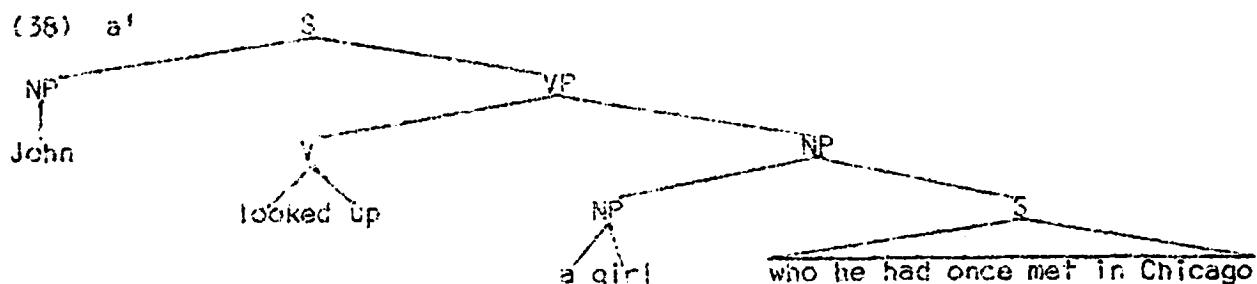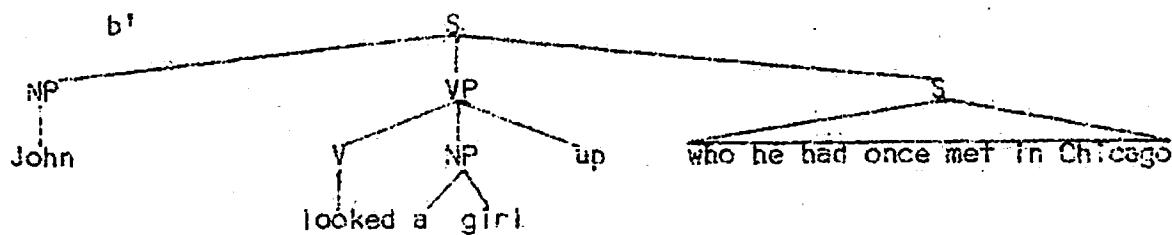
(a') and (b') have very different surface structure constituents. Assuming
that main stress falls on "Chicago" in both cases, Chomsky and Halliday would
predict that these sentences should be different in focus possibilities and in
corresponding presuppositions, and that therefore they should answer different
questions, and have quite different semantic representations. But it is clear
that they do not answer different questions and do not make different presup-
positions. Thus, it is clear that focus cannot be defined purely in terms of
surface structure constituents. Rather it seems that derived structure at
some earlier point in derivations is relevant.

These difficulties notwithstanding, it is clear that the phenomenon of
focus does involve global derivational constraints of some sort involving
derived structure. Halliday certainly deserves credit for the detailed work
he has done in this area, despite the limitations of working only with surface
structure. Generative semantics should provide a natural framework for con-
tinuing Halliday's line of research.

Another notion which can be handled naturally within the framework of
generative semantics is that of 'topic'. Klima has observed that sentences
like the following differ as to topic.

(39) a. It is easy to play sonatas on this violin.
      b. This violin is easy to play sonatas on.
      c. Sonatas are easy to play on this violin.

(a) is neutral with respect to topic. (b) requires "this violin" to be topic,
while (c) requires "sonatas". There are of course predicates in English which
relate topics to the things they are topics of. For example,

(40) a. My story is about this violin.
      b. That discussion concerned sonatas.

The predicates "be about" and "concern" are two-place relations, whose argu-
ments are a description of a proposition or discourse and the item which is
the topic of that proposition or discourse. Thus, the (a), (b), and (c) sen-
tences of (41) and (42) are synonymous with respect to topic as well as to the
rest of their content.

(41) a. Concerning sonatas, it is easy to play them on this violin.
      b. Concerning sonatas, they are easy to play on this violin.
      c. Sonatas are easy to play on this violin.

(42) a. About this violin, it is easy to play sonatas on it.
      b. About this violin, it is easy to play sonatas on.
      c. This violin is easy to play sonatas on.

If the topics mentioned in the clause containing "concern" or "about" differ
from the superficial subjects in these sentences, then there is a conflict of

topics and ill-formedness results, unless it is assumed that the sentence can have more than one topic.

(43) ?*About sonatas, this violin is easy to play them on.

(44) ?*Concerning this violin, sonatas are easy to play on it.

These are well-formed only for those speakers who admit more than one topic in such sentences.

These considerations would indicate that the notion 'topic' of a sentence is to be captured by a two-place relation having the meaning of "concerns" or "is about". If the set of presuppositions contains such a two-place predicate whose arguments are P_i and some NP, then it will be presupposed that that NP is the topic of P_i. Thus, the notion 'topic' may well turn out to be a special case of a presupposition. Since a semantic specification of "concerns" and "is about" is needed on independent grounds, it is possible that the special slot for TOP in semantic representation is unnecessary. Whether all cases of topic will turn out to be handleable in this way remains, of course, to be seen. Whichever turns out to be true, it is clear that the facts of (39) can be handled readily by derivational constraints. Assume that there is a rule which substitutes "this violin" and "sonatas" for "it" in (39). Let (C_1, C_2) describe this operation. Let C_3 describe the topic relation obtaining between P_i and the NP being substituted. Then the facts of (39) can be described by the following derivational constraint:

(45) $(P_i/C_1 \ \& \ P_{i+1}/C_2) \supset PR/C_3$

Note that this has exactly the form of the constraint discussed above describing the deletion of future will. Global derivational constraints linking transformations and presuppositions have this form. Of course, it may turn out to be the case that a more general characterization of the facts of (26) is possible, namely, that surface subjects in some class of sentences are always topics. In that case, there would be a derivational constraint linking presuppositions and surface structure. In any event the theory of generative semantics seems to provide an adequate framework for further study of the notion 'topic'.

Another sort of phenomenon amenable to treatment in the basic theory is lexical presupposition. As Fillmore ([15]) pointed out, "Leslie is a bachelor" presupposes that Leslie is male, adult, and human and asserts that he is unmarried. Similarly, "Sam assassinated Harry" presupposes that Harry is an important public figure and asserts that Sam killed him. Thus, lexical insertion transformations for "bachelor" and "assassinate" must be linked to presuppositional information. This is just the sort of linkage that we discussed above.

Thus far, most of the examples of global derivational constraints we have discussed mention semantic representations in some way. However, this is not true in general. For example, Ross' [65] constraints are purely syntactic. Another example of a purely syntactic global derivational constraint has been discussed by Harold King [38]. King noted that contraction of auxiliaries as in "John's tall", "The concert's at 5 o'clock", etc. cannot occur when a constituent immediately following the auxiliary to be contracted has been deleted. For example,

(46) a. Max is happier than Sam is these days.
      b. *Max is happier than Sam's these days.

(47) a. Rich though John is, I still like him.
     b. *Rich though John's, I still like him.

(48) a. The concert is this afternoon.
     b. The concert's this afternoon.
     c. Tell John that the concert is in the auditorium this afternoon.
     d. Tell John where the concert is this afternoon.
     e. *Tell John where the concert's this afternoon.
     f. Tell John that the concert's this afternoon.

In (e) the locative adverb has been moved from after is; In (f) no such movement has taken place.

Since contraction is an automatic consequence of an optional rule of stress-lowering, the general principle is that stress-lowering on an auxiliary cannot take place if at any point earlier in the derivation any rule has deleted a constituent immediately following the auxiliary. Let (C_1, C_2) be the rule of stress-lowering for Aux^j. Let $C_3 = X^i - Aux^j - A - X^k$ where A is any constituent and $C_4 = X^i - Aux^j - X^k$. The constraint is:

(49) $\sim(\exists X)\ (\exists Y)\ (P_x/C_1\ \&\ P_{x+1}/C_2\ \&\ P_y/C_3\ \&\ P_{y+1}/C_4)$

A wide range of examples of global derivational constraints not mentioning semantic representation will be discussed in (Lakoff, [43]). The redundancy rules discussed by R. Lakoff ([44]) are further examples of this sort. The exact nature and extent of global derivational constraints is, of course, to be determined through future investigation. It should be clear, however, that a wide variety of such constraints do exist. Thus, the basic theory, in its account of global derivational constraints, goes far beyond the standard theory and the Aspects theory, which included only a very limited variety of such constraints.

The basic theory is, of course, not obviously correct, and is open to challenge on empirical issues of all sorts. However, before comparing theories of grammar, one should first check to see that there are empirical differences between the theories. Suppose, for example, one were to counterpose to the basic theory, or generative semantics, an 'interpretive theory' of grammar. Suppose one were to construct such an interpretive theory in the following way. Take the class of sequences of phrase-markers $(P_1, \ldots, P_i, \ldots \ldots, P_n)$ where all lexical insertion rules occur in a block between P_1 and P_i and all upward-toward-the-surface cyclic rules apply after P_i. Call P_i 'deep structure'. Assume that $P_i \ldots P_n$ are limited only by local derivational constraints, except for those global constraints that define the cycle and rule ordering. Call $P_i, \ldots, P_n$ the 'syntactic part' of the derivation. Assume in addition that semantic representation $SR = (P_{-m}, PR, TOP, FOC, \ldots)$, where P_{-m} is a phrase-marker in some "semantically primitive" notation, as suggested by Chomsky ([10]) in his account of the 'standard theory'. Then a full derivation will be a sequence of phrase-markers:

$$P_{-m}, \ldots \ldots \ldots P_{-j}, \ldots \ldots \ldots \ldots P_0, P_1, \ldots \ldots \ldots \ldots P_i, \ldots \ldots \ldots \ldots P_n$$

Call $P_0 \ldots P_{-m}$ the semantic part of the derivation. Assume that the sequences of phrase-markers $P_0 \ldots P_{-j}$ are defined by local derivational constraints and global derivational constraints that do not mention any stage of the derivation after P_i, the 'deep structure'. Call these constraints 'deep structure interpretation rules'. Assume that the sequences of phrase-markers $P_{-j} \ldots P_{-m}$

are defined by local derivational constraints paired with global derivational constraints that may mention P_i and P_n as well as $P_{-i}, \ldots, P_{-n}$, FR, TOP, FOC, etc. Call these constraints 'surface structure interpretation rules'. (If such global constraints may mention not only P_i and P_n, but all points in between, then we will call them 'intermediate structure interpretation rules.') It should be clear that such an 'interpretive theory of grammar' is simply a restricted version of the basic theory. One can look at the deep structure interpretation rules as operations 'going from' P_i to P_{-i}, which are able to 'look back' only as far as P_i. And we could look at the surface structure interpretation rules as operations going from the 'output' of the deep structure interpretation rules P_{-i} to SR, while being able to 'look back' to P_i and P_n. However, as Chomsky ([10]) points out, the notion of 'directionality' is meaningless, and so there is no empirical difference between these operations and derivational constraints. Thus, such 'interpretive theories' are no different in empirical consequences than the basic theory, restricted in the above way, provided that such interpretive theories assume that semantic representations are of the same form as phrase-markers or are notational variants thereof. The only empirical differences are the ways in which the basic theory is assumed to be constrained, for example, the question as to whether levels like P_i, P_i, and P_{-i} exist. As we saw above, there is reason to believe that a level P_i does not exist, and no one has ever given any reasons for believing that a level P_{-i} exists, that is, that 'deep structure interpretation rules' are segregated off from 'surface structure interpretation rules'.

So far, no interpretive theory this explicit has been proposed. The only discussion of what might be called an interpretive theory which goes into any detail at all is given by Jackendoff ([33]), who discusses both surface and intermediate structure interpretation rules. However, Jackendoff explicitly refuses to discuss the nature of semantic representation and what the output of his interpretive rules is supposed to look like, so that it is impossible to determine whether his interpretive theory when completed by the addition of an account of semantic representation will be simply a restricted version of the theory of generative semantics. Jackendoff claims that semantic representations are not identical to syntactic representations ([33], page 2), but he does not discuss this claim. However, the empirical nature of the issue is clear: Are Jackendoff's interpretation rules simply notational variants of derivational constraints? (The only examples he gives do, in fact, do this.) Will the output of his rules be phrase-markers, or notational variants thereof? Of course, such questions are unanswerable in the absence of an account of the form of his rules and the form of their output.

Although Jackendoff does not give any characterization of the output of his envisioned interpretation rules, he does discuss a number of examples in terms of the vague notions 'sentence-scope' and 'VP-scope'. Many of the examples he discusses overlap with those discussed above in connection with global derivational constraints 1 and 2. For example, he discusses sentences like "Many of the arrows didn't hit the target" and "The target wasn't hit by many of the arrows", claiming that the difference in interpretation can be accounted for by what he calls a difference in scope, which boils down to the question of whether the element in question is inside the VP or not. If he were to provide some reasonable output for his rules, then his scope-difference proposal might be made to match up with those subcases of constraint 1 where L^i is in subject position and L^j is dominated by VP. The overlap is due to the fact that the subject NP precedes VP. However, there are certain crucial cases which decide between constraint 1 and the extended Jackendoff proposal,

namely, cases where the two elements in question are both in the VP. Since
they would not differ in VP-scope, the Jackendoff proposal would predict that
the relative order of the elements should not affect the meaning. Constraint
I, however, would predict a meaning difference just as in the other cases,
where the leftmost element in shallow structure commanded the other element in
semantic representation. We have already seen some examples of cases like
this.

(50)   a.   John talked to few girls about many problems.
       b.   John talked about many problems to few girls.

(51)   a.   I talked to few girls about only those problems.
       b.   I talked about only those problems to few girls.

These sentences show the meaning difference predicted by constraint I, but not
by Jackendoff's scope-difference proposal. Other examples involve adverbs
like carefully, quickly, and stupidly, which he claims occur within the scope
of the VP when they have a manner interpretation (as opposed to sentence ad-
verbs like evidently, which he says have sentence-scope and are not within the
VP). Since Jackendoff permits some adverbs like stupidly to have both VP-scope
and sentence-scope with differing interpretations, all of the following exam-
ples will contain the sentence adverb evidently just to force the VP-scope in-
terpretation for the other adverbs, since a sentence may contain only one sen-
tence adverb.

(52)   a.   John evidently had carefully sliced the bagel quickly.
       b.   John evidently had quickly sliced the bagel carefully.

(53)   a.   John evidently had carefully sliced few bagels.
       b.   John evidently had sliced few bagels carefully.

