SOME RULES OF SEMANTIC INTERPRETATION FOR ENGLISH

by

RAY S. JACKENDOFF

B.A., Swarthmore College
(1965)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF
PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF
TECHNOLOGY

June, 1969

Signature . . . .
Department of Modern Languages   June 13, 1969
and Linguistics   *i.e. Foreign Literatures + Linguistics*

Certified by . . . .
Thesis Supervisor

Accepted by . . . .
Chairman, Departmental Committee
on Graduate Students

SOME RULES OF SEMANTIC INTERPRETATION FOR ENGLISH

by

RAY S. JACKENDOFF

Submitted to the Department of Modern Languages and Linguistics on June 13, 1969 in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

ABSTRACT

This thesis critically examines the hypothesis that semantic readings are determined solely on the basis of deep structure. Several examples are presented to show that certain semantic generalizations actually involve derived structure, and a theory of the semantic component is developed which allows for certain aspects of semantic interpretation to be determined from derived structure. Four grammatical phenomena--pronominalization and reflexivization, complement subject deletion, negation and quantifiers, and adverbs--are analyzed from the point of view of this theory, and it is shown that it has significant advantages over the standard approach to these problems.

Thesis Supervisor: Noam Chomsky
Title: Professor of Linguistics

ACKNOWLEDGEMENTS

Doing the heretical work in this thesis has been, to say the least, traumatic (although perhaps no more so than anybody else's thesis is). During the three years I have been messing around in the semantic component, many people have helped me clarify issues, find and/or solve potential problems, and just plain let off steam. Here are some that stand out in my mind, in no particular order: Joe Emonds, Bill Watt, Janet Dean, Peter Culicover, Adrian Akmajian, John Bowers, Steve Anderson, Dick Stanley, Jay Keyser, and Jerry Katz. To them, my profound thanks.

In addition, I think it only fair to thank my opponents Haj Ross, George Lakoff, Paul Postal, and Guy Carden for doing the work that made me mad enough to start working on most of the topics in this thesis, and for providing an inexhausible mine of facts for long winter evenings of contemplation.

Deepest gratitude goes to my teachers Noam Chomsky, Morris Halle, and Edward Klima, who by now must be getting tired of being thanked. It can't be helped--I owe to them whatever grasp I may have of how to do linguistics.

Finally, thanks to Katy Freeman, who typed the thing.

TABLE OF CONTENTS

CHAPTER I

A THEORY OF THE SEMANTIC COMPONENT

1. The problem of semantics in grammar

In the early days of transformational grammar, it was not
clear what relation, if any, the generative rules were supposed
to bear to meaning. With the publication of Katz and Fodor's
"The Structure of a Semantic Theory" (1963) it became one of
the explicit goals of generative grammar to take meaning into
account. In this paper Katz and Fodor argued that the grammar
should be thought of as a system of rules relating the sentences
of a language in their externalized form to their meanings.

If a grammar is to relate sound to meaning, it must have
some way to talk about meanings. Katz and Fodor postulated
that meanings are to be expressed in a universal semantic repre-
sentation, in the same way that sounds are expressed in a
universal phonetic representation. Universality is necessary
so that representations are language-independent: we must be
able to compare meanings of sentences across languages. Put
more strongly, to suppose a universal semantic representation
is to make a strong claim about the innateness of semantic
structure. Presumably the semantic representation is very closely
integrated into the whole cognitive apparatus of the mind.

Of course, unlike phonetic representations, semantic repre-

sentations are only very indirectly accessible. It is fairly easy to talk about sameness and difference of meaning, but meaning itself, as generations of philosophers have known, is elusive. Thus the study of semantics has always been somewhat derivative, indirect, and fuzzy. It was Katz and Fodor's hope that by making semantics an explicit part of generative grammar, one could make more incisive studies of meaning.

It has been generally assumed that semantic representations are not identical to syntactic structures (although some recent work, especially by McCawley, denies this). There must therefore be a new part of the grammar, a semantic component, to relate meanings to the syntax and phonology. This is the content of the slogan in Katz and Fodor, "Linguistic description minus grammar equals semantics" (where "linguistic description" and "grammar" are used in my sense of "grammar" and "syntax and phonology," respectively). This phraseology, however, is unfortunate: the inaccessibility of semantic information coupled with the availability of sophisticated syntactic formalism has tended to bias research heavily in favor of syntactic solutions to problems. Thus Katz and Fodor's slogan very rapidly acquired the negative connotation "Semantics is whatever you have to shove under the rug."

It is the intent of this investigation to begin to right this imbalance. We will discuss a number of problems, heretofore treated as syntactic, and try to justify treating them as part of the semantic component, giving analyses comparable in

rigor to current transformational formulations. To lay the groundwork, we will first go over the development of the currently accepted framework and then propose an alternative.

2. The Katz-Fodor framework for a semantic component

The semantic component proposed by Katz and Fodor consists of two parts, a lexicon and a set of projection rules. The lexicon contains a list of all the lexical items of the language together with their semantic readings. In subsequent work, such as Aspects of the Theory of Syntax (Chomsky (1965)), the lexicon has been taken to include syntactic and phonological information about lexical items as well. Thus the lexicon can be thought of as a repository of all the idiosyncratic information about lexical items.

Much of the discussion in Katz and Fodor's paper is concerned with the proper way to represent the internal structure of lexical items with multiple readings, such as bachelor, and with formalisms to capture facts about analyticity and anomaly. Their proposals have been extensively criticized by Bolinger (1965), Weinreich (1966), and Bierwisch (1967), and defended by Katz (1967). Since these claims were the most substantive ones made in the original paper, discussion has continued to center about them, to the detriment of the rest of the semantic component. In the present investigation, we will remain largely outside of this discussion; we will just need the assumption that the lexicon makes some specification for each reading of

each lexical item. In addition, we will assume that there is some way of capturing regularities within the lexicon (see §9).

The other part of the Katz and Fodor semantic component is the set of projection rules. These rules contribute to the semantic representation of a sentence those parts of its content and organization not due to the lexical items in the sentence, i.e. the part of the interpretation traceable to the syntactic structure. The projection rules are the semantic component proper: since the lexicon, in current theory, contributes phonological and syntactic information as well as semantic, it really belongs set off by itself.

Katz and Fodor assume that the projection rules are interpretive, that is, that they operate on the syntactic structures generated by the base structures and transformations to produce semantic readings. This is the conception of projection rules which almost all subsequent work on the semantic component has assumed. Some recent work which proposes that the semantic component be the generative part of the grammar will be mentioned in §6.

The approach to projection rules in Katz and Fodor and in subsequent work (until the topic ceased to be discussed around 1965 or so) is something like the following: the projection rules describe the contribution of syntactic structure to meaning. The syntactic structure of a sentence is generated by the application of a sequence of rules--first phrase-structure rules, then

transformations. The differences in structure between two sentences are produced by differences in the sequences of rules which generate the sentences. Since two sentences containing the same lexical items can differ in meaning only if different sequences of rules have applied, it seemed natural to Katz and Fodor to attribute the meaning contributed by the structure to the operations of the rules themselves. That is, for each phrase structure rule and transformation in the grammar, there would be an associated projection rule, telling how the phrase structure rule or transformation contributes to the meaning of sentences in which it operates. Thus projection rules can be divided into two classes, those that are associated with phrase structure rules (called Type 1 projection rules or P1) and those associated with transformations (called Type 2 projection rules or P2).

The Type 1 projection rules create readings for a tree by starting at the lowest constituents in the tree and successively amalgamating readings of sister constituents to produce a reading for the mother constituent. Eventually when the readings of all constituents have been amalgamated, there is a reading associated with the highest node, S. For each phrase structure rule telling how to expand a node into its constituents, there is a projection rule telling how to amalgamate the readings of the constituents to form a reading for the higher node.

There are two examples of P1 rules in Katz and Fodor. They have been elaborated in subsequent work (for example Katz

(1966)), but no essentially different P1 rules have been proposed in the literature. The first rule is the rule of attribution--the projection rule that functions in the case of modifier-head relationships such as adjective-noun combinations and possibly across the copula. Essentially this rule takes the union of the semantic markers of the modifier and the head to be the set of markers for the higher node.

The other projection rule discussed is that involved in combining the interpretation of a verb with the interpretation of its subject or object. We can think of a transitive verb, for example, as being semantically represented as a two-place function $f(x, y)$, where f represents the semantic content of the verb itself and x and y are variables. The operation of the projection rule for subjects and objects has the effect of substituting for x and y the constants, say a and b, which represent the semantic readings of the subject and object. The resulting constant, $f(a, b)$, is thus the semantic representation for the entire sentence.

Type 2 projection rules are much less substantively discussed. Presumably a P2 rule shows how a given transformation changes meaning. At the time of the Katz and Fodor article, there was one place where P2 rules were absolutely essential. At that time, the combining of kernel sentences into complex structures was regarded as a transformational operation; thus P2 rules were needed to explain how the meaning of an embedded

kernel affected the meaning of the kernel into which it was embedded. Since then, the theory has been modified so that recursion is a property of the base, so this particular type of P2 rule is no longer necessary.

It should be noticed that all the projection rules actually proposed by Katz and Fodor are purely combinatorial in nature--they do little but insure that all the features of the lexical items eventually get taken into account in the interpretation of the sentence. As a result of this, linguists and philosophers alike have tended to regard the semantic component as a rather trivial affair, merely a simple notational convention for treating phrase-markers as semantic representation. Interesting discussions of the semantic analysis of lexical items have taken place, but investigation of the rich possibilities inherent in the semantic component itself has not been forthcoming, and discussion of projection rules has largely ceased.

3. The Katz-Postal hypothesis

Whatever desultory discussion of projection rules there may have been after the Katz and Fodor article was published, it seems to have been cut off by the publication of Katz and Postal's An Integrated Theory of Linguistic Descriptions in 1964.

Katz and Postal set out to show that P2 rules are unnecessary, in other words, that there are no changes of meaning induced by performing a transformation. First consider obligatory

transformations. Under the Katz-Fodor conception of projection rules, differences in meaning must be associated with choice-points in the derivation. Since there is no option as to whether to perform an obligatory transformation or not, there can never be two sentences differing in meaning solely because of the operation of the transformation. Therefore there can be no P2 rules associated with obligatory transformations. This leaves two cases: optional singulary transformations and generalized, or embedding transformations.

Optional singularies fall into three types: deletion, insertion, and order-changing rules. Deletion rules can be kept from changing meaning if we propose the condition on transformations of recoverability of deletion. The known insertion rules all insert meaningless particles such as do and number morphemes on verbs, so they don't change meaning. Order-changing rules preserve understood grammatical relations. Since it is an inherent assumption of the Katz-Fodor style semantic component that grammatical relations (as generated by phrase-structure rules and interpreted by P1 rules) are the only kind of structural information relevant in semantic interpretation,[1] order-changing transformations preserve meaning as well. (We will turn to Katz and Postal's discussion of counterexamples shortly.) As for generalized transformations, the condition of recoverability of deletion guarantees that a kernel will be embedded in place of a dummy symbol, so no meaning is added by the embedding transform-

ation other than the pure combining of features in the proper manner. The generalized transformations can be dispensed with altogether if we assign recursion to the base instead, in a fashion which has since Katz and Postal become standard and is exemplified in, for example, Aspects of the Theory of Syntax.

Thus empirically it is likely that we need no P2 rules. Furthermore, Katz and Postal show that stating a P2 rule is tantamount to restating the operation of the transformation with which it is associated, and so the whole concept of P2 is rather suspicious. Therefore, given the three logically possible alternatives (1)-(3),

(1) No correctly formulated transformations change meaning (i.e. there are no P2 rules).
(2) All transformations change meaning.
(3) Some transformations change meaning and some don't and (perhaps) there is some feature of a transformation which tells you whether it changes meaning or not.

the first is of course best on the basis of the evidence given thus far. If there are no P2 , then all the burden of semantic interpretation falls on P1, which operate on phrase-markers.

Which phrase-markers in the derivation should P1 operate on? Again there are three possible alternatives which suggest themselves:

(4) P1 only operate on the final P-markers.
(5) They operate on all levels of P-markers.
(6) They operate on the underlying P-markers.

Katz and Postal then proceed through a series of arguments (pp. 34-45) to show that the effect of transformations is to reduce the structure of underlying P-markers and to distort underlying

grammatical relations, making it implausible for Pl to operate
on derived structure. On p. 45 we find the conclusion

> (7)   If Pl cannot operate on derived P-markers produced
>   by singular transformations, and if all the infor-
>   mation required for the operation of Pl is found in
>   underlying P-markers, then the most efficient way
>   to characterize operation of Pl is simply to require
>   in such cases that Pl operate on all, and only, the
>   underlying P-markers.

This statement is indeed correct, taken as an implication. We
will shortly call into question the premises of this implication,
which Katz and Postal spend most of the book trying to justify.

Assuming the correctness of their arguments, we reach the
following conclusion:

> Katz-Postal Hypothesis, weak form (KP1):
> Semantic projection rules operate exclusively on under-
> lying phrase-markers; hence transformations do not
> change meaning.

From the conclusion that all information required for the operation
of projection rules is present in underlying structure, it is a
simple rhetorical step to

> Katz-Postal Hypothesis, strong form (KP2):
> All semantic information is represented in underlying
> structure.

KP2 follows from KP1 only if we assume that the projection rules
cannot in themselves contribute any meaning. From the kinds of
rules proposed at that time, this seemed likely.

Thus we see that given the Katz-Fodor concept of projection
rules as associated with phrase-structure rules and transformations,
the choice-points in the generation of sentences, and given the
limited power proposed for projection rules, Katz and Postal make

a very good case for even the strong form of their hypothesis.
The hypothesis was universally adopted by generative grammarians,
and subsequent work until very recently used the hypothesis as
a very powerful constraint on possible analyses.

4. Does KP2 follow from KP1?

KP2 is a very strong claim. Just how strong it is can
be seen from the following paraphrase: Every non-lexical semantic
difference must be represented as a deep structure difference.
But even this strong a claim seems to have been nearly universally
accepted without question.

KP2 does indeed follow from KP1 if the semantic component
is composed entirely of very simple rules of the kind proposed
by Katz and Fodor. Suppose, however, that some projection rules
actually _add_ meaning, that is, they perform more than the purely
combinatorial function of the two rules proposed by Katz and
Fodor. Suppose in addition that there are two projection rules
whose ranges of application partially overlap. Such a situation
is not ruled out _a priori_ if we assume projection rules to
operate on structures and not derivations. If such a situation
arose, there would be sentences which are structurally unambiguous
but with two semantic interpretations. This situation would be
consistent with KP1 but not with KP2, showing that the two hy-
potheses are not equivalent.

One case of such a system of projection rules might be the
rules that interpret a sentence as generic or non-generic. Usually

generic sentences can be identified by the use of a peculiar aspect of the verb.

(8) A beaver builds dams. (generic)
(9) A beaver is building dams. (non-generic)

However, there are cases when the two forms coincide.

(10) The King of Andorra was an important figure.

This sentence can be read as referring to a particular King of Andorra, or to whoever might be the King of Andorra over a long period of time. Yet there is no obvious lexical or structural ambiguity.

If one accepts KP2, one is compelled to hunt for two separate deep structures for (10), with no motivation other than the semantic interpretation. But then the forms (8) and (9) must also be put into underlying forms parallel to those of (10), and at least one of them, say for argument (8), must undergo a transformation before coming out in its surface form. The phrase structure rules as they now stand will generate in addition a deep structure D which only has to undergo verb agreement and other trivial rules to become the surface form (8). What is the semantic interpretation of D? To prevent ambiguity of (8), D must be ruled out on some grounds. In proposals on generics that I know of,[2] non-obvious underlying structures are given for generic sentences, but no mention is ever made of how one gets rid of the "natural" deep structures for these sentences.

Assume instead that there is a semantic rule R_1 which interprets a sentence as generic under certain conditions, and

another semantic rule R_2 which interprets a sentence as non-generic under certain conditions, and that the conditions for application of R_1 and R_2 overlap here and there. Then we would expect ambiguous sentences like (10) to crop up somewhere in the language. With this solution, there is no need to go beyond the "natural" deep structures for either (8) or (10).

The two solutions will take approximately the same amount of machinery to state: in both proposals it must be specified which particular combinations of tense and aspect are interpreted as generic. This is the aspect of English generics that is accidental. However, if a transformation produces generic surface structures, it must only be an accident that they look like non-generic surface structures; the grammar would be no more expensive if the subject wound up someplace else or some new aspect marker were introduced. But this similarity, one feels, is no accident, but rather a natural outcome of the syntactic distribution of subjects and aspects. In the solution with projection rules, on the other hand, we can capture the syntactic generality of the position of the subject and of the distribution of tense and aspect directly, in the base. Thus in a formulation with projection rules, one can foresee the possibility of separating the accidental from the motivated aspects of generic sentences, which seems impossible in the transformational solution. In the absence of actual proposed solutions, the case must rest here. What has been shown is that there is a conceivable solution to generics which violates KP2 but which on important independent grounds is

superior to a solution which conforms to KP2. Hence KP2, the universally accepted form of the hypothesis, does not follow from KP1, the form Katz and Postal state and claim to prove.

5. Counterexamples to KP1

After presenting their hypothesis, Katz and Postal devote the rest of the book to dealing with apparent counterexamples. The first two cases they discuss deserve attention here.

The first case is the passive transformation. It was pointed out very early in the history of generative grammar that if the passive is performed on a sentence containing quantifiers, the reading (or the _preferred_ reading) of the sentence changes:

(11) Everyone in this room knows at least two languages.
(12) At least two languages are known by everyone in this room.

The difference in meaning is simply that the understood order of quantifiers is reversed in the passive sentence.

Katz and Postal first assert that both readings are present in both sentences, thus denying that the passive has changed meaning. Then (pp. 72-73) they argue that even if passive does change meaning, the difference can be attributed to the presence of the independently motivated marker by+Passive in the underlying structure.

Their observations about the readings of (11) and (12) are correct. But the following examples show more effectively that active and passive need not be synonymous.

(13) No one in this room knows many languages, but
{(a) one or two of us know a few languages.}
{(b) *one or two of us know many others.}

(14)   Many languages are known by no one in this room, but
$\left\{\begin{array}{ll}\text{(a)} & \text{a few languages are known by one or two of us.}\\ \text{(b)} & \text{many others are known by one or two of us.}\end{array}\right\}$

(13a) and (14a) are definitely not synonymous, and (13b) is very strange, whereas its passive (14b) is quite acceptable and understandable.

A different type of example is the following pair of sentences.[3]

(15)   John saw two pieces of paper.
(16)   Two pieces of paper were seen by John.

Now assume that John is hallucinating, and that there is only one piece of paper, or none at all, or an umbrella, in John's field of vision. For many speakers, only (15) can describe this situation; (16) implicitly assumes that there really are two pieces of paper in John's field of vision. These examples appear to show that there are cases where passive changes meaning. Still more examples will appear in a moment.

As will be shown in Chapter 4, the difference in understood order of quantifiers is a surface structure phenomenon. The fact that the passive can change order of quantifiers is not something peculiar to the passive; rather this is a general property of movement rules. Other order changing rules, such as adverb preposing and subject-aux inversion will be shown to have similar effects. Thus the argument that the difference in meaning induced by the passive can be blamed on the by+Passive marker misses the point: the passive changes meaning in ways that cannot be characterized by a more or less arbitrary mark in deep structure.

The second case Katz and Postal discuss is the analysis
of negation. In the grammar of <u>Syntactic Structures</u>, negation
was merely added by an optional transformation, which obviously
changed meaning. To answer this, Katz and Postal remark (p. 74)

> It is interesting that, quite independently of sem-
> antic considerations, certain more recent descriptions
> of English have found motivations for descriptions of
> some of these facts which are not incompatible with the
> view that projection rules operate exclusively on under-
> lying P-markers. In particular, both Lees [(1960)] and
> Klima [(1964)] have found it necessary to describe negative
> sentences by generating a negative morpheme in the phrase
> structure. But under this interpretation, the projection
> rules which operate on the readings for the negative mor-
> pheme need refer only to the underlying P-markers.

Katz and Postal neglect the fact that even with the negative
morpheme, Klima's system of transformations is not meaning-pre-
serving: he derives (17) and (18) from the same underlying form.

(17) Not much shrapnel hit the soldier.
(18) Much shrapnel didn't hit the soldier.

As will be shown in Chapter 4, subsequent papers by Lakoff (1965)
and Carden (1968), which purport to rectify this problem, do not
succeed.

Another case where transformations affect meaning is re-
lations of coreference among noun phrases. This aspect of grammar
had hardly been investigated at all at the time of Katz and
Postal's book, so it is no wonder that they do not discuss it.
The most obvious cases of coreferentiality are pronominalization
and reflexivization; it is easy to conceive of transformations
that, at least superficially, account for these processes. How-
ever, pronominalization and reflexivization are only part of a

much more general phenomenon which is difficult to conceive of as transformational. For example, words such as _else_, _other_, _remainder_, and epithets such as _the bastard_ and _the poor guy_ take part in processes that are very similar to pronominalization, yet they cannot be derived by simple transformational processes, as can pronouns.

It is easy to see that the proper reading of these referential words is crucially dependent on derived structure. Again I will pick on the passive as the meaning-changing culprit:

(19) One lion attacked four other lions.
(20) *Another lion attacked four lions.

(21) *Four other lions were attacked by one lion.
(22) Four lions were attacked by another lion.

Many examples of this type, and some much less transparent ones, are given in Dougherty (1968b). We will only point out here that if these referential terms cannot be produced transformationally, then they must be generated by the base in their surface forms. But then the operation of an order-changing rule such as the passive determines whether the sentence has a semantic reading or not. If this is the case, not all aspects of semantic interpretation can be determined from the underlying structure, and so KP1 is falsified.

Another counterexample to KP1 has to do with the focus and presupposition of sentences. Chomsky (to appear) discusses focus and presupposition in yes-no questions, calling the part understood as being questioned the _focus_ of a question. Typical of his examples is the following: consider the range of possible

natural responses to these questions, read with normal intonation:

(23) Was it an ex-convict with a red shirt that he was warned to look out for?

(24) Was it a red-shirted ex-convict that he was warned to look out for?

(25) Was it an ex-convict with a shirt that is red that he was warned to look out for?

There are many sentences which are natural responses to all three of these, such as

(26) No, he was warned to look out for an automobile salesman.

But other sentences are more restricted: (27) seems natural only as a response to (23), (28) to (24), and (29) to (25).

(27) No, he was warned to look out for an ex-convict with a red tie.

(28) No, he was warned to look out for a red-shirted automobile salesman.

(29) No, he was warned to look out for an ex-convict with a shirt that is green.

It turns out that the focus must be a constituent containing the stress center of the sentences. Thus the difference between (23), (24), and (25), which are generally assumed to be transformationally related, is that the regular rules of stress assignment place the stress center on different words because of the difference in surface structure order. As Chomsky shows, it is only by extremely artificial means that this difference can be represented in the deep structure. Any attempt to deal with the semantic interpretation of focus entirely on the basis of underlying structure, in accordance with KP1, must miss generalizations or resort to introducing elements into deep structure of extremely dubious status.

One respect in which all of these counterexamples are
similar is that the aspect of semantic interpretation changed
by transformations has nothing to do with grammatical relations.
Katz and Postal are apparently correct in their claim that
transformations do not change understood grammatical relations
(e.g. understood subjects and objects of verbs). But what is
incorrect is their assumption that all semantic interpretation
is based on grammatical relations. In §7 we will present an
alternative.

The three examples just presented--quantifiers and nega-
tion, reference, and focus--show that not all aspects of semantic
interpretation can be derived from underlying phrase-markers.
However, the variety of ways in which a transformation, for ex-
ample the passive, can change meaning indicates that we cannot
specify a Type 2 projection rule that states just how the rule
changes meaning. Instead, as will be shown, the correct general-
izations have to do with derived structure configurations.

In the Katz-Fodor conception of the semantic component,
such a generalization cannot be stated. In that framework, pro-
jection rules are associated with syntactic rules of the grammar--
phrase-structure rules and transformations, each projection rule
telling how performing the associated syntactic rule affects the
interpretation of the sentences. Thus the generalizations about
meaning are predicted to be about rules rather than structural
configurations. In the case of underlying (base) structures, the
two views are entirely equivalent, because corresponding to each

structural configuration is a sequence of phrase structure rules
describing how to generate the configuration, and there is an
isomorphism between derivations and structures. In the case of
derived structures, however, there is no such correspondence,
since there are no phrase structure rules in the grammar that
generate derived structures. Thus if some semantic rules involve
derived structures, and none of them involves transformational
derivations, we can capture a generalization about semantic rules
if and only if we associate all of them with structural configu-
rations and not derivations.[4]

6. Is deep structure significant?

If one accepts both parts of the Katz-Postal hypothesis,
one must represent all meaning differences in deep structure.
Since the kind of projection rules implied by KP2 cannot intro-
duce meaning or restructure intermediate semantic representations,
deep structures must represent the logical structure of sentences
rather than the syntactic structure; all presuppositions must be
explicitly represented, and numerous hypothetical lexical items
such as causative and inchoative must be introduced. Meaning
differences, being numerous and easy to find, become the focus of
study: proposing an adequate deep structure must always come
first in the analysis, regardless of the possibly horrendous con-
sequences in the transformational component.

And suddenly one realizes that one is doing semantics, and
the deep structures one is looking for are in some sense semantic

representations. This has to be the case, given the rudimentary
possibilities inherent in a Katz-Fodor semantic component con-
strained by KP2. This leads directly to the idea of dispensing
with deep structure altogether and generating instead semantic
representations, proceeding directly to surface structure. Such
a proposal has been made by McCawley (1968a, 1968b, 1969), and
some of the ramifications are already being explored extensively.

The possibility of doing generative semantics at all is
of course crucially dependent on the assumption of KP1, which
guarantees that there is a single straight continuous line from
semantic interpretation to surface structure. The direction of
derivation is irrelevant: the point of issue is the existence
of an independent level of deep structure somewhere on the line
of derivation between semantic interpretation and surface struc-
ture. But if it is true that some aspects of semantic inter-
pretation are dependent only on derived structure, we arrive at
a non-straight line conception of the grammar: instead of a
schematic like (30), we have (31).

$$
\begin{array}{c}
\text{Semantic interpretation} \\
\uparrow \quad \begin{array}{l}\text{Semantic} \\ \text{component}\end{array} \\
(30) \quad \text{Base rules} \longrightarrow \quad \text{Deep structures} \\
\downarrow \quad \begin{array}{l}\text{Transformational} \\ \text{component}\end{array} \\
\text{Surface structures}
\end{array}
$$

Base rules $\longrightarrow$ Deep structures $\longrightarrow$

(31)   Transformational
         component

         Surface structures

Semantic
interpretation
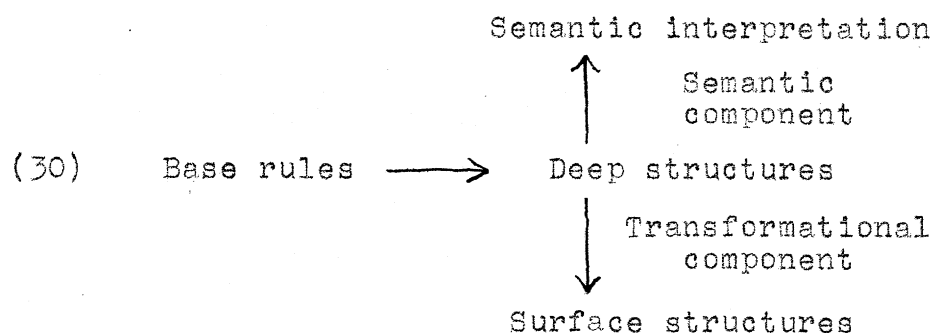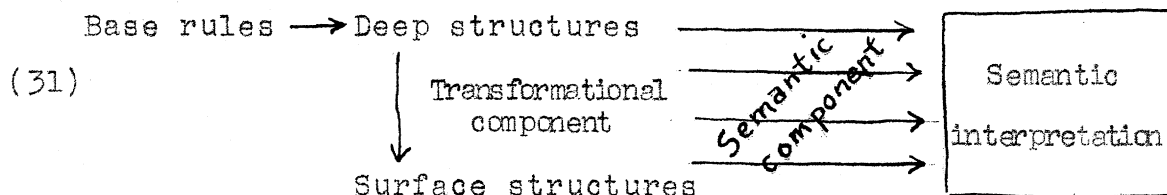
If (31) represents the organization of the grammar, it cannot be claimed that the "deep structures" we have been searching for are actually semantic representations, and thus that the level of deep structure can be dispensed with. For (31) this claim is incomprehensible, since the output of the base cannot contain all semantic information.

There are a couple of intuitive biases which many people seem to have about the organization of the grammar, which it might be well to dispel at this point. (31) seems to be somehow a re-pugnant way of organizing the grammar. To be sure, it is more complicated than (30), but its correctness is an empirical question, not one of intuition.

The straight-line model of grammar is admittedly very appealing in that it makes a performance model look easier to construct, particularly if we reverse the direction of the upper arrow in (30). Such a proposal has been made explicitly in Chafe (1967). (31), on the other hand, requires a performance model in which some sort of parallel processing is going on in constructing a sentence. To be sure, a hypothetical performance model is easier to construct with a competence model like (30). But the correct performance model is again an empirical question.

From what is known about performance even in the domain of phonetics, it is clear that the correct performance model will involve undreamed-of subtleties. So much more should we expect them in areas as abstract as syntax and semantics. The conceptual difference between (30) and (31) is undoubtedly trivial compared to the complexity of any adequate theory of performance.

7. Some different aspects of semantic representation

Let us assume, then, that the general plan of the grammar looks like (31) and not like (30). Is it possible to be more specific about what kinds of things semantic rules can do? In fact, the semantic rules to be discussed in this investigation fall into a few neat categories. In order to discuss the categories, it is necessary at the same time to say a few words about semantic readings.

First of all, it is not even clear that one can construct a formal object which corresponds to the intuitive notion "semantic reading," because of the infinite divisibility of semantic properties and the (perhaps undecidable) problem of choosing what information is part of the reading and what merely follows from the reading. But I think it is possible to separate out certain discrete aspects of the meanings of sentences and deal with them coherently. Some of these aspects are more closely linked to syntactic structures than others.

The aspect of semantic representation that is perhaps most closely linked to syntactic structure is something we might call

the functional structure of a semantic reading. If we think
of verbs as being represented as semantic functions (as Katz
and Fodor suggest), then for each verb in the deep structure
of a sentence there will presumably be a function in the semantic
representation. The embedding relations of functions in the
semantic representation will presumably mirror the embedding
relations of verbs (and other functional words) in deep struc-
ture. This is the part of semantic representation that the
Katz-Fodor semantic component is trying to capture.

A refinement of their proposal might try to find some
partial analysis of verbs into semantic subfunctions such as
causative, direction, and so forth, giving the semantic repre-
sentation of a verb some internal functional structure. This
has been done explicitly by Katz (1966) and Gruber (1965, 1967a,
1967b). The "higher pro-verbs" of Lakoff (1965 and later) and
McCawley (1968b) are an attempt to represent this internal
structure externally. But basically there is no disagreement
that these semantic properties can be represented structurally,
and that it is the deep structure which determines them.

Other aspects of semantic representation, however, do not
lend themselves to being represented in trees or functional re-
presentations, coreference properties among noun phrases, for
example. Although the determination of coreference relations
does depend on syntactic structure, the semantic notion "x is
(non-) coreferential with y" has nothing to do with the functional

structure of sentences. Rather, an independent device is neces-
sary to express coreferentiality relations. Referential indices,
introduced in Aspects, are one such device. Unfortunately the
individual index has a dangerous tendency to take on a life of
its own in the hands of the unwary investigator; it is tempting
to begin to talk about reference instead of coreference. (McCawley
(1968a) comes close to this view.) For this reason we will treat
coreference relations here as represented in a table outside of
the functional structure, with entries containing pairs of NPs
and a relation "coreferential" or "noncoreferential" marked between
them (cf. Chapter 2, § 3). But however coreference is marked, it
is clearly not the same kind of semantic information as functional
structure.