(54)   a.   John evidently had stupidly given none of his money away.
       b.   John evidently had given none of his money away stupidly.

Each of the pairs of underlined words would be within the scope of the VP ac-
cording to Jackendoff, and so, according to his theory, the (a) and (b) sen-
tences should be synonymous. They obviously are not, and the difference in
their meaning is predicted by constraint I. Thus, constraint I handles a
range of cases that Jackendoff's scope-difference proposal inherently cannot
handle.

But Jackendoff's proposal is inadequate in another respect as well. In
order to generate a sentence like "Not many arrows hit the target" he would
need a phrase-structure rule expanding Determiner as an optional negative
followed by a quantifier (Det → (NEG) Q). The reaing of the NEG relative to
the quantifier would be given by his sentence-scope interpretation rule, since
"not many" is part of the subject "not many arrows" in the above sentence.
This interpretaion rule makes no use of the fact that not happens to precede
many in this example (and is interpreted as commanding many), since Jacken-
doff's interpretation rule would depend in this case on subject (sentence-
scope) position, not the left-to-right order of negative and quantifier. Thus,
in Jackendoff's treatment, it is an accident that NEG happens to precede the
quantifier with this meaning. Jackendoff's scope rule would give exactly the
same result if the NEG had followed the quantifier within the subject, that is,
if the impossible *many not arrows existed. Thus, Jackendoff's phrase-structure

rule putting the NEG in front of the quantifier misses the fact that this order is explained by constraint 1.

On the whole I would say that the discussion of surface and intermediate structure interpretation rules found in Chomsky ([10]), Jackendoff ([33]) and Partee ([55]) do not deal with the real issues. As we have seen, such rules are equivalent to transformations plus global derivational constraints, given the assumption that semantic representations can be given in terms of phrase-markers. We know that transformations are needed in any theory of grammar, and we know that global derivational constraints are also needed on independent grounds, as in rule ordering, Ross' constraints ([65]), R. Lakoff's redundancy rules ([44]), Harold King's contraction cases ([38]), and the myriad of other cases discussed in Lakoff [43]. Thus, surface and intermediate structure interpretation rules are simply examples of derivational constraints, local and global, which are needed independently. The real issues raised in such works are (i) Can semantic representation be given in terms of phrase-markers or a notational variant? (ii) Is there a level of 'deep structure' following lexical insertion and preceding all cyclic rules? and (iii) What are the constraints that hold at the levels of shallow structure and surface structure? These are empirical questions. (i) is discussed in Lakoff (forthcoming), where it is shown that, to the limited extent to which we know anything about semantic representations, they can be given in terms of phrase-markers. (ii) was discussed in the previous section, and will be discussed more thoroughly in the following section. (iii) has been discussed in some detail by Perlmutter (56) and Ross ([65]). It seems to me that many of the regularities concerning nominalizations that have been noted by Chomsky and other low-level regularities noted by Jackendoff ([33]) and Emonds ([14]) are instances of constraints on shallow or surface structure.

4. Autonomous and Arbitrary Syntax

A field is defined by certain questions. For example,

(i) What are the regularities that govern which linear sequences of words and morphemes of a language are permissible and which sequences are not?

(ii) What are the regularities by which the surface forms of utterances are paired with their meanings?

Early transformational grammar, as initiated by Harris ([28],[29]) and developed by Chomsky ([8],[9]), make the assumption that (i) could be answered adequately without also answering (ii), and that the study of syntax was the attempt to answer (i). This assumption defined a field which might well be called 'autonomous syntax', since it assumed that grammatical regularities could be completely characterized without recourse to meaning. Thus, early transformational grammar was a natural outgrowth of American structural linguistics, since it was concerned primarily with discovering the regularities governing the distribution of surface forms.

However, the main reason for the development of interest in transformational grammar was not merely that it led to the discovery of previously unformulated and unformulable distributional regularities, but primarily that, through the study of distributional regularities, transformational grammar provided insights into the semantic organization of language and into the relationships between surface forms and their meanings. If transformational grammar had not led to such insights--if its underlying syntactic structures

had turned out to be totally arbitrary or no more revealing of semantic orga-
nization than surface structures--then the field would certainly, and justi-
fiably, have been considered dull. It may seem somewhat paradoxical, or per-
haps miraculous, that the most important results to come out of a field that
assumed that grammar was independent of meaning should be those that provided
insights as to how surface grammatical structure was related to meaning. In-
tensive investigation into transformational grammar in the years since 1965
has shown why transformational grammar led to such insights. The reason is
that the study of the distribution of words and morphemes is inextricably
bound up with the study of meanings and how surface forms are related to their
meanings. Since 1965, empirical evidence has turned up which seems to show
this conclusively. Some of this evidence will be discussed below. Conse-
quently, a thorough-going attempt to answer (i) will inevitably result in pro-
viding answers to (ii). The intensive study of transformational grammar has
led to the abandonment of the autonomous syntax position, and with it, the
establishment of a field defined by the claim that (i) cannot be answered in
full without simultaneously answering (ii), at least in part. This field has
come to be called 'generative semantics'.

To abandon the autonomous syntax position is to claim that there is a con-
tinuum between syntax and semantics. The basic theory has been formulated to
enable us to make this notion precise, and to enable us to begin to formulate
empirically observed regularities which could not be formulated in a theory of
autonomous syntax. Perhaps the empirical issues can be defined more sharply
by considering the basic theory vis-a-vis other conceptions of transformational
grammar. Suppose one were to restrict the basic theory in the following way.
Let PR, TOP, F, ... in SR be null. Limit global derivational constraints to
those which specify rule order. Limit local derivational constraints to those
specifying elementary transformations, as discussed in Aspects. Assume that
all lexical insertion transformations apply in a block. The resulting restric-
ted version of the basic theory is what Chomsky in [10] describes as a version
of the "standard theory".

Of course, not all versions of the theory of grammar that have been assum-
ed by researchers in transformational grammar are restricted versions of the
basic theory, nor versions of the standard theory. For example, the theory of
grammar outlined in Aspects of the Theory of Syntax [9] is not a version of
either the basic theory or the standard theory. The principal place where the
theory of Aspects deviates from the standard theory and the basic theory is in
its assumption of the inclusion of a non-null Katz-Fodor semantic component,
in particular, their conception of semantic readings as being made up of amal-
gamated paths,[8] which are strings of semantic markers and of symbols which are
supposed to suggest Boolean operations.[9] They nowhere say that readings are
to be defined as phrase-markers made up of the same nonterminal nodes as syn-
tactic phrase-markers, nor do they say that projection rules are operations
mapping phrase-markers onto phrase-markers, and I am sure that no one could
legitimately read such an interpretation into their discussion of amalgamated
paths and Boolean operations on markers. Thus a derivation of a sentence,
including the derivation of its semantic reading, would be represented in the
Aspects theory as a sequence $A_m, \ldots, A_0, P_1, \ldots, P_n$, where the P_i's are
phrase-markers defined as in the basic theory and the standard theory while
the A_i's are amalgamated paths of markers, which are not defined in either the
basic theory or the standard theory. (Chomsky ([10], p. 12) says "Suppose
further that we regard S as itself a phrase-marker in some "semantically prim-
itive" notation....Suppose now that in forming Σ, we construct P_1, which is,
in fact, the semantic representation S of the sentence." In allowing for

semantic representations to be phrase-markers, not amalgamated paths. Chomsky is ruling out a nonnull Katz-Fodor semantic component in his new "standard theory".) In the Aspects model it is assumed that the P_i's are defined by well-formedness constraints (base rules). It is not assumed that either the P_n's (surface structures) or the A_m's (semantic readings) are constrained by any additional well-formedness conditions; rather it is assumed that they are completely characterized by the application of transformations and projection rules to the base structures.

Thus, the Aspects theory differs in an important respect from the basic theory and the standard theory in the definition of a derivation.

(1) (a) Aspects theory: $A_m, \ldots\ldots\ldots\ldots\ldots\ldots, A_0, P_1, \ldots\ldots\ldots\ldots\ldots P_n$

(b) "Standard" theory: $P_1, \ldots\ldots\ldots\ldots\ldots\ldots\ldots, P_i, \ldots\ldots\ldots\ldots\ldots P_n$

<div style="text-align:center">

↑   ↑   ↑

semantic   deep   surface

representation   structure   structure

</div>
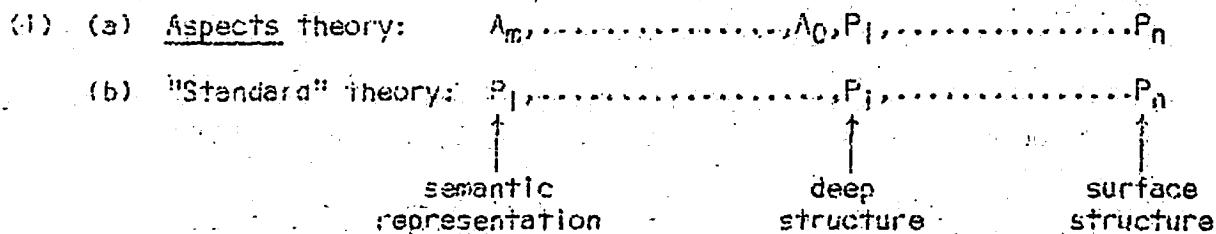
The Aspects theory assumes that semantic readings are formal objects of a very different sort than syntactic phrase-markers and that projection rules are formal operations of a very different sort than grammatical transformations. One of the most important innovations of generative semantics, perhaps the most fundamental one since all the others rest on it, has been the claim that semantic representations and syntactic phrase-markers are formal objects of the same kind, and that there exist no projection rules, but only grammatical transformations. In his discussion of his new "standard theory", Chomsky has therefore adopted without justification one of the most fundamental innovations made by the basic theory.

The "standard" theory is a considerable innovation over the Aspects theory in this sense, since it represents an implicit rejection of Katzian semantics and since the difference between having amalgamated paths and phrase-markers as semantic representations is crucial for Chomsky's claim that there exist surface structure interpretation rules. Suppose this were a claim that there are rules that map surface structures onto amalgamated paths containing strings of semantic markers and symbols for Boolean operations. If it were, then such rules would be formal operations which are of an entirely different nature than grammatical transformations. Then such rules could not have those properties of grammatical transformations that depend crucially on the fact that both the input and output of the transformations are phrase-markers. But it has been shown (Lakoff, [42]) that, at least in the case of surface interpretation rules for quantifiers and negatives proposed by Partee ([55]) and Jackendoff ([33]), such interpretation rules must obey Ross' constraints on movement transformations (Ross, [65]). Since Ross' constraints depend crucially on both the input and output of the rules in question both being phrase-markers (cf. the account of the coordinate structure constraint in § I), it can be demonstrated that, if the outputs of surface interpretation rules are not phrase-markers, then the surface interpretation rule proposals for handling quantifiers and negation are simply incorrect. Thus, although Chomsky doesn't give any reasons for adopting this innovation of generative semantics, his doing so is consistent with his views concerning the existence of surface structure rules of semantic interpretation.[10]

The assumption that there exists a level of deep structure, in the sense of either the Aspects or standard theories defines two possible versions of

the autonomous syntax position: to my knowledge, these are the only two that
have been seriously considered in the context of transformational grammar. We
have already seen in § 2 that there is evidence against (1b). In what follows,
I will discuss just a few of the wide range of cases that indicate that both
the Aspects and standard-theory versions of the autonomous syntax position are
open to very serious doubt.

Though Chomsky does not mention the cycle in his discussion of the stan-
dard theory, we saw in section 2 above that it interacts crucially with the
claim of the standard theory that all lexical insertion rules occur in a
block, since it is shown that there can be no level of 'deep structure' if
such is defined as following all lexical insertion rules and preceding all up-
ward-toward-the-surface cyclic rules. The argument of section 2 also showed
that there exist some cases of post-transformational lexical insertion, as was
conjectured by McCawley ([48]) and Gruber ([25]). Postal ([58]) has found a
rather remarkable case to confirm McCawley's conjecture. Postal considers
sentences like John strikes me as being like a gorilla with no teeth and John
reminds me of a gorilla with no teeth. He notes that both sentences involve
a perception on my part of a similarity between John and a gorilla with no
teeth. This is fairly obvious, since a sentence like John reminds me of a
gorilla with no teeth, though I don't perceive any similarity between John and
a gorilla with no teeth is contradictory. Postal suggests that an adequate
semantic representation for remind in this sense would involve at least two
elementary predicates, on of perception and one of similarity. Schematically,
SU strikes IO as being like O and SU reminds IO of O would have to contain a
representation like:

(2)   IO [perceive] (SU [similar] O)

where [perceive] is a two-place predicate relating IO and (SU [similar] O)
and [similar] is a two-place predicate relating S and O. (2) might be repre-
sented as (3).

(3)
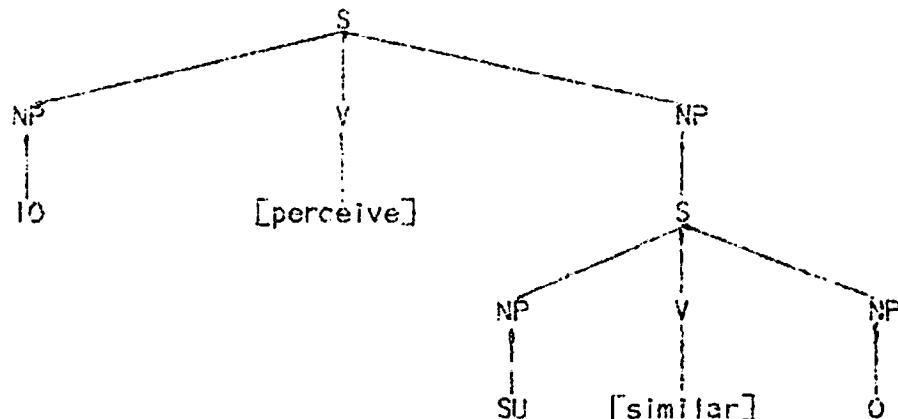
Postal suggests that the semantic representation could be related to the sur-
face structure by the independantly needed rules of subject raising and psych-
movement, plus McCawley's rule of predicate lifting (McCawley, [48]). Subject-
raising would produce (4).

(4)

```
              S
   NP         V         NP              NP
   |                    |                S
   IO    [perceive]     SU         V          NP
                              [similar]        O
```
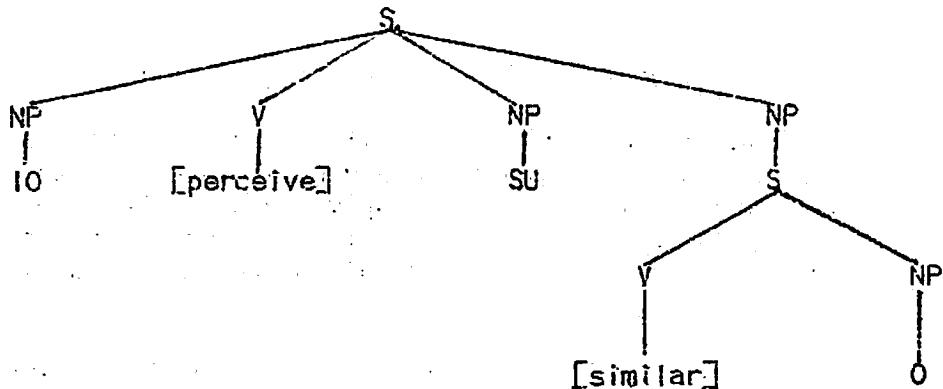
Psych-movement would yield (5).

(5)

```
              S
   NP         V         NP              NP
   |          |         |                S
   SU   [perceive]      IO         V          NP
                              [similar]        O
```
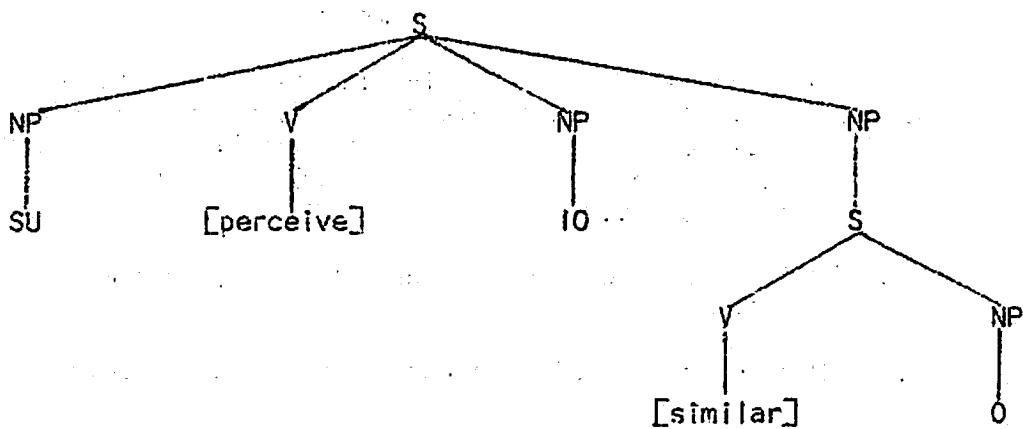
Predicate lifting would yield (6).

(6)

```
              S
   NP         V              NP           NP
   |        /   \            |             |
   SU  [perceive]  V         IO            O
                   |
              [similar]
```
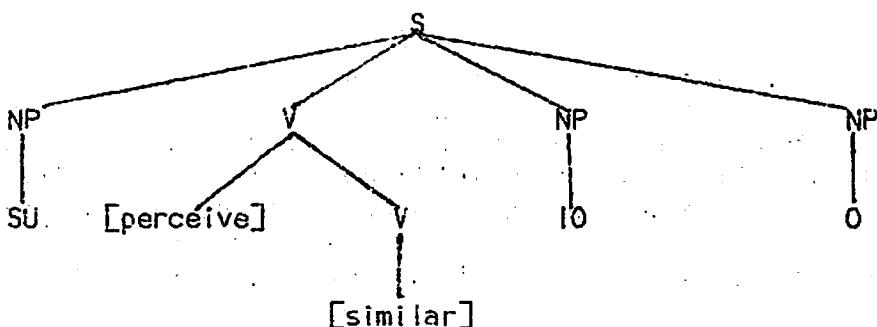
Remind would substitute for [[PERCEIVE] [SIMILAR] $_V$].[11] The question to be asked is whether there is any transformational evidence for such a derivation. In other words, is there any transformational rule which in general would apply only to sentences with a form like (3), which also apply to remind sentences. The existence of such a rule would require that remind sentences be given underlying syntactic structures like (3), which reflect the meaning of such sentences. Otherwise, two such rules would be necessary--one for sentences with structures like (3) and one for remind sentences.

Postal has discovered just such a rule. It is the rule that deletes subjects in sentences like:

(7) To shave oneself is to torture oneself.

(8) Shaving oneself is (like) torturing oneself.

Postal observes that the rule applies freely if the subject is the impersonal one (or the impersonal you). However, there are other, rather restricted circumstances where this rule can apply, namely, when the clause where the deletion takes place is a complement of a verb of saying or thinking and when the NPs to be deleted are coreferential to the subject of that verb of saying or thinking.

(9) Bill says that to shave { himself / *herself } is to torture { himself / *herself }.

(10) Bill feels that shaving { himself / *themselves } is (like) torturing { himself / *themselves }.

The presence of the reflexive indicates what the deleted NP was. Sentences like (11) show that this rule does not apply in relative clauses in addition to complements, and (12) shows that it does not apply if the deleted NPs are identical only to the subject of a verb of saying or thinking more than one sentence up.

(11) *Bill knows a girl who thinks that shaving himself is like torturing himself.

(12) *Mary says that Bill thinks that shaving herself is torturing herself.

Postal notes that this rule also applies in the cases of remind sentences.

(13) Shaving { himself / *herself } reminds John of torturing { himself / *herself }.

If remind is derived from a structure like (3), then this fact follows automatically, since (3) contains a complement and a verb of thinking. If (13) is not derived from a structure like (3), then a separate rule would be needed to account for (13). But that would be only half of the difficulty. Recall that the general rule applies when the NPs to be deleted are subjects of the next highest verb of saying or thinking and the clause in question is a complement of that verb, as in (9) and (10). However, this is not true in the case of remind.

(14) *Mary says that shaving herself reminds Bill of torturing herself.

If remind is analyzed as having an underlying structure like (3), this fact follows automatically from the general rule, since then Mary would not be the subject of the next-highest verb of saying or thinking, but rather the subject of the verb two sentences up, as in (12). Thus, if remind is analyzed as having an underlying syntactic structure like that of (3), one need only state the general rule given above. If remind, on the other hand, is analyzed as having a deep structure like its surface structure, with no complement construction as in (3), then one would (i) have to have an extra rule just for remind (to account for (13)), and (ii) one would have to make remind an exception to the general rule (to account for (14)). Thus, there is a rather strong transformational argument for deriving remind sentences as Postal suggests, which requires lexical insertion to take place following upward-toward-the-surface cyclic rules like subject-raising and psych-movement. This is but one of a

considerable number of arguments given for such an analysis by Postal [58].

Postal's claims about remind, if correct, provide crucial evidence of one sort against the existence of a level of deep structure in the sense of both the 'standard theory' and the Aspects theory, since such a level could be maintained only by giving up linguistically significant generalizations. This would be similar to the argument made by Halle ([26]) against phonemic representation. Another such argument has been advanced by McCawley, [49]. McCawley discusses the phenomenon of respectively-sentences, rejecting the claim that respectively-sentences are derived from sentence conjunction. Chomsky, in [10], gives a particularly clear description of the position McCawley rejects. Chomsky discusses the following examples. The angle-bracketed numbers correspond to Chomsky's numbering. The square-bracketed numbers are McCawley's; where McCawley gives no number, the page is listed. Note that not all the examples have square-bracketed numbers or page references, since Chomsky gives more sentences than McCawley does. <9>; <10>, <15>, and <18> are Chomsky's examples, not McCawley's.

[p. 164-5]   <8> $\underline{Ax}:\underline{x}$ {John, Harry} [$\underline{x}$ love's $\underline{x}$'s wife]

                 <9>   John loves John's wife and Harry loves Harry's wife.

                 <10>   John and Harry love John's wife and Harry's wife, respectively.

[142]   <11>   John and Harry love their respective wives.

[164]   <12> $Ax:\underline{x}$ ε m [x love's x's wife]

[149]   <13>   Those men love their respective wives.

[159]   <14>   That man(x) loves Mary and that man(y) loves Alice.

        <15>   That man(x) and that man(y) love Mary and Alice respectively.

[158]   <16>   Those men love Mary and Alice respectively.

[159]   <17>   That man(x) loves Mary and that man(x) loves Alice.

        <18>   That man(x) and that man(x) love Mary and Alice respectively.

[157]   <19>   That man loves Mary and Alice.

Chomsky's reconstruction of position McCawley rejects is given in a diagram which he numbers <20>.

<20>   <8> $\xrightarrow{\text{I}}$ <9> $\xrightarrow{\text{R}}$ <10> $\xrightarrow{\text{R}'}$ <11>

              <12> $\xrightarrow{\text{I}'}$ <13>

              <14> $\xrightarrow{\text{R}}$ <15> $\xrightarrow{\text{C}}$ <16>

              <17> $\xrightarrow{\text{R}}$ <18> $\xrightarrow{\text{C}}$ <19>

R' is the rule that converts respectively to respective in the appropriate
cases, and C is a rule conjunction collapsing. I, R, and R' are the crucial
part of Chomsky's reconstruction of the argument. I and I' are not mentioned
by McCawley at all, and are entirely due to Chomsky. R is what Chomsky refers
to as the "respectively-transformation" as it is discussed in the transforma-
tional literature. Such a rule would map <9> into <10>, <14> into <15>, and
<17> into <18>. On pages 163 and 164, McCawley shows that grammars incorpo-
rating a rule such as R are inadequate because of their inability to handle
sentences containing both plurals and respectively. He then remarks (p. 164):

> Thus, in order to explain 141-149, it will be necessary to
> change the formulation of the respectively transformation so as to
> make it applicable to cases where there is no conjunction but there
> are plural noun phrases, or rather, noun phrases with set indices:
> pluralia tantum do not allow respectively unless they have a set
> index, so that
>
> 156.   The scissors are respectively sharp and blunt.
>
> can only be interpreted as a reference to two pairs of scissors and
> not to a single pair of scissors.

McCawley then goes on to outline what he thinks an adequate rule for stat-
ing the correct generalization involved in respectively sentences might look
like. (Incidentally, Chomsky describes <20> as the position McCawley accepts
rather than the one he rejects. He then proceeds to point out, as did McCawley,
that such a position is untenable because of its inadequate handling of plurals.
On the basis of this, he claims to have discredited McCawley's position in par-
ticular and generative semantics in general, though in fact he had described
neither.)

McCawley gives a rather interesting argument. He begins his discussion of
what an adequate account of the respectively phenomena might be like as follows
(pp. 164-5).

> ...The correct formulation of the respectively transformation must
> thus involve a set index. That, of course, is natural in view of
> the fact that the effect of the transformation is to 'distribute'
> a universal quantifier: the sentences involved can all be repre-
> sented as involving a universal quantifier, and the result of the
> respectively transformation is something in which a reflex of the
> set over which the quantifier ranges appears in place of occur-
> rences of the variable which was bound by that quantifier.

He then continues:

> ...For example, the semantic representation of 149 is something

$\forall$

like x ε M [x loves x's wife], where M is the set of men in question, and 142 can be assigned the semantic representation $\forall$

$$x \; \varepsilon \; (x_1, x_2)$$

[x loves x's wife], where x_1 corresponds to John and x_2 to Harry: the resulting sentence has those men or John and Harry in place of one occurrence of the bound variable, and the corresponding pronominal form they in place of the other occurrence. Moreover, wife takes a plural form, since after the respectively transformation the noun phrase which it heads has for its index the set of all wives corresponding to any x in the set in question. The difference between 142 and 145 is that the function which appears in the formula that the quantifier in 142 binds is one which is part of the speaker's linguistic competence (f(x) = x's wife), whereas that in 145 is one created ad hoc for the sentence in question $(f(x_1) = Mary, \; f(x_2) = Alice)$.

Basically, McCawley is saying the following. The sentences John and Harry love their respective wives and Those men love their respective wives have certain things in common semantically which can be revealed by a common schema for semantic representation, namely,

(15) $\qquad \underset{x \, \varepsilon \, M}{\forall}$ [x loves x's wife]

The differences between the two sentences come in the specification of the set M. In the former case, M is given by the enumeration of its elements, John and Harry, whereas in the latter case, M would be specified by a description of the class (the members each have the properties of being a man). Given that these two sentences have a common form, McCawley notes that "the result of the respectively transformation is something in which a reflex of the set over which the quantifier ranges appears in place of the occurrences of the variable which was bound by that quantifier." Note that he has not proposed a rule; rather, he has made the observation that given the open sentence in the above expression

(16) $\qquad$ x loves x's wife

the surface form of the respectively sentences is of essentially this form with the nonanaphoric x's filled in in the appropriate fashion--by a description (those men) if the set was defined by a description and by a conjunction (John and Harry) if the set was given by enumeration. McCawley does not propose a characterization of the necessary operation. He merely points out that there is a generalization to be stated here, and some such unitary operation is needed to state it.

Now McCawley turns to a more interesting case, namely, John and Harry love Mary and Alice respectively. He notes that the form of (15) is not sufficiently general to represent this sentence, and observes that there does exist a more general schema in terms of which this sentence and the other two can be represented.

(17) $\qquad \underset{x \, \varepsilon \, M}{\forall}$ [x loves f(x)]

In cases like (15), f(x) is specified generally in terms of the variable which binds the open sentence, that is, f(x) = x's wife. In cases like the one mentioned above, the function is specified by an enumeration of its values for each of the members of the set M over which it ranges, that is, $f(x_1)$ = Mary and $f(x_2)$ = Alice, where M = $\{x_1, x_2\}$. As before, McCawley notes that all three respectively sentences have the surface form of the open sentence,

(18)   x loves f(x)

where x and f(x) have been filled in as specified. Again there is a generalization to be captured, and McCawley suggests that in an adequate grammar there should be a unitary operation that would capture it, though he proposes no such operation.

He then makes the following conclusion.

> I conclude from these considerations that the class of representations which functions as input to the <u>respectively</u> transformation involves not merely set indices but also <u>quantifiers</u> and thus consists of what one would normally be more inclined to call semantic representations than syntactic representations.