Generative grammarians up until now have treated the
marking of coreferentiality as an independent property of the
semantics generated in the base. Then, it is assumed, transforma-
tions can also make use of coreferentiality to change the appear-
ance of sentences, for example, to produce pronouns and reflexives
and to delete complement subjects. It will be assumed here,
however, that coreferentiality is a purely semantic property to
which transformations cannot refer. Under this assumption, it
turns out that coreferentiality must be marked not by the base,
but by an entirely new kind of semantic rule, one that operates
on derived structure. The three important rules to be proposed
in Chapters 2 and 3 all operate cyclically, applying at the end

of each cycle of transformations to the derived tree which has
just been produced by the transformations.

Another aspect of the semantic reading which has nothing
to do with the functional structure is focus and presupposition.
In effect, marking part of a sentence _focus_ locates a piece of
the semantic reading of arbitrary structure, draws a circle
around it, and says "the material within this circle is focus
and the rest is presupposition." The rule which marks focus is
again of an entirely new type: it operates very late in the
grammar, even after the stress assignment rules have taken
place.

There are some other semantic properties of sentences
that seem merely to be special markings on various parts of the
structure: specificty of indefinite NPs (cf. Baker (1966) and
Dean (1968)), genericness, and **referential opacity (as discussed
by many philosophers, including Frege, Quine, Russell, and Carnap).**
These three are similar in a number of respects and seem to be
somehow related: they all have to do (on the syntactic side)
with the interpretation of NPs (genericness and specificity with
indefinite NPs in particular); they involve (on the semantic side)
the difference between an NP as description and as identification
of an individual; and (on the methodological side) they have all
defied rigorous analysis in generative grammar so far. I have
no proposals for rules to handle them, but I suspect that they,
like focus, will be dependent on derived structure.

An aspect of semantic representation which will be dis-
cussed here is the interpretation of negation and quantifiers.
These appear at first glance to be part of the functional struc-
ture of the semantic representation, since we can set up express-
ions in the predicate calculus which seen to mean the same as
sentences containing them. But in natural language, the scope
of negation and quantifiers can cut across the functional struc-
ture, negating or quantifying something that is not a deep
structure constituent, as will be seen in Chapter 4. It turns
out that an adequate semantic account of negation and quantifi-
cation must introduce a semantic rule operating on derived
(perhaps surface) structure, the Scope Rule. This rule has the
effect of creating ambiguities by optionally enlarging the under-
stood scope of negation and quantifiers.

As with negation and quantifiers, the interpretation of
adverbs and modals appears to be part of the functional structure
of the semantic representation. But, as will be shown in Chapter
5, a more interesting account results if we attribute their inter-
pretation to a set of derived structure rules of interpretation,
one of which is formally very similar to the Scope Rule for
negation and quantifiers.

It must be emphasized that the statement of these rules
does not require a total theory of semantic representation, and
nowhere will I assume that I have any idea what a complete semantic
representation looks like. However, it is perfectly possible to

isolate those particular aspects of interpretation in which we are interested, to use our knowledge as speakers of the language to tell us when such-and-such a semantic property is present, and to formulate theories about these properties. Sometimes it will of course happen that a complete account is beyond our ability to isolate and formalize relevant factors. At that point we either try harder or leave the analysis on an informal basis.

8. What is there to gain?

If we merely give up the Katz-Postal hypothesis and let in the possibility of all these different kinds of semantic rules, we are making a weaker hypothesis about the nature of language, in that there are many more grammars possible within the theory. Therefore, it might be asked, what is there to gain by adding all this extra machinery?

There are three ways in which this question can be answered. First, if it can be shown that a theory including the Katz-Postal hypothesis is incorrect, this enlargement of the theory may let in the correct grammar. That alone would justify the addition. In practice this answer is not so easy to give, because of the infinite ability of any theory to accommodate awkward points. Indeed, as Chomsky has pointed out (class lectures, fall 1968), one can always simulate derived structure rules of interpretation in a theory obeying the Katz-Postal hypothesis, by generating constituents of arbitrary structure in the base, filtering them in the base for the semantic property to be des-

cribed, then using a filtering transformation later on to match
these arbitrary structures with the surface structure. Surely
such an analysis misses the point; nevertheless, many current
proposals in syntax approximate it.

Another thing we may gain is that significant generaliza-
tions can be expressed by these rules that cannot be expressed
in the standard treatment. It usually turns out that the amount
of machinery required by the two analyses is virtually identical
for the easy cases. When we dig deeper, though, it often turns
out that they make slightly different but crucial predictions.
For example, it turns out that in an interpretive theory of re-
ference, the three rules Pronominalization, Reflexivization, and
Complement Subject Interpretation can be ordered together, enabling
us to capture the generality in their environments; this cannot
be done in the transformational theory. Several such differences
arise in the course of this investigation. They may be relatively
small points in the entire description of the language, but
given two theories as sophisticated as those we are comparing, it
is going to be the small points of generality that must decide.

A third answer that can be given is that the addition of
several new kinds of rules may actually enable us to reduce the
class of possible grammars. This will be the case if we turn out
to be able to put much heavier constraints on the power of trans-
formations. For example, Emonds (1969) proposes (roughly) that,
with a certain specificable class of exceptions, the output of a
transformation must be a structure that could have been produced

by a base rule. The exceptional transformations, which only operate in a special class of clauses (primarily main clauses), each perform one of a very small class of possible operations. Clearly this hypothesis puts very strong constraints on the notion "possible transformation." It seems much more likely to be true of the transformations needed by the theory proposed here than it is of those needed by a theory incorporating the Katz-Postal hypothesis.

Another constraint on transformations that appears to be permitted by the interpretive theory and excluded by the Katz-Postal hypothesis has to do with the integrity of lexical items. Chomsky (1969) proposes the Lexical Hypothesis, roughly, that transformations do not perform derivational morphology. This is consistent with the interpretive theory proposed here. In fact, something stronger appears to be possible: we can also prohibit deletion under identity, deletion under positive absolute exception (cf. Lakoff (1965)), and perhaps _all_ deletion (and all positive absolute exceptions). This means that perhaps the only changes transformations can make to lexical items is to add inflectional affixes such as number, gender, case, person, and tense. We will refer to this set of constraints occasionally, under the name Extended Lexical Hypothesis.

Of course, these restrictions must be concomitant with equal restrictions on possible semantic rules, if the number of possible grammars is to be reduced. Before such restrictions can

be even guessed about, it is necessary to have enough proposed
semantic rules to be able to say something general about them.
In this light Lakoff's accusation (in Lakoff (1969)) that no
constraints have been proposed seems to me a trifle premature.
It is interesting, however, that the semantic rules to be dis-
cussed here fall into several well-defined and restricted types;
so perhaps some tentative constraints are just over the horizon.

9. A remark on the lexicon

Under the Extended Lexical Hypothesis, transformations
cannot do derivational morphology. How then can we capture the
semi-productivity of morphological processes? At the time of
The Grammar of English Nominalizations (1960), there was no
possibility but a transformational solution, since the concept
of lexicon had not yet been proposed. But even after the theoreti-
cal framework was available, research tended to follow along the
lines laid down by Lees. Thus at present most well-known pro-
posals about derivational morphology are couched in transformational
formalisms, for example, Lakoff (1965) and Chapin (1967). One
notable exception is Gruber (1967b).

Such solutions are not available to us. Rather, it is
necessary to list, for example, both a verb and its nominalization
in the lexicon. We could express regularities by making the
evaluation measure for the lexicon such that it is cheaper to list
a noun if it is related to a verb by a regular process.

Suppose, then, that part of the lexicon is a set of deri-

vation rules. For example, the nominalization system would
contain the following morphological rules for -tion and -ment,
along with more for other productive affixes such as -al, -ism,
-er, etc.

M1: V + tion → N
M2: V + ment → N

Associated with these would be at least the following semantic
derivation rules, which give the range of possible interpretations
of nominalizations (underlined words are to be taken to stand for
their semantic interpretations):

S1: act of Ving → N
S2: thing Ved → N
S3: thing transferred in act of Ving → N

Under the usual assumptions about the lexicon, a lexical
entry will have the general form

$$
\left[
\begin{array}{l}
\text{Location number} \\
\text{PHONOLOGICAL REPRESENTATION} \\
\text{syntactic features} \\
\underline{\text{semantic representation}}
\end{array}
\right]
$$

The location number is simply a way of identifying the item in
the lexicon.

To show how the derivation rules are used in the lexicon,
take the verbs compensate and consign, which will have lexical
entries as follows (using arbitrary location numbers):

$$
\left[
\begin{array}{l}
73. \\
\text{COMPENSATE} \\
+\begin{bmatrix} +V \\ NP\underline{\quad} NP \text{ for } NP \end{bmatrix} \\
\underline{\text{compensate}}
\end{array}
\right]
\qquad
\left[
\begin{array}{l}
67. \\
\text{CONSIGN} \\
+\begin{bmatrix} +V \\ NP\underline{\quad} NP \text{ to } NP \end{bmatrix} \\
\underline{\text{consign}}
\end{array}
\right]
$$

The cost of these lexical items would be (more or less) the number of features it takes to build up the lexical entry. Now for their nominalizations <u>compensation</u> and <u>consignment</u>, it is not necessary to spell everything out. All we need is the following:

$$
\begin{bmatrix}
540. \\
\text{derived from 73} \\
\text{using M1, } \begin{Bmatrix} S1 \\ S2 \end{Bmatrix}
\end{bmatrix}
\qquad
\begin{bmatrix}
765. \\
\text{derived from 67} \\
\text{using M2, } \begin{Bmatrix} S1 \\ S2 \end{Bmatrix}
\end{bmatrix}
$$

The crossreferences to the verbs supply the basic phonological, syntactic, and semantic information. Then the crossreferences to the derivational rules tell how to alter the verb's entry to produce the noun. The cost of these items will simply be the cost of the crossreferences, which will presumably be low.

For a verb with no nominalization, there will simply be no related entry containing a crossreference. For a nominalization such as <u>agression</u>, where there is no parallel verb <u>aggress</u>, we want the lexical entry to cost more than if there were such a verb, but less than if it did not contain the regular affix and regular semantic interpretation associated with that affix. Accordingly, its lexical entry will be as follows:

$$
\begin{bmatrix}
42. \\
\text{derived from} \quad
\begin{bmatrix}
\text{AGGRESS} \\
+V \\
+[\text{NP}\underline{\quad}\text{against NP}] \\
\text{aggress}
\end{bmatrix} \\
\text{using M1, S1}
\end{bmatrix}
$$

The cost of this entry will be the cost of the sub-entry <u>aggress</u>
plus the cost of the crossreferences to the derivation rules.
Hence its cost is between the costs of totally regular entries
like <u>compensation</u> and totally irregular ones of approximately
equal complexity such as <u>story</u>. This is exactly the result we
want.

It is fairly easy to see how this general approach can
be extended to handle all sorts of word-derivation. It will be
explored in more detail in Jackendoff (forthcoming). What is
important for the present purposes is that it is possible to
capture the notion "separate but related lexical items," using
an evaluation measure on the lexicon rather than transformations.

10. A remark on motivating rules

In a theory of grammar which includes the Katz-Postal
hypothesis, particularly in the stronger form KP2, there is only
one way in which similarity in meaning or co-occurrence restric-
tions between two constructions can be captured: a transformation.
In a theory permitting a wide variety of different kinds of
rules, such as the theory to be explored here, there are many
ways of capturing generalizations. In addition to transformations,
there are all the different kinds of semantic rules, operating
at different levels of the derivations. Generalizations can also
be captured within the lexicon, by means of the redundancy rules
discussed in §9, which express the notion of separate but re-
lated lexical items. Furthermore, certain generalizations can be

expressed by treating the nodes for lexical categories as feature
complexes, then stating base rules, transformations, and semantic
rules so as to refer to more than one major category at a time.
This is the essence of the position arrived at in Chomsky (1969).

With all these different kinds of rules at our disposal,
several very different analyses will often come to mind for the
same phenomenon, each of which seems equally capable of expressing
the proper generalization. How do we decide which account is to
be preferred? There can be no sort of principle that says, "Al-
ways choose an $\underline{X}$ if you have a chance": it is not difficult to
construct algorithms to reduce all rules to any chosen type, given
exception machinery of sufficient power, such as Lakoff's (1965).
Rather the decision will be made on the basis of how the rules
interact with each other most naturally and how appropriate the
power already proposed for a particular type of rule is for
handling something new. Similar processes should be handled by
similar kinds of rules.

Also of prime importance in motivating a particular treat-
ment of a phenomenon is how it is reflected in the lexicon. If
a process only takes place for certain lexical items, or varies
over several classes of lexical items, we should choose the way
of handling the process that least increases the cost of the
lexicon. The use of exception features is the worst possible
solution, in that it represents an arbitrary bifurcation of the
lexicon, and so every marked feature counts. If the difference

has something to do with the meaning of the items in question, that is the best possible case, since the rule has only to refer to properties already present.

Unfortunately, this latter case is also the least formalizable, since we don't yet have a way of expressing the meaning. For the sake of stating a rule, however, it seems to me perfectly all right to provisionally adopt an arbitrary feature, if we have clear intuitions about when this feature is present, and if it is fully understood that it has no life independent of the total meaning of the lexical item.

NOTES

1. See § 7 for another view.

2. Mostly by word of mouth, but explicitly, for example, in Perlmutter (1969).

3. This example was brought to my attention by Stephen Anderson, who found them somewhere in the works of J. L. Austin.

4. Katz and Postal seem to implicitly accept this view of projection rules in giving the alternatives (4)-(6), since they admit the possibility of Pl rules operating on derived P-markers, though at no point is the switch in viewpoint ever mentioned.

PRONOUNS AND REFLEXIVES

Generative grammarians have always assumed that pronouns and reflexives are the product of a transformation which substitutes them for a more fully specified noun phrase under certain conditions. In the early framework of Syntactic Structures, this approach was the only way of expressing the relations between pronouns and their coreferents. But with the advent of the semantic component, another possibility opened up, namely that the properties of coreference are due to a rule of semantic interpretation. This possibility is strongly suggested by the examples in Dougherty (1968a, b) on referential terms other than pronouns and reflexives, such as each other and coworker.

1. The usual treatment of reflexives

The standard transformational account of reflexives is that given by Lees and Klima in "Rules for English Pronominalization." This analysis gave the following rule:

(1) (Lees and Klima Reflexive Rule)

$$X \quad NP^1 \quad Y \quad NP^2 \quad Z$$
$$1 \quad 2 \quad 3 \quad 4 \quad 5 \Longrightarrow 1 \quad 2 \quad 3 \quad \begin{bmatrix} 4 \\ +\text{refl} \end{bmatrix} \quad 5$$

Conditions: 1. NP^1 and NP^2 are referentially and morphologically identical

2. NP^1 and NP^2 are in the same simplex S.

OBLIGATORY

This rule accounts for the simple reflexive sentences such as (2).

(2) John shaved himself.

In addition, with the mechanisms of the complement system and the transformational cycle, this rule can account for the reflexives in (3) and (4).

(3) Mary forced Bill to shoot himself.
(4) John expected himself to be able to abstain from eating.

In (3), reflexivization takes place in the first cycle; then, on the second cycle, <u>Bill</u> is deleted from the complement by Equi-NP deletion, leaving <u>himself</u> still standing. In (4), reflexivization takes place in the second cycle, after <u>John</u> has been moved up into the main clause by IT-replacement. The restriction to simplex S prevents sentences like (5) and (6).

(5) *John forced Mary to shoot himself.
(6) *John saw the girl who hated himself.

The first doubts about the rule (1) began to arise upon investigation of the so-called picture-nouns, first discussed in a transformational framework in Warshawsky (1965). Sentences like (7) and (8) are handled by rule (1) without difficulty; two different NPs can fulfill the structural condition of NP^1, and so there are two possible reflexivizations.

(7) I told Bill a story about myself.
(8) I told Bill a story about himself.

Thus we correctly predict an ambiguity in (9).

(9) John told Bill a story about himself.

Following this line of analysis, we would expect three readings

for (10), but unfortunately there is only one, with <u>himself</u> referring to <u>Harry</u>.

    (10) Tom told Dick Harry's story about himself.

This fact becomes clearer when we substitute for <u>Harry</u> a noun that cannot be coreferential with <u>himself</u>.

    (11) *Tom told Dick $\left\{ \begin{array}{l} \text{Mary's} \\ \text{the newspaper's} \end{array} \right\}$ story about himself.

These examples show that even within the same simplex sentence there are times when reflexivization does not work.

Any number of people have suggested introducing another S node into the underlying form of (10) and (11), using the justification that <u>story</u> can be related to an underlying abstract verb "to story." This extra S, it is alleged, would prevent reflexivization because of the simplex S condition. Then some pruning argument must be resorted to in order to handle (9), where reflexive does take place.

The immediate trouble with this approach is that it requires all kinds of abstract verbs for which there is no justification. To be sure, <u>description</u> and <u>picture</u> and <u>photograph</u> have corresponding verbs, but what about <u>story</u>, <u>poem</u>, <u>novel</u>, <u>biography</u>, and many others which take part in the same construction?

As it turns out, even such a drastic approach to the picture nouns cannot save the reflexive rule (1). The following sentences illustrate some violations of the rule which are in fact perfectly acceptable.

    (12) Tom believes that there is a picture of himself hanging in the post office.

(13) Tom made the claim that the picture of himself
      hanging in the post office is a fraud.
(14) That the picture of himself in the newspaper is
      ugly enrages John.
(15) The fact that there is a picture of himself hanging
      in the post office frightens Tom.
(16) The picture of himself that John saw hanging in the
      post office was ugly.
(17) The description of himself that John gave the police
      over the phone was a pack of lies.
(18) The king was hit over the head by a massive portrait
      of himself which had fallen off the wall.
(19) Unflattering descriptions of himself have been
      banned by LBJ.

All of these sentences except for the last two demonstrate re-
flexives in a different simplex sentence from corresponding co-
referential NPs. In (12) and (13), the reflexive is to the right,
as in normal simplex S cases. However, in (14) and (15) the re-
flexive is to the left, violating even the basic structural descrip-
tion of (1). In these first four examples the reflexive is in a
lower S than its coreferential NP. However, in (16) and (17),
the reflexive is not only to the left of the coreferential NP,
but also in a higher simplex sentence.

        (13) and (15) also violate the Complex NP Constraint,
which has been proposed by Ross to explain the impossibility of
questioning or relativizing NPs that are within a relative
clause or a <u>the fact that</u> S construction. This constraint in
its most recent form (Ross 1967a) says that except for pronomi-
nalization, no transformation may change features within such a
construction; Ross gives numerous examples to illustrate this
point. However, (13) and (15) seem to show that reflexivization
is not subject to this constraint either.

(18) and (19) illustrate reflexives in passive sentences, although passive reflexives without picture nouns are ungrammatical:

(20) *Himself was shaved by John.
(21) *John was shaved by himself.

In (19) the reflexive is on the left, which is generally also a violation.

2. An interpretive approach

One could try to account for these problems by just patching up the standard theory of reflexives. For example, Postal (1968a) alludes to a rule of Late Reflexivization (Chapter 17) to take care of the sentences with reflexives on the left. It would be more interesting, though, if all the cases could be handled under a single rule of reflexivization. To this end, I will start afresh with a new approach.

The difference between my approach and the standard generative grammar approach is fundamental: instead of accounting for the properties of pronouns and reflexives by deriving them from underlying more fully specified noun phrases, I will assume that they are generated as lexical items, inserted into base structures. I will then try to show that all their properties can be explained in terms of rules of semantic interpretation.

In this approach, which I will call the interpretive theory, noun phrases in general will be unmarked for a reference in the base. Rules of semantic interpretation establish relations between pairs of noun phrases, marking them as coreferential or noncoreferential with each other.[1]

Pronouns and reflexives will be generated by the base as lexical items containing features +pro and +refl, but, like other noun phrases, unmarked for reference. Since pronouns such as someone must be generated in the base anyway, the use of the feature pro in the lexicon cannot be objectionable. The only innovation in the lexicon is the feature refl.

An important prediction arises already at this stage of definition. If we have interpretive rules which mark coreference, it should be as easy to mark a pair of noun phrases noncoreferential as to mark them coreferential. And in fact this provides a ready way to handle such phrases as someone else and another, where the reference is unspecified but different from previous references. A transformational approach to pronouns provides no easy way to take care of items like these.

Another immediate advantage of this approach is that the infinite recursion of deep structures in sentences such as (22) (first pointed out by Emmon Bach and Stanley Peters) does not arise.

(22)   The man who deserves it will get the prize he
       wants.

If the underlying structure of pronouns is a fully specified NP identical with its coreferent, both it and he in (22) must have infinite deep structures:

(23)   The man [who deserves the prize [which the man
       [who...] wants]] will get the prize [which the man
       [who deserves the prize [which...] wants]]

The drastic way out in the transformational approach to reference

is to generate only indices in NP positions, then bring in lexical material from outside clauses (cf. McCawley (1969), Postal (1967), Bach (1968)). But such a solution obviously violates the Extended Lexical Hypothesis, since it requires proforms which consist of just an index. (In fact the standard transformational treatment of pronominalization violates the Extended Lexical Hypothesis too. A fully interpretive theory of reference will enable us to further restrict the kinds of changes transformations can make.)

If the reference of pronouns is determined by a rule of semantic interpretation, the deep structure of (22) contains the pronouns themselves, so there is no recursion. Furthermore, in the process of semantic interpretation, a pronoun need not be replaced with a duplicate of the noun phrase with which it is coreferential (which would again bring up the problem of recursion), but rather it may just be marked coreferential with another NP.

3. The structural change of the pronominalization and reflexivization rules.

It is very important to see what it means to apply a semantic rule of coreference. There are three relevant points that must be made clear. First, coreference is an exclusively semantic property that cannot be referred to by transformations. Second, as I mentioned in Chapter 1, coreference is an aspect of semantic interpretation that has nothing to do with the struc-

ture of the sentence (although the structure does play an important part in establishing it). Third, coreferentiality is to be conceived of as a binary relation holding between two NPs (or their semantic readings). Three or more NPs can be understood as mutually coreferential only if they have been marked pairwise coreferential.

In order to emphasize these three points, we will abandon the fairly standard device of indices of coreference. Instead, we will think of the rules of coreference as constructing entries in a table of coreferents and noncoreferents, off to the side of the part of the reading that expresses the functional structure. This table is probably only a notational variant of indices, but it expresses more directly the three important points above.

Leaving aside the environments for pronominalization and reflexivization for the moment, let us see what kinds of entries the interpretive rules will make in the table of coreference. First consider the simple cases of reflexivization.

(24) John washed himself.
(25) John washed him.
(26) John washed John.
(27) John washed Bill.

According to the interpretive theory, (24)-(27) are all possible deep structures, and the noun phrases are not marked for reference in the deep structure. The rule for interpretation of reflexives must enter in the table that the object of (24) is coreferential with the subject, and it must mark all the other

objects noncoreferential with the subject. A plausible way
to express this is (28).

(28) $\begin{bmatrix} NP^2 \\ \alpha \text{ reflexive} \end{bmatrix}$ is α coreferential with NP^1 in the
environment....

OBLIGATORY

(28) says that NP^2 can be coreferential with NP^1 if and only if
it is reflexive.

One convention of application must be that if two possible
environments for interpretation of a reflexive crop up at once,
either reading is possible. This convention is necessary in
order to permit the ambiguity of (29), treating either Bill or
John as NP^1.

(29) Bill told John about himself.

Note that (28) does not say anything about agreement in
person or number. At first glance, this would seem to lead to
trouble in sentences such as (30).

(30) *John shot yourself.

However, closer examination shows that this sentence can be
blocked. The reflexive rule does indeed mark John and yourself
coreferential. But then the blocking can be accomplished by the
obvious general convention that coreferential noun phrases must
be able to have the same reference and thus must agree in number,
person, and gender as well as animacy, humanness, abstractness,
and myriad other semantic properties.[2]

In addition to (28), we need a principle which says that a sentence blocks if it contains a reflexive that has not been assigned coreference with anything else. This principle blocks sentences like (31).

(31) *Himself was sick.

Turning to pronominalization, we have the following basic data to consider.

(32)   John looked at him.   noncoreferential
(33)   John said that he was sick.   ambiguous
(34)   John said that John was sick   noncoreferential
(35)   John said that Bill was sick.   noncoreferential

Pronouns are at the most optionally coreferential with another NP. Any two non-pronominal noun phrases in the same sentence, morphologically identical or not, are always noncoreferential. We can therefore state the pronominalization rule as follows:

(36) $\begin{bmatrix} NP^2 \\ + \text{pronoun} \end{bmatrix}$ is coreferential with NP^1 in the environment....

OPTIONAL

We will also need the following rule, which will apply at the end of the rules of reference:

(37)   (Noncoreferentiality of unmarked NPs)
       NP^2 is noncoreferential with NP^1 if it has not been marked coreferential with NP^1.

OBLIGATORY

To illustrate how these rules interact with the reflexive rule and with each other, consider (32)-(35). In (32), the reflexive rule obligatorily applies, entering in the table of coreference that John and him are distinct. If the pronoun rule

were then to apply, it would enter in the table that <u>John</u> and
<u>him</u> are coreferential. But then the reading of the sentence
would be anomalous, since the table of coreference would say
that <u>John</u> and <u>him</u> do and do not refer to the same individual.
Therefore the only consistent reading is produced by refraining
from applying pronominalization. In (33), reflexivization does
not apply, but pronominalization may, marking <u>John</u> and <u>he</u> co-
referential. If we choose not to apply pronominalization, (37)
will apply, marking them as distinct. In (34) and (35), only
(37) is applicable, so the two NPs are marked distinct.

    Next let us consider a more complicated example, (38).

    (38) John said that Bill had shot him.

The reflexive rule first applies to <u>Bill</u> and <u>him</u>, marking them
non coreferential. As in (32), the pronoun rule cannot apply to
<u>Bill</u> and <u>him</u>, but it may apply between <u>John</u> and <u>him</u>, marking them
coreferential. Finally (37) applies, marking <u>John</u> and <u>Bill</u>
distinct.

4. The environment for pronominalization

    In the transformational account of pronominalization, the
transformation can be stated roughly as follows, following Ross
(1967b):

    (39) $NP^2 \Longrightarrow$ +<u>pro</u> if NP^1 is identical with NP^2 and if
           either $\overline{NP^1}$ is to the left of NP^2 or NP^2 is domin-
           ated by a subordinate clause which does not
           dominate NP^1.

           OBLIGATORY

Since the environment for this rule seems to be essentially correct, it would be nice to preserve it as the environment in an interpretive theory of pronominalization. However, in the transformational theory, it has been argued (Ross 1967b) that the pronominalization rule is crucially ordered with respect to many other rules, and that it takes part in the cycle. To keep the environment intact, therefore, we will have to give up any hope of having the interpretive rule for pronominalization operate on deep structures.

We can preserve the environment of (39) if we place the interpretive rule for pronominalization at exactly the point in the grammar where (39) takes place in the transformational theory of pronouns. The complete form of the rule (36) will then look like this:

(41) $\begin{bmatrix} NP^2 \\ + \text{pronoun} \end{bmatrix}$ is coreferential with NP^1 if either NP^1 is to the left of NP^2 or NP^2 is dominated by a subordinate clause which does not dominate NP^1.

OPTIONAL

Each time (41) is reached in the cycle, new pairs of noun phrases will be available for interpretation, and the rule will determine the coreferentiality of each pair as it appears. In addition, (37), the noncoreferentiality rule, will take place after (41).

5. The ordering of pronominalization

Lakoff has shown ("Pronouns and Reference") that it is plausible that pronominalization is ordered after all the trans-

formations. This would be convenient if it were a semantic rule.
On the other hand, Ross's argument for the cyclical nature of
pronominalization, in which a complex set of facts is handled
by the rule (39) and the principle of the transformational cycle,
is very appealing. How can these be reconciled?

Lakoff's arguments fall into two parts. First, it can be
shown that pronominalization must take place after all rules like
Adverb Preposing and Wh-Preposing, which move NPs in a particular
cycle. This shows that pronominalization must be the last rule
in the cycle. Second, he purports to show that pronominalization
must be last-cyclic, since it must always take place after Wh-
preposing, which, it is claimed, is last-cyclic. (This argument
also appears in Postal (1968a).)

The reason pronominalization must follow Wh-preposing in
a particular cycle is shown by examples like these (she and Mary
to be read coreferential):

    (42) Who that Mary knows did she visit?
    (43) Who that she knows did Mary visit?
    (44) Mary visited a guy she knows.
    (45) *She visited a guy Mary knows.

Both directions of pronominalization are possible only when the
subordinate clause is on the left, i.e. after Wh-preposing.

If Wh-preposing is last-cyclic, pronominalization must be
also, as shown by the following example, discovered by Postal.

    (46) Who that Mary knew do you think she visited?

(47) is the underlying structure of this, and (48) is the surface
structure.

(47)

```
                        S¹
          Q    NP            VP
               you      V          NP
                       think    IT      S²
                              that  NP          VP
                                   {Mary}    V        NP²
                                   {she}  visited
                                              wh-someone      S³
                                                           NP          VP
                                                         {Mary}    V        NP
                                                         {she}   knew   wh-someone
```

(48)

```
                              S¹
        NP²           ?    NP          VP
      who    S³      do   you     V         NP
          /that Mary     think    think     S²
           knew\                         /she visited\
```
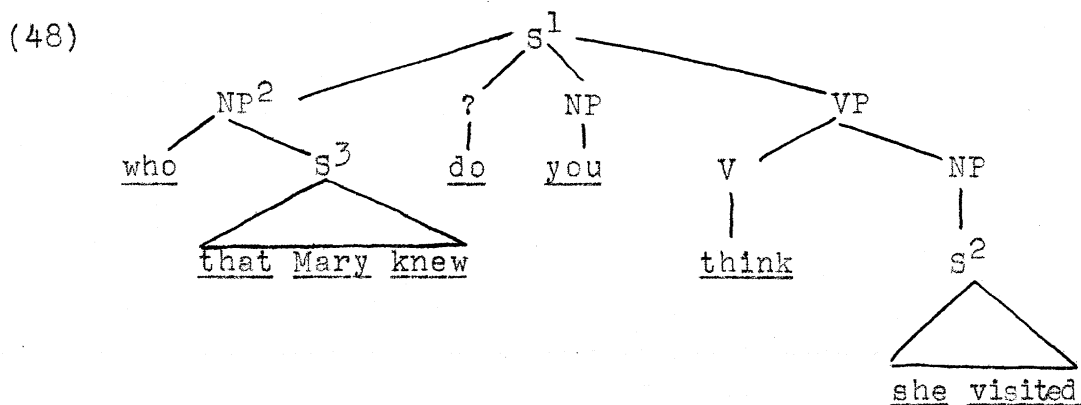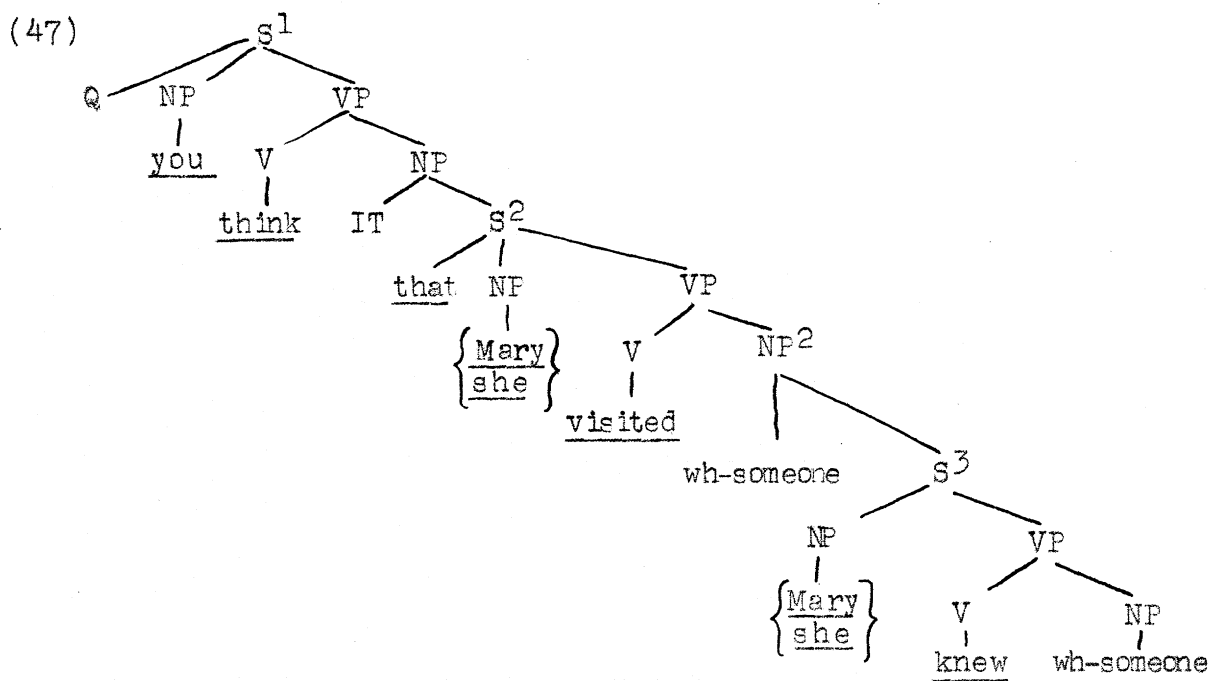
Under the transformational theory, using Ross's rule and last-cyclic
Wh-preposing, we must start with <u>Mary</u> in both S^2 and S^3. <u>Mary</u> in
S^3 must obligatorily become <u>she</u> on the S^2 cycle. Thus when NP^2
is fronted on the S^1 cycle by the question transformation, we can
only get (49).