Recall that he is arguing against the <u>Aspects</u> theory, not against the standard theory, and in the following paragraph, he goes on to propose what the standard theory, but not the <u>Aspects</u> theory, assumes, namely, that semantic representations are given in terms of phrase-markers. In the <u>Aspects</u> theory, it is assumed that semantic representations and phrase-markers are very different kinds of objects, and McCawley goes on to suggest that if there is a unitary operation relating semantic objects like (17) to the phrase-markers representing <u>respectively</u> sentences <u>and</u> if, as Postal has suggested, ordinary conjunction reduction (which is assumed to map phrase-markers onto phrase-markers, is just a special case of <u>respectively</u> formation (see pages 166-7), then <u>respectively</u> formation must <u>be a rule that</u> maps phrase-markers into phrase markers, and hence semantic representations like (17) musts be given in terms of phrase-markers. If McCawley's argument goes through, then it would follow that the concept of deep structure given by the <u>Aspects</u> theory (though not necessarily that of the "standard theory") would be wrong because of its inadequate concept of semantic representation as amalgamated paths. Chomsky's claim in [10] that McCawley has not proposed anything new in this paper is based on an equivocation in his use of the term "deep structure" and collapses when the equivocation is removed. With respect to the issue of whether or not semantic representations are given by phrase markers, the notion of "deep structure" in the <u>Aspects</u> theory is drastically different than the notion of "deep structure" in the "standard theory"; thus, it should be clear that McCawley's argument, if correct, would indeed provide a Halle-type argument against the <u>Aspects</u> notion of "deep structure", as was McCawley's intent.

But McCawley's proposal is interesting from another point of view as well, for he has claimed that the requirement that one must state fully general rules for relating semantic representations to surface structures may have an effect on the choice of adequate semantic representations. In particular, he claims that an adequate semantic representation for <u>respectively</u> sentences must have a form essentially equivalent to (17). Such a claim is open to legitimate discussion, and whether it turns out ultimately to be right or wrong, it raises an issue which is important not only for linguistics but for other fields as well. Take, for example, the field of logic. Logic, before Frege, was the

study of the forms of valid arguments as they occurred in natural language. In the twentieth century, logic has for the most part become the study of formal deductive systems with only tenuous links to natural language, although there is a recent trend which shows a return to the traditional concerns of logic.[12] In such logical systems, even the latter sort, the only constraints on what the logical form of a given sentence can be are given by the role of that sentence in valid arguments. From the generative semantic point of view, the semantic representation of a sentence is a representation of its inherent logical form, as determined not only by the requirements of logic, but also by purely linguistic considerations, for example, the requirement that linguistically significant generalizations be stated. Thus, it seems to me that generative semantics provides an empirical check on various proposals concerning logical form, and can be said in this sense to define a branch of logic which might appropriately be called "natural logic".

The imposition of linguistic constraints on the study of logical form has some very interesting consequences. For example, McCawley (in a public lecture at M.I.T., spring 1968) made the following observations: Performative sentences can be conjoined but not disjoined.

(19) a. I order you to leave and I promise to give you ten dollars.
     b. *I order you to leave or I promise to give you ten dollars.

This is also true of performative utterances without overt performative verbs.

(20) a. To hell with Lyndon Johnson and to hell with Richard Nixon.
     b. *To hell with Lyndon Johnson or to hell with Richard Nixon.

The same is true when conjunction reduction has applied.

(21) a. To hell with Lyndon Johnson and Richard Nixon.
     b. *To hell with Lyndon Johnson or Richard Nixon.

McCawley then observes that universal quantifiers pattern in these cases like conjunctions and existential quantifiers like disjunctions.

(22) a. To hell with everyone.
     b. *To hell with someone.

Ross has pointed out that the same is true of vocatives.

(23) a. John and Bill, the pizza has arrived.
     b. *John or Bill, the pizza has arrived.

(24) a. (Hey) everybody, the pizza has arrived.
     b. *(Hey) somebody, the pizza has arrived.

McCawley points out that it is no accident that existential quantifiers rather than universal quantifiers pattern like disjunctions,[13] given their meanings. McCawley argues that if general rules governing the syntactic phenomena of (19) - (24) are to be stated, then one must develop, for the sake of stating rules of grammar in general form, a system of representation which treats universal quantifiers and conjunctions as a single unified phenomenon, and correspondingly for existential quantifiers and disjunctions.

Further evidence for this has been pointed out by Paul Postal (personal

communication). It is well-known that repeated coreferential noun phrases are excluded in conjunctions and disjunctions. Thus, (25) and (26) are ill-formed.

(25) a. *Harry$_i$, Sam, and $\begin{Bmatrix} he_i \\ Harry_i \end{Bmatrix}$ are tall.

b. *Harry, I, Max, and I are tall.

(26) a. *Either Harry$_i$, Sam, or $\begin{Bmatrix} he_i \\ Harry_i \end{Bmatrix}$ will win.

b. *Either Harry, I, Max, or I will win.

Postal notes that conjunctions with everyone and disjunctions with someone act the same way.

(27) *Everyone and Sam left.

(28) *Someone or Sam left.[14]

(27) and (28) are excluded if Sam is assumed to be a member of the set over which everyone and someone range, though of course not, if other assumptions are made. Postal points out that this is the same phenomenon as occurs in (25) and (26), namely, conjuncts and disjuncts may not be repeated. If there is to be a single general rule covering all of these cases, then the rule must be stated in some notation which treats quantifiers and conjunctions as a single unified phenomenon.

A further argument along these lines has been provided by Robin Lakoff. It has long been known that in comparative constructions a conjunction may be expressed by a disjunction. Thus the meaning of (29) may be expressed by (30).

(29) Sam likes lox more than herring and whitefish.

(30) Sam likes lox more than herring or whitefish.

(Of course, (30) also has a normal disjunctive reading.) Let us assume that there is a transformation changing and to or in such comparative constructions. Lakoff notes that the same phenomenon occurs with quantifiers.

(31) Sam likes canned sardines more than everything his wife cooks.

(32) Sam likes canned sardines more than anything his wife cooks.

The meaning of (31) can be expressed by (32), in which any replaces every. Again, as she argues, we have the same phenomenon in both cases, and there should be a single general rule to cover both. Thus, the same transformation that maps and into or must also map every into some/any. This can only be done if there is a single unified notation for representing quantifiers and conjunctions.

These facts also provide evidence of the sort brought up by Postal in his discussion of remind, evidence showing that indicates that certain transformations must precede the insertion of certain lexical items. Consider prefer,

which means _like more than_.

(33) Sam likes lox more than herring.

(34) Sam prefers lox to herring.

As we saw in (29) - (32) above, _and_ optionally changes to _or_ and _every_ to _any_ in the _than_-clause of comparative constructions. The same thing happens in the corresponding place in _prefer_ constructions, namely, in the _to_-phrase following _prefer_.

(34) Sam prefers lox to herring and whitefish.

(35) Sam prefers lox to herring or whitefish.

(36) Sam prefers canned sardines to everything his wife cooks.

(37) Sam prefers canned sardines to anything his wife cooks.

If _prefer_ is inserted for _like-more_ after the application of the transformation mapping conjunctions into disjunctions, then the fact that this mapping takes place in the _to_-phrase following _prefer_ follows as an automatic consequence of the meaning of _prefer_. Otherwise, this phenomenon must be treated in an ad hoc fashion, which would be to make the claim that these facts are unrelated to what happens in comparative constructions.

Further evidence for such a derivation of _prefer_ comes from facts concerning the "stranding" of prepositions. The preposition _to_ may, in the general case, be either "stranded" or moved along when the object of the preposition is questioned or relativized.

(38) a. Who did John give the book to?
      b. To whom did John give the book?

(39) a. Who is Max similar to?
      b. To whom is Max similar?

(40) a. What city did you travel to?
      b. To what city did you travel?

The preposition _than_, on the other hand, must be stranded, and may move along only in certain archaic-sounding constructions like (43).

(41) a. What does Sam like bagels more than?
      b. *Than what does Sam like bagels more?

(42) a. Who is Sam taller than?
      b. *Than whom is Sam taller?

(43) ?God is that than which nothing is greater.

The preposition _to_ following _prefer_ does not work like ordinary occurrences of _to_, but instead works just like _than_: it must be stranded where _than_ is stranded and may move along in just those archaic-sounding constructions where _than_ may.

(44) a. What does Sam prefer bagels to?
      b. *To what does Sam prefer bagels?

(45) ?God is that to which I prefer only bagels.

Unless prefer is derived from like-more, these facts cannot be handled in a unified way, and the correlation must be considered accidental. Such facts seem to provide even more evidence against the Aspects conception of deep structure.

A particularly strong argument for post-transformational lexical insertion has been made by David Perlmutter, on the basis of certain facts concerning adjective and participle agreement in Spanish. There, as in many languages, adjectives and participles agree with their derived subjects in gender and number. For example,

(46) Mi padre es perseguido por mi madre.
      My father is being pursued (masc. sg.) by my mother.

Thus it is obvious that the rule of adjective-participle agreement must follow passivization. Now consider the sentences:

(47) a. Mi madre es rica y mi padre es rico.
          My mother is rich (fem. sg.) and my father is rich (masc. sg.).

      b. Mi madre y mi padre son ricos.
          My mother and my father are rich (masc. pl.).

      c. Mis padres son ricos.
          My parents are rich (masc. pl.).

When an adjective has a conjoined subject with two different genders and that adjective is predicated of both members of the conjunctions, then the adjective is in the masculine plural. This is also true if the subject is a plural noun which is understood as referring to persons of different natural gender, as in (47 c). The cases in (47) are typical of the situation where a single adjective is predicated of two persons. Now consider (48), where two different adjectives are predicated of two different persons. Adjectives of opposite meaning have been chosen so that both cannot be understood as being predicated of both persons at once.

(48) a. Mi madre es alta y mi padre es bajo.
          My mother is tall (fem. sg.) and my father is short (masc. sg.).

This parallels (47 a). In such cases there is a corresponding respectively-construction.

(48) b. Mi madre y mi padre son respectivamente alta y bajo.
          My mother and my father are tall (fem. sg.) and short (masc. sg.) [respectively].

Note that the adjectives are in the singular, each agreeing with the noun phrase that is its derived subject in the unreduced sentence. Masculine plural agreement parallel to (47 b) is impossible.

(48) b'. *Mi madre y mi padre son respectivamente altos y bajos.
        *My mother and my father are tall (masc. pl.) and short (masc. pl.)
        respectively.

In such cases adjectives cannot agree with the entire conjoined noun
phrase, but only with the individual noun phrases making up the conjunction.
Now consider the case with the plural noun of mixed gender.

(48) c. Mis padres son respectivamente alta y bajo.
        My parents are tall (fem. sg.) and short (masc. sg.) respectively.

(48) c'. *Mis padres son respectivamente altos y bajos.
        *My parents are tall (masc. pl.) and short (masc. pl.) respectively.

The plural noun of mixed gender works exactly like the corresponding conjunction:
plural adjective agreement is impossible: the adjectives must agree with the
singular noun phrases that they are predicated of. However, in (48 c), there
is no such conjunction after lexical insertion. There is only a single noun,
which is grammatically masculine plural, but which is understood semantically
as being a conjunction of two singular noun phrases of different natural gen-
der: male parent and female parent. The adjectives of (48 c) agree not with
the single-lexical-item plural noun, but with the component noun phrases that
constitute its meaning. This would be impossible if agreement applied after
lexical insertion, since the singular noun phrases with which the two adjec-
tives agree are not present after the lexical insertion of "padres". Thus,
if there is a general agreement transformation in the syntax of Spanish, it
must apply before the insertion of the lexical item "padres". Moreover, we
know that such a transformation must apply to syntactic (not semantic) repre-
sentations for two reasons: (i) agreement is based on grammatical gender,
not natural gender, though natural gender will determine grammatical gender in
a certain limited range of cases; (ii) agreement applies to derived, not un-
derlying subjects (e.g., the output of passivization). Thus the lexical in-
sertion of "padres" must have followed the syntactic agreement transformation,
which in turn follows passive, which is a cyclic rule. Thus, it is false that
all lexical insertion must precede all upward-toward-the-surface cyclic rules,
a result which is in accord with all of the above-mentioned results and which
seems to provide clear and incontrovertible evidence against the concept of
'deep structure' in both the Aspects and the standard theories.
        Since the claim that a level of deep structure exists in natural languages
requires that it exist in all dialects of all natural languages, such a claim
can be disproved by the existence of one dialect of one natural language which
cannot have such a level. The above examples are from one such dialect of
Spanish. However, since these facts are not true for all Spanish speakers,
variations should be discussed. Some speakers find respectively sentences of
the sort discussed above stilted, and, although they understand them, they
would never use any such sentences in their normal speech. To my knowledge,
no speakers find (48 b') and (48 c'), where both antonymous adjectives are
plural, to be grammatical at all. However, some speakers also find (48 b) and
(48 c), where the adjectives have mixed genders to be unacceptable. In their
place, they must have both adjectives in the masculine singular: Mi madre y
mi padre son respectivamente alto y bajo (in place of (48 b)) and Mis padres
son respectivamente alto y bajo (in place of (48 c)). Here too, plural NPs
act like conjoined singular NPs. What is interesting about this dialect is
that the gender agreement is determined by the gender of the derived conjoined

(50) a. *I want to know where he went.
　　b. I want to know where did he go.

If there is no request for information, inversion does not occur, even with the same verbs.

(51) a. Bill told me where he went.
　　b. *Bill told me where did he go.

(52) a. I know where he went.
　　b. *I know where did he go.

This phenomenon, though it does not occur in standard English, does have its counterpart there in cases like:

(53) Where did he go, I want to know.

(54) *Where did he go, I know.

(55) Where did he go, tell me.

(56) *Where did he go, tell Harry.

In this dialect, even such a late syntactic rule as subject-auxiliary-inversion, must be stated not in terms of superficial syntactic structure but in terms of the meaning of the sentence; that is, if the generalization is to be captured the subject-auxiliary inversion rule must have in its structural index the information that the sentence in question describes or is a request for information. This is obviously impossible to state in either the Aspects or "standard" theories.

Given the rather considerable array of evidence against the existence of a level of 'deep structure' following all lexical insertion and preceding all upward-toward-the-surface cyclic rules, it is rather remarkable that virtually no arguments have ever been given for the existence of such a level. The arguments that one finds in works of the Aspects vintage will usually cite pairs of sentences like "John ordered Harry to leave" and "John expected Harry to leave," show that they have very different properties, and claim that such properties can be accounted for by assuming some 'higher' level of representation reflecting the different meanings of the sentences ("order" is a three-place predicate; "expect" is a two-place predicate). Such arguments do seem to show that a 'higher' or 'more abstract' level of representation than surface structure exists, but they do not show that this level is distinct from the level of semantic representation. In particular, such arguments do not show that any intermediate level of 'deep structure' as defined in the precise sense given above exists. It was simply assumed in Aspects that this 'higher' level contained all lexical items and preceded all transformations: no arguments were given.