(49)   Who that she knew do you think Mary visited?

Since pronominalization is obligatory, there is no way to hold
it off until the S^1 cycle to produce (45).

With a cyclic rule of Wh-preposing, however, the Wh-phrase
is fronted in the S^2 cycle, giving an intermediate stage where
pronominalization can take place in either direction (42)-(43).
Then the final form is reached by a further fronting on the S^1
cycle.

In the interpretive theory, we can generate _she_ instead of
Mary in S^2 or S^3. Then after fronting in the S^2 cycle, pronomi-
nalization can mark them coreferential, and the further fronting
will give the right reading of (45).

What is the argument that Wh-preposing is last-cyclic?
The only argument Postal gives (Chapter 10, §B) is that preposi-
tions that move up with Wh-words cannot be stranded at some
intermediate point, as in (52)-(53), although the movement of
prepositions with Wh-words is generally optional (50)-(51).

    (50)   Who do you believe Bill saw Mary talk to?
    (51)   To whom do you believe Bill saw Mary talk?
    (52)   *Who do you believe Bill saw to Mary talk?
    (53)   *Who do you believe to Bill saw Mary talk?

Chomsky has suggested (class lectures, winter 1969) that option-
ality of preposition movement is due to the optional attachment
of the feature _wh_ to the PP instead of the NP. Then Wh-preposing
applies to whatever node is marked _wh_, and there is no way for
a step-by-step process of Wh-preposing to make a mistake like
(52) or (53). Postal himself (Chapter 12, §C) suggests that such
a feature (called by him [+Pied Pipe]) is necessary, so his earlier

argument is nullified. Therefore it can be maintained that Wh-preposing is cyclic, and hence that pronominalization is also.[3]

There are two other rules that should be mentioned in connection with the cyclic nature of pronominalization: extraposition, and extraposition from NP. Like Wh-preposing, they must both precede pronominalization, and furthermore, they have all been claimed to be last-cyclic, which would force pronominalization also to be last-cyclic.

<u>Extraposition</u> is the rule that produces, for example, (54) from (55).

    (54)   It bothers John that Bill is sick.
    (55)   That Bill is sick bothers John.

This rule changes pronominalization possibilities, as in (56)-(57), so it must precede pronominalization.

    (56)   It bothers him that John is sick.   non-coreferential
    (57)   That John is sick bothers him   coreferential

In Ross (1967a), chapter 5, two arguments are given to show that this rule must be last-cyclic. The first argument (pp. 271-276) is based on the necessity of producing correct derived constituent structure and hence correct intonation breaks in sentences where extraposition takes place on lower cycles, such as

    (58)   It appears to be true that Harry likes girls.

But, as Ross admits, the intonation contours may be produced by late adjustment rules of the sort that give the derived structure the correct form in sentences such as

    (59)   This is the dog that chased the cat that caught
           the rat that ate the cheese.

Ross's second argument (pp. 276-281) is based on the fact that extraposition must follow <u>Particle Movement</u>. Particle movement, he argues, must be last-cyclic in order for the derivation of action nominalizations to come out right. This however assumes that action nominals are produced by a transformation from sentences, which cannot be the case in the version of grammar presented here. Such a transformation would violate the Extended Lexical Hypothesis, since for example, it would require changing adverbs into adjectives in examples like (60).

(60) She rapidly looked up the information $\xrightarrow{\text{action nominalization}}$
Her rapid looking up of the information

If action nominals are generated as NPs in the base (perhaps as VP dominated by $\bar{N}$), particle movement can be put back in the cycle, and hence so can extraposition.

<u>Extraposition from NP</u> is the rule that produces, for example, (61) from (62).

(61) A man came into the room who was ten feet tall.
(62) A man who was ten feet tall came into the room.

This rule also changes pronominalization possibilities, as in (63)-(64), so it precedes pronominalization.[4]

(63) A man spoke to him who hated John. noncoreferential
(64) A man who hated John spoke to him. coreferential

Ross argues (pp. 281-285) that extraposition from NP is last-cyclic, again on the grounds that it must follow particle movement, which he claims must be last-cyclic in order to handle nominalization. Since we are giving up the transformational derivation of action nominalizations, it is still possible to maintain

Particle Movement and hence Extraposition from NP as cyclic rules.
This allows pronominalization to still be in the cycle too.

There are several rules, however, which move phrases
forward over a variable, but which do not appear to reduce to a
step-by-step movement as easily as does Wh-preposing. Four con-
structions produced by such rules are

(65) How brave Bill is!
(66) Beans I'll never eat. (Topicalization)
(67) Handsome though Melvin is,... (Though-movement)
(68) They said Bill would pay up, and pay up he did.

As Ross shows (Chapter 6), these rules all involve crucial use of
a variable; the phrases can be moved up from indefinitely deep
embeddings, within Ross's constraints. For example,

(69) How brave John thinks Mary said Bill was!
(70) Beans you'll never persuade me to force my kids
      to eat.
(71) Handsome though it is said by many people that
      Melvin is, ...
(72) They said Bill would pay up, and pay up everyone
      knows Bill claims he did.

These rules all increase pronominalization possibilities. In
the untransformed versions (73)-(76), he and John are distinct,
but in the transformed versions (77)-(80), he and John may be
coreferential.

(73) Mary thinks he is fond of the girl John kicked
      yesterday.
(74) Mary thinks he secretly loves the girl who kicked
      John.
(75) Though Mary thinks he is fond of the girl John
      kicked yesterday, ...
(76) ??We all bet that Mrs. Provolone would kick someone,
      and Mary thinks his mother did kick the girl John
      hates.

(77) How fond of the girl whom John kicked yesterday Mary
      thinks he is!
(78) The girl who kicked John Mary thinks he secretly loves.

(79)   Fond of the girl John kicked yesterday though
       Mary thinks he is, ...
(80)   We all bet that Mrs. Provolone would kick someone,
       and kick the girl John hates Mary thinks his mother
       did. (where John is understood as Mrs. Provolone's
       son)

With an obligatory pronominalization transformation, (77)-(80)
cannot be produced unless pronominalization takes place _after_
the fronting. But in these sentences, if pronominalization is
cyclic, it will take place in a cycle before fronting, because
of the intervening cycle with _Mary thinks that..._ as its main
clause. It is difficult to see how all these fronting rules
could be step-by-step processes like Wh-fronting, so that way
out is not available.

     As we have stated it so far, the interpretive theory runs
into exactly the same problems as the transformational theory.
However, a fairly simple modification enables it to handle these
cases. Recall that the interpretive rule of pronominalization
is optional, and that if pronominalization does not apply between
two NPs, then the non-coreferentiality rule (37) obligatorily
marks them distinct. We assumed that (37) was in the cycle. In-
stead, let us make (37) operate only on the last cycle. Then,
even if pronominalization does not operate between two NPs on one
cycle, it has a second chance on a subsequent cycle. Only when
it has not operated by the end of the derivation does (37) mark
the NPs distinct. With this modification we can hold off pro-
nominalization in (77)-(80) until after the cycle in which fronting
has taken place, even though the pronominalization rule is cyclic.

     With this modification it is conceivable that (37) is not

even an actual rule, but rather a convention on interpretation of the table of coreference. It can be then stated as follows:

(81) (Noncoreferentiality Convention)
If a pair of NPs (NP^1, NP^2) does not appear in the table of coreference, they are noncoreferential.

Our conclusion is then that we can maintain pronominalization in the cycle if we make it optional and if the Noncoreferentiality Rule or Convention is last-cyclic.

6. Two further cases of pronominalization

The first case is related to the one discussed in Ross (1967b).[5]

(82) The fact that he realized John was sick bothered him.

S^1

NP^1 VP

Det N S^2 NP^2 bothered him

the fact that he realized NP^2

S^2

that John was sick
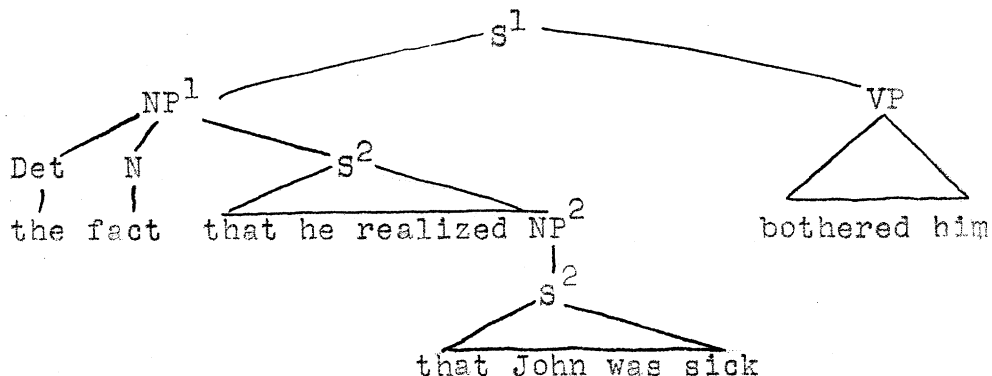
On the S^3 cycle, nothing of interest happens. On the S^2 cycle, pronominalization cannot apply to mark he and John coreferential, since the conditions for backward pronominalization are not met. On the S^1 cycle, it is theoretically possible that him is marked coreferential with both he and John. However, since no rule has entered the pair (he, John) in the table of coreference, (81)

marks them distinct. Therefore the table of coreference finally
looks like (82).

(82)   he   coref   him
       John coref   him
       he   noncoref John

But this table is an anomalous reading, since him is coreferential
with two distinct individuals. This leaves open only the possi-
bility of marking him coreferential with one or the other of the
other two noun phrases. In fact the predominant reading is the
one in which him is coreferential with he and distinct from John.
The other reading is difficult to get unless contrastive stress
is placed on he. This suggests that the usual reduced stress on
he plays a part in making one reading predominate over the other.
The set of readings predicted by the interpretive rule in this
case is exactly the same as that predicted by Ross's pronominali-
zation transformation.

The second case I will discuss in this section is one that
is conceptually impossible within a transformational framework,
but very natural in an interpretive system. Consider sentences
like the following:

(83)   I wanted Charlie to help me, but the bastard wouldn't
       do it.
(84)   Irving was besieged by a horde of bills that the
       poor guy couldn't pay.
(85)   Although the bum tried to hit me, I can't really
       get too mad at George.

There are many noun phrases such as the bum, the bastard, and the
poor guy, which can be used coreferentially with another noun
phrase if they are reduced in stress. These "pronominal epithets"

can occur in all the configurations that a normal pronouns can:
(83), (84), and (85) illustrate pronominal epithets in subordi-
nate clauses on both the left and the right. (86) shows one in
a main clause, with the coreferential NP in a subordinate clause
on the left. (87) shows that the other NP cannot be on the right
and in a subordinate clause.

(86) The fact that Charlie is unpopular irritates the
      bastard.
(87) It irritates that bum that Charlie is unpopular.

Since the pronominal epithets obey similar structural conditions
as pronouns (probably a subset of those conditions), we would
obviously be missing a generalization if we did not handle them
by the same rule.

In a transformational framework, however, the generalization
cannot be captured. The pronominalization rule changes NPs into
pronouns. We will have to add the condition that it can optionally
change an NP into a pronominal epithet instead. But then which
pronominal epithet should the NP be changed into? The meaning
is obviously changed if we substitute an epithet for a pronoun or
one epithet for another.

In an interpretive framework, we can mark epithets as
special lexical items which may function as pronouns, adding their
lexical meaning to the intended attributes of the person they refer
to. Then the pronominalization rule requires no changes at all,
since the optional feature pro on the epithets automatically
brings them under the domain of the rule.

7. The environment for reflexivization

Since most of my examples contain picture nouns, I must first take a stand on their structure. I will assume the type of structure given in Chomsky (1969), putting the possessive phrase in the determiner and the head noun and the of-phrase under a node $\bar{N}$. Thus (88) will have the structure (89).

(88) Mary's picture of Bill

(89)

```
                NP
          Det        N̄
           |
          NP      N        NP
           |      |         |
         Mary   picture    Bill
```
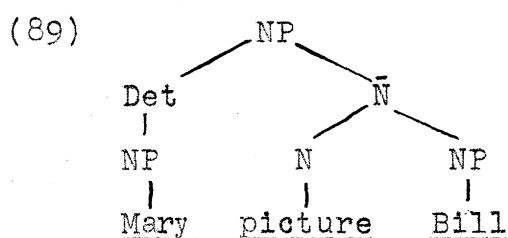
The possessive ending is put on Mary by a late rule which applies to any NP in the determiner; another late rule inserts of between N and NP.

This structure is very close to the structure of a simple sentence such as Mary saw Bill, with the highest NP in (89) corresponding to the S node of the sentence, the N corresponding to the NP, and the N corresponding to the V. Chomsky shows that this analysis leads to various desirable conclusions, including a generalization of the passive transformation to include within its domain NPs such as (89) as well as Ss.

Given the fact that derived nominals like (88) behave like sentences with respect to certain transformations, it is not unreasonable to assume that the transformational cycle has as its domain not only Ss, but also NPs containing derived nominals.

This assumption enables us to account for a number of hitherto unexplained reflexives, using the rule (90).

(90) $\begin{bmatrix} NP^2 \\ \alpha \text{ reflexive} \end{bmatrix}$ is α coreferential with NP^1 if

    a) NP^2 has not been marked for coreferentiality with any other NP (i.e. has not yet served as NP^2 in a coreference rule); and

    b) NP^2 is immediately dominated (except for a possible preposition) by VP or $\bar{N}$; and

    c) NP^1 is not in a subordinate clause relative to the present cycle; and

    d)   either NP^1 is to the left of NP^2 or NP^2 is dominated by a clause subordinate to the clause immediately dominating NP^1.

OBLIGATORY

Condition (a) and the obligatoriness of the rule insure that a reflexive will be assigned coreference with the first NP encountered that stands in the proper structural relation with the reflexive. Note the asymmetry of the use of "marked for coreferentiality with" in condition (a). We will make use of this asymmetry in Chapter 3, § 4. The structural relation, condition (d), is identical to the structural relation for pronominalization. If this environment is correct, it is a significant generalization.

All the standard cases of reflexives not involving picture-nouns will undergo the reflexive rule on the first cycle they are encountered, since the subject of a sentence will always fulfill the conditions for interpretation as coreferential with a reflexive in the predicate. However, in sentences containing derived nominals there are two cycles to be considered.

(91)   John showed Bill a picture of himself.
(92)   *John showed Bill Mary's picture of himself.

In (91), the first cycle deals with the phrase <u>a picture</u> <u>of himself</u>. <u>Himself</u> obeys conditions (a) and (b) of the reflexive rule, but there is no NP^1 in the domain of this cycle, so nothing happens. The second cycle includes the whole sentence. On this cycle, <u>John</u> and <u>Bill</u> are eligible to be NP^1 in the reflexive rule, and so one or the other of them must be assigned as coreferential with himself, producing the two possible readings.

On the other hand, the first cycle in (92) deals with <u>Mary's picture of himself</u>. <u>Mary</u> is eligible to be NP^1 in the reflexive coreferential with <u>Mary</u>. Of course, the reading is then anomalous because of gender disagreement, and so the sentence is ungrammatical. Because of the presence of <u>Mary</u> on the first cycle, the reflexive rule does not get a chance to pair <u>himself</u> with <u>John</u> or <u>Bill</u> on the second cycle.

It is interesting to note that there is no natural way to state a transformation which is cognate to this cyclic reflexive rule. To see this, assume a transformational theory of reflexives, and consider the underlying structure (93).

(93) $John_i$ showed $Bill_j$ $Mary_k$'s picture of $John_i$.

On the first cycle, <u>Mary's picture of John</u>, nothing can happen since there are no two identical noun phrases. On the second cycle, the whole sentence, we face the old problem of how to prohibit reflexivization. The only way to do it is to mark on the first cycle that <u>John</u> cannot be subject to reflexivization because it and <u>Mary</u> fulfill all of the conditions for reflexivization

except identity. The reflexive transformation will thus have
to look like (94).

$$(94) \quad NP^2 \Rightarrow \begin{cases} \text{[+reflexive] if } NP^2 \text{ is} \\ \quad \text{identical with } NP^1 \\ \text{[-reflexivization rule]} \\ \quad \text{if } NP^2 \text{ is not identi-} \\ \quad \text{cal with } NP^1 \end{cases} \quad \text{in the environ-} \\ \text{ment...}$$

The second half of this rule, which introduces a rule feature
under conditions of non-identity, is very strange and constitutes
strong evidence against a transformational theory of reflexives.

Let us discuss some of the sentences mentioned earlier
which blatantly violate the simplex sentence condition of Lees
and Klima's reflexive rule.

(12)   Tom believes that there is a picture of himself
       hanging in the post office.
(15)   The fact that there is a picture of himself hanging
       in the post office frightens Tom.

On the first cycle, only <u>a picture of himself</u> is involved, and
there is no way for the reflexive rule to apply. The second cycle
involves <u>there is a picture of himself hanging in the post office</u>.
<u>There</u> is a dummy noun phrase, totally non-referential; this seems
sufficient reason not to consider it as a possible coreferent of
<u>himself</u>. <u>Post office</u> is excluded for reasons I don't understand.
I will show some relevant examples in a moment. Since there is
no pair of noun phrases meeting the conditions for reflexivization
on the second cycle, we go on to the third cycle, where <u>Tom</u> is
encountered, providing the correct reading.

Returning to the question of excluding <u>post office</u> in these
examples, we notice the following contrasts.

(95) The newspaper printed a story about itself.
(96) *A story about itself appeared in the newspaper.
(97) A portrait of himself hangs in the president's
      bedroom.
(98) *A picture of itself hangs in the Pru's main lobby.

Apparently the conditions for backward reflexivization differ

depending on the animacy of the noun phrases in question. In

similar contexts the same constraints hold for pronominalization:

(99)   The newspaper$_i$ printed Harry's story about it$_i$.
(100) *Harry's story about it$_i$ appeared in the newspaper$_i$.
(101) Mary's portrait of him$_i$ hangs in the president$_i$'s
       bedroom.
(102) *Mary's painting of it$_i$ hangs in the Pru$_i$'s main
       lobby.

These examples show that the generalization of pronominalization

and reflexivization does not break down; but I can offer no

suggestions as to what the proper alteration in the rule might be.

The passive reflexive examples (18) and (19) are no

problem.

(18) The king was hit over the head by a massive por-
      trait of himself which had fallen off the wall.
(19) Unflattering descriptions of himself have been
      banned by LBJ.

Like the normal picture-noun sentences such as (91), the rule

operates on the second cycle to produce coreferentiality. The

difficulty however lies in preventing passive reflexives from

working on the first cycle: how do we prevent (20) and (21)?

(20) *Himself was shaved by John.
(21) *John was shaved by himself.

(20) doesn't work for two reasons: first, because the reflexive

is on the left in the main clause, and second, because the re-

flexive is in the subject, which is not allowed because of condition

(b) of the reflexive rule (43). This condition is independently
needed to prevent (103).

(103) *John thought that himself was sick.

(21) unfortunately is not excluded by any of the conditions
on the reflexive rule. A solution will be proposed in §12.

8. The ordering of reflexivization

Remembering that the environments of pronominalization and
reflexivization are very similar, it would be an important result
if they could be ordered together and hence partially collapsed.
However, Postal claims (1968a, chapter 10, §B) that reflexivi-
zation and pronominalization must be ordered very differently.
Let us examine his arguments.   Postal cannot collapse the
rules at all for two reasons: reflexivization is cyclic, whereas
pronominalization is non-cyclic; and Wh-preposing intervenes be-
tween the two rules. The first argument we have already disposed
of: by making Wh-preposing cyclic, pronominalization can be
made cyclic as well.

The argument that reflexivization precedes Wh-preposing
is based on the fact that in (104) the application of Wh-preposing
on the S^1 cycle causes the S^2 node to prune, since it dominates
only a VP.[6]

(104)

```
                S¹
         /            \
        NP              VP
        |           /        \
       you         V          NP
                   |          |
                  saw         S²
                          /        \
                        NP          VP
                        |        /      \
                   wh-someone   V        NP
                                |        |
                              stab      you
```
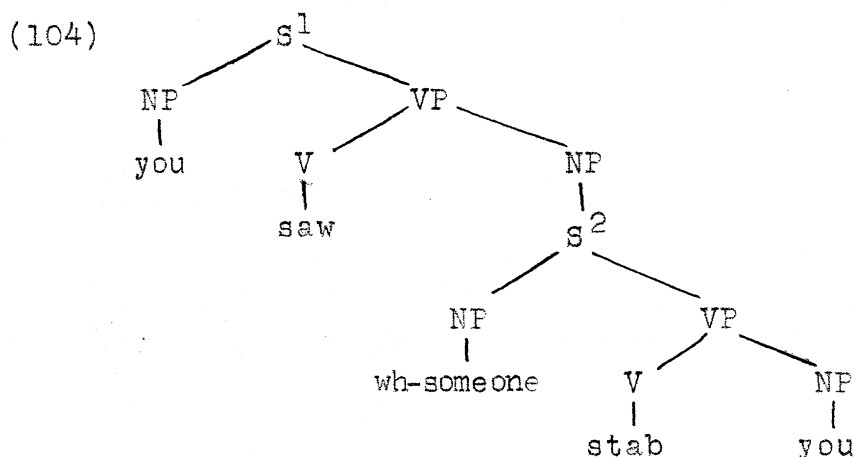
The result is the derived structure (105).

(105)

```
                    S¹
         /          |            \
        NP          NP            VP
        |           |         /        \
   wh-someone      you       V          NP
                            saw         |
                                        VP
                                    /        \
                                   V          NP
                                   |          |
                                 stab        you
```
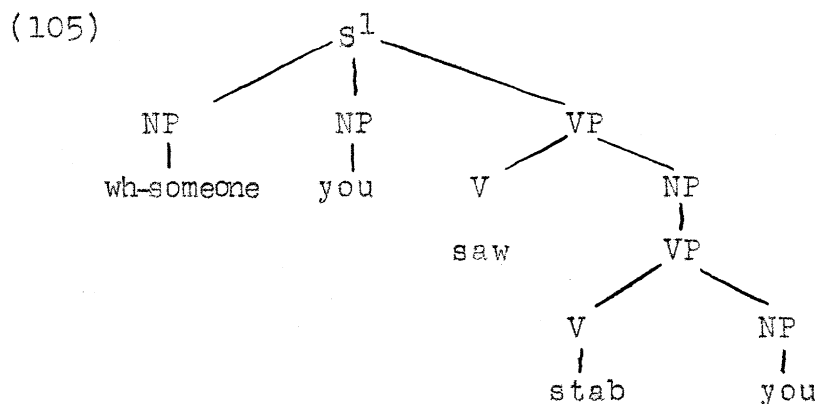
If the reflexivization transformation took place after the
formation of (105), the surface structure would be (106) rather
than the correct (107).

(106)   *Who did you see stab yourself?
(107)   Who did you see stab you?

Therefore, Postal concludes, reflexivization must take place before
S^2 is pruned by the application of Wh-preposing.

In the interpretive theory, however, this problem does
not arise. On the S^2 cycle, you in the object of S^2 will be
marked noncoreferential with wh-someone. Then, in the S^1 cycle,

even if the node S^2 has been pruned, this <u>you</u> does not meet the structural description for reflexivization, since it violates condition (a) of the rule, having been marked once already for coreferentiality. Thus it can undergo normal pronominalization and be marked coreferential with the other <u>you</u>, even though it is now in the same simplex sentence. On the other hand, if <u>yourself</u> is substituted for <u>you</u> in S^2, it is anomalously marked coreferential with <u>wh-someone</u> on the S^2 cycle, so (106) cannot occur.

Thus, even with the pruning of the S^2 node, the interpretive theory of reflexives permits reflexivization to be ordered after Wh-preposing and hence to be collapsed with pronominalization, a significant generalization. Condition (a) of the interpretive reflexivization rule now explains two different phenomena: the picture-noun cases and the Wh-preposing cases. We will see yet one more application in Chapter 3, in connection with the rule that raises complement subjects. On the other hand, the transformational theory, in order to collapse reflexivization and pronominalization, must either give up the generality of the pruning convention or put the very strange condition (94) on the reflexivization transformation.

9. <u>The relative clause cases of reflexivization</u>

This section will tend to throw a certain amount of doubt on the analysis I have presented. However, the evidence will show that, if anything, reflexivization is <u>more</u> dependent on semantics

than I have claimed so far. Thus the transformational theory of reflexives will be made to seem even less plausible.

One pair of sentences which I gave at the beginning is still unaccounted for by the reflexivization rule (90): the sentences involving relative clauses.

(16) The picture of himself that John saw hanging in the post office was ugly.
(17) The description of himself that John gave the police over the phone was a pack of lies.

The structure of these sentences at the time of reflexivization looks like (108). (I assume, with Chomsky, that restrictive relative clauses are daughter-adjuncts to the NP at this stage of the derivation, not Chomsky-adjuncts.)

(108)

```
                              S¹
              NP¹                              VP
   Det          N̄                   S²      was ugly
    |                                 
   the     N         NP    that John gave the police
           |        /  \
        picture  (of) himself
```
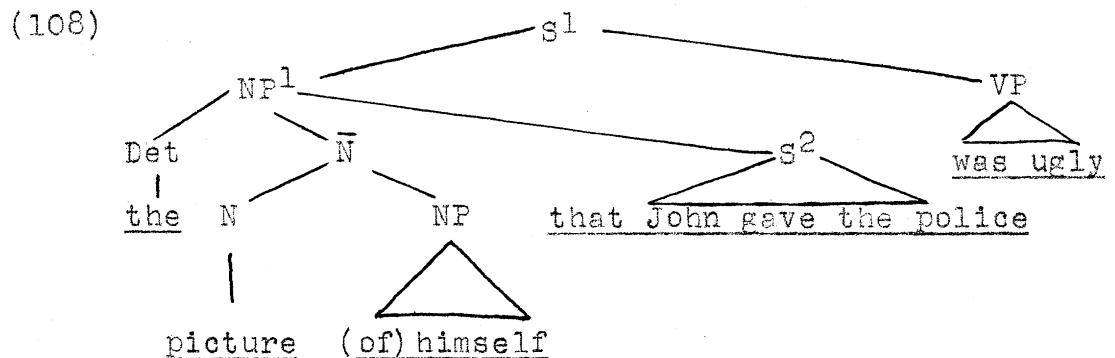
The difficulty is that on the $NP^1$ cycle, when himself is first encountered, John is in the subordinate clause $S^2$, violating condition (c) of the reflexive rule.

Condition (c) was motivated by the nonexistence of sentences like (109)-(111).

(109) *It bothered himself that John was sick.
(110) *The book which I showed John impressed himself.
(111) *I gave a book about John to himself.

(In (111), <u>a book about John</u> counts as a subordinate clause be-
cause it merits a cycle of its own.)  The relevant case to
compare with (16) is (109), which has the structure (112).

(112)

```
                      S¹
         _____/ |  _____
        /             |              \
      NP             VP               S²
       |            /  \             / \
      IT           /    \           /   \
                  V      NP    that John was sick
                  |       |
              bothered  himself
```
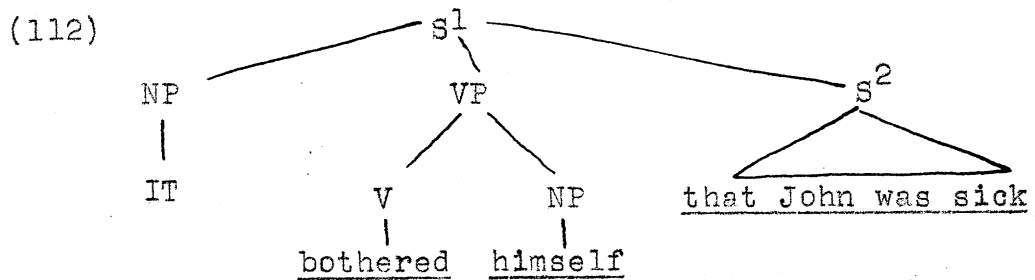
Comparing the whole of (112) to NP^1 in (108), we notice a total
parallelism of structure. Any change in condition (c) that
allows (108), then, will also allow (112), which must be excluded.

One difference between (108) and (112) is that there is
a possibility of a possessive NP in the determiner of (108),
but the subject of (112) admits of no such addition. In other
words, (113) is possible, but (114) is hash.

    (113)   John's picture of himself, which he gave the
              police, was ugly.
    (114) *John bothered himself that he was sick.

<u>John</u> in the determiner would be sufficient to produce the re-
flexive. Therefore one's immediate reaction is to try to say
that all relative clause reflexivizations start out like (113),
with an NP in the determiner which gets deleted later. Unfor-
tunately, this won't work, as there are sentences where putting
the necessary NP in the determiner destroys the correct inter-
pretation:

    (115)   The unflattering descriptions of himself which
            LBJ has banned from the public press still
            circulate underground.

(116)   LBJ's unflattering descriptions of himself which
        he has banned from the public press still circu-
        late underground.

Thus we see that in general a possessive NP cannot be added to
the underlying form without producing drastic results.

Another possible suggestion is that the relative clause
transformation actually moves the relativized noun and its com-
plements from the relative clause into the main sentence. Thus
(108) would be explained by saying that <u>John gave the police a
picture of himself</u> is produced on the S^2 cycle, then <u>picture
of himself</u> is moved up by the relative clause formation. This
explanation fails, too, because reflexivization does not invari-
ably happen in these clauses, particularly with indefinite
articles.

(117)   *I painted a picture of himself that John saw
        yesterday.

Since we probably don't want to have two separate relative clause
transformations, one moving the noun up from the relative clause
and one not doing so, we have to reject this solution too.

Exploring more data, we notice the startling fact that
the choice of verbs in the main clause and the relative clause
affects the acceptability of reflexives in the relativized noun
phrase. We get paradigms like these:

(118)
        I hate the story about $\begin{Bmatrix} *\text{him} \\ \text{himself} \\ \text{me} \\ \text{myself} \end{Bmatrix}$ that John

        always tells.

(119)   I told the story about $\begin{Bmatrix} *\text{him} \\ *\text{himself} \\ *\text{me} \\ \text{myself} \end{Bmatrix}$ that John likes

to hear.