The only attempt to provide such an argument that I have been able to find is rather recent one. Chomsky ([10]) cites the context

(57) Bill realized that the bank robber was _____.

and considers the sentences formed by inserting

(58) John's uncle

and

(59) The person who is the brother of John's mother or father or the husband
       of the sister of John's mother or father.

into the blank in (57). He claims that these sentences would have different
semantic representations, and this claim is based on the further claim that
"what one believes, realizes, etc., depends not only on the proposition expres-
sed, but also on some aspects of the form in which it is expressed." These
claims are not obviously true, and have been disputed.[17] But let us assume for
the sake of argument that Chomsky is right in this matter. Chomsky does not
propose any account of semantic representation to account for such facts. He
does, however, suggest (albeit with reservations and without argument) that
such examples refute any theory of grammar without a level of "deep structure",
but not a theory with such a level. He continues,

> Do considerations of this sort refute the standard theory as well?
> The example just cited is insufficient to refute the standard theo-
> ry, since [(57) - (59)] differ in deep structure, and it is at
> least conceivable that "realize" and similar items can be defined
> so as to take account of this difference.   --page 29

But this is a non sequitur, even given Chomsky's assumptions. (57) - (59) would
show, under Chomsky's assumptions, only that truth values of sentences depend
in part on the particular phonological form in which semantic information is
expressed. It does not follow that the correlation between phonological forms
and the corresponding semantic information must be made at a single level of
grammar, and certainly not at a level preceding all cyclic rules. It only
follows that such correlations must be made at some point or other in the de-
rivations defined by a grammar. So long as such correlations are made some-
where in the grammar, "it is at least conceivable that "realize" and similar
items can be defined so as to take account of" (57) - (59). Of course, this
is not saying much, since anything that pretends to be a grammar must at the
very least show how semantic information correlates with phonological form.
All that Chomsky's argument shows is that his examples do not refute any the-
ory of grammar that defines the correlations between semantic information and
phonological form, that is, any theory of grammar at all.

      This is, as far as I know, the present state of the evidence in favor of
the existence of a level of 'deep structure' which contains all lexical items
and precedes all cyclic rules. Since the burden of proof must fall on someone
who proposes a level of 'deep structure', there is at present no good reason
to believe in the existence of such a level and a number of good reasons not
to. This of course does not mean that there is no intermediate level at all
between semantic representations and surface structures. In fact, as we have
seen, it is not unreasonable to believe that there exists a level of 'shallow
structure', perhaps following all cyclic rules. I think it is fair to say
that at present there is a reasonable amount of evidence disconfirming the
autonomous syntax position and none positively confirming it. This is, of
course, not strange, since virtually no effort has gone into trying to prove
that the autonomous syntax position is correct.

      Going hand-in-hand with the position of autonomous syntax is what we might
call the position of 'arbitrary syntax'. We might define the arbitrary syntax
position as follows: Suppose there is in a language a construction which bears
a meaning which is not given simply by the meanings of the lexical items in the

sentence (e.g., the English question, or the imperative). The question arises as to what the underlying structure of a sentence with such a 'constructional meaning' should be. Since the meaning of the construction must appear in its semantic representation (P_1, in the standard theory) and since this meaning is represented in terms of a phrase-marker (at least in the standard theory), one natural proposal might be that the 'deep structure' phrase-marker of the sentence contain the semantic representation corresponding to the construction directly. Call this the 'natural syntax' position. The arbitrary syntax position is the antithesis of this. It states that the deep structure corresponding to a construction of the sort described never contains the phrase structure configuration of the meaning of the sentence directly. Instead, the deep structure corresponding to the configuration must contain some arbitrary marker. Consider the English imperative as an example. The meaning of the imperative construction in a sentence like Come here must be given in terms of a 3-place predicate relating the speaker, the addressee, and a sentence describing the action to be performed, as expressed overtly in the sentence I order you to come here. Any adequate theory of semantic representation must say at least this much about the meaning of Come here. The arbitrary syntax position would maintain in this case that the 'deep structure' of Come here would not contain such a 3-place predicate, but would instead contain an arbitrary marker. In recent studies such a marker has been given the mnemonic IMP, which may tend to hide its arbitrariness. A good name to reveal its true arbitrary nature would be IRVING. Under the arbitrary syntax position, the deep structure of Come here would contain IRVING.[18]

It is possible to show in certain instances that the arbitrary syntax position is incorrect. One of the most telling arguments to this effect has been given by Robin Lakoff in Abstract Syntax and Latin Complementation [44].

The Lakoff argument concerns the distribution of the Latin subjunctive and of two morphemes indicating sentence negation. She begins by considering the Latin sentence:

(60) Venias. (Form: 2nd person singular subjunctive of venio, 'to come')

(60) is an example of what is called an "independent subjunctive" in Latin, and it is at least three ways ambiguous, as shown in (61).

(61) (i) Come! I order you to come.
      (ii) May you come! I want you to come.
      (iii) You may come. It is possible that you will come.

There is also in Latin a dependent subjunctive which functions as a complementizer with verbs of certain meaning classes. Some typical examples are:

(62) (i) Impero ut venias. 'I order you to come'
      (ii) Volo ut venias. 'I want you to come'
      (iii) Potest fieri ut venias. 'It is possible that you will come'

She argues that the sentences of (62) should have underlying structures roughly like (63).

(63)   (i)

```
                              S
        ┌──────────┬──────────┴──────────────┐
      NP₁          V         NP₂             NP
        │          │          │               │
       ego       imper-      tu               S
       'I'       'order'    'you'        ┌────┴────┐
                                        NP₂        V
                                         │         │
                                        tu        ven-
                                       'you'      'come'
```

(ii)

```
                      S
        ┌─────────────┼─────────────┐
      NP₁             V             NP
        │             │              │
       ego          vol-            S
       'I'         'want'      ┌────┴────┐
                              NP₂        V
                               │         │
                              tu        ven-
                             'you'      'come'
```

(iii)

```
                      S
           ┌──────────┴──────────┐
          NP                     V
           │                     │
           S               potest fieric
      ┌────┴────┐          'is possible'
     NP₂        V
      │         │
     tu        ven-
    'you'      'come'
```
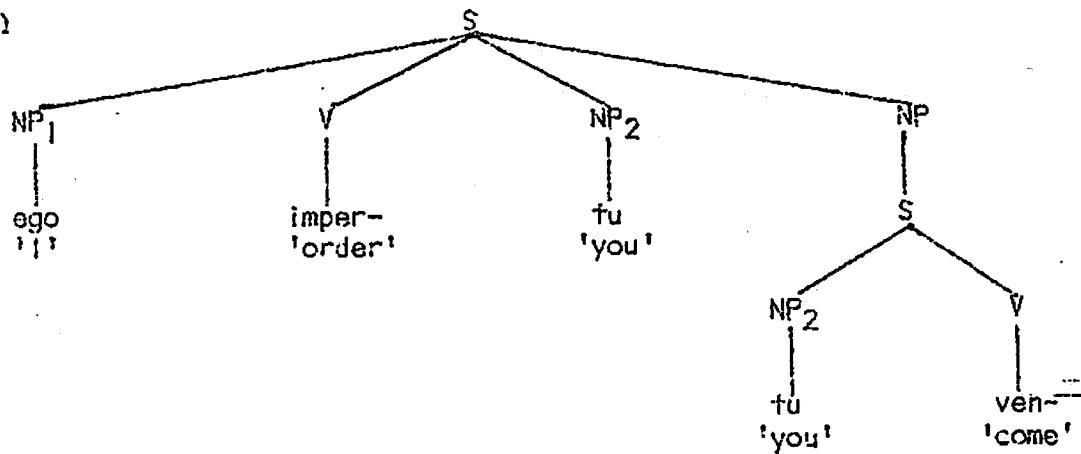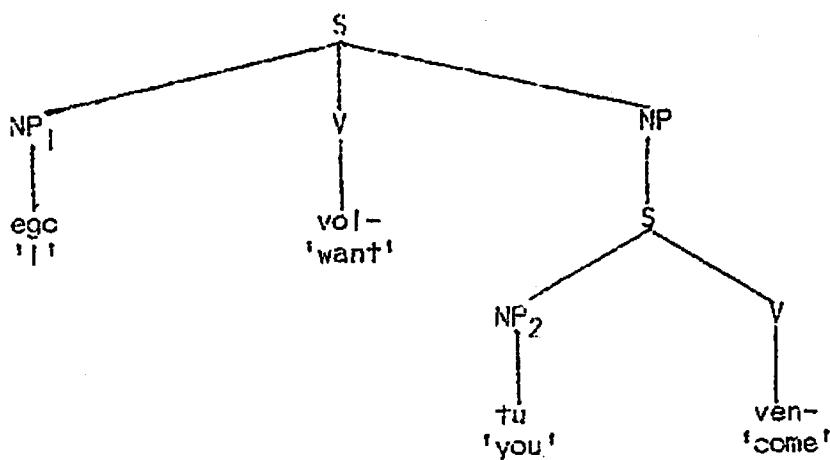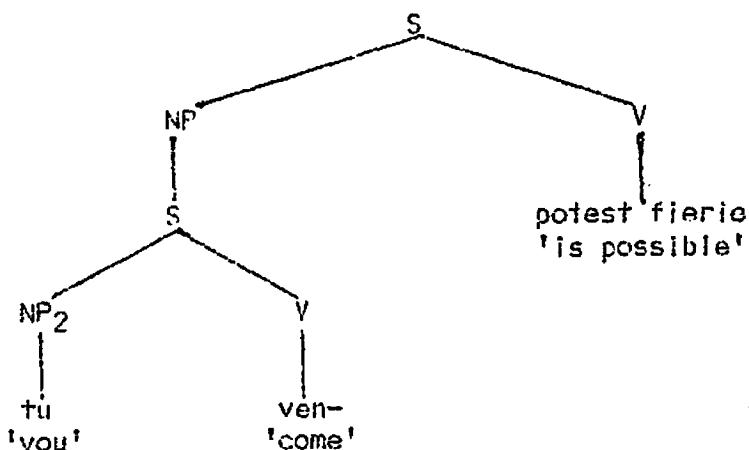
The rule of complementizer-placement will in each case mark the main verb inside a complement structure, that is, one of the form

```
        NP
         │
         S
        / \
       /___\
```
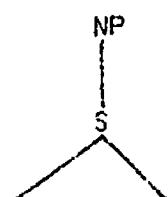
as subjunctive, given the meaning-class of the next-highest verb.

Lakoff then suggests that it might not be an accident that the sentence of (60), which is odd in that it has a subjunctive in the main clause, has the same range of meanings as the sentences of (62), where the normal rule of subjunctive complementation for verbs of those meanings has applied. She proposes that if it were hypothesized that (60) had three underlying structures just like those in (63), except that in the place of the real predicates impero, volo, and potest fieri there were 'abstract predicates', or nonlexical predicates bearing the corresponding meanings, then the subjunctive in (60) would be derived by the same, independently motivated, rule that derives the embedded subjunctives of (62). Since the structures of (63) reflect the meanings of the three senses of (60), such a solution would provide an explanation of why a subjunctive should show up in a main clause with just those meanings. But the arbitrary syntax position would rule out such an explanation. In terms of the arbitrary syntax position, (60) would have three different deep structures, all of them with venio as the main verb and with no complement constructions. The difference between the three deep structures could only be given by arbitrary symbols, for example, MARCUS, PUBLIUS, AND JULIUS (they might be given mnemonics like IMP, VOL, and POSS, though such would be formally equivalent to three arbitrary names.). In such a theory, the three deep structures for (60) would be:

(64)   (i)

```
                        S
               _____|_____
          MARCUS        NP        V
          (or IMP)      |          |
                        tu        ven-
```

(ii)

```
                        S
               _____|_____
          PUBLIUS       NP        V
          (or VOL)      |          |
                        tu        ven-
```

(iii)

```
                        S
               _____|_____
          JULIUS        NP        V
          (or POSS)     |          |
                        tu        ven-
```
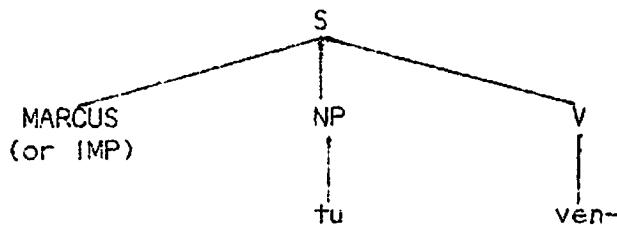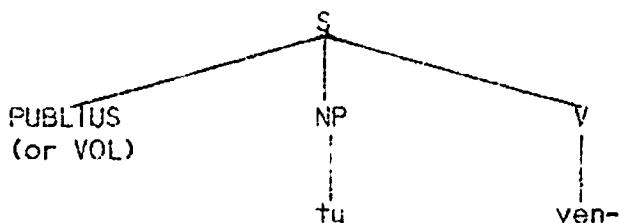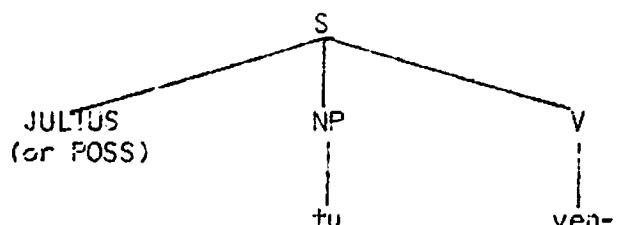
There would then have to be a rule stating that verbs become subjunctive in the environment

(65) $\left\{ \begin{array}{l} \text{MARCUS} \\ \text{PUBLIUS} \\ \text{JULIUS} \end{array} \right\}$ NP ____

or equivalently,

$$\begin{bmatrix} IMP \\ VOL \\ POSS \end{bmatrix} \cdot NP \underline{\hspace{1.5cm}}$$
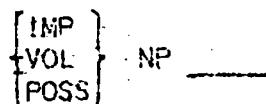
Such a rule would be entirely different than the rule that accounts for the subjunctives in (62), and having two such different rules is to make the claim that the appearance of the subjunctive in (60) is entirely unrelated to the appearance of the subjunctive in (62), and that the fact that the same endings show up is a fortuitous accident. To claim that it is not a fortuitous accident is to claim that the arbitrary syntax position is wrong in this respect.[19]

Lakoff then goes on to discuss negatives. In Latin, sentence negation may be expressed by one of two morphemes, ne and non. These are, of course, in complementary distribution, and she shows that in sentential complements there is a completely general rule governing their distribution: ne occurs in object complements where the verb inside the complement is subjunctive; non occurs elsewhere, e.g., in nonsubjunctive complements and in subject complements where the main verb is subjunctive. For example, the negatives corresponding to (62) would be (66): The sentences of (67) would be ungrammatical in Latin. Ut is optional before ne.