(118) and (119) look the same as far as noun phrase relationships
are concerned:

(120)
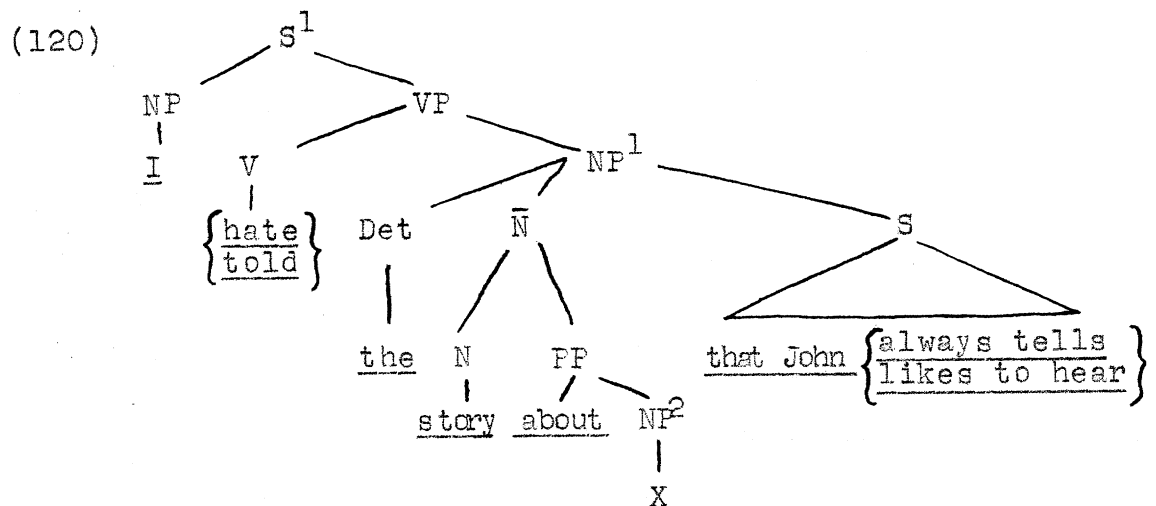
Assume that there is an optional semantic rule that dupli-
cates the subject of a sentence in the determiner of the object,
and that this rule is ordered before reflexivization. If this
rule puts John in the determiner of NP^1 in (120), we will get
the paradigm of (118): himself would be coreferential with John,
me would be noncoreferential, and myself would be coreferential
and thus anomalous. If the rule puts I in the determiner of NP,
we will get the paradigm (119): him cannot refer to John because
the conditions for backward pronominalization are not met, him-
self is anomalously marked coreferential with I, me is anomalously
marked distinct from I, and myself is correctly marked corefer-
ential with I. If the rule does not operate, leaving only the
in the determiner of NP^1, we also get (119): again backwards

pronominalization cannot operate to make $\underline{him}$ coreferential with $\underline{John}$; and on the S^1 cycle, NP^2 relates to $\underline{I}$ just as it did in the NP^1 cycle in the previous case. Thus the postulation of this rule gives exactly the right results.

The conditioning of this rule, as we have noticed, seems to depend on some semantic property of the verb. I suggest that the property in question is related to the subject's performing some sort of direct action on the object. $\underline{Tell}$ certainly implies more direct action than $\underline{hate}$ or $\underline{like\ to\ hear}$, and so in (118) and (119) the sentence containing $\underline{tell}$ overrides the other sentence in connecting its subject with $\underline{story}$. The correct notion may be Agent (see § 11).

The only evidence I have for such a semantic principle aside from these reflexives is contrasts of the following type:

(121) Today I shot my first lion.
       *Today I was scared of my first lion.
(122) Yesterday I told my first Polish joke.
       *Yesterday I heard my first Polish joke.
(123) Today I performed my first Mozart symphony.
       *Today I hated my first Mozart symphony.

In none of these sentences is $\underline{my}$ connected in any way with possession; its semantic relation to the head noun is extremely unclear to me. However, the acceptability of the $\underline{my}$ seems to be correlated again to whether the verb expresses some sort of direct action performed on the object. The same rule that makes the reflexives work could allow the genitive only when connection can be established through the verb.

10. Some unexpectedly bad cases of reflexivization

We still have not accounted for the fact that reflexives are not allowed as the NP in the by-phrase of passives:

(21) *John was shaved by himself.

This cannot be accounted for by ordering reflexive before passive, since (20), the then predicted result, is much, much worse.

(20) *Himself was shaved by John.

Postal (1968a) cites four more sets of cases of unexpectedly bad reflexivization like (21).

Tough Movement:
  (Postal, chapter 3)

(124)   It is $\begin{Bmatrix} \text{easy} \\ \text{tough} \\ \text{a breeze} \\ \vdots \end{Bmatrix}$ for Jack to hit Tony.

(125) Tony is tough for Jack to hit.
(126) It is tough for Tony to shave himself.
(127) *Tony is tough for himself to shave.
(128) *Himself is tough for Tony to shave.

It Replacement
  (Postal, chapter 4)
(129) It seems to me that Schwarz is clever.
(130) Schwarz seems to me to be clever.
(131) It seems to me that I am clever.
(132) ?I seem to myself to be clever.
(133) *Myself seems to me to be clever.

About Movement
  (Postal, chapter 5)
(134) I talked to Mary about Louise.
(135) I talked about Louise to Mary.
(136) I talked to Thmug about himself.
(137) *I talked about Thmug to himself.
(138) *I talked about himself to Thmug.

Psychological Predicates
  (Postal, chapter 6)
(139) I regard myself as pompous.
(140) ?I strike myself as pompous.

(141) I like myself.
(142) ?I please myself.

(143) I smelled myself.
(144) ?I smelled funny to myself.
(145) I am amused with myself.
(146) ?I am amusing to myself.

> ...and a host of similar contrasts: look at/ ?look
> funny to, disgusted with/ ?disgusting to, horrified
> at/ ?horrifying to, irritated at/ ?irritating to,
> loathe/ ?be loathsome to, accept/ ?be acceptable
> to, familiar with/ ?familiar to, be impressed with/
> ?impress, annoy/ ?be annoying to, frighten/
> ?be frightening to, be pleased with/ ?be pleasing to

Postal proposes accounting for these facts by means of a
constraint on movement transformations, called the Crossover
Principle, which says roughly that a transformation may not move
an NP over another NP with which it is coreferential. (20) is
then ruled out because the passive moves the deep subject over
the object into the by-phrase; since the two NPs are coreferential,
the Crossover Principle rules out the result. Likewise, (127)
and (128) are ruled out by the movement of the object of the
subordinate clause over the subject of the subordinate clause;
(132) and (133) are out because of the movement of the complement
subject over the to-phrase (assuming the account of It-Replace-
ment in Rosenbaum (1967)); (137) and (138) are out because of
the movement of the about-phrase over the to-phrase; and (140),
(142), (144), and (146) are out because of an alleged rule of
Psych Movement which derives them from underlying forms in which
the positions of the subject and object are the opposite of their
surface forms.

In an interpretive theory of reference, the Crossover
Principle cannot be stated, since coreferentiality, a purely

semantic concept, cannot be referred to in transformations, and conversely, semantic rules cannot depend on what transformations have taken place, but only on the resulting structural configurations. Furthermore, if we want to maintain the Extended Lexical Hypothesis, we cannot simultaneously maintain the Crossover Principle, since the transformation Psych Movement is just the kind of rule the Extended Lexical Hypothesis seeks to exclude: if (140), (142), (144), and (146) are not the output of a movement rule, then they cannot be ruled out by a constraint on movement.

Therefore these cases of bad reflexivization must be taken care of in some way other than a constraint on movement of coreferential NPs. As there does seem to be a generalization lurking somewhere in the data, Postal is right in claiming that merely naming the particular cases is not enough. Accordingly, I will propose a constraint, involving the semantic functions of reflexives and their antecedents with respect to the verb of the clause. This constraint will handle not only the cases cited above, but a couple of cases which Postal cannot handle conveniently.

11. Thematic relations--a brief discussion

Jeffrey Gruber's dissertation (Gruber 1965) discusses a system of semantic functions that NPs and PPs may perform in a clause. These functions, which I will refer to here as thematic relations, are independent of the notions of subject, object,

indirect object, and so forth. They correspond more closely to
some traditional intuitions about agent and patient, which have
not been taken into account in generative grammar except by
going to a case grammar-formulations, as in Fillmore (1968).
Gruber's formulation provides a possible way of describing the
intuitions of case grammar while retaining standard grammatical
relations in the base. The case-like thematic relations are
treated as properties of semantic interpretations, related to
the grammatical relations (subject, object, etc.) by projection
rules or (in Gruber's formalization) the process of lexical
insertion.

The fundamental semantic function in Gruber's analysis
is the theme. In every sentence there is an NP functioning as
theme. Gruber does not give explicit criteria for determining
which NP is the theme, but some overall considerations emerge
from the dissertation and some subsequent work (Gruber 1967a).
With verbs of motion the theme is defined as the NP understood
as undergoing the motion.

(147) The rock moved away.
(148) John rolled the rock from the dump to the house.
(149) Bill forced the rock into the hole.
(150) Harry gave the book away.
(151) Will inherited a million dollars.
(152) Charlie bought the lamp from Max.
(153) Dave explained the proof to his students.

In the first three examples, involving physical motion, the rock
is obviously theme. Note that depending on the main verb and
the presence of other NPs, theme can be either subject or direct

object. The other four examples illustrate non-concrete types
of motion, to which the definition of theme is extended by
analogy. The book, a million dollars, and the lamp are themes
undergoing change in possession rather than physical position.
The proof, or rather information about it, is undergoing change
in some sort of abstract position. This last kind of motion is
the least conceptually transparent, but it is important because
it will be of greatest interest in subsequent discussion here.

With verbs of location, the theme is defined as the NP
whose location is being asserted.

        (154)   The rock stood in the corner.
        (155)   John clung to the window sill.
        (156)   Herman kept the book on the shelf.
        (157)   Herman kept the book (instead of selling it).
        (158)   The book belongs to Herman.
        (159)   Max owns the book.
        (160)   Max knows the answer.

In the first three examples, we are dealing with physical location,
and the rock, John, and the book are theme. The next three in-
volve possessional location, and the book is theme in each case.[7]
The last example is an abstract analogue of possession, so the
answer is theme.

An important principle to notice about the extension of
the notion theme to abstract verbs is that when a verb can be
used to express motion or location in different domains, the
theme occupies the same syntactic position. For example, keep
in (156) expresses physical location, and in (157) possessional
location, but the book is theme in both cases. This principle

follows from the belief that a verb is fundamentally the same in its different uses. A large number of such correspondences are discussed in Gruber's thesis.

A further principle that emerges from Gruber's work is that the themes of morphologically related words are in semantically parallel positions. For example, consider the following two sentences.

(161) The circle contains the dot.
(162) The dot is contained in the circle.

In (161) it is not clear which NP is the theme and which is the location. But (162) has the preposition in, an unmistakable mark of a location phrase, so the dot must be the theme. Therefore, according to the above claim, the dot must also be the theme in (161). Arguments similar to this are used throughout Gruber's work.

Besides the theme, Gruber works with several other thematic relations. I will discuss only four more here. The first three of these are the expression of Location, Source, and Goal. Location is defined as the thematic relation associated with the NP expressing the location, in a sentence with a verb of location. It is often, but not always, in a PP: (154), (155), (156), and (158) have a preposition, and (157) (Herman), (159), and (160) (Max) have none. Adjectives can function as abstract locations, as if they meant "in the abstract domain (of 'quality space') containing those things which are Adj." For example, stay can express either a physical or an abstract location:

(163) John stayed in the room.
(164) John stayed angry.

Coresponding to Location with verbs of location, we have the thematic relations Source and Goal with verbs of motion. Like Location, these are often expressed with a PP, but not always. The dump in (148) and Max in (152) are clear expressions of source. The house in (148), the hole in (149), and his students in (153) are expressions of goal marked by a PP. Away in (147) and (150) is analyzed by Gruber (§5.3) to mean something like "to another place" so it is also an expression of goal. Will in (151) and Charlie in (152) are expressions of goal in the subject. As with abstract Location, abstract Source and Goal may be also filled by an adjective: compare the following examples.

(165) George got to Philadelphia.
(166) George got angry.
(167) Harry went from Bloomington to Boston.
(168) Harry went from elated to depressed.

The last thematic relation I will discuss is Agent. The Agent NP is to be identified by a semantic reading which attributes to the NP will or volition towards the action expressed by the sentence. Hence only animate NPs can function as agents.[8]

The agent is generally in the subject, but the subject can simultaneously take part in other thematic relations. For example, in (147) there is no theme, but if we change the rock to John, there is a reading in which John deliberately moved away, so John is functioning as both agent and theme. In (148), John is only the agent, and in (149), Bill, since in both cases

it is the rock that moves, i.e. is theme.[9] In (150) (and possibly (153)) the subject is functioning as both source and agent, whereas in (152) the subject is goal and agent.[10] In (151), however, there is no agent: an act of volition is not being attributed to Will since one cannot say, for example Will inherited the money intentionally.

The Agent subject correlates with the possibility of using purposive constructions like in order to and so that, and purpose adverbials such as intentionally, accidentally, on purpose.

(169) *The rock deliberately rolled down the hill.
(170)   John deliberately rolled down the hill.

(171) *John received the book from Bill in order to
          read it.
(172)   John took the book from Bill in order to read it.

(173) ?John lost the money so that he could get sympathy.
(174)   John gave the money away so that he could win
          his friend's admiration.

(175) ?John intentionally struck Bill as funny.
(176)   John intentionally made Bill think of him as funny.

The first two of these examples show the difference between an inanimate subject acting as theme only and an animate subject functioning both as theme and agent. (171) is an example of an animate subject which is not an agent. Contrast this with (172), which expresses the same semantic content but with the added proviso that the subject is agent, permitting the in order to phrase. (173) and (174) form a similar pair. Finally, (175) is a type of example we will return to in the next section; the inappropriateness of the adverb indicates that the subject is not an agent. In (176), however, the volition on John's part is

expressed explicitly, and so the adverb is permissible.

Likewise, imperatives are permissible only for agent subjects:

(177) *Receive the book from Bill.
       ?Lose the money.
       ?Strike Bill as funny.

This follows naturally from the fact that successfully carrying out an order requires that the order allow for volition (i.e. agenthood) on the part of the hearer.

The lack of an agent subject in a sentence generally correlates with the possibility of embedding it as a gerund under such verbs as resent and accept:

(178) John resented { inheriting the money.
                       *hitting Bill
                       having to hit Bill.

As usual, I am only giving a skeleton of Gruber's lengthy analysis.

12. Interaction of thematic relations and reflexivization

Set up the following hierarchy of thematic relations:

(179) 1. Agent
       2. Location, Source, Goal
       3. Theme

I propose the following restriction on reflexives:

(180) (The Thematic Hierarchy Condition)
       A reflexive may not be higher on the hierarchy
       (179) than its antecedent (i.e. NP^1 in (90)).

Violation of this condition will result in sentences which are not fully grammatical, yet not nearly as bad as sentences which violate the structural conditions of the reflexive rule.

The Thematic Hierarchy Condition is a rather curious condition, in that it simultaneously refers to properties of the

semantic reading (the thematic relation and coreference), and the feature _reflexive_, which is presumably not a part of the reading. Therefore either it must be a condition involving two levels of structure, or else it must be taken as a statement about some inherent semantic content of the feature _reflexive_, heretofore undiscovered. I do not know how to decide between these alternatives. Whatever the meaning of the condition, it is consistent with an interpretive theory of pronouns, whereas the Crossover Principle is not. If second alternative interpretation of the Thematic Hierarchy Condition should turn out to be correct, this would argue against maintaining both a transformational theory of reflexives and the Katz-Postal Hypothesis, since the reflexivization transformation would add meaning. I don't know whether or not the first interpretation of the condition is consistent with a transformational theory of reflexives.

I will attempt to show the plausibility of using the Thematic Hierarchy Condition to account for the facts given in §10. Doubtless a full discussion would be at least as long as Postal's 235-page monograph on the Crossover Principle; obviously I must defer such a work to a later date. However I hope the outlines of an argument will be clear from the present brief exposition.

a. Passive

First look at (20) and (21), the passive reflexives.

(20) *Himself was shaved by John.
(21) *John was shaved by himself.

John is the deep structure object and the theme of the sentence. Himself is the deep structure subject and the agent. Therefore himself is higher on the hierarchy (179) than John, and so (180) is violated. (21) obeys (180) but violates the structural condition for the reflexivization rule, so is much more crashingly ungrammatical. In the active sentence, the agent is to the left of the theme, so both the structural condition of the reflexivization rule and (180) can be met, permitting good sentences.

What about the actives and passives of sentences that do not necessarily have an agent subject? Take for example sentences with touch or hit, which have theme in the subject and location or goal in the object, but only optionally an agent in the subject.

(181) The tree was touching the wire.
(182) The falling rock hit the car with a crash.
(183) John was touching the bookcase.
(184) John hit the car with a crash.

(183) and (184) are ambiguous between an agent and non-agent reading of John. However, the parallel reflexive sentences seem only to have the agent reading.

(185) John was touching himself.
(186) John hit himself (??with a crash).

This is predicted by (180). For if John were only the theme, the hierarchy would be violated. Therefore the only acceptable reading has John as agent as well.

We would expect then that the passives of (185) and (186) would be acceptable, in the interpretation that the deep subject is only the theme. Unfortunately this is not the case.

(187)   *John was being touched by himself.
(188)   *John was hit by himself.

However, there is a fact about the passive by-phrase we have not

taken into account. If the NF in the by-phrase is animate,

i.e. a potential agent, it seems to be invariably understood as

an agent (or at least as non-theme). For example, the passives

of (183) and (184) are unambiguous, having only the agent

reading.

(189)   The bookcase was (being) touched by John.
(190)   The car was hit by John (?with a crash).

Since the passive forces an agent reading on the by-phrase of

(187) and (188), the **reflexives violate (180)**, and the sentences

sound funny as usual. We will have more evidence for this special

effect of the passive in a moment.

What would disprove this alleged explanation of the re-

flexive passive? If there were a sentence with an inanimate (and

therefore non-agent) subject, but with a semantically plausible

reflexive object, and with a verb like touch, which puts theme

in the subject, such a sentence would violate (180), since the

reflexive would have to be higher on the hierarchy than its

antecedent. Such a sentence is the tree was touching itself,

which sounds to me somewhat worse than (185). According to the

analysis so far, this feeling should be due to either a violation

of (180) or a dubious attribution of volition to trees. At this

point I would rather not speculate any further, though; I leave

this line of questioning for subsequent research.

b. Psychological predicates

Actually, there are some verbs which potentially yield violations of (180), for example, impress NP as, strike, and please. To see this, notice that in (191) the adjective, functioning as an abstract location, is attributed to the subject, whereas in (192) the adjective is attributed to the object.

(191) Bill $\begin{Bmatrix} \text{strikes} \\ \text{impresses} \end{Bmatrix}$ Harry as pompous.

(192) Bill regards Harry as pompous.

This means that the subject is theme with strike and impress, but the object is theme with regard. With strike and impress, we have the alternative forms (193).

(193) Harry is $\begin{Bmatrix} \text{striking} \\ \text{impressive} \end{Bmatrix}$ to Bill.

The presence of to here shows that the object of strike and impress is some sort of goal. Presumably the subject of regard is also a goal and possibly an agent, although we cannot verify this quite so directly.

Now notice what (180) predicts about reflexives with these verbs. Since with strike and impress the theme is on the left, but the reflexivization rule only permits reflexives on the right, the condition should be violated and we predict (correctly) the ungrammaticality of (140). Regard, with the theme on the right, should function normally with respect to reflexivization, and as expected, (148) is all right. The passive of (139) then is out for violation of (180). We would expect the passive of (140) to be acceptable, but it doesn't exist:

b. Psychological predicates

Actually, there are some verbs which potentially yield violations of (180), for example, impress NP as, strike, and please. To see this, notice that in (191) the adjective, functioning as an abstract location, is attributed to the subject, whereas in (192) the adjective is attributed to the object.

(191) Bill $\begin{Bmatrix} \text{strikes} \\ \text{impresses} \end{Bmatrix}$ Harry as pompous.

(192) Bill regards Harry as pompous.

This means that the subject is theme with strike and impress, but the object is theme with regard. With strike and impress, we have the alternative forms (193).

(193) Harry is $\begin{Bmatrix} \text{striking} \\ \text{impressive} \end{Bmatrix}$ to Bill.

The presence of to here shows that the object of strike and impress is some sort of goal. Presumably the subject of regard is also a goal and possibly an agent, although we cannot verify this quite so directly.

Now notice what (180) predicts about reflexives with these verbs. Since with strike and impress the theme is on the left, but the reflexivization rule only permits reflexives on the right, the condition should be violated and we predict (correctly) the ungrammaticality of (140). Regard, with the theme on the right, should function normally with respect to reflexivization, and as expected, (148) is all right. The passive of (139) then is out for violation of (180). We would expect the passive of (140) to be acceptable, but it doesn't exist:

strike and impress do not passivize. But we have accounted for
this already, in noting that the passive by-phrase must be read
as agent or non-theme if the NP is animate. Since the subjects
of strike and impress can only be theme,[11] the reading of the
passive is anomalous. Regard can passivized, since the by-
phrase may be an expression of goal.

Please also has the theme in the subject and a goal in
the direct object, as is pointed out by the variant is pleasing to.
Thus it also cannot have reflexive objects without violating
(180). Note, too, that its passive sounds a bit strange unless
the by-phrase is either inanimate or construed as an agent:

(194) Bill was pleased by $\begin{cases} \text{?Harry;} \\ \text{Harry's performance.} \end{cases}$

To fill the semantic gap, there is the alternative
pleased with.

(195) Bill was pleased with Harry.

I am not sure what the thematic relations of pleased with are.
By the usual arguments from parallelism, the subject should be
goal and the with-phrase theme. This correctly predicts the
good reflexives with pleased with, amused with (145), disgusted
with, and impressed with, so this analysis and (180) support
each other. But Gruber does not discuss thematic relations in-
volving this with-phrase, so we are off into further research
again.

Returning to the adjectives with to-phrases (e.g. striking
impressive, pleasing, amusing, disgusting, irritating, horrifying,

familiar, annoying, and frightening),[12] we remember that the to-phrase marks an expression of goal. This means that the theme is in the subject, which is what we predict if the adjective attributes an abstract location to the theme. If the subject is purely theme, then a reflexive in the to-phrase will violate (180).

Can the subject of these adjectives be agent as well? If so, reflexives in the to-phrase should be permissible. The use of progressive aspect with these adjectives, where at all felicitous, seems to give a reading of volition to the subject:

(196) John is being $\begin{cases} \text{disgusting} \\ \text{amusing} \\ \text{frightening} \\ \text{impressive} \end{cases}$

But in this usage the expression of goal seems to be ruled out: adding to Bill in (196) makes the sentence sound funny. We predict from (180) that the addition of to himself in (196) should not make it any worse than to Bill does. But my judgments fail me at this point.

The last of the psychological predicates are the ones that occur in patterns like (144) (NP V Adj to NP): look, sound, taste, and smell. But these fall into a by now familiar pattern: the to-phrase is a mark of the expression of goal, the adjective is an abstract location attributed to the subject, and so the subject is the theme. The subject is not an agent: *I smelled funny to Bill in order to get rid of him.[13] Therefore these verbs cannot take part in reflexive sentences either, without violating

(180). On the other hand, the usage of <u>smell</u> in (143) allows

an agent subject (<u>I smelled myself in order to see if I needed</u>

<u>a shower</u>) so (180) can be met, permitting reflexivization.

c. About Movement

```
(134)   I talked to Mary about Louise.
(135)   I talked about Louise to Mary.
(136)   I talked to Thmug about himself.
(137)  *I talked about Thmug to himself.
(138)  *I talked about himself to Thmug.
```

<u>Talk</u>, <u>tell</u>, <u>speak</u>, etc. can be analyzed as verbs of motion, where

the thing undergoing the motion is the information being conveyed

by the speech-act (see Gruber, Chapter 7, § 2). This means that

the subject is agent and source, and the <u>to</u>-phrase is the goal.

Then the <u>about</u>-phrase must be the theme. Gruber does not mention

any instances of themes in PPs, but an extension of his arguments

about the theme of <u>look</u> and <u>see</u> (in Gruber (1967)) might lead

one to conclude that, for example, in (134), the theme is some

abstract instantiation of <u>Louise</u>. If the <u>about</u>-phrase is theme,

then we can explain (137) as follows. In (136) the goal is to

the left of the theme, so condition (180) and the structural

description of the reflexive rule can be met simultaneously.

But in (137), the theme is to the left of the goal, so (180) must

be violated. (138) of course violates the reflexivization rule.

Consider also the following sentences.

```
(197)   I talked to Thmug about myself.
(198)   I talked about myself to Thmug.
(199)   I talked to myself about myself.
(200)   I talked about myself to myself.
```

(197) and (198) are both all right, since <u>I</u>, the agent and source,

is to the left of the theme whether or not about movement takes place. (199), which has the to and about-phrases in underlying order, has reflexives in both goal and theme, with the agent as antecedent. (200), which for Postal should be a violation of the Crossover Principle, is unexpectedly good, and he goes to a lot of trouble to explain it away. The present account here takes care of (200) easily, though. Since both reflexives have the agent as antecedent, no violation of (180) occurs and so the sentence is acceptable.

d. It-replacement

(129) It seems to me that Schwarz is clever.
(130) Schwarz seems to me to be clever.
(131) It seems to me that I am clever.
(132) ?I seem to myself to be clever.
(133) *Myself seems to me to be clever.

(133) is of course out because it violates the reflexive rule. But (132) is more difficult. Let us make the usual assumption that Schwarz in (130) and I in (132) arise from a rule moving them from the complement subject into the position of the it in (129) and (131) respectively. The Thematic Hierarchy Condition as stated will not account for all cases of this sort, since it is easy to find a sentence in which the subject is an agent, yet the reflexive to-phrase (a goal) is still bad.

(201) ?John appears to himself to have hit Bill.

Since there is nothing we can do about the thematic relations in (201), we will have to modify the Thematic Hierarchy Condition. One possibility that suggests itself is to say that the condition

only applies to thematic relations <u>induced by the verb of the</u> <u>clause</u> that the reflexive and its antecedent are in when reflexi-vization takes place. It is reasonable that a raised NP such as the derived subject of <u>seem</u> has no thematic relation at all with respect to its new clause, since thematic relations are re-lated to deep structure grammatical relations; <u>John</u> in (201) is an agent only with respect to the complement. Then, to prevent (132) and (201), we need only assume that lack of thematic relation counts as the lowest position on the thematic hierarchy (179). Under this assumption, the derived subject, having no thematic relation, is lower on the hierarchy than the reflexive, which is in an expression of goal.

e.   Tough Movement

```
(124)   It is tough for Jack to hit Tony.
(125)   Tony is tough for Jack to hit.
(126)   It is tough for Tony to shave himself.
(127)   *Tony is tough for himself to shave.
(128)   *Himself is tough for Tony to shave.
```

Postal claims that there is a rule Tough Movement that takes the object out of the complement sentence in (124) and puts it in the subject of the main clause to produce (125). This, he says, causes a crossover violation in cases like (127). In a prefatory note to the monograph, Postal remarks that he no longer believes that a movement rule is involved in producing (125), but rather a deletion rule under coreference. The reason is probably (he does not say) that a movement rule would change the meaning, pro-ducing a semantic contrast in examples like those pointed out by

Klima:

    (202)   Sonatas are easy to play on this violin.
    (203)   This violin is easy to play sonatas on.

Postal realizes that giving up Tough Movement throws the whole Crossover Principle into doubt, but in his brief note he gives no hint of what he thinks might be a possible alternative. In Chapter 3, § 8, however, we will show that the difference in meaning between these two is a natural consequence of the derived structure interpretation rules for Focus, making it possible to retain Tough Movement in the theory of grammar developed here.

The crucial factor in these sentences, I claim, is that the for-phrase is a deep structure constituent of the main clause, but the subject is a deep structure constituent of the complement. If this is the case, the subject has no thematic relation in the main clause, whereas the for-phrase does. Applying the modification of the Thematic Hierarchy Condition we have just proposed to handle It-Replacement, we see that whatever the thematic relation of the for-phrase, it will be higher on the hierarchy than the subject. Hence a reflexive in the for-phrase will violate the condition and the sentence will be unacceptable.

To establish the claim that the for-phrase in (127) is a constituent of the main clause, consider the following sentences.

    (204)   The problem is easy (for John).
    (205)   Shaving Bill is easy (for John).
    (206)   For John, shaving Bill is easy.

(204)-(206) show that there is an independently motivated for-

phrase in the complement of easy. Now we must show that when
Tough Movement takes place, the for-phrase that appears is this
main clause for-phrase and not the for-phrase of the for-to
complement.

    (207)   *John's shaving Bill is easy.
    (208)   ?To shave Bill is easy (for John).
    (209)   It is easy (for John) to shave Bill.
    (210)   *For John to shave Bill is easy.

These examples show that usually a subject must be deleted from
the complement of easy. (207) shows that a gerund subject cannot
itself have a subject. With an infinitive subject (208), extra-
position is preferred, as in (209), but an infinitive in subject
position is much worse if it has an overt subject (210).

        There are a few cases where both for-phrases show up on
the surface.

    (211)   It would be easier for me for John to do the job
            than for me to do it myself.
    (212)   It is a waste of time for me for John to try to
            help with this job.

However, these sentences do not have the corresponding form with
the object fronted from the subordinate clause.

    (213)   *This job would be easier for me for John to do
            (than for me to do myself).
    (214)   *This job is a waste of time for me for John to
            try to help with.

        For further evidence that the for-phrase is part of the
main clause, notice that in the cases where the complement object
is fronted, the for-phrase can still prepose as in (206).

    (215) For John, Bill is easy to shave.

These examples argue that Tough Movement can take place only if the complement subject has been deleted, since every time there is an acceptable _for_-phrase in such sentences, it turns out to be the main clause _for_-phrase. If the _for_-phrase is in the main clause, the Thematic Hierarchy Condition can account for the violation (127) in the manner shown above.

f. Dative Shifts

In addition to the cases given so far, there are two cases which Postal cites (in Chapter 15) as "non-counterevidence" to the Crossover Principle. These actually turn out to be evidence against the Crossover Principle and for the Thematic Hierarchy Condition. The rules are To-Dative Shift and For-Dative Shift, the effects of which we now illustrate.

(216) Dave sold a book to Pete. } (To-Dative)
(217) Dave sold Pete a book. }

(218) Dave bought a book for Pete. } (For-Dative)
(219) Dave bought Pete a book. }

Each of these rules permutes the two objects and deletes or inserts a preposition, depending on one's assumptions about the underlying order.

Which member of the pairs above has the underlying order of objects? I maintain that the upper member is more primitive, since the order of complements NP-PP is widespread in English and hence a plausible base rule, whereas the order PP-NP or NP-NP is unknown except in these two constructions. Hence for economy in the base it seems wiser to assume that the dative rules pro-

duce (171) from (170), and (173) from (172).

Furthermore, in the nominalized form, the direct object always directly follows the head and is followed by the to or for-object.

(220) Dave's sale of a book to Pete
(221) ?Dave's sale to Pete of a book
(222) *Dave's sale Pete of a book

(223) Dave's painting of a picture for Mary
(224) *Dave's painting for Mary of a picture
(225) *Dave's painting Mary of a picture

Under our usual assumptions about the base rule schema, as in Chomsky (1969), these facts can best be captured by supposing (216) and (218) to have the underlying order of objects in sentences, corresponding to (220) and (223). The two dative rules, however, do not generalize to NP constructions, but occur only in sentences; thus they only produce (217) and (218), not (221), (222), (224), or (225).

Postal points out that the direct and indirect objects cannot be coreferential, no matter what their order.

(226) ?*I sold the slave to himself.
(227) *I sold the slave himself.
(228) ?*I bought the slave for himself.
(229) *I bought the slave himself.

Since only (227) and (229) can be instances of crossover violations, Postal alludes to "some mysterious, independent constraint which prevents the direct and indirect objects from being coreferential in such cases" (p. 114).

Actually, these examples are a great deal more transparent in light of the present treatment of reflexivization. Consider

the thematic relations of these sentences. With <u>sell</u>, the subject is agent and source, the direct object is theme, and the indirect object is goal. With <u>buy</u>, the subject is agent, the direct object is theme, and the <u>for</u>-phrase is goal. The <u>from</u>-phrase, if it occurs, is source. In (226), then, the reflexive is in the goal, but its antecedent is theme, so (180) is violated. Likewise, in (228), the reflexive is in the goal, but its antecedent is theme, so (180) rules the sentence out.

What then of (227) and (229), where the reflexive is theme and its antecedent goal, so that (180) is met? We can account for these by appeal to the well-known constraint that Dative Shifts may not move pronouns out of post-verbal position:

(230)   I gave it to John.
(231)   *I gave John it.
(232)   I bought it for John.
(233)   *I bought John it.

Thus we see that the combination of (180) and this condition on Dative Shift rules out both possible orders, accounting nicely for (226)-(229). Notice also that the relative force of the two conditions we have invoked accounts for the difference in acceptability, correctly predicting (227) and (229) to be worse than (226) and (228).

The Crossover Principle, on the other hand, can only rule out (227) and (229), since no movement takes place in (226) and (228). (226) and (228) cannot be ruled out by a condition on pronouns like that which rules out (231) and (233), since pronouns

are acceptable in <u>to</u> and <u>for</u>-phrases:

(234)   I gave the book to him.
(235)   I bought the book for him.

g.   Emphasis

The only reflexive cases Postal explains that I haven't
yet are emphatic reflexives, which are often acceptable even when
they violate the Crossover Principle:

(236)   John was shaved by <u>himself</u>.

Postal explains these by getting them from Pseudo-Cleft sentences
by a derivation that somehow manages not to violate the Crossover
Principle. Going to such great lengths seems to me unnecessary.
Notice that (236) can only be used as an answer to some question
like

(237)   Who was John shaved by?

Thus in this particular case, there is no way to answer which is
more appropriate than (236). Also, for me at least, (238) sounds
better than (236).

(238)   John was shaved by <u>Bill</u>.

This suggests that (236) in fact is a violation of the Crossover
Principle or the Thematic Hierarchy Condition, but that the exi-
gencies of the discourse permit one to override the relatively
weak force of the violation.

What evidence is there to decide whether the Crossover
Principle or Thematic Hierarchy Condition is the better way of
accounting for all these facts of reflexivization? It is of
course impossible to compare them in relative complexity. But we

can compare them with respect to the number of phenomena they
can explain, and the Thematic Hierarchy Condition seems to ex-
plain more, if the arguments presented here can be tightened up.
It explains the Dative Shift case, which the Crossover Principle
cannot. It explains the Tough Movement case, which the Cross-
over Principle cannot, since under the Katz-Postal Hypothesis,
Tough Movement cannot be a movement rule. It explains the double
reflexive cases of About Movement, which the Crossover Principle
cannot explain without resort to a more complex deep structure.

The other cases Postal discusses in connection with the
Crossover Principle have to do with pronouns. We will turn to
these cases now.

13. Some unexpectedly bad cases of pronominalization

Postal (Chapter 10) brings up a class of sentences in
which understood coreference apparently should be possible, but
is not. Contrast the following pairs of questions, in which he
or himself and who are to be understood as coreferential.

(239) Who shaved himself?

(240) *Who did $\begin{Bmatrix} \text{he} \\ \text{himself} \end{Bmatrix}$ shave?

(241) Who does Mary think hurt himself?
(242) *Who does Mary think he hurt?
(243) Who did the police accuse of trying to enrich himself?
(244) *Who did the police accuse him of trying to enrich?

(245) *Who did the girl $\begin{Bmatrix} \text{he hated} \\ \text{who hates him} \end{Bmatrix}$ describe?

(246) *Who did you talk to the girl $\begin{Bmatrix} \text{he likes} \\ \text{who likes him} \end{Bmatrix}$ about?

Exactly the same judgments obtain for parallel relative clause con-

structions. Postal suggests that the Crossover Principle can take care of these bad sentences, since in each case, Wh-preposing has moved who over the pronoun. Then, to account for the fact that (247) is good,

(247) Who that Charley knows did he criticize?

Postal incorporates Ross's suggestion that the Crossover Principle be operative only when that one of the coreferent NPs that moves is the NP "mentioned" in the structural description of the rule that moves it. Then, to account for the fact that (248) is bad, even though the wh-word is only a subpart of the NP being moved,

(248) *Evidence for whose claim did he deduce?
                        (contrast with
            Who deduced evidence for his claim?)

Postal goes through a long discussion of the Pied Piping Convention, first proposed in Ross (1967a), which enables the wh-fronting rule to "mention" only the wh-word but still carry along larger NPs.

I won't discuss Postal's analysis in detail because it seems to me to be unable to handle the whole range of relevant facts. Consider the following sentences. (Who and he to be coreferential, as usual.)

(249) *Whose mother did you talk to the girl who likes him about?
(250) *Whose mother did you talk about to the girl who likes him?
(251) *Whose mother did you talk about the girl who likes him to?
(252) *Whose mother did you talk to about the girl who likes him?

The Crossover Principle can handle (249) and (251), since whose

is "mentioned" by the fronting rule and it crosses over <u>him</u>.
But (250) and (252) cannot be covered, since <u>whose mother</u> has
been moved away from the left of <u>him</u>. It might be argued that
the violation in these sentences is produced by the crossing in
About Movement, but this proposal can only handle one of the
two: whichever one has the underlying order of <u>to</u> and <u>about</u>-
phrases still has no crossover violation. Furthermore, an appeal
to About Movement would be specious anyway, since About Movement
would have to "mention" the whole NP <u>whose mother</u> in moving the
PP.

Another counterexample concerning About Movement is the
fact that (253) seems good, to me at least, in fact better than
the non-question version (254).

(253) Whom did you talk about himself to?
(254) ?I talked about himself to John.

It is only by virtue of the fact that <u>whom</u> has crossed over that
(253) is good. These examples also argue for ordering reflexi-
vization after wh-preposing, as we claimed in §8, since only if
this is the case can the two sentences be distinguished in accepti-
bility.

In a search for another way to deal with these examples,
note first that in the theory given here, the anomalies of (242)
and (244) can be reduced to that of (240): with a cyclic rule
of wh-fronting, there will be an intermediate stage in their deri-
vation where they will look just like (240). Therefore we have
at least two cases: (a) no coreference between <u>wh</u>- and the subject

(240, 242, 244), and (b) no coreference between <u>who</u> and pronouns which are certain in relative clauses (245, 246, 248-252).

Going to a wider range of data, we find some other constructions which have paradigms at least partially like questions.[14]

 (255) In Mary's apartment, a thief assaulted her.
 (256) ?In Mary's apartment, she was assaulted by a thief.
 (257) It was John's dog that bit him.
 (258) *It was John's dog that he bit.

Note that in none of these examples has a movement rule "mentioned" one of the coreferential NPs. In (255)-(256) this is obvious, since the adverb preposing rule moves adverbials. In the cleft sentences, there is a relative clause movement involved, but the NP "mentioned" by the relative clause rule is <u>John's dog</u>, not <u>John</u>. But notice that if the crossover condition were altered to cover (256), it would also cover (255), since the adverbial moves from the extreme right-hand end of the sentence. Furthermore, Lakoff shows (1968b) that neither the adverbial cases nor the cleft cases can be handled by ordering a rule of adverbial fronting or cleft formation <u>after</u> pronominalization, since pronouns within relative clauses contained in the adverbial or clefted NP exhibit the possibilities one would expect if the rule were ordered <u>before</u> pronominalization. Since neither the Crossover Principle nor ordering of rules can explain these examples, we conclude that there must be some additional specifications in the environment for pronominalization, or else some other principle not yet discovered. I won't try to discuss here what modifications must be made. But whatever the principle is,

it can be seen to easily generalize to cover case (a) of the wh-
anomalies, since the relevant distinction is subject versus
non-subject.

Thus we have found an independently needed constraint that
handles case (a) of pronominalization violations allegedly due
to the Crossover Principle. What about case (b)? There seems
to be a variety of factors interacting in these sentences. For
one thing, some sentences of this form sound bad even without
an anaphoric pronoun: (259) doesn't seem appreciably worse than
(260).

(259) Who did the man who hated him see?
(260) Who did the man who hated John see?

This seems to be only a stylistic violation of some sort, but it
does serve to confuse the issue of coreferentiality in (259).

Second, there seems to be a great deal less freedom in
the way pronouns can relate to whose than there is in their use
with definite or even some indefinite (generic) NPs.

(261) John's mother hit him in the nose.
(262) ?Whose mother hit him in the nose?
(263) In $\begin{Bmatrix} \text{a pedant's} \\ \text{John's} \end{Bmatrix}$ mind, no possible argument against
him can ever be justified.
(264) ?In whose mind can no possible argument aginst him
ever be justified?

The difference is not just a result of the "mention" clause in the
Crossover Principle, since there isn't even any movement taking
place in (262). This factor plays a part in the anomalies of
(249)-(252).

A third factor is whatever it is that distinguishes (265) from (266), and (267) from (268).

     (265) ?During what movie do you cry every time you see it?
     (266) During what movie do you cry every time you get drunk?
     (267) ?Who do you talk about to people that respect his work?
     (268) Who do you talk about to people that respect Bill's work?

(265) and (267) seem to me to be good only as echo questions. This corresponds precisely to my feelings about the anomalies of (245), (246), (248), and perhaps (249)-(252), (259), (262), and (263). Yet again there is no possibility for the questioned phrase to cross over the anaphoric pronoun. Impressionistically speaking, it seems as though in understanding these as questions, one "gets into a loop": the reference of the pronoun must be known before the question can be asked. But if the reference of the pronoun is known, what is the sense of asking the question? I have no idea how to characterize this set of questions, but it should be clear by now that the generalization has nothing to do with the Crossover Principle.

Thus it seems that virtually all the pronominalization cases adduced as evidence for the Crossover Principle should actually be handled by other processes that will capture a wider range of phenomena. Since these other processes are not known to conflict with the interpretive theory of pronouns, whereas the Crossover Principle does, the discussion in this section can be taken as a vindication of the interpretive theory, even though no positive proposals have been made.

NOTES

1. In the case of plural noun phrases, additional specification will have to be made of set coreference versus individual-by-individual coreference, in order to explain, for example, the difference between themselves and each other. Also, a distinction has to be made between type and token coreference in order to explain the difference between one and it. These distinctions, which must be made by any theory of reference in language, have been discussed somewhat by Dougherty (1968a).

2. There may be a conflict here between grammatical and natural gender and number. The problem is complex, and I do not propose to resolve it here.

3. For the intermediate stage of derivation in sentences like What do you know how to do, where two wh-phrases are fronted in the lower clause, we must appeal to J. Emonds's highly motivated notion of doubly filled node, discussed in his dissertation.

4. (63) is marginal for many speakers, but the pronominalization facts themselves seem secure.

5. We will discuss Ross's actual example and Lakoff's putative disproof of his solution in Chapter 3, § 6.

6. The pruning of S nodes is discussed in Ross (1967a), Chapter 3.

7. There is not time in the present discussion to moti-
vate the analogy of possessional location to physical location
in these examples. Gruber goes through a great deal of dis-
cussion (Chapter 4) to establish such an analysis, based on
correspondences in prepositional patterns.

8. Or perhaps it is the other way around, i.e. that the
semantic notion "potential agent" is a defining criterion for
the feature [+animate].

9. Gruber notes that "causative" sentences are sentences
like (148), where the subject is only the agent.

10. Notice that thematic relations can express the re-
lationship holding in pairs like buy and sell. The two verbs
express the same transition being made by the theme (object).
For both verbs, the subject is agent, but with buy the subject
is goal and the from-phrase is source, whereas with sell the
to-phrase is goal and the subject is source.

11. For an account of sentences like Bill tried to strike
Harry as pompous, in which volition is attributed to the subject,
see Chapter 3, § 11.

12. There is an interesting correlation here: the adjectives
ending in -ed take with, and those ending in -ing take to, with
concomitant correlation in induced thematic relations. I have no
explanation to offer for this fact.

13. For the sense of volition in sentences like John tried
to look sick, see Chapter 3, § 11. Note also the strangeness of

<u>John tried to seem sick to Bill</u>. As with the adjective construc-
tions just discussed, the agent subject and the expression of
goal seem to be mutually exclusive.

14. These are mentioned in Akmajian and Jackendoff (1969)
and in Lakoff (1968b).

CHAPTER 3

THE COMPLEMENT SYSTEM

1. The problem

One of the more thoroughly explored problems in English syntax is the system of sentential complements of verbs. The work in Rosenbaum (1967) gives an extensive range of data and proposes rules to account for most of it. The rules have subsequently been refined by, among others, Ross (1967), Lakoff (1966), Kiparsky and Kiparsky (1969), and Perlmutter (1968). However, there has been general agreement that there are three basic transformational processes which take place in the complement system: extraposition of the complement, possibly leaving an _it_ behind; deletion of complement subject; and raising of complement subject, probably connected with extraposition of the complement verb phrase.

The current accounts of the complement system all are incompatible with the interpretive theory of reference presented in the last chapter. The fundamental problem is that the rule of complement subject deletion (called in most recent work Equi-NP Deletion) requires identity of referential indices in its statement. The interpretive theory, however, claims that co-referentiality is not a syntactic property and hence that no syntactic rule can be contingent on it. In order to maintain

the interpretive theory it will therefore be necessary to treat complement subject deletion in such a way that the referential identity is handled as part of the semantic component.

It should be noted, by the way, as justification for even wanting to handle subject deletion interpretively, that the identity requirement in the Equi-NP transformation has led to a certain amount of difficulty in syntactic theory. In particular, it was pointed out by Lakoff (1965) that for a verb such as try, Equi-NP is not only obligatory, but the structural description of Equi-NP must be met in order for a sentence to be acceptable.

(1) I tried $\begin{cases} \text{to go.} \\ \text{*for Bill to go.} \end{cases}$

This kind of exception, called by Lakoff a positive absolute exception, is a very strong addition to the theory; Lakoff can only cite four kinds of examples. Perlmutter (1968) attempts to eliminate the necessity for positive absolute exceptions by arguing that there is a deep structure constraint of referential identity on the subject and complement subject of try. It will be argued below that this claim, too, entails some fairly dubious syntactic moves. We will investigate the like-subject requirement from a more semantic orientation.

2. An Interpretive Rule for the complement subject

As in the case of the pronominalization transformation, the kind of identity required for Equi-NP Deletion is referential identity. Morphological identity need not play a role, since the morphological form of the deleted NP seems to be irrelevant,

provided the reference is correct. Thus the move to an inter-
pretive theory of pronominalization suggests a parallel move in
the analysis of the deleted complement subject.

Let us try to state an interpretive rule for the comple-
ment subject. For the sake of avoiding complications due to
complement subject raising, we will work with verbs which take
subject complements in which subject raising does not take place,
for example, bother and disturb.

(2) For Harold to eat his peas with a knife would
     bother his mother.
(3) (extraposed)
     It would bother his mother for Harold to eat his
     peas with a knife.
(4) $\begin{Bmatrix} \text{To eat} \\ \text{Eating} \end{Bmatrix}$ his peas with a knife would bother Harold.
(5) (extraposed)
     It would bother Harold to eat his peas with a knife.
(6) (subject raised)
     *Harold would bother his mother to eat his peas with
     a knife.

Let us look at the coreference relations in sentences
with these complements. We observe that only when the complement
subject is absent on the surface ((7)-(8)) is it interpreted as
coreferential with the matrix object. In particular, if the
complement subject is morphologically identical to the matrix
object ((10)-(11)), or if the complement subject is an appropriate
pronoun (12), it is still distinct from the matrix object.

(7)   To have to leave so soon would bother Bill.
(8)   To have to leave so soon would bother some of the
      men.
(9)   For Sam to have to leave so soon would bother Bill.
(10) For Bill to have to leave so soon would bother Bill.
(11) For some of the men to have to leave so soon would
      bother some of the men.
(12) For him to have to leave so soon would bother Bill.

The rule here will have to look rather like reflexivization. In reflexivization, we recall, there is given a particular structural relation between two NPs, say NP^1 and NP^2. Assuming the structural relations hold, the NPs are coreferential if and only if NP^2 is a reflexive. In the case of complements the structural relation is that obtaining between the complement subject and some NP in the clause above it; the effect of the rule is then similar. The problem of which NPs in the upper clause are permissible as coreferents of the complement subject has been termed the control problem; the NP selected as coreferent is called the controller (these terms are introduced in Postal (1968b)). We will discuss the control problem later; but for the moment, we will merely say that the complement subject is coreferential with some NP in the upper clause.

What is the nature of the NP in the complement subject when it is interpreted as coreferential with the controller NP? It clearly cannot be any normal pronoun or NP, as we can see from (7)-(12). The worst solution we might have to accept would involve a special pro-NP (call it DEL) which has no phonological interpretation and serves the sole semantic function of functioning as a coreferential subject complement. Then the interpretive rule for complement subject would like something like (13).

(13) (Complement Subject Coreference Rule)
      NP^1 is $\propto$ coreferential with NP^2 if NP^2 is the subject
      of a for-to or poss-ing complement, and NP^1 is in
      the main clause, and NP^2 is $\propto$ equal to DEL.

      OBLIGATORY

In addition to this rule, we will need a rule at the end of the semantics, rejecting any sentence in which there remains an uninterpreted DEL. This system of rules is formally just like the system of rules for reflexivization proposed in Chapter 2.

If we assume that the kind of coreferentiality specified by this rule is the same as that specified by pronominalization and reflexivization, we can account for the behavior of plural and quantified NPs as controllers. In (11), the two occurrences of some of the men are understood as unrelated; in (8) each in-dividual is understood as being bothered by having to leave. These are exactly the interpretations in the case where pronominalization (14) and reflexivization (15) are the appropriate rules of co-reference.

(14) Some of the men knew that some of the men had lost.
Some of the men knew that they had lost.
(15) Some of the men hit some of the men.
Some of the men hit themselves.

(9) and (11) are the types of examples adduced by Carden (1967, 1968) as evidence for treating quantifiers as higher verbs. I have argued in detail against this proposal in Jackendoff (1968a,b). It is interesting to note, however, that the examples with quantifiers work exactly like the examples with simple NPs. The fact that Carden is forced to account for this parallelism in the surface by a bizarre non-parallelism in the deep structure argues against the standard account of complement subject deletion, based on morphological and referential identity.

One objection that might be made to (13) is that it requires

a special pro-NP which is required nowhere else in the grammar, and never shows up on the surface. This pro-form would of course violate the Extended Lexical Hypothesis. Perhaps a deeper objection to the use of DEL is possible as well. Suppose that it were desirable to extend the interpretive theory to other systems of rules presently handled with transformations of deletion under identity. To give some sort of consistent approach, one would want to use the same sort of formalism for all of the interpretive rules. The postulation of a special proform such as DEL presupposes that the missing elements in the surface structure form a constituent. This is not always the case; a comparative construction such as (16) has a missing subject and verb.

(16) Stanley ate more beans than cabbage.

I do not intend to propose an interpretive analysis of comparatives at this point, but it would be nice to have an interpretive theory of complement subjects that leaves open the possibility of doing comparatives with similar mechanisms.

Chomsky has suggested (personal communication) a more general way of handling rules of this sort. Suppose that the phrase structure rules and lexical insertion are optional, so that potential deep structures can be generated containing non-terminal nodes at the end of one or more branches. It was assumed in Aspects that such structures would block at the end of the transformations. Observe, however, that these structures would be semantically ill-formed since no semantic information would be

available for branches ending in a non-terminal symbol. I propose that the semantic blocking rather than surface structure blocking be taken as the criterion for rejecting unexpanded nonterminal nodes. Suppose then that rules of semantic interpretation give readings to unexpanded non-terminal nodes (or combinations of them) under certain conditions. Then the semantic blocking of a non-terminal node can be prevented, just in case it is interpreted.

Unexpanded nodes can undergo transformations like any other node. If such a node is either deleted or else interpreted by convention as phonologically null, without blocking, a legal surface structure will result. If, however, an unexpanded non-terminal node is not given a semantic interpretation, the sentence it is in will be blocked semantically.

In this theory, the semantic rules which interpret unexpanded non-terminal symbols will correspond to the transformations of the standard theory which delete items leaving no trace. Possible examples of this kind of rule are the comparative rules, the gapping rule (cf. Ross (1967c)), perhaps the deleted passive by-phrase, the rule for deletion of a noun phrase in the complements of the degree modifiers too and enough, and of course the rule for the subject of for-to and poss-ing complements.

The revised version of (13) will simply substitute Δ for DEL, where Δ is the symbol (used in Aspects) for an unexpanded non-terminal node. All interpretive rules of this type likewise

mention Δ rather than some specialized pro-form.

3. Postal's generalization

Postal (1968b) gives several arguments to show that the coreference relations induced by deleted complement subjects share several very interesting aspects with pronominalization. The most interesting examples are cases in which an indefinite NP and a complement subject on its left cannot be coreferents. (All following examples from Postal (1968b), pp. 35-37)

(17) *Finding out Greta was a vampire worried somebody.
(18) *Discovering that their daughters were pregnant worried some old ladies.
(19) *Kissing was fun for some kids.

(unstressed some in all these sentences)

Contrast these with the acceptable sentences below, where the complement subject is distinct from the object of the main clause.

(20) Bill's finding out that Greta was a vampire worried somebody.
(21) My discovery that their daughters were pregnant worried some old ladies.
(22) Tony and Betty's kissing was fun for some kids.

The same paradigms seem to hold for backwards pronominalization to indefinites: contrast (23)-(24), with anaphoric pronouns, to (25)-(26), without them.

(23) *The man who lost it$_1$ needs to find something$_1$.
(24) *It was their$_1$ strength that made some gorillas$_1$ famous. (unstressed some again)
(25) The man who lost the camel needs to find something.
(26) It was my strength that made some gorillas famous.

That the conditions are the same is further borne out by derived nominal constructions, in which complement subject deletion

is optional (at least in Postal's dialect--I'm not sure about mine).

(27)　His$_i$ realization that you knew Greta disturbed Tony$_i$.
(28)　The realization that you knew Greta disturbed Tony.

In neither case can the object of the main clause be indefinite.

(29)　*His$_i$ realization that the Earth was exploding
　　　worried somebody$_i$.
(30)　*The realization that the Earth was exploding
　　　worried somebody.

Some more similar examples:

(31)　My discovery that Johnson was a puppet scared some
　　　congressmen.
(32)　Schwarz's realization that God was dead didn't worry
　　　anybody.
(33)　$\begin{Bmatrix} *\text{Their} \\ *\text{The} \end{Bmatrix}$ discovery that Johnson was a puppet scared
　　　some congressmen. (where the congressmen discovered it)
(34)　$\begin{Bmatrix} *\text{His}_i \\ *\text{The} \end{Bmatrix}$ realization that God was dead didn't worry
　　　anybody$_i$.

Although I don't agree entirely with Postal's data, of which (17)-(34) constitute only a small sample, it is evident that there is a fair correlation between complement subject deletion and pronominalization constraints on indefinite NPs.

There is not space here to go into the rest of Postal's examples. Suffice it to say that he has shown that a theory which combines complement subject deletion and pronominalization in some interesting way is to be preferred to a theory which does not do so.

In Postal's theory, since pronominalization is non-cyclic but complement subject deletion is cyclic, it is necessary to introduce a special device called the Doom Marker, which is placed

as a feature on a complement subject by a cyclic rule. Then if,
in the last cycle, Doom is also marked [+Pro] by pronominalization,
it can delete, producing a correct surface form. Otherwise, the
derivation blocks.

There is, I believe, a fundamental metatheoretical flaw in
Postal's solution. In effect, what his solution does is to cir-
cumvent the ordering of rules by use of an arbitrary, otherwise
unmotivated diacritic feature: an NP is allowed to delete at
the time of pronominalization just in case at some stage earlier
in the derivation it satisfied a certain structural relation.
Since one of the primary insights of transformational grammar is
that rules are ordered, a trick to get around ordering can only
be seen as a breach of the ground rules of the investigation.
Even if we have no clearly articulated principle in the theory
to prevent such a proposal (such a principle is being currently
discussed for phonology in such papers as Kiparsky (1968) and
Vetter (1968)), one would hope that there is a more interesting
way of capturing the facts than Postal points out.

The interpretive theory of coreference proposed in Chapter
2 appears to potentially have a little more promise. First of
all, pronominalization is a cyclic rule, as is the complement
subject rule, so the immediate conflict presented by Postal's
theory does not arise. It remains to be shown that the comple-
ment subject rule can take place at the end of the cycle, so
that whatever it shares with pronominalization can be collapsed
into a single statement, as has been done with the reflexivization

rule.

An interesting observation emerges from considering the possibility of collapsing the complement subject rule with reflexivization and pronominalization. The complement subject rule and the reflexive rule, it will be recalled, look formally very similar, both being obligatory α rules of coreference. But reflexive pronouns and $\triangle$, which govern the two rules, have mutually exclusive distributions: reflexives are allowed only under VP or $\bar{N}$, and $\triangle$ is allowed only in subjects. Thus the two most important rules of coreference aside from pronominalization itself are in some sense complementary.

4. Interaction of the Complement Subject Rule with other rules

We want to show that the complement subject rule (13) can take place at the end of the cycle. Its transformational equivalent, Equi-NP Deletion, in fact is usually assumed to be ordered at the beginning of the cycle. The reason for this ordering has to do with selection of the controller. In the accounts I know of, e.g. Rosenbaum (1967), Perlmutter (1968) it is assumed that the selection of the controller NP in the higher clause is based on structural principles--either the grammatical relations of subject and object, or something equally dependent on the structure, like Rosenbaum's distance principle (see §9 for more discussion of this). It is important that these structural relations be captured before transformations start distorting the main clause, because of pairs like (35)-(36).

(35) Bill forced John to wash the dishes.
(36) John was forced by Bill to wash the dishes.

If the controller is selected on the basis of grammatical rela-
tions, i.e. object in this case, it must be the deep object in
(36), since there is no surface object, and the controller is
the surface subject. If the controller is selected on the basis
of a distance principle, then it must be the underlying structure
distance, since in (36) Bill is in the VP and hence nearer to the
complement than John.

We will show, however, in §10, that probably a better
principle for selecting the controller NP can be based on the
thematic relations we discussed in Chapter 2. Thematic relations
are not altered by transformations, since they are properties of
the semantic reading which correlate to the deep structure
grammatical relations. Hence it doesn't matter for selection of
controller whether or not transformations have distorted the main
clause, and so examples like (35)-(36) do not argue against order-
ing the complement subject rule at the end of the cycle, along
with pronominalization.

The one distortion we will still have to worry about is
the one that moves the complement subject itself, the Complement
Subject Raising transformation (often called It-Replacement).
This transformation breaks up a for-to complement, moving the VP
of the complement to the end of the VP of the main clause, and
raising the complement subject to the position originally occupied
by the complement. Thus, for example, (37) becomes (38) by the

complement subject raising rule.

(37)

```
                    S
              ／         ＼
           NP            VP
            |         ／     ＼
         Steve      V         NP
                    |          |
                 believes      S
                           ／      ＼
                        for NP       VP
                             |      ／\
                            Tom    /    \
                                  to be sick
```

(38)

```
                 S
            ／        ＼
         NP             VP
          |       ／     |      ＼
       Steve    V       NP       VP
                |        |      ／\
             believes   Tom    /    \
                              to be sick
```
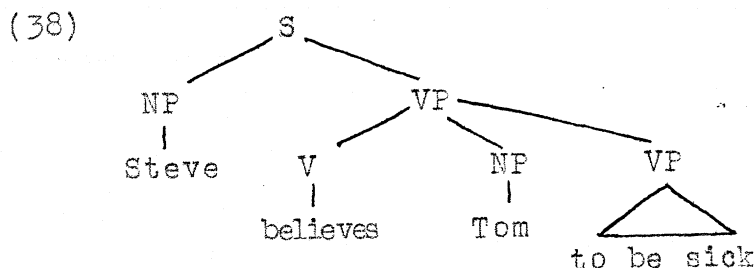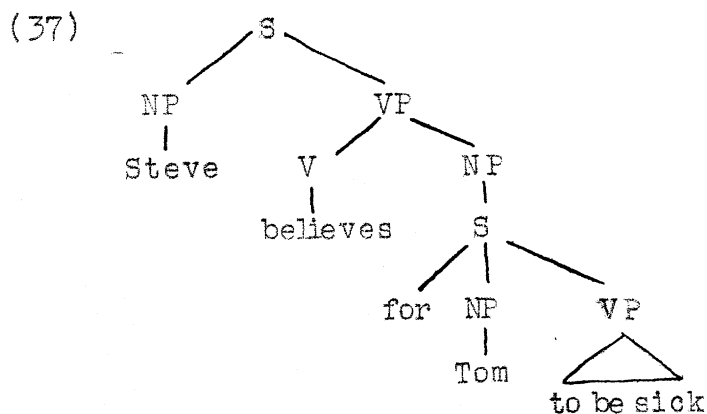
Various ways of accomplishing this change have been proposed
(Rosenbaum, Kiparsky and Kiparsky, Lakoff), but for this discus-
sion all that is relevant is the difference in structure due to
the disappearance of the S node.

How will raising affect the interpretation of $\triangle$? The
Complement Subject Reference Rule only applies to $\triangle$ if it is in
complement subject position. But inadvertently, Raising may be
applied to a complement containing $\triangle$. Contrast the two derived
structures (39) and (40).

(39) John expects $[_S \triangle$ to go]
(40) John expects $\triangle [_{VP}$ to go]

In (39) Raising has not applied, and the reference rule can operate

properly. In (40) Raising has applied, and now the reference rule is not applicable. Therefore, if this is the end of the derivation, the sentence must be thrown out, since there is an uninterpreted node present.

However, if (40) is embedded in another clause, as in (41), the derivation can be saved.

(41) $[_{S^1}$ Sam hoped $[_{S^2}$ for John to expect $\Delta[_{VP}$ to go$]]]$

Assume (41) is a stage in the S^2 cycle, after Raising. Then the Passive can take place in S^2 to form (42).

(42) $[_{S^1}$ Sam hoped $[_{S^2}$ for Δ to be expected by John $[_{VP}$ to go$]]]$

Now on the S^1 cycle, the reference rule can mark Δ coreferential with Sam, giving the correct reading of (43).

(43) Sam hoped to be expected by John to go.

Although Δ has moved around in the tree quite a bit by the time its antecedent is established, it can still be identified in the semantic reading as the subject of go, since this is a property of the reading which is identified in the deep structure.

To point out the independence of coreference relations from grammatical relations, let us see what happens if (39), which has not undergone Raising, is embedded.

(44) $[_{S^1}$ Sam hoped $[_{S^2}$ for John to expect $[_{S^3} \Delta$ to go$]]]$

This time, the complement subject reference rule operates in the S^2 cycle, making John the antecedent of Δ. Then in the S^1 cycle, the rule marks Sam and John distinct. Thus we get the correct reading for (45).

(45) Sam hoped for John to expect to go.

The difference in meaning between (43) and (45) is purely one of coreference, and hence not represented in the deep structure, which only says that Δ is the subject of go. The difference in coreference is produced by the application of the optional rules Raising and Passive; if only Raising applies (if, for example (41) is the final derived structure), then still another variant is produced, this one without a valid semantic reading.

One other thing should be mentioned about this kind of derivation. We mentioned that since the selection of controller is based on thematic relations, it does not matter where the controller ends up in its clause. But since the Complement Subject Rule mentions an NP[1] "in the main clause," it is important that the controller actually be in the main clause at the time of the rule, and not moved into another clause by, say, Raising. For example, how is the reference of Δ established in (46), where its antecedent has been raised?

(46) $[_{S1}$Bill expected John $[_{VP2}$to try $[_{S3} \Delta$ to go$]]]$

This is in fact very simple. The Complement Subject Rule takes place at the end of the S^2 cycle, before John has been raised out of S^2.