(66) (i)   Impero (ut) ne venias.   'I order you not to come'
      (ii) Volo (ut) ne venias.   'I want you not to come'
      (iii) Potest fieri ut non venias.   'It is possible that you won't come'

(67) (i)   *Impero ut non venias.
      (ii) *Volo ut non venias.
      (iii) *Potest fieri (ut) ne venias.

In main-clauses without subjunctive main verbs we find non, not ne, just as in the corresponding complement clauses. But in main clauses with subjunctive main verbs, namely, cases like (60), referred to as 'independent subjunctives', we find both ne and non. That is, both Ne venias and Non venias are grammatical in Latin. However, they do not mean the same thing. Their meanings are distributed as in (68).

(68) (i)   Ne venias.   'Don't come! I order you not to come'
      (ii) Ne venias.   'May you not come. I want you not to come'
      (iii) Non venias.   'You may not come. It is possible that you won't come'

The distribution of meanings and negatives in (68) corresponds exactly to the distribution in (66). Lakoff argues that this too is no accident. She notes that if the underlying structures for (60) are those of (53), with the appropriate abstract predicates, then the distribution of negatives in (68) follows the ordinary rules specifying the occurrence of ne in object complements, not subject complements. This would explain the facts of (68). However, if the deep structures of (60) are those of (64), then an entirely different rule would have to be stated, namely: In main clauses with subjunctive main verbs the negative appears as ne if either MARCUS (or IMP) or PUBLIUS (or VOL) is present, and non otherwise. Such a rule would be entirely different from the rule for negatives inside complements, and to have two such different rules is to make the claim that there is no generalization governing the distribution of negatives in (68) and (66), and that the fact that the distribution of forms correlates with the distribution of meanings in these cases is a fortuitous

accident. To say that it is not an accident is to say that any theory which rules out abstract predicates and forces such rules to be stated instead in terms of arbitrary markers like MARCUS, IRVING, Q, and IMP is wrong, since linguistically significant generalizations cannot be stated in such a theory.

Now consider what the semantic representations of the three senses of (60) would look like. Sense (i), which is an order, would involve a three-place predicate, specifying the person doing the ordering, the person to whom the order is directed, and the proposition representing the order to be carried out. If one conceives of semantic representation as being given along the lines suggested in Lakoff ([43]) such a semantic representation would look essentially like (63 i). Similarly, sense (ii) of (6C) expresses a desire, and so its semantic representation would have to contain a two-place predicate indicating the person doing the desiring and the proposition expressing the content of the desire. That is, it would essentially have the structure of (63 ii). Sense (iii) of (60) expresses a possibility, and so would have in its semantic representation a one-place predicate containing a proposition, as in (63 iii). In short, if semantic representations are given as in Lakoff ([43]), then the semantic structures are exactly the structures required for the formulation of general rules introducing the subjunctive complementizer and ne in Latin. This seems to me to support the claim that semantic representations are given in terms of syntactic phrase markers, rather than, say, the amalgamated paths of the Aspects theory. It also seems to support the generative semantics position that there is no clear distinction between syntactic phenomena and semantic phenomena. One might, of course, claim that the terms "syntactic phenomena" and "semantic phenomena" are sufficiently vague so as render such a statement meaningless. But I think that there are enough clear cases of "syntactic phenomena" to give the claim substance. It seems to me that if anything falls under the purview of a field called "syntax" the rules determining the distribution of grammatical morphemes do. To claim that such rules are not "syntactic phenomena" seems to me to remove all content from the term "syntax". Thus the general rules determining the distribution of the two negative morphemes ne and non in Latin and the subjunctive morpheme in Latin should be "syntactic phenomena" if there are any "syntactic phenomena" at all. Yet, as we have seen, the general rules for stating such distributions must be given in terms of structures that reflect the meaning of the sentence rather than the surface grammar of the sentence.

From the fact that the arbitrary syntax position is wrong, it does not necessarily follow that the natural syntax position is right. It is logically possible to hold a 'mixed' position, to the effect that for some such constructions there must be arbitrary markers and for others not. However, since semantic representations for such constructions must be given independently in any adequate theory of grammar, the strongest claim that one could make to limit the class of possible grammars would be to adopt the natural syntax position and to say that there are no arbitrary markers of the sort discussed above. It is conceivable that this is too strong a claim, but it is perhaps the most reasonable position to hold on methodological grounds, for it requires independent justification to be given for choosing each proposed arbitrary marker over the independently motivated semantic representation. To my knowledge, no such justification has ever been given for any arbitrary marker, though of course it remains an open question as to whether any is possible. In the absence of any such justification, we will make the strongest claim, namely, that there exist no such markers.

It should be noted that this is a departure from the methodological assumptions made by researchers in transformational grammar around 1965, when Aspects

was published. At that time it had been realized that linguistic theory had to make precise claims as to the nature of semantic representations and their relationship to syntax, but existence of semantic representations continued to be largely ignored in syntactic investigations since it was assumed that syntax was autonomous. Moreover, in Katz and Postal [36] and Aspects, a precedent had been set for the use of arbitrary markers, though that precedent was never justified. Given such a precedent, it was widely assumed that any deviation from the use of arbitrary markers required justification. However, as soon as one recognizes that (i) semantic representations are required, independent of any assumptions about the nature of grammar, and that they can be represented in terms of phrase-markers and (ii) that the autonomous syntax position is open to serious question, then the methodological question as to what needs to be justified changes. From this point of view, arbitrary markers must be assumed not to exist until they are shown to be necessary, and the autonomous syntax position can no longer be assumed, but rather must be proved.

Lakoff's argument for the existence of abstract predicates was one of the earliest solid arguments not only against arbitrary syntax but also for the claim that the illocutionary force of a sentence is to be represented in underlying syntactic structure by the presence of a performative verb, real or abstract. She has more recently (R. Lakoff, [45]) provided strong arguments for the existence of an abstract performative verb of supposing in English. Arguments of essentially the same form have been provided by Ross ([66]) for the existence of an abstract verb of saying in each declarative sentence of English. Thus, the importance of the argument given above for Latin goes far beyond what it establishes in that particular case, since it provides a form of syntactic argumentation in terms of which further empirical evidence for abstract predicates, performative or otherwise, can be gathered. The basic argument is simple enough:

If the same syntactic phenomena that occur in sentences with certain overt verbs occur in sentences without those verbs, and if those sentences are understood as though those verbs were there, then we conclude (1) a rule has to be stated in the cases where the real verbs occur; (2) since the same phenomenon occurs with the corresponding understood verbs, then there should be a single general rule to cover both cases; (3) since we know what the rule looks like in the case of real verbs, and since the same rule must apply, then the sentences with understood verbs must have a structure sufficiently like that of those with the overt verbs so that the same general rule can apply to both.

If one wishes to avoid the consequences of the Lakoff argument and of other similar arguments, there are two possible ways out. First, one can deny that the form of the argument is valid. Second, one can claim that the generalization is spurious. Let us start with the first way out. Arguments of the above form are central to generative grammar. The empirical foundations of the field rest to a very large extent on arguments of just this form. Take for example the argument that imperative constructions in English are not subjectless in underlying structure, and that they in fact have second person subjects. The sort of evidence on which this claim rests is the following:

(69) I shaved $\begin{Bmatrix} *me \\ you \\ him \end{Bmatrix}$ $\begin{Bmatrix} myself \\ *yourself \\ *himself \end{Bmatrix}$

(70) You shaved $\begin{Bmatrix} me \\ *you \\ him \end{Bmatrix}$ $\begin{Bmatrix} *myself \\ yourself \\ *himself \end{Bmatrix}$

etc.

(71) Shave $\begin{Bmatrix} me \\ *you \\ him \end{Bmatrix}$ $\begin{Bmatrix} *myself \\ yourself \\ *himself \end{Bmatrix}$

. . . .

(72) I'll go home, won't $\begin{Bmatrix} I \\ *you \\ *he \end{Bmatrix}$

(73) You'll go home, won't $\begin{Bmatrix} *I \\ you \\ *he \end{Bmatrix}$

etc.

(74) Go home, won't $\begin{Bmatrix} *I \\ you \\ *he \end{Bmatrix}$

The argument is simple enough. In sentences with overt subjects, we find reflexive pronouns in object position just in case the subjects and objects are coreferential, and nonreflexive pronouns just in case the subjects and objects are noncoreferential. We hypothesize that there is a rule of reflexivization which reflexivizes object pronouns that are coreferential with their subjects. Similarly, in tag questions, we find that as a general principle the pronominal form of the subject of the main clause occurs as the subject of the tag. In imperative sentences we find that a second person subject is understood and that a second person reflexive, but no other, shows up in object position. and that a second person nonreflexive pronoun is excluded in object position. Similarly, we find that the tags for imperative sentences contain second person subjects. We assume that all this is no accident. In order to be able to conclude that imperatives have underlying second person subjects, we need to be able to argue as follows:

If the same syntactic phenomena that occur in sentences with certain overt subjects occur in sentences without those subjects, and if those sentences are understood as though those subjects were there, then we conclude (1) a rule has to be stated in the cases where the overt subjects occur; (2) since the same phenomenon occurs with the corresponding understood subjects, then there should be a single general rule to cover both cases; (3) since we know what the rule looks like in the case of real subjects, and since the same rule must apply, then the sentences with the understood subjects must have a structure sufficiently like that of those with the overt subjects so that the same general rule can apply to both.

If arguments of this form are not valid, then one cannot conclude on the basis of evidence like the above that imperative sentences have underlying second person subjects, or any subjects at all. A considerable number of the results of transformational grammar are based on arguments of just this form.

If this form of argument is judged to be invalid, then these results must all be considered invalid. If this form of argument is valid, then the results of R. Lakoff ([44],[45]) and Ross ([66]) concerning the existence of abstract performative verbs must be considered valid, since they are based on arguments of the same form. A consistent approach to empirical syntactic evidence requires that abstract performative verbs be accepted or that all results based on arguments of this form be thrown out. Whether much would be left of the field of transformational grammar if this were done is not certain.

Another way out that is open to doubters is to claim that the generalization is spurious. For example, one can claim that the occurrence of a second person reflexive in Shave yourself has nothing whatever to do with the occurrence of the second person reflexive pronoun in You will shave yourself, and that the fact that the imperative paradigms of (71) and (74) above happen to match up with the second person subject paradigms of (70) and (73) is simply an accident. If the correspondence is accidental, then there is no need to state a single general rule---in fact, to do so would be wrong. And so it would be wrong to conclude that imperative sentences have underlying subjects. If someone takes a position like this in the face of evidence like the above, rational argument ceases.

Let us take another example, the original Halle-type argument. Suppose a diehard structural linguist wanted to maintain that there was a level of taxonomic phonemics in the face of the argument given against such a position by Halle ([26]) (cf. Chomsky's discussion in Fodor and Katz, [16], p. 100). Halle points out that there is in Russian a rule that makes obstruents voiced when followed by a voiced obstruent. He then observes that if there is a level of taxonomic phonemics, this general rule cannot be stated, but must be broken up into two rules, the first relating morphophonemic to phonemic representation (1) obstruents except for c, č and x become voiced before voiced obstruents, and the second relating phonemic to phonetic representation (2) c, č and x become voiced before voiced obstruents. As Chomsky states in his discussion, "the only effect of assuming that there is a taxonomic phonemic level is to make it impossible to state the generalization." The diehard structuralist could simply say that the generalization was spurious, that there really were two rules, and that there was no reason for him to give up taxonomic phonemics. In such a case, rational argument is impossible. It is just as rational to believe in taxonomic phonemics on these grounds as it would be to maintain the arbitrary or autonomous syntax positions in the face of the examples discussed above.

It should be noted in conclusion, that transformational grammar has in its theoretical apparatus a formal device for expressing the claim that a generalization does not exist. That formal device is expressed by the curly-bracket notation. The curly-bracket notation is used to list a disjunction of environments in which a rule applies. The implicit claim made by the use of this notation is that the items on the list (the elements of the disjunction) do not share any properties relevant to the operation of the rule. From the methodological point of view, curly-brackets are an admission of defeat, since they say that no general rule exists and that we are reduced to simply listing the cases where a rule applies.

Let us take an example. Suppose one wanted to deny that imperative sentences have underlying subjects. One would still have to state a rule accounting for the fact that the only reflexive pronouns that can occur in the same clause as the main verb of an imperative sentence are second-person reflexive pronouns. A natural way to state this is by the use of curly-brackets. Thus, such a person might propose the following reflexivization rule:

$$\text{SD:} \quad \begin{Bmatrix} \text{IMP} \\ \text{NP}_1 \end{Bmatrix} - X - \text{NP}_2$$

$$1 \quad - 2 - 3$$

(75)

$$\text{SC:} \quad 1 \quad - 2 - \begin{bmatrix} +\text{REFL} \\ +\text{PRO} \end{bmatrix}$$

Conditions: a) 1 commands 3 and 3 commands 1.
             b) If 1 = IMP, then 3 is second person.
             c) If 1 = NP$_1$, then NP$_1$ is coreferential with NP$_2$.

Such a rule makes a claim about the nonexistence of a generalization. It claims, in effect that the occurrence of yourself in Shave yourself has nothing to do with the occurrence of yourself in You will shave yourself, since the former would arise because of the present of an IMP marker, while the latter would arise due to the occurrence of a coreferential NP. This, of course, is a claim to the effect that it is an accident that this construction, which allows only second person reflexives, happens to be understood as though it had a second person subject.