How does the Reflexivization rule interact with the rules we have just discussed? First of all, there is the fairly trivial observation that a reflexive can be marked coreferential with Δ, as in (47).

(47) Frank tried $[_\Delta$ to scratch himself$]$

More interesting things happen if we put a reflexive in place of Δ in examples like (39)-(46). Take the analogues of (39)-(40), which differ by the application of Raising.

(48) John expects $[_S$ himself to like Bill$]$
(49) John expects himself $[_{VP}$ to like Bill$]$

What happens to these on the lower cycle? The Reflexivization rule does apply in this cycle, with himself as NP^1 (the antecedent) and Bill as NP^2, marking them distinct. However, himself has not been used as NP^2 in the rule, so it is still without an antecedent: recall the exact statement of condition (a) of the reflexive rule, which we use here crucially.

Now what happens on the upper cycle? If himself is not raised (48), the Complement Subject Rule applies, marking it non-coreferential with John. But this leaves himself without an antecedent, and so the sentence is thrown out. On the other hand, if himself is raised (49), the reflexive rule applies, marking himself coreferential with John to give a good reading.

If, however, the passive takes place on the upper cycle of (49), himself will be moved into subject position and the reflexivization rule again cannot find it an antecedent. But if (49) is embedded,

(50) $[_{S1}$ Sam believed $[_{S2}$ for John to expect himself $[_{VP}$ to like Bill$]]]$

then passive is applied in S^2,

(51) $[_{S1}$ Sam believed $[_{S2}$ for himself to be expected by John $[_{VP}$ to like Bill$]]]$

the day can be saved by raising <u>himself</u> on the S^1 cycle.

(52) $[_{S1}$ Sam believed himself $[_{VP}$ to be expected by John $[_{VP}$ to like Bill$]]]$

Now Reflexivization can apply on the S^1 cycle, marking <u>Sam</u> as the antecedent of <u>himself</u>.

Again in the case of Reflexivization, we see the independence of coreference relations from deep structure grammatical relations. For if (50) does not undergo any further transformations, <u>himself</u>, the deep subject of <u>like</u>, will be coreferential with <u>John</u>; but by application of Passive and Raising, <u>himself</u> comes to be coreferential with <u>Sam</u>, though it is still understood as the subject of <u>like</u>.

The discussion of examples illustrates clearly the complementarity of the rules of Complement Subject Reference and Reflexivization. We have followed Δ and <u>himself</u> through derivations involving cyclic iterations of Raising and Passive, sometimes getting good sentences, sometimes bad ones. It is interesting, however, that those acceptable sentences containing <u>himself</u> are bad if Δ is substituted, and vice versa. This is because a reflexive can get a good interpretation only in <u>object</u> position, and Δ only in <u>subject</u> position.

Note also that in this discussion we make crucial use of the asymmetry of condition (a) of the reflexive rule, i.e. that a reflexive only counts as having been marked if it has served as NP^2 in the rule. This is important just for the case where a

reflexive is generated as a deep subject, as in (48), where we want it to be interpreted on the following cycle, by which time it has become an object, as in (49).

Incidentally, it should be noticed that the Thematic Hierarchy Condition (Chapter 2, § 12) causes no difficulties with these raised reflexives. Recall that in the final form of the condition (Chapter 2, § 12.e), it applies to the thematic relations defined by the verb of the clause in which reflexivization takes place. A raised reflexive, having no thematic status in its new clause, is considered lowest on the hierarchy. Hence whatever the thematic relation of the subject, reflexivization will be permitted. Furthermore, passives such as (53) are prohibited by the condition,

(53) *Tom is believed by himself to have hit Norman.

since Tom comes from the lower clause and therefore is lower on the hierarchy than himself.

5. Reflexives, raising, and pruning

In the transformational account of reflexivization, the derived structure after Complement Subject Raising poses an interesting problem. Because raising leaves only the VP dominated by the subordinate S node, the pruning convention proposed in Ross (1967a) dictates that the derived structure contain only the VP of the complement sentence attached at the right-hand end of the matrix VP, the S node having been pruned.

(54)

```
                    S
         NP                  VP¹
         |            V      NP      VP²
        John          |      |       |
                    expects Bill   .to go
```

Since now VP² is in the same simplex sentence as the subject, we would expect reflexivization to take place, producing, for example, (56) from (55).

(55) I expect [for Bill to examine me]
(56) *I expect Bill to examine myself.

The only way to stop this in the transformational theory is to somehow retain the S node above VP², by restricting pruning in some way that will throw the generality of pruning into doubt.

However, in the interpretive theory of reflexives, this problem does not arise. We can deal with this example much as we dealt with sentences like (57) and (58).

(57) *John showed Bill Mary's picture of himself.
                                    (Chapter 2, §7)
(58) *Who did you think stabbed yourself?
                                    (Chapter 2, §8)

(56) will have the underlying structure (59).

(59) I expect [for Bill to examine myself]

On the inner cycle, before the complement is broken up, myself and Bill are marked coreferential by the reflexive rule; since the reflexive rule is obligatory, the sentence is anomalous.

The problem really arises for the interpretive theory in handling the interpretation of what starts out as (55) and ends

up as (60).

(60) I expect Bill to examine me.

Why, once the complement S node has pruned, cannot the reflexive rule apply to mark I and me as noncoreferential, ruining the interpretation? The key lies in the first cycle. In the first cycle, the reflexive rule marks Bill and me noncoreferential, which is correct. However, this marking on the first cycle prevents the structural description of reflexivization from being met on the second cycle, after the S node has pruned: condition (a) of the rule is not met since me has already served once as NP^2 in the reflexive rule. Therefore the reflexive rule cannot apply in the second cycle between I and me; and so pronominalization is free to apply, marking them coreferential even though at this point they are in the same simplex sentence.

The interdependence of proposals on reflexivization and proposals on complement subject raising is thus as follows: the interpretive theory of reflexives is consistent with the standard complement subject raising transformation whether or not pruning takes place. The transformational theory of reflexives, on the other hand, is consistent only with a derived structure for subject raising that retains an S node above the postposed complement VP. Such a derived structure is possible only if the pruning convention is modified in some ad hoc fashion so as to prevent the S from disappearing; the interpretive theory requires no such modification.

Note also the complete parallelism of this argument and that presented for the interaction of reflexivization with wh-fronting in Chapter 2, § 8. There also, the pruning convention caused difficulty for the transformational theory. In that case, it could be circumvented by ordering reflexivization before wh-fronting, but this of course made it impossible to capture naturally any generalizations between pronominalization and reflexivization. In the present case, a reordering of the two is impossible, since reflexivization must follow raising to get John believes himself to be beautiful. Thus here the plight of the transformational theory is more acute. It is because of condition (a) of the interpretive reflexivization rule combined with the fact that it is an α rule that these cases are both handled nicely. These conditions, it will be recalled, cannot be stated naturally in a transformational framework. It is cases like these that really decide between the theories.

6. Ross's example

We can finally do Ross's crucial example in "On the Cyclic Nature of English Pronominalization" (Ross 1967b)),

(61) Realizing that John was sick bothered him,

in which John and him cannot be the same person. This has the structure (in the present theory)

(62) $[_{S1} [_{S2} \Delta \text{Realizing} [_{S3} \text{ that John was sick}]_{S3}]_{S2}$
        $\text{bothered him}]_{S1}$

On the S^3 cycle, nothing of interest happens. On the S^2 cycle, pronominalization applies between $\triangle$ and John, marking them dis-dinct. On the S^1 cycle, the Complement Subject Rule applies, marking $\triangle$ coreferential with him. Now if pronominalization also applies, between John and him, a paradoxical reading will result, since we will have formed a table of coreference

(63) 　John　coref　him
　　　him　coref　$\triangle$
　　　$\triangle$　noncoref　John

Thus the only alternative is not to apply pronominalization be-tween John and him, so that they will be marked distinct by the Noncoreferentiality rule (Chapter 2, § 3). This gives the correct reading for (61).

On the other hand, John and him can be the same person in (64).

(64) Bill's realizing that John was sick bothered him. In this sentence, application of the Complement Subject Rule on the S^1 cycle results in him and Bill being marked noncoreferential because the rule is an α rule. Therefore him and John can be marked coreferential by pronominalization without producing a strange reading like (63).

Lakoff (1968b) gives examples which purport to show that the principle of the cycle cannot handle all cases like (61) and that it therefore must be replaced by an ad hoc constraint on a last-cyclic pronominalization rule. The first two examples are based on the last-cyclic nature of the rules Extraposition from

NP and Wh-preposing; we disposed of these in Chapter 2, §5.

His third argument is based on the assumption that in
(65), _Mary_ is understood as the subject of _realizing_; he is trying
to construct an example which behaves like Ross's but for which
Ross's argument will not work.

(65) *Realizing that John had cancer seemed to him to
have been bothering Mary.

But Lakoff's assumption about the subject of _realizing_ seems false
to me; I can't find any subject for _realizing_, whether or not
John and _him_ are coreferential. This can be more readily seen
from the simpler example (66), which involves no pronouns at all.

(66) *Realizing that John had cancer was bothering Mary.

Note that if _the realization_ is substituted for _realizing_ in
these sentences, they are acceptable, and in (65) _John_ and _him_
can be coreferential. Or, the sentences can be fixed up by sub-
stituting _bother_ for _have been bothering_. Thus the violation in
(65) has to do with something like sequence of tenses, and not at
all with pronominalization. This redeems Ross's argument, at
least for the interpretive theory.

7. Three aspects of the control problem

There are two problems about the selection of a controller
that immediately strike the eye: finding which NP in the main
clause is Δ coreferential with, and deciding whether coreferent-
iality is obligatory or optional. These two factors interweave
in interesting ways, varying more or less independently from one

class of verbs to another. Here are some examples of different combinations.

A. VP dominates only complement, subject controller, optional control: <u>wait</u>, <u>pray</u>, <u>hope</u>, <u>decide</u>
John prayed to leave.
John prayed for Bill to leave.

B. VP dominates only complement, subject controller, obligatory control: <u>try</u>, <u>learn</u>, <u>condescend</u>, <u>be lucky</u> <u>be wise</u>
John tried to leave.
*John tried (for) Bill to leave.

C. VP dominates only complement, raising obligatory, so no control: <u>believe</u>, <u>imagine</u>
*John believed to be going.
John believed himself to be going.
John believed Bill to be going.

D. VP dominates only complement, raising optional; if no raising, subject controller, obligatory: <u>expect</u>, <u>want</u> (in some dialects):
John expects to leave.
*John expects for Bill to leave.
John expects himself to leave.
John expects Bill to leave.

E. VP dominates NP and complement, object controller, obligatory control: <u>permit</u>, <u>persuade</u>, <u>force</u>
*John permitted to leave.
John permitted Bill to leave.
*John permitted Bill (for) Harry to leave.

F. VP dominates complement and optional NP or PP, subject controller, obligatory control: <u>promise</u>, <u>vow (to NP)</u>, <u>agree (with NP)</u>, <u>learn (from NP)</u>
John agreed to leave.
John agreed with Bill to leave.

G. VP dominates complement and optional NP; object controller obligatory if present, otherwise subject controller obligatory: <u>get</u>, <u>keep</u>
John got Bill to leave.
John got to leave.
*John got there to be an explosion.

H. VP dominates complement and optional NP, object con-
    troller optional if present: <u>shout</u>, <u>scream</u>
        John shouted for Bill to leave.
        John shouted to Bill to leave.
        John shouted to Bill for Harry to leave.
    *John shouted to leave.

I. Subject complement, VP dominates NP, object controller,
    optional control: <u>bother</u>, <u>benefit</u>
        (For John) to leave so early would bother Bill.

Lakoff (1965) and Perlmutter (1968) attempt to account
for the optional-obligatory distinction in different ways. La-
koff uses the device of <u>positive absolute exceptions</u>: verbs can
be marked as positive absolute exceptions to Equi-NP Deletion,
which means that the structural description of Equi-NP Deletion
must be met and the transformation must take place. This kind
of exception is illustrated in Lakoff (1965) with three other
kinds of examples. The first involves verbs that must have re-
flexive objects, such as <u>behave</u> and <u>perjure</u>. The second is the
adjective <u>gala</u>, which must undergo <u>wh-be</u> deletion to get <u>gala</u>
<u>affair</u> but not *<u>the affair was gala</u>. This may be a true exception
(but see Chapter 5, § 3). The third example involves "verbs"
which must nominalize (in his theory of nominalizations), such
as <u>aggress</u> and <u>king</u>; this kind of case will not arise under the
theory of nominalizations assumed here, i.e. that of Chomsky
(1969). Thus the need for a device as powerful as the positive
absolute exception rests on only three cases. It would be nice
to find a more constrained way to talk about them, particularly,
in view of the fact that the verbs with obligatory control hardly

seem "exceptional" among the types listed above. We will deal
here only with the examples involving coreference.

Perlmutter tries to do away with the need for positive
absolute exceptions for Equi-NP Deletion by showing that obliga-
toriness of complement subject deletion can be expressed as a
contraint on coreference of deep structure matrix and complement
subjects. To maintain this, he must show that all cases of
deleted derived structure subjects of complements are deep struc-
ture subjects as well. An apparent counterexample that comes
immediately to mind is a sentence with a passive complement
like (67).

(67) John tried to be examined by the doctor.
Perlmutter argues that this really has a deep structure like (68),

(68) John tried $[_s$ John $\left\{ {get \atop let} \right\} [_s$ the doctor examine John$]_s]_s$

which satisfies his proposed deep structure constraint. Then a
transformation deletes the get or let. The argument is based on
claim that the extra sentence in (68) reflects the actual meaning
of (67), and that all the constraints on these passive complements
follow from constraints on passive complements of get and let.

But even if this somewhat counter-intuitive analysis of
the deep structure of (67) works, there are plenty of verbs and
adjectives, including get itself, which, like try, have obligatory
control but which permit passive complements freely. Furthermore,
these passive complements are not paraphrasable with an extra
sentence containing get or let.

(69) John was lucky to be examined by the doctor.
      John got to be examined by the doctor.
      Bill permitted John to be examined by the doctor.

(70) John was lucky to $\left\{\begin{matrix} \text{let} \\ \text{get} \end{matrix}\right\}$ the doctor (to) examine him.

      John got to $\left\{\begin{matrix} \text{get} \\ \text{let} \end{matrix}\right\}$ the doctor (to) examine him.

      Bill permitted John to $\left\{\begin{matrix} \text{get} \\ \text{let} \end{matrix}\right\}$ the doctor (to) examine him.

In (70) there is a sense of volition on John's part which need not be present in (69). Thus a deep structure constraint on subjects cannot account for the obligatoriness of complement subject deletion with these verbs.

This extra sense of volition is the crucial factor here. The normal cases of complement-subjects of try are agents of their sentences. Recall the discussion of agents in Chapter 2, § 11. We noted there that sentences with agent subjects are also those which permit adverbials like intentionally, on purpose, in order to VP, so that S, and so forth. The meanings of these adverbials require that the subject be "in control of the situation"; thus it would make no sense to say, for example, *John received the book intentionally. Likewise, I claim, the meanings of try, condescend and intend (and note the relation here to intentionally) require that their complement subjects be "in control of the situation," i.e. it must be conceptually possible for them to bring about the event expressed by the complement. However, lucky, get, and permit place no such requirements on their complement subjects.

Thus we have uncovered a third independent aspect of the control problem--whether or not the verb requires the complement

subject to be able to bring about the event expressed by the complement. For verbs which don't have this restriction, Perlmutter's analysis fails. We will return to how to specify this restriction in § 11. First let us find a way of handling the other two aspects of the control problem.

8. Networks of obligatory coreference

Among Lakoff's four examples of absolute exceptions, two have to do with coreference. This suggests the use of a much less powerful device than absolute exceptions, one dealing only with properties of coreference. What seems to be the case is that verbs can require certain of the NPs around them to be coreferential. Besides the case of obligatory complement subject deletion, we have the reflexive cases mentioned by Lakoff,

(71) Harry behaved $\begin{cases} \text{himself.} \\ \text{*Bill.} \end{cases}$
Teddy perjured $\begin{cases} \text{himself.} \\ \text{*Mac.} \end{cases}$

one with both object and complement subject required to be coreferential to the subject,

(72) Frank prides $\begin{cases} \text{himself} \\ \text{*Jack} \end{cases}$ on (*Bob's) being intelligent.

and many expressions with obligatory pronouns or possessive pronouns.

(73) Bill took Mary at $\begin{cases} \text{her} \\ \text{*his} \\ \text{*Walt's} \end{cases}$ word.

Gerald blew $\begin{cases} \text{his} \\ \text{*Tom's} \end{cases}$ stack.

Ron knows Sue for what $\begin{cases} \text{she} \\ \text{*he} \\ \text{*Dave} \end{cases}$ is: a liar.

Millie has a wart on $\begin{cases} \text{her.} \\ \text{her nose.} \\ \text{*Bill's nose.} \end{cases}$

and a host of others.

It seems to me that in these expressions the meaning of the verb or idiom requires that the two NPs be coreferential. In a sense the second NP is redundant semantically, but it is present because of the syntactic vagaries of the expressions. If the two NPs are referentially distinct, the semantic reading comes out wrong, in that there are two individuals in evidence where the verb presupposes only one.

My proposal, then, is this: an optional part of the lexical (hopefully semantic) representation of a verb will be a network of obligatory coreference relations among its arguments (and perhaps noncoreference relations as well).[1] A network of coreference relations is violated if two NPs, independently determined noncoreferential by the rules of coreference, fill semantic places which are required by the network to be coreferential. A violation of a network results in rejection of the offending sentence.

This proposal obviously can handle cases like (71)-(73). Its extension to obligatory complement subject identity depends on the validity of one assumption, namely that "surface complement subject" is an identifiable position in the representation of the complement at the point in the grammar where the network of coreference is applied, presumably semantic representation. If this were not so, there would be no way for a verb such as <u>try</u>

to specify that the surface complement subject is identical to the matrix subject. This does not seem too extravagant an assumption: we will show in §11 that the surface subject of a clause is semantically a privileged position in other respects.

In connection with networks of coreference, let us look at two other rules which operate on NPs in complement clauses. The rule of Tough Movement (which produces John is tough to please), discussed briefly in Chapter 2, §2, could operate in two possible ways. It either moves the object out of the complement clause, in which case it is a transformation like Complement Subject Raising, or else it deletes the object because of obligatory coreference with the matrix subject, in which case it is (in our theory) an interpretive rule like Complement Subject Reference. If it is the latter, we would be forced to say that networks of coreference can also extend to complement surface objects, which, as far as I know, have no particularly privileged status in the semantic representation. This would argue that the transformational derivation is the correct one.

The only objection I know of to the transformational solution is that it does not preserve meaning, since (74) can be transformed into either (75) or (76) by Tough Movement.

(74) It is easy to play sonatas on this violin.
(75) Sonatas are easy to play on this violin.
(76) This violin is easy to play sonatas on.

But this argument can be answered. Similar differences seem to appear when contrastive stress is added to (74):

(77) It is easy to play sonatas on this <u>violin</u>.
(78) It is easy to play <u>sonatas</u> on this <u>violin</u>.

(77) is (for me at least) close in meaning to (75), and (77)
is close to (76).

In all of the sentences (75)-(78), the understood gram-
matical relations are identical. Therefore the sentences do
not have to differ in deep structure. The semantic differences
between them have to do with focus and presupposition--what ques-
tions the sentences are possible answers to. Chomsky (to appear)
has convincingly argued that focus and presupposition are elements
of semantic interpretation that are determined on the basis of
a level of derived structure in which stress levels of the sen-
tence have been assigned--certainly not the base structure.[2]
Since the meanings of the sentences differ in a way that has
nothing to do with deep structure, the transformational solution
for Tough Movement is consistent with the assumptions of the
present investigation. The semantic version of the rule would
require a special addition to the proposal on networks of corefer-
ence, so we will assume the transformational version to be correct.

The other rule that affects NPs in complements is the rule
for the complements of degree modifiers like <u>too</u> and <u>enough</u>.

(79) The pig is $\begin{Bmatrix} \text{old enough} \\ \text{too young} \end{Bmatrix}$ for Harry to kill. (object missing)

(80) The pig is $\begin{Bmatrix} \text{old enough} \\ \text{too young} \end{Bmatrix}$ to climb the fence. (subject missing)

(81) The pig is $\begin{Bmatrix} \text{fat enough} \\ \text{too fat} \end{Bmatrix}$ to eat. (ambiguous)

For this rule, a coreference assignment is obviously called for, since there is no sentence like *It is too fat to eat the pig, which could then undergo something like Tough Movement.

Sometimes pronouns are needed instead of Δ in the position of coreference:

(82) The pig is fat enough that we can eat it.
      Saskatchewan is too far away for you even to start
       to conceive of getting there in one day by unicycle.

If there are networks of coreference in the representations of too and enough, they will have to refer to an NP in the complement which need not be the subject. But in fact, the restriction on the complement is weaker than coreference, for we have sentences like

(83) This room is too chilly to turn on the air-conditioning.
      The weather is warm enough for us to go swimming.

From these sentences we see that the restriction is something more like "the complement sentence must be directly relevant to the matrix," where "directly relevant" is of course a euphemism for a semantic relation I don't understand.[3] What is important here is that no network of coreference need be established for too and enough, so the generalization still stands that the subject is the only position in the complement to which a network of coreference can apply.

9. When there is a unique controller

Rosenbaum (1967) proposes a general principle for the selection of the controller NP. He says that the choice of con-

troller can be made purely on the basis of structural considerations, by means of a "minimal distance principle." According to this principle, the controller is always the noun phrase "nearest" to the complement subject, where distance is measured by counting nodes along the path in the tree joining the complement subject to the noun phrase. This principle successfully predicts the controller in all cases listed in § 7 except the promise class, and furthermore it correctly predicts that the controller of sentence adverbial clauses (in order to and by -ing clauses) is always the subject.

This principle, if correct, would be a very surprising result. Nothing like it has been found elsewhere in the grammar; the device of node-counting somehow seems foreign to our formalisms. Furthermore, the glaring exception promise throws doubt on its correctness; it seems no accident that the meaning of promise correlates with its formal behavior under complement subject deletion.

An underlying assumption of Rosenbaum's principle is that there is never more than one NP in a sentence that can serve as controller. Yet, as Postal points out (1968b, § V.D), this assumption is invalid, since the following sentences are ambiguous as to the reference of the deleted complement subject.

(84)   Harry talked to Bill about kissing Greta.
        Harry wrote to Bill about not voting for Humphrey.
        Harriet argued with Betty about visiting you.

There are in fact three branches of the ambiguity: the complement subject can be coreferential with either upper NP or with both

jointly. As Postal points out, this is exactly parallel to the
ambiguity of the pronoun in (85).

   (85)   Mary argued with John about their getting married
          in a church.

Their can denote Mary and someone else, John and someone else,
or Mary and John. (The first two readings are more apparent if
Bill is substituted for Mary.) And in general the environments
for joint interpretations of Δ are just like those for joint
interpretation of pronouns (see Postal (1968b, §V.E), for all
the evidence).

   The ambiguities of (84) show that it is not always possible
to pick a unique NP which will serve as controller. The proposal
made in the last section suggests when it should be possible to
pick a unique NP: just when the matrix verb establishes a network
of coreference involving the complement subject, i.e. when the
verb requires the complement subject to be missing. If there is
n o network of coreference imposed by the verb, selection of a
controller should be free within the constraints imposed by pro-
nominalization.

   To test this claim, let us look through the classes of
verbs listed at the beginning of §7. Only three classes there
have optional control. Two of these--wait, hope, pray, decide
and bother, benefit--only have one possible NP in the clause
other than the complement, so they don't test the present claim.
The other class includes shout and scream. In the examples
given, only object controller was illustrated. But in fact

subject controller is also possible, as Perlmutter (1968) pointed out in his discussion of these verbs:

(86)   I screamed (to Bill) to be allowed to go.

In this case and in the examples given in §7, there is no ambiguity in the complement subject. What we have established, however, is that the controller is not always the same NP in the matrix, and that the variation depends on the content of the complement clause. We will show in §11 why these sentences are unambiguous; the extra readings of coreference will be blocked by an independent constraint.

In addition to the examples in §7, there are the constructions (84) we originally cited as ambiguous. But these, too, have merely optional coreference between the complement subject and something in the main clause:

(87)   Harry talked to Bill about Sam's kissing Greta.
       Harry wrote to Bill about Walt's not voting for
        Humphrey.
       Harriet argued with Betty about Jerry's visiting
        you.

Thus, assuming that we can account for the lack of ambiguity in sentences with scream and shout, the generalization seems to be that there is a unique identifiable controller position if and only if control is obligatory.

10. The position of the obligatory controller

Having decided when we must specify a controller, we must now decide how to specify where it is.

Compare get and promise.

(88)   John $\begin{Bmatrix} got \\ promised \end{Bmatrix}$ to leave.

(89)   John $\begin{Bmatrix} got \\ promised \end{Bmatrix}$ Bill to leave.

We cannot say that the controller for a particular verb is always its subject or always its object, because of get. Nor can we use the distance principle, restricting its application to conditions of obligatory control, because of promise.

It would not be very nice to have to say that for promise, whether or not there is a direct object, the subject is controller; but with get, the object is controller if present and otherwise the subject is controller. A more interesting theory would predict this difference on the basis of independently motivated differences between promise and get.

Towards such a theory, notice the similarity in the following sets of examples.

(90)   Joe got to Philadelphia.
Frank got Joe to Philadelphia.

(91)   Joe got furious at Henry.
Frank got Joe furious at Henry.

(92)   Joe got to wash the dishes.
Frank got Joe to wash the dishes.

(93)   ?Joe kept $\begin{Bmatrix} to \\ in \end{Bmatrix}$ his room.
Frank kept Joe in his room.

(94)   ?Joe kept angry.
Frank kept Joe angry.

(95)   Joe kept at the job.
Frank kept Joe at the job.

(96)   Joe kept working on the problem.
Frank kept Joe working on the problem.

The interesting thing about these examples is that the switching
of understood complement subjects in (92) and (96) is exactly
parallel to the switching of attribution of the adjectives and
locatives in the rest of the examples. And this switching in
turn is exactly parallel to the switching of attribution of motion
in (97).

> (97) The rock rolled away.
> Bill rolled the rock away.

In other words, we appear to be dealing with a manifestation of
the system of thematic relations discussed in Chapter 2, §11.

The thematic relations in (90) and (93) are transparent.
In each case, Joe is the theme; with get he is asserted to have
undergone a motion resulting in achievement of the goal Philadelphia;
with keep his location over a certain period of time is asserted
to be his room. In the transitive members of the pairs, Frank
is acting as an agent to bring about this state of affairs. By
the usual analogies, we can consider the adjectives and complements
in the other examples to be abstract goals and locations.

We could explain the variable controller position for get
by claiming that networks of coreference are defined on thematic
relations rather than grammatical relations. Then control with
get could always be associated with the theme, no matter what its
position. Since the fact that the theme of get occurs in various
positions must be expressed in the grammar anyway, there will be
nothing special about the way get behaves under the Complement
Subject Rule.

It still must be shown, however, that thematic relations can distinguish get from promise. Promise unfortunately cannot describe concrete events, so we have no direct analogy as we do with get. However, compare promise with sell.

(98) John sold a bathtub.
      John sold Bill a bathtub.

(99) John promised to go straight.
      John promised Bill to go straight.

In both pairs, a transfer is effected away from John--in (98), a bathtub is transferred; in (99), information (that John will go stright) is transferred. The thing transferred is obviously theme in both pairs. John is in all four sentences the source (and agent). With both verbs, the inclusion of a goal, in these examples Bill, is optional. The fact that Bill is the goal is made clearer by the variant with to, just in case the theme is placed next to the verb; this is of course impossible if the theme is a clause.

(100) John sold a bathtub to Bill.
       John would promise nothing to a scoundrel like Bill.

If this analysis of promise is correct, we see that control can always be associated with the source, whether or not the goal is present. Thus the difference between get and promise is that promise belongs to a class of verbs in which the thematic relations map isomorphically into grammatical relations, whereas get belongs to a class of verbs for which this is not the case.

Next compare give and permit.

(101) Sylvia gave Joe a tricycle.
(102) Sylvia permitted Joe to cross the street.

In (101) <u>a tricycle</u> is being transferred; in (102) the information that Joe may cross the street is being transferred. These are the themes of their respective sentences. In both cases, <u>Sylvia</u> is the source and agent, and <u>Joe</u> is the goal. As with <u>sell</u> and <u>promise</u>, the goal can be marked in certain constructions with <u>to</u>, identifying it more clearly:

(103) Sylvia gave a tricycle to Joe.
       Sylvia's permission to Joe to cross the street

The interpretation of the complement subject in (102) indicates that the controller must be the <u>goal</u> of the matrix sentence. Thus <u>permit</u> differs from <u>promise</u> only in that goal rather than source is controller; the correspondence of thematic relations to grammatical relations is identical.

Another example: the positional and abstract forms of <u>force</u> have the same syntax, and by the usual assumptions, the same thematic relations.

(104) George forced the ball into the hole.
(105) George forced Bob into selling his car.
(106) George forced Bob to sell his car.

Here <u>George</u> is agent, <u>the ball</u> and <u>Bob</u> are themes, and <u>the hole</u> and <u>sell his car</u> are goals, as shown by the directional preposition <u>into</u>, which deletes in (106) by the well-known rule deleting prepositions before <u>that</u> and <u>for-to</u> complements.

Examples could be multiplied, but the point is clear. It is possible, using thematic relations, to explain the position of controller NP by means of a single marking, and to differentiate the various classes of verbs without any special exception appar-

atus. The crucial cases to show that thematic relations are the correct device are _get_ and _promise_, which have an optional element in the VP. If all verbs with an optional NP behaved like _promise_, maintaining a fixed control position when the optional object is added, this would argue that grammatical relations determine control. If all such verbs behaved like _get_, switching the control position when the object is added, this would argue that some restricted variant of the distance principle determines control. But the fact that both types occur, correlated exactly with differences in thematic relations, argues that thematic relations are the factors determining control.

The linking of controller position to thematic relations offers some insight into a well-known aspect of the control problem. (107) and (108) differ in interpretation of complement subject, but there is no difference either in distance or in grammatical relations.

(107) Mary gave Alex permission to go.
(108) Mary received permission to go from Alex.

As we showed earlier, the goal of _permit_ controls complement subject coreference. However, in the nominal form _permission_ the goal is optional, and not present in these examples.

However, examine the thematic relations in the main sentences. In (107), _Mary_ is source and _Alex_ goal; in (108) it is the other way around. In each sentence, it turns out that the goal of the main clause is interpreted as complement subject. If some semantic rule could establish an understood identity between the source-goal

patterns in the sentence and the NP, it could be inferred that in (107), <u>Alex</u> is the goal of <u>permission</u> and in (108), <u>Mary</u>. Then the network of coreference associated with <u>permission</u> would give the correct result.[4]

Such an analysis would also predict correctly the complement subject in (109) and (110), since control with <u>promise</u> goes with the source instead of the goal.

(109) Mary gave Alex a promise to go.
(110) Mary received from Alex a promise to go.

I don't know the conditions under which this rule requiring identity of source-goal patterns is enforced, but the assumption of such a rule gives beautiful predictions in these heretofore refractory examples.

If control is established, then, by networks of coreference which refer to thematic relations, we have solved the ordering problem brought up in §3. The Complement Subject Rule can now be ordered at the end of the cycle, since it is free to assign coreference between the complement subject and any NP at all in the main clause, subject to pronominalization constraints. When there is a restricted choice of coreferents for the complement subject, the restriction is imposed independently by a network of coreference acting as a filter. Like the Thematic Hierarchy Condition, a network of coreference can be thought of as a well-formedness condition on semantic readings, having nothing to do with the derivation of the semantic reading. The existence of two separate kinds of conditions utilizing thematic relations to

constrain coreference is good evidence for the need to represent thematic relations in the grammar.

11. Agents not conditioned by the verb of their clause

In §7 we discovered that a third independent aspect of the control problem is whether or not the complement subject is required to be "in control of the situation." In general, if the subject is marked as an agent by the verb of the complement sentence, it will satisfy this requirement, since the defining semantic property of an agent is the exertion of will or volition towards the accomplishment of the event described by the sentence. However, there are sentences in which some individual other than an agent defined by the thematic relations of the verb can deliberately influence the event.

(111) Tom intentionally struck Bill as rude.
(112) Willy was examined by the doctor in order to
        prove to his mother that he didn't have TB.

Strike (as Adj) does not normally mark its subject as an agent, yet in (111) the adverb intentionally implies that Tom went out of his way to cause Bill to consider him rude. Likewise, since agents are generally marked by the verb in (deep) subject position, the surface subject of a passive such as (112) cannot have been marked by the verb. Yet the in order to clause implies that Willy went out of his way to make sure he underwent examination.

Various factors other than the verb can mark NPs as agents in a sentence. (111) and (112) illustrate the ability of adverbials

to mark the surface subject of their clause as an agent. As we
noted in §7, <u>try</u> requires the subject of its complement to be
an agent. Some interpretations of modals mark their subjects
as agents:

> (113) Billy won't (i.e. refuses to) be examined by Dr.
>   Gronk.
> (114) You $\begin{Bmatrix} \text{must} \\ \text{should} \end{Bmatrix}$ (i.e. are obliged to) be examined by
>   Dr. Schlep.

As remarked in Chapter 2, the imperative requires volition on the
part of the deleted subject, suggesting that it, too, introduces
the marking <u>agent</u> on the subject. Also, the requirement that
the passive <u>by</u>-phrase be interpreted (usually) as an agent may
be expressible as an agent marking on the object of <u>by</u> (assuming
an <u>Aspects</u> account of the passive with an underlying <u>by</u> $\triangle$).

Note, by the way, that all these devices that mark agents
within their own clause (except the passive <u>by</u>-phrase, if it is
one of these) mark the <u>surface subject</u>. This is evidence that
the surface subject does indeed play some special role in semantic
interpretation, as was intimated in §8. It is unclear to mé,
however, what this special role is.

Of course, there are many cases where the introduction of
an agent marking by something other than the verb causes a sen-
tence to become unacceptable.

> (115) ?John intentionally knew the answer.
> (116) ?Harold tried to be small.
> (117) ?You must be judged inadequate.

Thus the question arises, What sentences can have agents and which

NPs in those sentences can be agents?

Fischer and Marshall (1969), in the course of discussing Perlmutter's (1967, 1968) theory of begin, give a promising start towards an answer to this question. They show that the ability of a subject to be an agent is a complex and subtle property, tied very closely to the meaning of the sentence. Of course, if the verb marks the subject as an agent, there can be no conflict if something else marks the subject as agent, too. So the interesting cases are sentences in which the verb does not mark the surface subject as an agent.

There appears to be a rough criterion for distinguishing agents marked by the verb from agents marked by other than the verb. If the subject is an agent not marked by the verb, there will be a plausible paraphrase with NP brought it about that...., whereas with agents marked by the verb such a paraphrase seems odd or redundant. Compare (118), which have agents marked by the verb, with (119), where the subject is marked agent by something else. Their paraphrases are (120) and (121) respectively.

(118) John (intentionally) moved away from the wall.
The doctor tried to examine John.

(119) John intentionally struck Bill as pompous.
John tried to be examined by the doctor.

(120) ?John (intentionally) brought it about that he moved away from the wall.
?The doctor tried to bring it about that he would examine John.

(121) John intentionally brought it about that he struck Bill as pompous.
John tried to bring it about that he would be examined by the doctor.

Fischer and Marshall show that the ability of a sentence to take an agent subject cannot be identified with the feature [-stative] on the verb (recognizable by the ability to take present progressive aspect), as Perlmutter claims. For example, know, a stative verb, may occur under try if accompanied by an appropriate adverb, but sweat, a non-stative verb, seems strange.

(122) John tried to know the answer by the next morning.
(123) ?John tried to sweat.

In fact, the ability to take a plausible agent subject is often a matter of factual knowledge about the world, as illustrated by their discussion (§ 4).

> As one might expect from the idiosyncratically oriented lexicon, there are wide differences among speakers as to what is self-controllable [i.e. can take an agent subject (R. J.)] and what is not, and even in the same speaker, one verb may be self-controllable, while its semantic opposite or a close synonym may not. For example, one of the authors (S. D. F.) can say
>> (109) a. Be taller by next year.
>>   b. I let myself be rumored to enjoy surfing.
> but not
>> (110) a. *Be shorter by Friday.
>>   b. *I let myself be said to enjoy surfing.
> The reason for this is that she knows of hormones that can produce growth, but none (discounting hookah-smoking caterpillars and mushrooms) that produce shrinkage, and that one can start or quash rumors about oneself, but one cannot stop people from saying things....

If the ability of a sentence is this intimately bound up with the meaning of the sentence and with the real-world conse- quences of the sentence, a formal criterion for appropriateness of agent subject is obviously unstatable within the scope of dis- cussion to which we are limited here. Indeed, it will probably be unstatable in any framework of the near future. This does

not, however, prevent us as linguists from using our knowledge
of the language to tell us when something is intended to be the
agent of a sentence. If we can describe their behavior informally,
we can discover interesting things about them even if we do not
know exactly how to represent them.

Besides the class containing try and condescend, we know
of several classes of verbs that mark their complement subjects
as agent. Persuade and force take direct objects as well as the
complement; the object and complement subject are coreferential,
and the complement subject must be an agent. Promise and vow
take optional indirect objects; the subject and complement subject
are coreferential, and the complement subject must be an agent.

In all the cases mentioned so far, since the complement
subject is always coreferential with some NP in the main clause,
we could have the verb mark either the complement subject or its
controller as the NP that is the agent of the complement. However,
consideration of another class of verbs shows that it is the NP
in the main clause which should be marked, not the complement
subject. Recall the discussion of § 9, in which we concluded
that there is no network of coreference invoked by verbs like
scream and shout. For these verbs, we get the following paradigms:

    (124)   *I screamed to go.
    (125)   I screamed to Bill to go.
    (126)   I screamed to Bill for Harry to go.
    (127)   I screamed to be allowed to go.
    (128)   I screamed to Bill to be allowed to go.
    (129)   I screamed to Bill for Harry to be allowed to go.

Note first of all that in (126) and (129) I am asking Bill to bring it about that Harry go or be allowed to go; it is only by virtue of action on Bill's part that anything will happen to Harry. Bill thus meets our intuitive criteria for being an agent of the complement clause, suggesting that it is indeed the NP in the main clause that is marked as agent, not the complement subject.

Consideration of the meaning of (126) reveals an interesting thing about sentences with two agents. Notice that Harry is the agent marked by the verb in the complement clause. The relation of Bill's action to Harry's action indicates that the agent defined in the higher clause has some kind of causal or temporal priority over the agent defined by the verb: the former must exert volition first. This is in fact also true in sentences like (130), in which Bill, the agent marked by the adverb, must act before the doctor, the agent marked by the verb, can do so.

(130) Bill was intentionally examined by the doctor.

Sentences like (126) also show the dependence of the ability to take an agent on the meaning of the sentence. Compare the following:

(131) *I shouted to Bill to be tall.
(132) *I shouted to Bill for Harry to be tall.
(133) I shouted to Bill for the next recruit to be tall.

Be tall does not normally allow an agent in the subject. In fact, because it is impossible for someone to make someone else tall, (132) is out: Bill cannot be an agent of the complement. Why is (133) good? The meaning of (133) is that Bill should exercise

choice in selecting the next recruit, and that the one he selects should be tall. The supposition that Bill is responsible for the choice of the next recruit is the only plausible way in which <u>Bill</u> can be understood as an agent over the sentence. I take it that this supposition is inferred from Bill's agenthood and not the other way around: what would be a conceivable deep structure that explicitly represented it?

R e t u r n to (124)-(129). In (128), since there is no network of coreference with <u>scream</u>, we would anticipate that the complement subject could be coreferential with either <u>I</u> or <u>Bill</u>. But assume <u>Bill</u> is chosen as antecedent of the complement subject. Then the complement subject would be agent of the complement. But since <u>be allowed to go</u> does not permit its subject to be an agent (as shown by, e.g. *<u>I intended to be allowed to go</u>), this reading would not be acceptable. Therefore <u>I</u> must be the antecedent of the complement subject. Similarly, in (127), <u>I</u>, not being an agent of the complement, may serve as antecedent of $\triangle$.

For (124) and (125), these considerations are not enough. What is to prevent <u>I</u> from being the antecedent of $\triangle$, even if <u>Bill</u> is the agent? If <u>I</u> could be the antecedent, we would expect a reading of (125) rather like that of (126). One plausible addition to the theory that could describe this situation might be some preference for identifying agents defined by the verb and by other factors as the same individual whenever possible.[5] This preference could explain the interpretation of the complement

subject in (125) as <u>Bill</u> exclusively, and the inability for the complement subject to refer to <u>I</u> in (124) because of the implicit agent. Admittedly, there is so far no other evidence for this addition to the theory, but it seems to me to point in the right direction.

Some support for this analysis comes from a slightly different class of verbs, including <u>beg</u> and <u>ask</u>. These have the same paradigm as <u>scream</u> except that the counterpart of (124) is acceptable.

(134)   I begged to go.
(135)   I begged Bill to go.
(136)   I begged Bill for Harry to go.
(137)   I begged to be allowed to go.
(138)   I begged Bill to be allowed to go.
(139)   I begged Bill for Harry to be allowed to go.

Note that (134) is synonymous with (137). They both para-phrase the direct question <u>May I go</u>?, whereas (135) paraphrases the direct question <u>Will you go</u>? In a theory accepting the strong Katz-Postal Hypothesis, we would thus be compelled to posit two complementary verbs <u>beg</u>, or two drastically dissimilar deep structures for (134) and (135).

In the present theory these interpretations make more sense. The similarity between the two sentences lies in the fact that I am requesting someone to act as agent over the event of the complement clause. In (135), he can do so directly by being the agent marked by the verb; in (134), the implicit recipient of of the plea can only act indirectly, by granting permission or otherwise making my going possible. But how the agency of the

indirect object is carried out is perhaps not specified in the
meaning of beg, but is rather a consequence of the meaning of the
sentence. Again, I don't know how to be more specific about this
analysis, or how to differentiate these verbs from those like
scream. However, the fact that this kind of account appears to
lead potentially towards a more sophisticated semantic analysis
and away from systems of arbitrary exception features seems to
me a strong incentive to continue work in the future.

Perhaps a few words should be said here about the so-called
Modal Constraints discussed in Postal (1968b). Postal observes
that with verbs of linguistic performance such as beg, scream,
shout, ask, and so forth, the subjects of certain that-complements
obey the same coreference conditions as the deleted subject of
for-to complements. For example, in the following sentences,
under the "obligation" interpretation of the modal, Postal says
that he can only be Harry.

     (140)   Bill told Harry that he ought to leave.
              Bill shouted to Harry that he should stop.

These parallel the coreference in

     (141)   Bill told Harry to leave.
              Bill shouted to Harry to stop.

Postal tracks down the exact senses of the that-complements for
which this correspondence holds; he discovers that they are the
ones that paraphrase direct discourse in which the speaker re-
quests some action on the part of the hearer. When all that is
happening is that information is being transmitted, the corres-
pondence does not hold.

In the analysis given here, this fact is meaningful: if
only information is being transmitted, the hearer is obviously
not being asked to act. If the hearer is being asked to act,
(i.e. asked to be an agent), the additional coreference constraints
follow from the semantic principles governing agents. Since for-to
complements with these verbs always request action, the coreference
constraints always hold.

Note, by the way, that in (142) the agent restriction
operates in the usual fashion to make Bill the antecedent of he.

    (142) Bill told Harry that he ought to be allowed to
            leave.
    (143) *Bill told Harry to be allowed to leave.

The same thing should happen in (143), but the network of corefer-
ence established by tell (over its for-to complements only) says
that only Harry is allowed to be the antecedent. Since neither
condition can be satisfied without violating the other, the sen-
tence is ruled out. This is a further illustration of the
interdependence of the network of coreference and agent constraints.

NOTES

1. Postal (1968b) hints at the need for this kind of restriction toward the end of the Appendix (p. 106).

2. Chomsky's rule for assigning focus says that the focus can be any constituent containing the stress center of the sentence; in the case of contrastive stress, only the stressed constituent may be focus. Observe the range of possible foci predicted by this rule for (74)-(78).

> It is easy to play sonatas on this violin.
>   Foci: entire sentence, entire VP, <u>to play sonatas on this violin</u>, <u>(on) this violin</u>.

> Sonatas are easy to play on this violin.
>   Foci: entire sentence, <u>are easy to play on this violin</u>, <u>to play on this violin</u>, <u>(on) this violin</u>.

> This violin is easy to play sonatas on.
>   Foci: entire sentence, <u>is easy to play sonatas on</u>, <u>to play sonatas on</u>, <u>sonatas</u>.

> It is easy to play sonatas on this <u>violin</u>.
>   Focus: <u>this violin</u>.

> It is easy to play <u>sonatas</u> on this violin.
>   Focus: <u>sonatas</u>

The five sentences are distinguished by their ranges of possible foci and presuppositions. In particular, the important difference is whether <u>sonatas</u>, <u>this violin</u>, or both are contained in the focus. See Chapter 5, §§ 7-8 for more discussion of focus.

3. We can see the same sort of semantic restriction between <u>of</u>-objects and <u>that</u>-complements of <u>believe</u> (pointed out by

J. R. Ross):

I believe of Holland that $\begin{cases} \text{Amsterdam is a swinging place.} \\ \text{the tulips are beautiful.} \\ \text{*New York is a drag.} \end{cases}$

4.   This rule may be related to the rule proposed in Chapter 2, § 9, which creates a surrogate subject (agent) for picture in <u>the picture of himself that John saw is ugly</u>.

5.   This rule bears some similarity to the rules of § 10 and Chapter 2, § 9 in requiring parallelism of thematic relations in two structurally related Ss or NPs. Preliminary investigation of this phenomenon ought to be possible, given these three (admittedly rather vague) examples.

CHAPTER 4
NEGATION AND QUANTIFIERS[1]

Edward S. Klima's paper <u>Negation in English</u> (Klima, 1964)
presents a system of rules for sentence negation in English. The
present chapter is an attempt to extend his principles so as to
handle verb phrase negation and multiple negation within a single
sentence. In the course of the argument it will become apparent
that, in order to capture the right generalizations, it is neces-
sary to give up the assumption that transformations do not change
meaning. Two rules of semantic interpretation will be proposed,
both of which refer to derived structure.

1. Klima's Rules for Negation

For the sake of presenting a certain range of data in a
systematic fashion, I will first discuss how Klima's rules account
for the facts of sentence negation. I will use (1) as a rough
semantic test for sentence negation.

(1) A sentence $[_S X\text{-neg-}Y]_S$ is an instance of sentence
negation if there exists a paraphrase <u>It is not so
that</u> $[_S X\text{-}Y]$.

This definition differs from Klima's in that it is based on
paraphrase properties rather than on the existence of negative
appositive tags such as <u>not even John</u> or negative conjoined sen-
tences such as <u>and neither did Bill</u>. It will be shown later on

that Klima's criteria follow from (1).

Klima generates sentence negation as an optional consti-
tuent neg attached to the main S as a daughter. He then has a
sequence of transformations which account for the occurrence of
the neg in a wide variety of surface structure configurations.

The first relevant transformation to apply is the passive.
Klima treats this as an optional rule which operates on simple
sentences with transitive verbs, taking, for example, John hit
Mary into Mary was hit by John.

The second rule is indef-incorporation, which I will call
informally the some-any rule. This rule involves the important
structural relation in construction with, which is defined as
follows:

(2) A node A is in construction with a node B if and
only if the node C directly dominating B also
dominates A.

To illustrate this concept, examine tree (3).

(3)
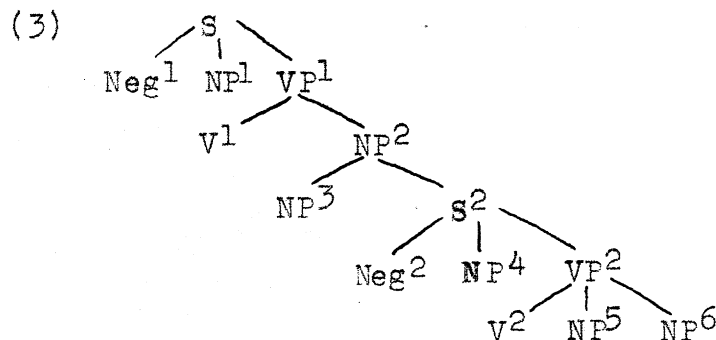
First looking at S^2, we see that NP^4, NP^5, and NP^6 are all in con-
struction with Neg^2 by virtue of being dominated by S^2, which
directly dominates Neg^2. NP^5 and NP^6 are also in construction
with NP^4, because S^2 directly dominates NP^4. On the other hand,

NP^4 is not in construction with NP^5 or NP^6, because VP^2, which directly dominates NP^5 and NP^6, does not dominate NP^4. In the larger sentence, all nodes except S^1 are in construction with Neg^1, but Neg^1 is only in construction with NP^1 and VP^1. Ross (1967) claims (§5.2.2.) that in construction with is not the proper structural relation to use to account for negation, but rather that the notion command, first suggested by Langacker (1969), is correct. Command is defined as follows:

> A node B commands a node A if and only if the lowest S node dominating B also dominates A.

We see that if A is in construction with B, B commands A. However, the converse is not true. For example, in (3), NP^5 commands NP^4, but NP^4 is not in construction with NP^5. I will show later on that in construction with rather than command is the structural relation relevant to the syntax of negation.[2]

Having defined in construction with, we can now state Klima's rule of indef-incorporation. The rule mentions two different classes of morphemes, Affectives and Indeterminates. The class Affective includes for example neg, wh of questions, reluctant, and too (the degree modifier on adjectives). Indeterminate includes for example too (also), once, sometime, somewhere, and quantifiers like many and some. The rule says that Indeterminate constituents undergo a morphological change to an "indefinite" form when they are in construction with an Affective. The morphological change is idiosyncratic for each indeterminate morpheme. The following examples illustrate the rule:

(4) John has some money.
John doesn't have any money.
    (<u>some</u> in construction with neg)

(5) John has been there once.
Has John ever been there?
    (<u>once</u> in construction with wh)

(6) Irving went to the movies, and Max went too.
Irving didn't go to the movies, and Max didn't go either.
    (<u>too</u> in construction with neg)

(7) John was eager to read something about the war.
John was reluctant to read anything about the war.
John was too lazy to read anything about the war.
    (<u>some</u> in construction with <u>reluctant</u>, <u>too</u>)

(8) I claim that someone will force John to do some work.
I'm not claiming that anyone will force John to do any work.
    (<u>some</u> changing to <u>any</u> in construction with neg in a higher sentence)

After indef-incorporation comes the rule that preposes adverbials such as <u>ever</u>, <u>either</u>, <u>once</u>, <u>somewhere</u>, and <u>for two years</u>. The indefinites among these preposed adverbials will later absorb neg to become <u>never</u>, <u>neither</u>, etc.

Next comes a set of rules which accounts for the surface structure placement of sentence negation. The first rule moves the neg from its position at the beginning of the sentence to a position directly before the auxiliary, where it remains in the case of infinitival and participial phrases. The second rule is neg-incorporation into indefinites: if there are any indefinites (the outputs of the some-any rule) before the neg, then neg is obligatorily incorporated into the first of these in the sentence; otherwise neg is optionally incorporated into the first following indefinite occurring in the same clause. The third rule, preverbal particle placement, moves neg to a position after the first

element of the auxiliary if it has not been incorporated into an indefinite, and if the auxiliary contains <u>Tense</u>.

To illustrate these rules, we first look at the various outputs from the underlying structure

(9) neg someone once gave John something

By indef-incorporation this becomes

(10) neg anyone ever gave John anything.

Neg then moves to the position in front of <u>gave</u>, by the first placement rule, then obligatorily incorporates into <u>anyone</u>, giving

(11) No one ever gave John anything.

On the other hand, if passivization had moved <u>anything</u> to the front, so instead of (10) we had

(12) neg anything was ever given to John by anyone,

then neg will attach to anything, giving

(13) Nothing was ever given to John by anyone.

If <u>ever</u> preposes in (10), it receives the neg, giving

(14) Never anyone gave John anything,

which by the later rules of subject-aux inversion and do-support becomes

(15) Never did anyone give John anything.

Likewise, if <u>John</u> is moved to the front by passive, <u>ever</u> will receive the neg, to give

(16) John was never given anything by anyone.

If <u>ever</u> is removed from (10), and <u>John</u> is again moved to the front by passive, we get the intermediate string

(17) neg John was given anything by anyone.

In (17) there is no indefinite before the auxiliary, so we can optionally attach neg to anything, the first indefinite to the right of the aux, giving

(18) John was given nothing by anyone.

If we choose not to attach neg to anything, then it must attach to the first element of aux to yield

(19) John wasn't given anything by anyone.

All the sentences (11), (13), (15), and (16) are synonymous and are paraphrases of

(20) It is not so that anyone ever gave John anything.

(18) and (19) are also paraphrases of this, with the exception of the missing adverb ever.

Following the rules of neg placement, there are two relevant transformations: subject-aux inversion and do-support. We have already seen one instance of their application in the derivation of (15). Inversion generalizes to take place either when a constituent containing wh of questions or one containing neg is in front of the subject. Do-support then operates in either case. In addition, do-support has one other application here. In case there is no element other than Tense in the aux, neg may end up between Tense and the main verb. This prevents Tense from attaching to the main verb and so do-support takes place, giving sentences like

(21) John didn't see anyone.

To sum up, then, Klima's array of rules looks like this:

    (22) 1. Passive
            2. Indef-incorporation (some-any)
            3. Adverb preposing
            4. Neg-placement rules
            5. Subject-aux inversion
            6. Do-support

2. Other than Sentence Negation

To a certain extent, Klima discusses sentences containing neg which are not sentence negation. For example, there are the ambiguous readings of

    (23) They're fighting about nothing.
    (24) I will force you to marry no one.

One reading of these sentence is synonymous with

    (25) It is not so that they're fighting about anything.
    (26) It is not so that I will force you to marry anyone.

But there is another reading for each sentence for which this paraphrase is not available. Klima calls this reading an instance of "constituent negation." In the case of (24), the second reading derives from sentence negation in the complement clause, i.e. from (27).

    (27) I will force you [$_s$neg you marry someone]$_s$

If this neg is not attached to the object of the complement, it shows up before the verb:

    (28) I will force you not to marry anyone.

On the other hand, the second reading of (23) cannot be derived from an underlying sentence negation, since there is no sentence in <u>nothing</u>. The neg appears to be semantically associated

with the NP. Thus we are forced to acknowledge the possibility of generating neg in positions other than as daughter of S.

A more interesting case of constituent negation is illustrated by the contrast of (29) and (30).

(29) Not many of the arrows hit the target.
(30) Many of the arrows didn't hit the target.

(29) can be derived from ordinary sentence negation of Many of the arrows hit the target; it is synonymous with
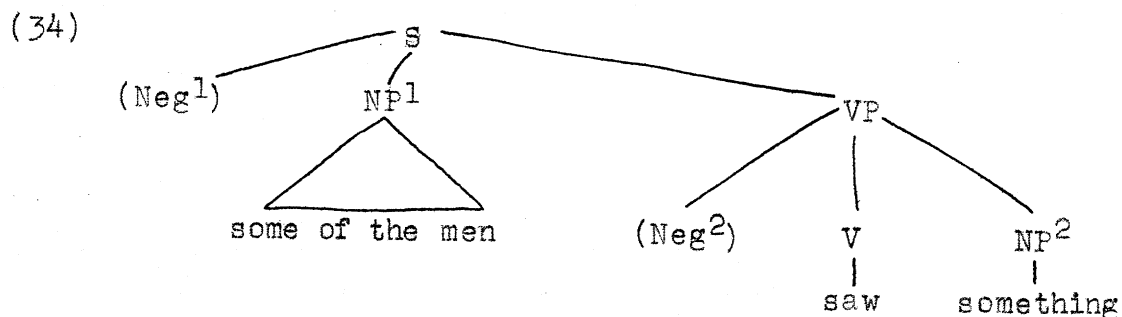
(31) It is not so that many of the arrows hit the target.

In (29), not is placed before many by the obligatory incorporation of neg into the first indefinite before the aux. What then of (30), which is not synonymous with (29) or (31)? We notice that (30) cannot be derived by the rules for sentence negation, since the neg is not attached to many by the obligatory neg-placement rule, but is rather associated with the aux.

Sentences like (30) show that command is the wrong structural relation to use to determine the some-any rule. Contrast (32) with (33), the sentence negation case.

(32) Some of the men didn't see anything.
(33) None of the men saw anything.

If we say that the neg in (30) and (32) is associated with the VP instead of the S, we can account for the behavior of the some-any rule by using in construction with, but not by using command.

(34)

```
                          S
           (Neg¹)      NP¹                    VP
                    /‾‾‾‾‾\
                   some of the men    (Neg²)    V        NP²
                                               saw    something
```

In (34), both Neg^1 and Neg^2 command both NP^1 and NP^2, and so the difference between Neg^1 and Neg^2 cannot provide the condition to distinguish (32) from (33) in deep structure, assuming that command is the correct structural relation. On the other hand, NP^2 but not NP^1 is in contruction with Neg^2, whereas both NPs are in construction with Neg^1. Thus if _in construction with_ is the correct relation to use for the some-any rule, we can derive (32) from the deep structure with Neg^2 and (33) from the deep structure with Neg^1. Similarly, (29) has an S negation and (30) a VP negation. We have seen already in (23) that we need to sometimes generate neg as a daughter of a constituent other than S, so generating it in the VP should not be particularly objectionable. Since the neg-placement rules only attach neg to indefinites, Neg^2 will not be moved into the subject.

The most remarkable aspect of sentences such as (30) and (32) is that they have no synonymous passive. If we form what intuitively looks like the passive of (30), we get (35), which is also the output we get from passivizing (29).

(35)   The target wasn't hit by many of the arrows.

Thus, we predict that (35) should be ambiguous. Unfortunately it is not.[3] Its only reading is synonymous with (29). In order to

make this clearer, append an additional clause to both (29) and (30):

(36) *Not many of the arrows hit the target, but many of them did hit it.

(37) Many of the arrows didn't hit the target, but many of them did hit it.

(36) expresses a contradiction since it asserts both a sentence and its negation. (37) is not a contradiction because (30), being an instance of VP negation, does not conflict with <u>many of them did hit it</u>. On the other hand, (38), where both clauses are passivized, can only be understood as a contradiction.

(38) *The target wasn't hit by many of the arrows, but it was hit by many of them.

This contradiction shows that the only possible reading of the passive (35) is the one synonymous with (29), the S negation.

There also exist pairs of passive sentences similar to (29) and (30), for which only one member has a synonymous active:

(39) Not many of the demonstrators were arrested by the police.

(40) Many of the demonstrators weren't arrested by the police.

Again, (39) is an instance of sentence negation and (40) is not; only (39) is synonymous with

(41) It is not so that many of the demonstrators were arrested by the police.

When we form what is intuitively the active of these sentences, we get in both cases

(42) The police didn't arrest many of the demonstrators.

However, this sentence has only the reading corresponding to (39), the sentence negation case. Again, we can add extra clauses to

bring out this fact more clearly.

(43) *Not many of the demonstrators were arrested by the
      police, but many were.
(44) Many of the demonstrators weren't arrested by the
      police, but many were.
(45) *The police didn't arrest many of the demonstrators,
      but they did arrest many of them.

Just as in (36)-(38), we get a contradiction only when we simul-
taneously assert both a sentence and its sentence negation; its
VP negation does not produce a contradiction.

Under the assumption that transformations do not change
meaning, these facts will be very difficult to account for. Klima
does not discuss sentences with VP negation anywhere in his paper,
and it is not clear whether his rules generate them or not. In
order for the some-any rule to operate correctly, the neg must be
in the VP at the time of this rule. If we add a transformation
to move neg into the VP before the some-any rule, this transforma-
tion will change meaning just in case there is a quantifier in the
subject, as in (29) and (30). If this transformation is ordered
before the passive, then the passive will change meaning just in
case it moves a quantifier into or out of the subject. Therefore
VP negation cannot be produced by adding a transformation to Klima's
system, since this transformation would change meaning.

Let us assume then that the VP negation is generated in
the base. Then in order to prevent (30) from passivizing, we must
add a restriction on the passive to the effect that it is inhibited
by VP negation when there is a quantifier in the deep subject.
What then is the deep structure of (40), the passive VP negation

with a quantifier in the derived subject? It would appear that
we must make passive obligatory when there is a VP negation and
a quantifier in the object, so that we will have no corresponding
active for (40). Thus we can, with a strange pair of conditions
on the passive, create a theory consistent with what we know about
VP negation, in which the rules preserve meaning. However, this
theory misses the fundamental similarity between the pairs (29-30)
and (39-40), namely that VP negation is significant exactly when
there is a quantifier in the derived subject.

It has been suggested by many people (personal communica-
tion) that this trouble can be avoided by using the notion,
introduced by Lakoff (1965), Carden (1967, 1968), and others,
that negation and quantifiers are generated in the base as verbs
of higher sentences.[4] Thus, for example, the structures of (46)
and (47) are supposed to be (roughly--the details don't matter
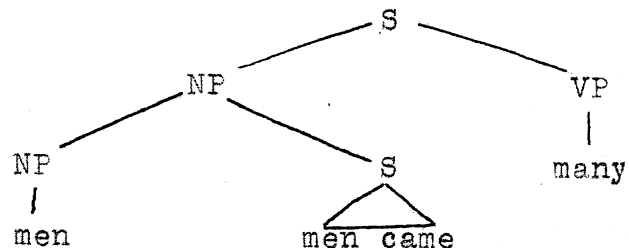for this argument) (48) and (49) respectively.

(46)   The men didn't come.
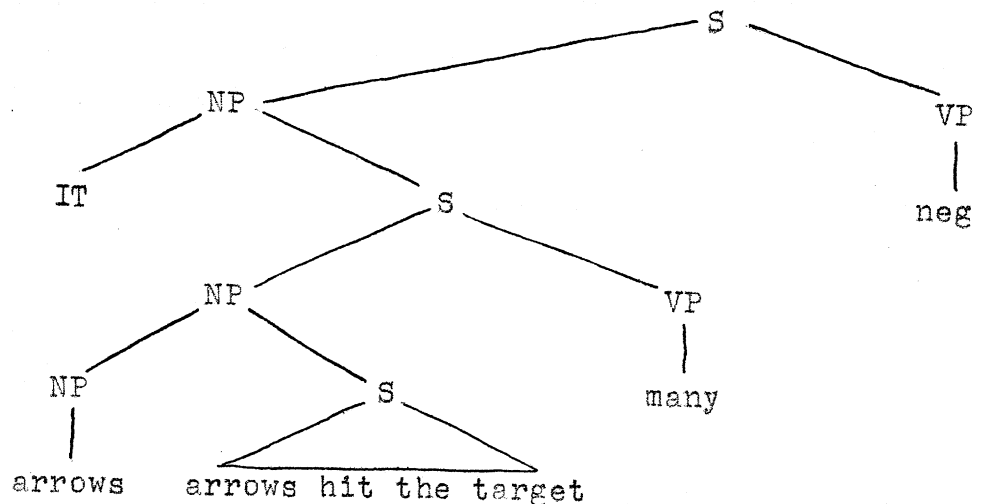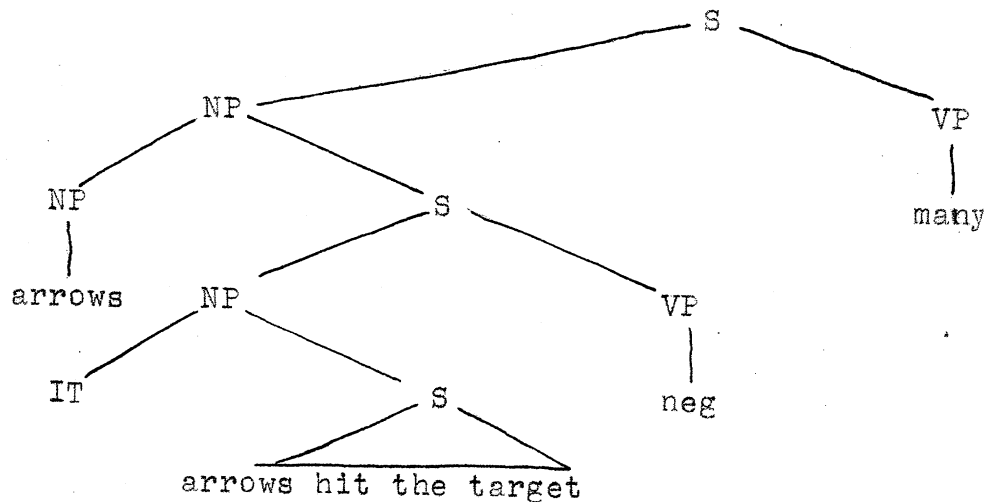(47)   Many men came.
(48)
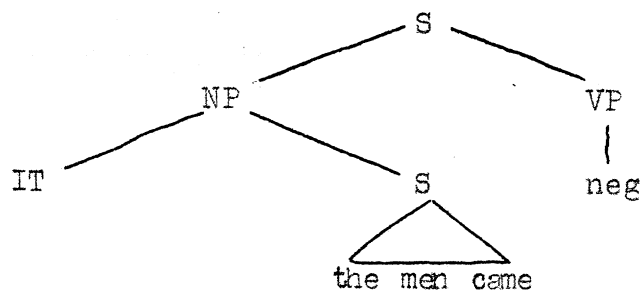
(49)

The understood order of quantifiers and negatives in a sentence is supposed to correspond to the hierarchy of upper sentences containing quantifiers and negatives. Thus the difference between (29) and (30) is reflected in their deep structures, (50) and (51).