This formalism for denying the existence of a generalization could, of course, also be used in the case of the Latin examples cited above. Suppose one wated to claim that the distribution of nē and nōn with independent subjunctives had nothing to do with the distribution of nē and nōn with dependent subjunctives. Then one might attempt (sloppily) to write a rule like the following to account for the occurrence of nē.

$$\text{SD:} \quad \begin{Bmatrix} \text{NP V (NP)} \\ \begin{Bmatrix} \text{IMP} \\ \text{VOL} \end{Bmatrix} \end{Bmatrix} - [_S \text{(ut)} - \underline{\text{nōn}} - \text{NP} \quad \begin{matrix} V \\ [+\text{SUBJ}] \end{matrix} \quad X \;]$$

(76)

$$1 \quad - 2 \quad - 3 - \quad 4$$

$$\text{SC:} \quad 1 \quad - 2 \quad - \underline{\text{nē}} - \quad 4$$

Such a rule would make the claim that the occurrence of nē in Nē venias has nothing to do with the occurrence of nē in Imperō nē venias. The former would arise due to the presence of IMP or VOL, while the latter would arise due to the presence of a verb that takes an object complement. The fact that IMP and VOL happen to mean the same thing as verbs that take object complements would have to be considered an accident.

An equivalent formalism for denying the existence of generalizations is the assignment of an arbitrary feature in a disjunctive environment. For example, suppose that instead of deriving nē by (76), we broke (76) up into two parts: (i) A rule assigning the arbitrary feature [+IRVING] as follows:

$$\text{SD:} \quad \begin{Bmatrix} \text{NP V (NP)} \\ \begin{Bmatrix} \text{IMP} \\ \text{VOL} \end{Bmatrix} \end{Bmatrix} - [_S \text{(nōn)} - \begin{matrix} V \\ [+\text{SUBJ}] \end{matrix} - X \;]$$

(77)

$$\text{SC:} \quad 1 \quad - 2 \quad - \begin{matrix} 3 \\ [+\text{IRVING}] \end{matrix} - 4$$

(ii) A rule changing nōn to nē in sentences with [+IRVING] verbs.

(78)

SD:　　X　-　nōn　-　NP　-　　V　　　-　Y
　　　　　　　　　　　　　　　　　　[+IRVING]

　　　　　1　-　2　-　3　-　　4　　-　5

SC:　　1　-　nē　-　3　-　　4　　-　5

One might claim then that one had a completely general rule predicting the oc-
currence of nē: nē occurs with [+IRVING] verbs. As should be obvious, assign-
ing arbitrary features in this fashion is just another way of claiming that no
general rule can be stated.

An excellent example of the use of such a feature-assignment occurs in
Klima's important paper "Negation in English" (in Fodor and Katz [16]) in the
discussion of the feature AFFECT (p. 313 ff.). Klima notes that two rules oc-
cur in certain disjunctive environments (negatives, questions, only, if, before,
than, lest), which seem vaguely to have something semantically in common. Not
being able to provide a precise semantic description of what these environments
have in common, he sets up a number of rules which introduce the "grammatico-
semantic feature" AFFECT, whose meaning is not explicated, in just those envi-
ronments where the rules apply. He then provides a "general" formulation of
the rules in terms of the feature AFFECT. He has, of course, told us no more
than that the rules apply in some disjunctive list of environments, those to
which the arbitrary feature AFFECT has been assigned. If these environments
do have something in common semantically (and I think Klima was right in sug-
gesting that they do), then the general formulation of these rules awaits our
understanding of just what, precisely, they do have in common.

Another example is given by Chomsky ([8], p. 39). In his analysis of the
auxiliary in English, Chomsky says [the numbering is his]:

(29) (ii)　Let Af stand for any of the affixes past, S, Ø, en, ing.
　　　　　　Let v stand for an M or V, or have or be ..... . Then:

$$\underline{Af} + \underline{v} \longrightarrow \underline{v} + \underline{Af} \#$$

Note that 'v' does not stand for the category 'verb', which is represented by
the capital letter 'V'. 'v' and 'V' in this framework are entirely different
symbols having nothing whatever in common, as different as '&' and 'Z'.

Af and v are arbitrary names and might equally well have been called SAM
and PEDRO, since they have no semantic or universal syntactic significance. In
(29) (ii), Chomsky is stating two rules assigning arbitrary names to disjunc-
tive lists of elements, and one rule which inverts the elements to which those
names have been assigned. Since assigning arbitrary features like [AFFECT] is
equivalent to assigning arbitrary names, (29) (ii) can be stated equivalently
in terms of feature-assignment rules. The following rules say exactly the same
thing as (29) (ii).

(79)

SD:　　X　-　⎰ M ⎱　-　Y
　　　　　　　　� V ⎰
　　　　　　　　⎰have⎱
　　　　　　　　⎰ be ⎰

　　　　　1　-　2　　-　3

SC:　　1　-　2　　-　3
　　　　　　　[+PEDRO]

$$\text{(80)} \quad \text{SD: } X - \left\{\begin{matrix} \text{past} \\ \text{s} \\ \emptyset \\ \underline{\text{en}} \\ \underline{\text{ing}} \end{matrix}\right\} - Y$$

$$1 - 2 - 3$$

$$\text{SC: } 1 - \underset{[+\text{SAM}]}{2} - 3$$

$$\text{(81)} \quad \text{SD: } X - [+\text{SAM}] - [+\text{PEDRO}] - Y$$

$$1 - 2 - 3 - 4$$

$$\text{SC: } 1 - 3 - 2 + \# - 4$$

Such a sequence of rules, in effect, makes the claim that there is no general principle governing the inversion of affixes and verbs. All that can be done is to give two lists, assign arbitrary names to the lists, and state the inversion rule in terms of the arbitrary names.

This, of course, is an absurd claim. There should be a general rule for the inversion of affixes and verbs. What should be said is that there is a universal syntactic category 'verb', of which M, V, have, and be are instances, and correspondingly, that there is a subcategory 'auxiliary verb' of which M (modals such as will, can, etc.), have, and be are instances. One would also need a general characterization of the notion 'affix' rather than just a list. The rule would then state that 'affixes' and 'verbs' invert. But this is not what (29) (ii) says. In the Syntactic Structures framework, M, V, have, and be are not all instances of the universal syntactic category 'verb'. They are entirely different entities, having no more in common than Adverb, S, windmill, and into. Since the same process applies to all of them, it is impossible to state this process in a nontrivial uniform way unless M, V, have, and be are instances of the same universal category 'verb', which is what Chomsky's analysis denies.[20] Of course, it is possible to state this process in a trivial uniform way, as in (81).

On the whole, I would say that the major insights of transformational grammar have not come about through embracing rules like (75), which claim that general statements do not exist, but through eschewing such rules wherever possible and seeking out general principles. Devices like curly-brackets may be useful as heuristics when one is trying to organize data at an early stage of ones work, but they are not something to be proud of. Each time one gives a disjunctive list of the environments where a rule applies, one is making a claim that there are no fully general principles determining the application of that rule. Over the years, curly-brackets have had a tendency to disappear as insights were gained into the nature of the phenomena being described. It may well be the case that they will turn out to be no more than artifacts of the methodological necessity of having to organize data in some preliminary fashion and of the theoretical assumption that syntax is autonomous and arbitrary.

REFERENCES

[1] Bach, E. and R. Harms, Universals in Linguistic Theory, Holt, Rinehart, and Winston, 1968.

[2] Bailey, C.-J., B. J. Darden, and A. Davison, Papers from the Fourth Regional Meeting of the Chicago Linguistic Society, Linguistics Department, University of Chicago, 1968.

[3] Belnap, Nuel, "An Analysis of Questions," TM-1287/000/00, Systems Development Corporation, Santa Monica, 1963.

[4] Binnick, R., A. Davison, G. Green, and J. Morgan (eds.), Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, Linguistics Department, University of Chicago, 1969.

[5] Carden, Guy, "Dialects in the Use of Multiple Negations," forthcoming.

[6] Castañeda, Hector-Neri, "'He: A Study in the Logic of Self-consciousness," Ratio, December, 1966.

[7] Chafe, Wallace, "Language as Symbolization," Language, 1967.

[8] Chomsky, Noam, Syntactic Structures, Mouton, 1957.

[9] Chomsky, Noam, Aspects of the Theory of Syntax, M.I.T. Press, 1965.

[10] Chomsky, Noam, "Deep Structure, Surface Structure, and Semantic Interpretation," in this volume.

[11] Curme, George O., Syntax, D. C. Heath, 1931.

[12] Davidson, Donald, "The Logical Form of Action Sentences," in Rescher [63].

[13] DeRijk, Rudolph, "A Note on Prelexical Predicate Raising," M.I.T. mimeo, 1968.

[14] Emonds, Joseph, "A Structure-preserving Constraint on NP-movement Transformations," in Binnick et al. [4].

[15] Fillmore, Charles J., "Types of Lexical Information," in this volume.

[16] Fodor, J. and J. Katz, The Structure of Language, Prentice-Hall, 1964.

[17] Føllesdal, Dagfinn, "Knowledge, Identity, and Existence," Theoria, 1967.

[18] Geach, Peter, Reference and Generality, Cornell University Press, 1962.

[19] Geach, Peter, "On Complex Terms," The Journal of Philosophy, January 7, 1965.

[20] Geach, Peter, "Logical Procedures and the Identity of Expressions," Ratio, 1965.

[21] Geach, Peter, "Intentional Identity," The Journal of Philosophy, October 16, 1967.

[22] Geach, Peter, "Quine's Syntactical Insights," Synthese, 19, 1968-69.

[23] Gruber, Jeffrey, "Functions of the Lexicon in Formal Descriptive Grammars," TM-3770/000/00 Systems Development Corporation, Santa Monica, 1967.

[24] Halliday, M.A.K., "Some Aspects of the Thematic Organization of the English Clause," RM-5224-PR, RAND Corporation, Santa Monica, 1967.

[25] Halliday, M.A.K., "Notes on Transitivity and Theme in English," Journal of Linguistics, 1967.

[26] Halle, Morris, The Sound Pattern of Russian, Mouton, 1959.

[27] Harrah, David, "A Logic of Questions and Answers," Philosophy of Science, Vol. 28.

[28] Harris, Zellig, "Discourse Analysis," Language, 1952. Also in Fodor and Katz, [16].

[29] Harris, Zellig, "Cooccurrence and Transformation in Linguistic Structure," Language, 1957. Also in Fodor and Katz, [16].

[30] Hintikka, Jaakko, "Modality as Referential Multiplicity," Eripianos A jatus, 1957.

[31] Hintikka, Jaakko, Knowledge and Belief, Cornell University Press, 1966.

[32] Jackendoff, Ray, "An Interpretive Theory of Negation," M.I.T. mimeo, 1968.

[33] Jackendoff, Ray, Some Rules for English Semantic Interpretation, M.I.T. dissertation, 1969.

[34] Jacobs, R. and P. Rosenbaum, Readings in English Transformational Grammar, Blaisdell, in press.

[35] Kaplan, David, "Quantifying In," Synthese, 19, 1968-69.

[36] Katz, J. J., and P. M. Postal, An Integrated Theory of Linguistic Descriptions, M.I.T. Press, 1964.

[37] Keenan, Edward, A Logical Base for English, University of Pennsylvania dissertation, 1969.

[38] King, Harold, "On Blocking the Rule for Contraction in English," Linguistic Inquiry, Vol. I, 1970.

[39] Kiparsky, C.A.S. and R.P.V., "Fact," in this volume.

[40] Lakoff, George, On the Nature of Syntactic Irregularity, Indiana University dissertation, 1965.

[41] Lakoff, George, "Pronouns and Reference," Harvard mimeo, 1968.

[42] Lakoff, George, "Repartee," Foundations of Language, to appear.

[43] Lakoff, George, Generative Semantics, Holt, Rinehart, and Winston, forthcoming.

[44] Lakoff, Robin, Abstract Syntax and Latin Complementation, M.I.T. Press, 1968.

[45] Lakoff, Robin, "A Syntactic Argument for Not-transportation," in Binnick et al. [4].

[46] Lakoff, Robin, "One or Two Things I Know about Conjunction," Proceedings of the First Annual Ohio State Semantics Conference, to appear.

[47] McCawley, James D., "Concerning the Base Component of a Transformational Grammar," Foundations of Language, 1968.

[48] McCawley, James D., "Lexical Insertion in a Transformational Grammar without Deep Structure," in Bailey et al. [2].

[49] McCawley, James D., "The Role of Semantics in a Grammar," in Bach and Harms, [1].

[50] McCawley, James D., "A Note on Multiple Negations," University of Chicago mimeo, 1968.

[51] Montague, Richard, "English as a Formal Language," UCLA mimeo, 1968.

[52] Morgan, Jerry L., "On the Treatment of Presupposition in Transformational Grammar," in Binnick et al., [4].

[53] Parsons, Terence D., "A Semantics for English," University of Illinois at Chicago Circle mimeo, 1968.

[54] Parsons, Terence D., "On the Logic of Mass Terms and Quantifiers," University of Illinois at Chicago Circle mimeo, 1968.

[55] Partee, Barbara, "Negation, Conjunction, and Quantifiers," Foundations of Language, to appear.

[56] Perlmutter, David M., Deep and Surface Structure Constraints in English, M.I.T.dissertation, 1968.

[57] Postal, Paul, Crossover Phenomena, IBM mimeo, 1968.

[58] Postal, Paul, "On the Surface Verb Remind," Linguistic Inquiry, Vol. I, 1970.

[59] Prior, A. N., Formal Logic, Oxford, 1955.

[60] Prior, M. and A. N., "Erotetic Logic," Philosophical Review, Vol. 64, 1955.

[61] Prior, A. N., Time and Modality, Oxford, 1957.

[62] Reichenbach, Hans, Elements of Symbolic Logic, Dover, 1947.

[63] Rescher, N. (ed.), The Logic of Decision and Action, University of Pittsburgh Press, 1966.

[64] Rescher, N., The Logic of Commands, Dover, 1966.

[65] Ross, John R., Constraints on Variables in Syntax, M.I.T. dissertation, 1967.

[66] Ross, John R., "On Declarative Sentences," in Jacobs and Rosenbaum [34].
[67] Ross, John R., "Auxiliaries as Main Verbs," in Todd [68].
[68] Todd, William, Philosophical Linguistics, Series I, Great Expectations (Evanston), 1969.

Footnotes

[1]Generative semantics is an outgrowth of transformational grammar as developed by Harris, Chomsky, Lees, Klima, Postal, and others. The generative. semantics position was arrived at through an attempt on the part of such linguists as Postal, Fillmore, Ross, McCawley, Bach, R. Lakoff, Perlmutter, myself, and others to apply consistently the methodology of transformational grammar to an ever-increasing body of data. We have not all reached the same conclusions, and those presented here are only my own. However, I think it is fair to say that there has developed in recent years a general consensus in this group that semantics plays a central role in syntax. The generative semantics position is, in essence, that syntax and semantics cannot be separated and that the role of transformations, and of derivational constraints in general, is to relate semantic representations and surface structures. As is the case of X generative grammar, the term "generative" should be taken to mean "complete and precise".

An earlier version of part of this paper has appeared in Binnick et al. [4]. The present paper is an early draft of some chapters of a book in progress to be called Generative Semantics [43]. I would like to thank R. T. Lakoff, J. D. McCawley, D. M. Perlmutter, P. M. Postal, and J. R. Ross for lengthy and informative discussions out of which most of the material discussed here developed. I would also like to thank Drs. McCawley, Postal, and Ross for reading an earlier draft of this manuscript and suggesting many improvements. Any mistakes are, of course, my own. I would also like to take this opportunity to express my gratitude to Professor Susumu Kuno of Harvard University, who has done much over the past several years to make my research possible. The work was supported in part by grant GS-1934 from the National Science Foundation to Harvard University.

[2]Some readers may have received an incorrect impression of the basic theory with regard to this issue from the confusing discussion of directionality by Chomsky [10]. Chomsky does not claim in that article that any advocates of the basic theory have ever said that directionality matters in any way. However, Chomsky's odd use of quotation marks and technical terms in that paper has led some readers to believe that he had made such a mistaken claim. A close reading of Chomsky's paper should clarify matters. On page 25, Chomsky says:

> ...Let us consider once again the problem of constructing a "semantically-based" theory of generative grammar that is a genuine alternative to the standard theory.

He then outlines a theory, in example (32), containing S (a semantic representation) and P (a phonetic representation), and he correctly notes that it makes no sense to speak of the 'direction' of a derivation from S to P or P to S. He concludes, on the same page:

> Consequently, it is senseless to propose as an alternative to (32) a "semantically-based" conception of grammar in which S is "selected first" and then mapped onto the surface structure P_n and ultimately P.

Here he gives 'directionality' as the defining characteristic of what he calls a "'semantically-based' theory of grammar". His next two sentences are (page 26):

> Consider once again a theory such as that proposed by McCawley in which P_1 is identified with S and condition (3) is dropped so that "deep structure" is undefined. Let us consider again how we might proceed to differentiate this formulation--let us call it "semantically-based grammar"--from the standard theory.

Having first used "'semantically-based' theory of grammar" as a technical term for the 'directionality' position, he then uses "'semantically-based grammar'" as a new technical term to describe McCawley's position. Of course, McCawley has never advocated the 'directionality' position and Chomsky has not said that he has. But one can see how such a bewildering use of technical terms might lead readers to such a mistaken conclusion. The only person that Chomsky cites as being a supporter of the 'directionality' position is Chafe [7], who has never been an advocate of the basic theory. However, Chomsky does not cite any page references, and in reading through Chafe's paper, I have been unable to find any claim to the effect that directionality has empirical consequences.

³It should be noted at the outset that all of the sentences discussed in this section are subject to dialect variation. At least one-third of the speakers I have encountered find (1) and (2) both ambiguous. The sentences to be discussed in the remainder of the section are subject to even greater variation, especially when factors like stress and intonation are studied closely. The data I will present below correspond to what I take to be the majority dialect. I take note of cases where my speech differs from that dialect, and a more thorough discussion of dialect differences will be given in the following section, after the discussion of the general constraints. It is especially important to remember throughout this section that the argument to be presented depends on the existence of a single dialect for which the data presented are correct. It is not even important that the dialect described by these data be the majority dialect, though. so far as I can tell, it is.

⁴I don't mean to suggest that P_1 would look like this in detail. The crucial point here is the relative height of many and few. For a detailed discussion, see Lakoff [43].

⁵I have ignored the role of stress in this discussion, though it is of course important for many speakers. Many people find that (i)

(i) Many men read few books.

where few has extra heavy stress can mean The books that many men read are few. Thus, the general principle here seems to be that where the asymmetric command relation is lost in derived structure, then either one or another of what Langacker calls 'primacy relations' must take over. One which I would propose is the relation 'has much heavier stress than'; the other is the relation 'precedes'. These relations seem to form a hierarchy with respect to this phenomenon in such dialects:

I. Commands (but is not commanded by)

2. Has much heavier stress than

3. Precedes

If one quantifier commands but is not commanded by another in surface struc-
ture, then it commands in underlying structure. If neither commands but is
not commanded by the other in surface structure, then the one with heavier
stress commands in deep structure. And if neither has much heavier stress,
then the one that precedes in surface structure commands in underlying struc-
ture. Letting $C_4 = Q^1$ has much heavier stress than Q^2, the constraint for such
a dialect could be stated as follows, though the notation is not an optimum one
for stating such a hierarchy:

$$P_i/C_i \supset (P_n/C_2 \supset (P_n/C_4 \bigvee P_n/C_3) \;)$$

Since the dialect with this condition is in the minority so far as I have been
able to tell from a very small amount of study, I will confine myself to the
normal dialect in the remainder of the discussion.

[6]Though these facts seem to hold for the majority of speakers I have
asked, they by no means hold for all. There is even one speaker I have found
for whom the crucial case discussed here does not hold. This speaker finds
that dissuade does not work like persuade not with respect to ambiguities. For
him, (42a) can mean not only (39), but also There were many girls that I dis-
suaded Bill from dating, even when (42a) does not contain stress on many. This
speaker seems to have a constraint that holds just at the level of surface
structure, and not at the level of shallow structure or above. Interestingly
enough, the same speaker has the No-double-negative constraint discussed below
in the same form as the majority of other speakers, only for him it holds at
surface structure, not shallow structure. For him, (51) below is ungrammati-
cal, but (54) is grammatical. Perlmutter has suggested that such variation in
the constraints is due to the fact that children learning their native language
are not presented with sufficient data to allow them to determine at which
level of the grammar various constraints hold, or exactly which classes of
items obey which constraints. Other cases of such variation are discussed in
the following section.

[7]I assume that either arises as follows. The underlying structure of a
disjunction is:

(i)
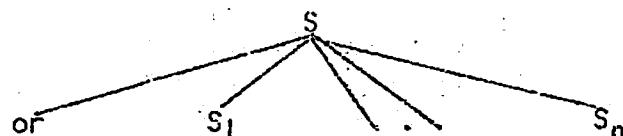
The rule of or-copying yields:

(ii)
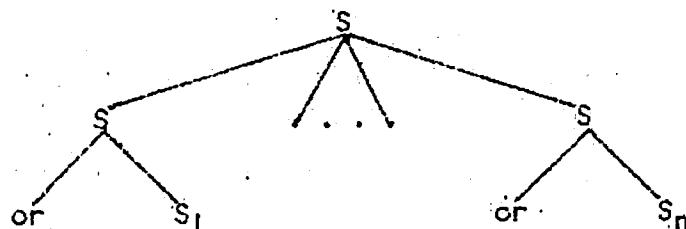
Then the leftmost or changes to either. In initial position in a sentence, either optionally deletes. All or's except the last optionally delete. And works in a similar fashion.

[8]See Fodor and Katz ([16]), p. 503 ff.

[9]See Fodor and Katz ([16]), p. 531 ff.

[10]The "standard theory" is quite different in this respect from the theory assumed by Jackendoff ([33]), who insists upon making no assumptions whatever about the nature of semantic representation. Moreover, it is not clear that Chomsky ever seriously maintained the "standard theory" as described in the passage quoted, since the main innovations of that theory—allowing semantic representations to be given in terms of phrase-markers (and thus ruling out Katzian semantics), allowing prelexical transformations, and allowing lexical semantic readings to be given as substructures of derived phrase-markers—were only made in the context of an argument to the effect that these innovations made by McCawley ([48]) and others were not new, but were simply variants of the "standard theory". Since Chomsky does not attempt to justify this innovation, and since he does not mention it outside the context of this argument, it is not clear that he ever took such an account of the standard theory seriously. In fact, it is not clear than anyone has ever held the "standard theory". Nonetheless, this theory is useful for pinpointing certain important issues in the theory of grammar, and we will use it for this purpose in subsequent discussion.

[11]It should be observed, incidentally, that arguments like Postal's do not depend on complete synonymy (as in the case of remind and perceive-similar) but only on the inclusion of meaning. As long as the remind sentences contain the meaning of (9), the appropriate rules would apply and the argument would go through. If it should turn out to be the case that remind contains extra elements of meaning in addition to (9), it would be irrelevant to Postal's argument. A mistake of this sort was made in an otherwise excellent paper by De-Rijk ([13]), who considered examples like John forgot X and John ceased to know X, with respect to the proposal made by McCawley, ([48]). DeRijk correctly notes that if X = his native language, nonsynonymous sentences result. He concludes that forget could not be derived from an underlying structure containing the meaning of cease to know. It would be true to say that forget cannot be derived from an underlying structure containing only the meaning of cease to know, since forget means to cease to know due to a change in the mental state of the subject. But McCawley's conjecture, like Postal's argument, only requires that the meaning of cease to know only be contained in the meaning of forget, which it is.

[12]For a small (and arbitrarily chosen) sample of such works see Reichenbach [62], Prior [59] - [61], Geach [18] - [22], Montague [51], Parsons [53], [54], Hintikka [30], [31], Davidson [12], Todd [66], Castañeda [5], Føllesdal [17], Rescher [63], [64], Belnap [3], Keenan [37], and Kaplan [35].

[13]The similarities between universal quantifiers and conjunctions on the one hand and existential quantifiers and disjunctions on the other hand have been recognized at least since Pierce, and various notations have been concocted to reflect these similarities. Thus, universal quantification and conjunction might be represented as in (1 a) and existential quantification and ·

disjunction as in (1 b).

(1) (a) (i) $\bigwedge_X$ fx

(ii) $\bigwedge$ $P_1 \ldots P_n$

(b) (i) $\bigvee_X$ fx

(ii) $\bigvee$ $P_1 \ldots P_n$

Such a similarity of notation makes clear part of the obvious relationship between the quantifier equivalence of (II a) and DeMorgan's Law (II b).

(II) (a) $\sim \bigwedge_X fx \equiv \bigvee_X \sim fx$

(b) $\sim \bigwedge P_1, \ldots, P_n \equiv \bigvee \sim P_1, \ldots, \sim P_n$

In the case where x ranges over a finite set, (II a) says the same thing as (II b). Yet despite the similarity in notation, the (ii) cases in (i) are not represented as special cases of the (i)'s and (II a) and (II b) are two distinct statements. There is no known notational system in which the (ii)'s are special cases of the (i)'s and in which (II a) and (II b) can be stated as a single equivalence, though it seems that the same thing is going on in (II a) and (II b).

It should also be clear that McCawley's observations are not unrelated to the fact that performative verbs cannot be negated and still remain performative. ("I do not order you to go" is not an order.) Since a conjunction of negatives is equivalent to the negative of a disjunction, by DeMorgan's Laws, it would seem that an adequate account of McCawley's observations should show how the impossibility of disjunctions of performatives follows from the impossibility of negatives of performatives.

[14]Note, however, that sentences like the following are possible:

Someone and Sam left.

Will everyone or just Sam come to the party?

[15]At the UCLA conference on historical linguistics, January, 1969.

[16]The same facts obtain for the author's native (Northern New Jersey) dialect. Curme ([11], p. 183) observes that the same phenomenon occurs in popular Irish English.

[17]The original argument was brought up by B. Mates, 'Synonymity', in Meaning and Interpretation, Berkeley, 1950, pp. 201-206. Replies were made by A. Church, 'Intensional Isomorphism and Identity of Belief', Philosophical Studies, 5, 1954, pp. 65-73 and by W. Sellars, 'Putnam on Synonymity and Belief', Analysis 15, 1955, pp. 117-20. I find it hard to see how one could know the meaning of the expression "man who has never been married" and the meaning of the expression "bachelor", and how one could believe the proposition expressed by "John is a bachelor" without also believing the proposition expressed by "John is a man who has never been married". I could see how one, upon hearing these sentences, might have perceptual or processing difficulties and so might not be able to figure out which propositions correspond to which sentences. But in the normal sense of believe and realize, one believes or realizes propositions expressed by sentences, not sentences themselves. There is, of course, another sense of believe which means roughly have blind faith in the truth of. In this sense, a speaker may fail to understand a sentence at all, but may believe whatever proposition is supposed to be expressed by that sentence, e.g., "God is good." In this sense, one may 'believe' all and only those sentences which have been uttered in a Southern accent by a military officer over the rank of colonel, who has red hair and is smoking a cigar at the time of the utterance. Thus, any aspect of phonetic or contextual difference may matter for this sense of the meaning of 'believe'. If one wants to consider such facts as part of the theory of meaning, ones theory of meaning will have to coincide with ones theory of language use.

[18]The arbitrary syntax position originated in the practice of Katz and Postal [36], and has been adopted by many investigators since then. It is interesting that Katz and Postal considered a position very close to the natural syntax position and saw that it had advantages over the descriptive practice that they decided to adopt. As they say on page 149, fn 9:

> On the basis of (41) - (44) plus the fact that there are no sentences like *I request that you want to go, *I request that you hope to be famous, a case can be made for deriving imperatives syntactically from sentences of the form I Verb$_{request}$ that you will Main Verb

by dropping at least the first three elements. This would account not only for (41)-(44) but also for the facts represented in (35)-(40). Such a derivation would permit dispensing with I and its reading RIM would simplify the semantic component by eliminating one entry. It would also eliminate from the syntax all the necessary heavy selectional restrictions on I and the rules that must introduce this element. Although we do not adopt this description here, it certainly deserves further study.

[19]Examples like this also occur in English. Compare

(i) Ah, to be able to insult my boss!

(ii) Ah, being able to insult my boss!

(i) presupposes that the speaker is not able to insult his boss, while (ii) presupposes that he is able to insult his boss. R. Lakoff has observed that this follows from the fact pointed out by Kiparsky and Kiparsky [39] that factive verbs take poss-ing complementizers, while for-to complementizers occur

only with nonfactive verbs. Since (i) expresses a wish and (ii), a liking, the occurrence of the complementizers could be predicted if abstract verbs bearing those meanings were hypothesized as occurring in the underlying syntactic structure of these sentences. As in the Latin examples, the distribution of the complementizing morphemes depends on the meaning of the understood predicates.

[20]Ross [67] has given an analysis of auxiliaries as instances of the category 'verb', in terms of which general rules mentioning M, V, have, and be can be stated as general rules, rather than as notational variants of lists. Moreover, according to Ross' analysis, the rule of affix-verb inversion can be eliminated in favor of the independently needed rule of complementizer placement.