(50)

```
                                              S
                              NP _____/ \____ VP
                            /    \                      |
                          IT      S                     neg
                                /   \
                              NP      VP
                            /    \       |
                          NP      S      many
                          |      / \
                       arrows  arrows hit the target
```

(51)

```
                                              S
                              NP _____/ \____ VP
                            /    \                      |
                          NP      S                     many
                          |      /  \
                       arrows  NP     VP
                              /  \      |
                            IT    S     neg
                                 / \
                            arrows hit the target
```

An immediate defect in this formulation is that no such difference in order of quantifiers and negation can be produced when the quantified NP is the <u>object</u> of the lowest sentence. For

example, if the lowest sentence in (50) and (51) is <u>John brought</u>
<u>arrows</u> instead of <u>arrows hit the target</u> the only possible vari-
ants are

(52)   John didn't buy many arrows.
(53)   John bought not many arrows.

But these two are synonymous and unambiguous; the negation is
interpreted as being outside the quantifier. Therefore we must
not allow a quantifier above negation just in case the quantified
NP is in the object of its sentence. This is a very strange re-
striction.

Next consider what happens when we introduce the passive.
If the passive, too, is a higher sentence, its relative configu-
ration with respect to quantifier and negation sentences must be
irrelevant to the understood order of quantifiers and negation.
Any restriction of the position of passive in the hierarchy only
requires further constraints. However the following constraints
must hold on the occurrence of passive. If we have a negation
and quantifier configuration such as (51), passive may not be
present just in case the quantified NP is a subject, because
otherwise the output would be understood the same as (50), which
has the opposite deep structure order of quantifier and neg. If
the quantified NP is an object, then the passive <u>must</u> be present,
giving the difference between (39) and (40). The single inter-
pretation of the corresponding active shows that this configuration
cannot exist without the passive.

Thus the hypothesis that quantifiers and negation are higher

verbs fails to explain naturally the facts about VP negation that we have noticed. In fact, the restrictions necessary in this theory are exactly cognate to the restrictions required by the traditional theory that quantifiers and negation are generated as part of the sentence in which they eventually occur.

Both of these approaches miss the significant generalization that S and VP negation differ in meaning exactly when there is a quantifier in the derived subject. Since the possibility of difference in meaning is dependent on the derived subject, and not on the underlying subject, a theory which requires that all semantic information be captured in deep structure cannot express this generalization.

If we give up the assumption that transformations do not change meaning and that all semantic information is represented in deep structure, it immediately becomes apparent how to go about explaining the interpretation of VP negation. We simply need a way to relate the understood order of quantifiers and negation to their position in the derived structure. We can do this informally by saying that the understood order is the same as the order in the surface structure. We then get the differences (29-30) and (39-40) because of the difference in surface structure order of neg and _many_. On the other hand, in the passive of (29-30) and in the active of (39-40) the quantifier is in the object and so there is no syntactic position to the right of it into which neg can fit. Since there is no such position, we cannot produce the expected second reading for the sentence.

The rest of the chapter will be devoted to developing a system of rules consistent with this line of speculation about interpretation of negation and quantifiers.

3. Scope--an intuitive discussion

Given a sentence containing a negation, let us define scope of the negation (intuitively) as that part of the sentence which is interpreted as being denied. Thus in cases of sentence negation, the scope of the negation is the whole sentence; this accounts for the paraphrase with it is not so that. In the cases of constituent negation we have discussed, this paraphrase is not possible, since not the whole sentence is being denied.

In cases of constituent negation, it is sometimes difficult to get a feel for what part of the sentence is being denied. Traditional symbolic logic is of no help in this respect, because it deals only with negations of propositions (sentences); it does not talk about the negation of predicates, constants, or variables.[5] However, if we identify the scope of negation with various subtrees in the derived structure of a sentence, we can achieve some intuitively satisfying conclusions.

For examples, let us return to the sentences of §2.

(24) I will force you to marry no one.

The two readings of this correspond to

(54) I won't force you to marry anyone.
(28) I will force you not to marry anyone.

In the reading (54) the entire sentence I will force you to marry someone is being negated; in the reading (28) I will force you

retains a positive sense, but <u>you marry someone</u> is negated. Thus
we can speak of the two readings being associated with two differ-
ent scopes of negation, the higher and the lower Ss. This corres-
ponds to the two positions in which we could generate the neg
according to Klima's rules in order to get the surface form (24).

(23) They're fighting about nothing.

In one reading, it is denied that they're fighting; in the other,
it is asserted that they're fighting, but denied that there is a
reason. In other words, we can associate that scope of negation
either with the entire sentence or with just the NP <u>nothing</u> (or
perhaps the PP <u>about nothing</u>). Hence again we can identify scope
of negation with two different places in which Klima's rules gen-
erate neg in order to produce the surface form (23).

Finally, let us turn to the examples of VP negation we
discussed.

(30) Many of the arrows didn't hit the target.
(32) Some of the men didn't see anything.
(40) Many of the demonstrators weren't arrested by the
police.

The reading of (30), for example, asserts of <u>many of the arrows</u>
that they failed to hit the target; in other words, the scope of
negation is the VP, exactly where we needed to generate neg in
order to produce the correct surface form.

If there is no quantifier in the subject of a sentence, the
VP negation of the sentence will be logically synonymous with the
S negation, but it will be understood as different in its topic-
comment relations. For example, in (55)

(55) The arrow didn't hit the target

the sense of the S negation is denying that the arrow hit the
target, while the VP negation asserts of the arrow that it failed
to hit the target. The truth values of these readings are equi-
valent; the only difference is in emphasis. However, as we have
seen, the presence of a quantifier in the subject does lead to a
difference in truth value between the S and VP negations.

The examples we have seen so far make it plausible to
assume that the interpreted scope of negation is always associated
with a particular node in the tree. Under this assumption we
would not expect to find cases where we can identify the scope of
negation, say, with the subject and verb of a sentence but not
the object, since any node dominating both the subject and the
verb must dominate the node VP and thus the object. If this scope
were possible, we would be able to get a reading of (35) synony-
mous with (30).

(35) The target wasn't hit by many of the arrows.
But in fact there is no such reading. As far as I know, none of
the examples in this chapter is susceptible to an interpretation
in which the scope of negation is not an entire constituent.

If the interpretation of a sentence depends on what material
is contained in the scope of negation, we would expect there to be
two possible ways to change meaning through scope of negation.
The first way, which we have just discussed, is to change the
scope of negation. The second way would be to leave the scope of
negation fixed on a particular constituent but move material in

and out of that constituent. This is in fact exactly what happens

when we form the passive of sentences with VP negation: the sub-

ject is moved into the scope of the negation, and the object is

moved out of the scope. If the scope of negation is the entire

sentence, the material is reshuffled by the passive, but nothing

moves in or out of the scope, so meaning is preserved.

If the preceding discussion is correct, we have found that

the structural notion "in construction with neg," involved in

the some-any rule, has the semantic correlate "within the scope

of neg." This lends support to the assertion that in construction

with rather than command is the structural relation in terms of

which the some-any rule should be stated. Furthermore, it suggests

that the change of some to any is not merely a capricious change

brought about by the presence of a neg nearby, but that there

may in fact be some semantic significance to the alteration. It

also suggests that we should search for phenomena of scope in

connection with other Affectives as well. Only the former sugges-

tion will be followed up here.

4. The some-any rule

Is it really desirable to state the some-any rule as a

transformation? If we do state it as a transformation, the rule

says that members of a certain class of lexical items are changed

in a single feature when in construction with an Affective. The

changing of this feature initiates a major change in the "phono-

logical spelling" of the lexical item. Some examples of the

alterations are

    (56)   some/any
           once/ever
           too/either

The most striking thing about this group is that it is irregular, involving pairs of totally dissimilar "spellings." This suggests immediately that we are perhaps not dealing with variant forms for single lexical items, but rather that affective environments permit different sets of lexical items than non-affective environments do. This suspicion is substantiated by the fact that there are a couple of "indefinites" for which there is no corresponding form in non-affective environments:

    (57)   any more
           at all
           any good

For these items a "spelling rule" would have to block derivations which end with the definite version. Such a rule should immediately raise eyebrows. Since it clearly violates the Extended Lexical Hypothesis, we are led to try to sate a lexicalist version of the some-any rule.

In this formulation, we will consider *some* and *any* (as well as the other pairs) as separate lexical items differing by a feature, say $[\pm X]$ (*some* is $[+X]$). There will be rules of semantic interpretation which specify which value of the feature must appear in what environments, much as a selectional restriction specifies features of NPs in relation to verbs.

The rule will be stated more precisely as follows:

$$(58) \quad [+\text{indeterminate}] \rightarrow \begin{cases} [-X] \text{ in construction with Affective} \\ [+X] \text{ elsewhere} \end{cases}$$

The convention for application of this rule is as follows:

(59) If an indeterminate is unspecified with respect to X, the rule fills in the feature according to the environment. If the indeterminate is already marked with respect to X, the sentence is marked semantically anomalous if the inherent feature and the feature assigned by the rule disagree.

This convention on application can be thought of as essentially what happens in the application of selectional restrictions such as those in Chomsky (1965).

To illustrate the rule, consider the following sentences.

(60) John bought some candy.
(61) John didn't buy any halvah.
(62) *John bought any grease.
(63) John bought a house.
(64) John didn't buy a monkey wrench.

In (60) and (61), the inherent feature S on some and any agree with the values assigned by the rule. In (62), any, being [-X], disagrees with the marking [+X] assigned by the rule, and so the sentence is anomalous. The indefinite article a is unmarked for the feature X, and so it receives the feature [+X] in (63) and [-X] in (64).

Let us compare the interpretive rule (58) to the transformational some-any rule with respect to postulated base forms. Both rules must take place after the passive in order to get the difference between, for example,

(65) Some of the men didn't do any jobs.
(66) Some jobs weren't done by any of the men.

Both of these are grammatical instances of VP negation. With
a transformational version of the some-any rule, both sentences
are derived from the base string (67) by Klima's rules. (This
assumes of course that the passive can change meaning. We have
already discussed the problems that arise if the passive may not
change meaning.)

(67) Some of the men neg did some jobs.

In a grammar with the interpretive some-any rule, (65) and (66)
derive from the base strings (68) and (69) respectively.

(68) Some of the men neg did any jobs.
(69) Any of the men neg did some jobs.

If passive takes place in (68), the some-any rule will mark the
sentence anomalous, since some will be within the scope of neg
and any will not. On the other hand, unless passive takes place
in (69), the sentence is anomalous because of incorrect use of
some and any.

In what sense is the rule (58) different from some sort
of filtering transformation that makes no change in the sentence
but blocks incorrect sentences? The difference is in the first
part of the convention on application (59). The claim is that
the some-any rule fills in a feature of semantic relevance, and
thus contributes to the reading of a sentence in which the fea-
ture X has been left unspecified on a noun phrase. A filtering
transformation could only use the second part of the convention,
blocking incorrect constructions but never contributing to the
reading.

There is in fact a semantic difference in the readings
of the indefinite article in (63) and (64). This difference
can be attributed to the presence of neg; it is therefore evi-
dence against a filtering transformation. In (63), the positive
sentence, a house is interpreted as specific, that is, it is
asserted that there is a particular house which John bought.
In (64), it is not asserted that there is a particular monkey
wrench that John didn't buy: the interpretation of a house is
non-specific. (For more complete discussions of specificity,
see Baker (1966) and Dean (1968).)

In other affective environments, a is interpreted as non-
specific also.

(70) Has John bought a garbage can?
(71) John is reluctant to buy a duck.
(72) John was too lazy to buy a newspaper.

In general, the feature [-X] seems to imply [-specific]. We can
capture this by a semantic redundancy rule. [-specific] does not
imply [-X], however, so the features are not identical: there
are non-specific environments in which an unqualified any is
impossible.

(73) I want to buy a basset-horn. (ambiguous as to
                                    specificity)
(74) I want to buy some fish.
(74) *I want to buy any fish.
(75) I want to buy any fish I can get my hands on.
      (OK with certain kinds of relative clauses added)

According to the interpretive version of the some-any rule,
a may be left unmarked for specificity when it is substituted into
a base tree; the rule provides the feature that permits interpre-

tation of specificity. In a transformational view, in which interpretation is made on deep structures, we must have two separate lexical items a, one specific and one non-specific. Then the incorrect uses will be blocked by the some-any filtering transformation. Klima's some-any rule, which treats the difference between some and any as merely a morphological change, cannot capture the difference in interpretation of a at all. Hence the variation in specificity of a due to negation strongly supports the lexicalist hypothesis, which treats some and any as separate lexical items with different semantic interpretations.

In fact, we can be more explicit about the semantic difference between some and any. Robin Lakoff (1968) presents a number of environments where some and any are both possible, but there is a distinct difference in meaning.

(77) Who wants some beans?
(78) Who wants any beans?

(79) Do you think those men want to do some work?
(80) Do you think those men want to do any work?

(81) I wonder if Bill would lend me some money.
(82) I wonder if Bill would lend me any money.

(83) If you eat some candy, $\left\{ \begin{array}{l} ?\text{I'll whip you.} \\ \text{I'll give you \$10.} \end{array} \right\}$

(84) If you eat any candy, $\left\{ \begin{array}{l} \text{I'll whip you.} \\ ?\text{I'll give you \$10.} \end{array} \right\}$

(85) Unicorns are mythical beasts: if John sees $\left\{ \begin{array}{l} ?\text{some} \\ \text{any} \end{array} \right\}$ unicorns out there, I'll eat my hat.

(86) If John sees $\left\{ \begin{array}{l} \text{some} \\ ?\text{any} \end{array} \right\}$ goldfish in that tank, it's not surprising: there are lots of them in there.

The differences in meaning are correctly explicated by Lakoff: "The beliefs or expectations of the speaker are reflected in his

choice of <u>some</u> or <u>any</u>, and the meaning of the sentence is correspondingly changed." <u>Some</u> appears in questions when the answer is hoped or supposed to be positive, and in conditionals when the expectation of the speaker is that the antecedent will be fulfilled. <u>Any</u> appears when the presupposition is negative or neutral. If we assume the difference in presupposition to be expressed by <u>some</u> and <u>any</u> themselves, this difference in meaning, it seems to me, is sufficient to explain the fact that, for example, (83) is a promise and (84) a threat, at least given the present degree of formality of this kind of semantic discussion.

Lakoff also explains why <u>some</u> cannot (except under very special conditions) occur in negative sentences: if <u>some</u> indicates a positive presupposition, it conflicts semantically with the negation. This is exactly the kind of explanation we are trying to make possible within the theory, by treating the some-any rule as a selection restriction. Again, it may be possible to use a filtering transformation to handle these cases, but such a treatment looks still more arbitrary than the rule to handle the specificity of <u>a</u>.

The decision as to how to handle the some-any rule thus boils down to a choice between the interpretive rule (58) and a filtering transformation. With a filtering transformation, we will have to make readings in the base and then rule them out later if incorrect, whereas with an interpretive rule we supply the reading later on. Since not very many cases are known where

these alternatives exist, I suspend judgment for the present.
In the next section there will be some stronger evidence in
favor of the interpretive rule.

5. The Scope Rule

Up to this point we have assumed that neg is generated
in the base at the position in the tree necessary for the proper
interpretation of scope. Then, we have assumed, Klima's rules
of neg-placement attach the neg to the position at which it
appears in surface structure.

However, since we have decided that the interpretation of
negation depends on derived structure and need not be represented
in the base, another logical possibility presents itself. It is
conceivable that neg is generated in its surface structure con-
stituent and that scope is determined by a semantic rule which
interprets derived structure. Since the some-any rule depends on
knowing the scope of negation, it must refer to the output of
this interpretive rule of scope. Now if the some-any rule is a
filtering transformation, we will have to assume not only that
semantic rules can depend on the output of transformations, but
that transformations can depend on the output of semantic rules.
On the other hand, if the some-any rule is semantic, we will have
to make no such addition to the theory of grammar. Since I know
of no other evidence that such an addition is necessary (though
this may be due to the way we have constructed grammars up until
now), this argues that if the determination of scope is a semantic

rule, then the some-any rule must be semantic as well. Of course, if the placement of neg is done by transformations, then we cannot conclude anything about the nature of the some-any rule.

The evidence for an interpretive rule again centers on the problem of doing morphology by means of transformations. Like the version of the some-any rule which treats the alternation as a morphological change, Klima's rule of neg-placement requires unsystematic and sometimes drastic changes in "spelling":

(87)   neg + any → no (obligatory)
       neg + ever → never (obligatory)
       neg + either → neither (obligatory)
       neg + so → not (optional) (as in I think not)
       neg + many → few (optional)
       neg + much → little (optional)

If we are to be consistent in our adoption of the Extended Lexical Hypothesis, we will want to treat these forms as separate lexical items which can be substituted into base trees. This will require an interpretive scope rule and thus an interpretive some-any rule.

An interpretive rule of scope essentially has to do Klima's rules of neg-placement in reverse. Instead of making neg move to the right and attaching to lower nodes in the tree, we have to generate neg in its surface structure position (as a feature of a lexical item, where appropriate) then move the neg up the tree by the scope rule. Since the scope rule is a semantic rule, it can make no changes in the surface form of the sentence; it rather produces the parts of the interpretations of the sentence having to do with scope.

The positions in which neg may be generated are on NPs, on adverbials, and on the auxiliary. The possible scopes of negation that we have found are S, VP, and NP. From the examples we have seen, and from the examples in Klima, we observe that the surface structure position of neg is always included in the interpreted scope. This suggests that the principle of the scope rule is expansion of the scope of neg to larger and larger constituents. Such expansion can be expressed by a rule that raises neg from the node on which it is generated to a dominating node. Though I do not have enough evidence about the general nature of semantic rules to be able to state such a raising rule precisely, I will discuss a number of examples in order to develop a general idea of how it might work.

(88) I didn't see anything.
(89) I saw nothing.
(90) John never left the house.

The neg can move from the auxiliary (which I will assume is dominated by VP), the object, or an adverbial to be interpreted as S negation. The raising is optional, but preferred. If the neg does not move in (88), the sentence is interpreted as VP negation; in (89), as NP negation. I do not know whether (90) is meaningful as Adv negation.

If the neg is moved up one node at a time, we would expect (89) also to be interpretable as VP negation. My feel for these sentences is not delicate enough to be able to tell whether such a reading is possible. If it is not, this would suggest that neg

is always moved to the node S if it is moved at all.

(91)   Nobody came.
(92)   Somebody didn't come.
(93)   *Anybody didn't come.
(94)   Nobody didn't come.

In (91) the neg (obligatorily?) moves from the subject quantifier to the S node, to be interpreted as S negation. In (92) the movement seems to be inhibited by the presence of the quantifier in the subject, and so only the VP negation reading is possible. Again in (93) the quantifier prohibits movement and the scope of neg must be VP. But then the some-any rule rejects the sentence because any is not within the scope of negation. In (94) the negation in the subject takes the scope S and the negation in the aux is read as VP negation.

The same principle seems to be operating in both (92) and (94) to prevent the neg in the aux from moving up. I conclude from this that quantifiers are subject to the same interpretive rule of raising as neg. Thus the reversal of preferred order of quantifiers under passive (noted many years ago by Chomsky) would be explained the same way as change of order of negation and quantifiers under passive.

Suppose then that the raising rule only raises the leftmost neg or quantifier (or perhaps the highest on the tree). Then in (92)-(94) the logical element in the subject gets raised to sentence scope, but the neg in the auxiliary must remain where generated.

Alternatively, suppose that all logical elements can be

moved up, one node at a time, but that preferably the leftmost one is moved first and that only one logical element (or a combination like <u>not many</u> generated at a single node) may occupy any given scope in the interpretation. Then when the element in the subject is raised, this precludes raising the neg in the aux. The evidence that would decide between these two alternatives seems to hinge on the (again) delicate distinction between object and VP negation in the interpretation of (95).

    (95) Nobody did nothing.

If a VP negation reading can be ferreted out of (95), originating in the object neg, this argues for the solution in which all logical elements may move up one node at a time but only one can occupy a given scope. If no such reading can be found, this permits the simpler solution where only one element moves, preferably the leftmost, and it always moves to S.

    (96) ?Nobody didn't do nothing.

    Assume (96) is not in the dialect where everything becomes negative within the scope of negation. Then there are three separate scopes of negation: S, arising from the subject, VP arising from the aux, and NP, from the object. The difficulty of this sentence, I suggest, is due to performance. One neg in a sentence is easy to interpret, two negs are more difficult, but three negs are almost impossible. This is exactly comparable to the classic case of a performance constraint, center-embedded relative clauses, where the difficulty is directly proportional to the number of embeddings, becoming virtually impossible by the time

three embeddings are reached.

(97)   ?Nobody didn't see anyone.
(98)   ?Nobody didn't see someone.

The overlapping scopes of negation seem to make both of these strange. They both mean to me "Everybody saw (at least) someone." It is not clear whether the two negs should cancel out within the overlapping part of the scope, VP, so as to allow (98) by the some-any rule, or whether they sould reinforce each other, so that the some-any rule allows (97) instead. And this unclearness seems to be reflected in performance.

(99)   That John didn't go bothered $\begin{cases} \text{someone.} \\ \text{*anyone.} \end{cases}$

(100)   I will force you to marry no one.
(101)   I will force you not to marry anyone.

If it were possible to use _anyone_ in (99), this would show that there is a reading in which neg comes out of the complement and takes as its scope the main clause. The fact that _anyone_ is impossible shows that the scope rule cannot move neg out of a full S.[6] On the other hand, (100) is ambiguous: there is a reading where the scope of neg is the main clause. The difference between (99) and (100) may be due to pruning[7] of the S node after equi-NP deletion in (100). Because of this pruning, the S to which the neg moves, if it moves at all, is the S of the main clause. (Recall that the scope rule operates on _derived_ structure, so pruning is relevant.)

The scope of negation in (100), however, is unambiguously the subordinate clause (_for you_) _to marry anyone_. The question arises as to where _not_ is attached in this sentence. I suggest

the following solution, for which there is additional support
in Jackendoff (1968a). Suppose that the particular neg which
ends up in the aux in the surface form is generated as a daughter
of S. Then, by the first and third rules of neg placement in
Klima's system (which we will continue to treat as syntactic
transformations), neg will be obligatorily moved into the aux
just in case the aux contains Tense. Otherwise neg will remain
as a daughter of S and be realized as _not_. In (101), Tense is
absent because of the complementizer _for-to_. Therefore _not_ re-
mains a daughter of S. Since the S of the subordinate clause
dominates neg as well as NP and VP, it will not prune when NP
is erased by equi-NP deletion. Because the S remains at the time
of the scope rule, neg cannot move up into the main clause.

Again, there is an alternate solution which will do the
job but is not quite as nice. We might simply not allow the _not_
in (101) to move up, due to some condition on the complementizer
and the position of _not_. However, this makes it a little more
difficult to explain the fact that neg can move all the way up
in (100). In particular, if movement is one step at a time, the
neg from _no one_ in (100) would at one point attain the position
of _not_ in (101) and could go no farther. Thus the solution seems
more plausible in which the movement, if there is any at all, is
to the next higher S.

(102) John didn't see anyone, and neither did Bill.
(103) None of the first-graders finished, and
       ⎰neither did any of the second-graders.
       ⎱none of the second-graders did, either.

(104)   John didn't see anyone, not even Bill.
(105)   None of the boys came, not even Bill.
(106)   *Some of the boys didn't see anyone, and neither
        did Mary.
(107)   *Many of the boys didn't see anyone, not even Bill.

The either/neither tags and the not even tags discussed
by Klima can be explained by saying that these tags require an
interpretation of sentence negation on the main sentence in order
to be grammatical. Thus in (102)-(105) the main clause can be
read as sentence negation, and so the tag is acceptable. In
(106) and (107) the quantifier in the subject prohibits a reading
of sentence negation, and so the tags are unacceptable.

(108)   Never before had none of his friends come to one
        of his parties.
(109)   None of his friends had never come to one of his
        parties before.
(110)   Never before had any of his friends not come to
        one of his parties.
(111)   Never before hadn't any of his friends come to
        one of his parties.

These examples show the ordering of the scope rule with
respect to certain transformations. (108) means "for the first
time, none of his friends came." Never before is associated
with S negation. Thus the neg in the subject does not move up
but remains as NP negation. On the other hand, in (109), which
means "all of his friends had at some time been to one of his
parties," S negation originates in the subject, and the neg of
never either remains as adverb negation or raises to become VP
negation--it is difficult to judge. These examples show that
the scope rule must take place after adverb preposing.

(110) means "for the first time, at least one of his friends

didn't show up." Again, S negation originates in never. Because we have chosen not to front neg along with the aux, it takes VP as its scope. (111) I take to be synonymous with (108). The fronting of neg along with the aux allows the subject to be included in its scope. Or, possibly, only the aux is negated. This pair of sentences shows that the scope rule must take place after subject-aux inversion.

(111), if it is grammatical (some people find it rather strange), and if the second neg includes the subject in its scope, raises some interesting questions about the derived structure after adverb preposing and subject-aux inversion. It suggests the possibility that scope has only to do with derived strings, and not with any structure other than the S node (which we have shown to be relevant in cases like (99)-(101). In this case the scope rule would involve moving neg to the left instead of up the tree, and the scope of negation would include everything to the right of neg within its clause instead of everything in construction with it. The alternatives seem equivalent in terms of all the examples so far except (111), where the evidence hinges on a subtlety of interpretation I am not prepared to make, namely, whether the subject is included in the scope of the neg of hadn't. The fact that there is no confusion in the use of any, as there is in (97) and (98), which have overlapping scopes, suggests that any in (111) is not in the scope of hadn't. But I don't want to commit myself further.

(112) Many of his friends never came.
(113) Never did many of his friends come.
(114) None of the boys kissed none of the girls.
(115) None of the girls were kissed by none of the boys.

These examples bring out the similarity of quantifiers and negation with respect to the scope rule. (112) and (113) are just like (108) and (109) with a quantifier instead of the second negation. Their difference in meaning shows that the interpreted scope of quantifiers is determined after adverb preposing. (114) and (115) show that passive affects interpreted order of multiple negations as well as order of quantifiers and negation, since these two sentences differ in meaning.

Finally, let us consider (108)-(111) a little more. They are extremely difficult to handle in a theory which requires that all meaning be represented in deep structure and that transformations do not change meaning. (108) will have to be generated with neg on S and on the subject. No surface form can be produced unless _ever_ is preposed in order to absorb the S neg, since the subject is already burdened with a neg and cannot absorb another. (109) will be generated with S negation and VP negation, then the subject absorbs the S neg and _ever_ absorbs the VP negation. If, however, _ever_ fronts, it will absorb the S neg, and we would predict that the VP neg would show up as in (110). However, (110) is not synonymous with (109), as we would then expect. The difference has to do with the order of the quantifiers contained in _ever_ and _any_ interacting with the two negs. In addition, one would expect (111) to be synonymous with (110), since the only

difference between them is the optional attachment of neg to
had before subject-aux inversion. Unfortunately, again a differ-
ence in meaning is produced because of the reordering of the neg
with respect to the quantifier _any_. Thus the three sentences
(109), (110), and (111) cannot be conveniently distinguished in
deep structure, if we use a system of rules like Klima's. Be-
cause of the lateness and optionality of the adverb preposing,
neg-attachment, and subject-aux inversion transformations, it is
hard to see how a theory of quantifiers and neg as higher sen-
tences would be of any help. On the other hand, a theory in
which the interpretation of negation and quantifiers is based on
derived structure can explain these differences quite naturally.

The clear superiority of the interpretive rule in account-
ing for multiple negation, and the generalization of quantifiers
and negation achieved by the rule argue strongly against the
first alternative presented in this section, namely generating
neg at each position in the tree where it is interpreted as having
scope. As we pointed out at the beginning of this section, the
interpretive version of the some-any rule is thus preferable to
a version of the some-any rule as a filtering transformation. The
scope rule will take as its input some fairly late level of de-
rived structure, perhaps even surface structure. Its output will
then go into the some-any rule. These rules will make no change
in the surface form, but will rather function as part of the
semantic component, providing partial interpretations of sentences.

The examples I have presented favor this interpretive theory of negation over a theory consistent with the hypothesis that transformations do not change meaning. Therefore it would appear necessary to give up this hypothesis.

NOTES

1.   This chapter has been published in substantially the same form as Jackendoff (1969).

2.   Ross's argument for command is based on sentences like That anyone at all came surprised Bill, in which the anyone is apparently due to surprise, even though anyone is not in construction with surprise. It is, however, commanded by surprise. Ross assumes that the usual some-any rule is operating here. If this were the case, we would expect also *That Bill came surprised anyone, just as we get any with Affective verbs such as deny and defy: I deny anyone the privilege of playing in tune; I defy anyone to contradict my grandmother. Thus apparently Ross's sentence involves a different rule, and so has no bearing on the issue.

3.   Some people seem to be able to get an ambiguity in (35). For me, this is not the case. Similar sentences involving numerals, such as The target wasn't hit by two of the arrows, manifest the second reading for me only if there is emphatic stress on the numeral. This interpretation can probably be made to follow from the rule of Neg Attraction to Focus, which will be discussed in Chapter 5.

4.   For other criticisms of this position, see Jackendoff (1968a, b).

5. At the 1969 La Jolla Syntax Conference, David Lewis (UCLA) proposed a categorial grammar for the base component and described how one might define intensions for various categories. The grammar he presented contained <u>quantifier</u> as a category that, when combined with a noun, yields a noun phrase (proper name). He showed that the intension of such a category could be defined reasonably. In the terminology of the present investigation, he has shown it possible to talk about the meaning of a quantifier whose scope is other than a proposition. Note that it is necessary to have additional principles in one's logic to decide in what order to evaluate multiple quantifiers (i.e. rules like those proposed here for the grammar) just in case one allows quantifiers to range over other than propositions, as in the system proposed by Lewis.

6. The feature <u>factive</u> (discussed in Kiparsky (1969)) might be imagined to be relevant here, but it turns out not to be: *<u>That John won't go seems likely to anyone</u> has a non-factive main verb, contrasted with the factive in (89).

7. For a discussion of pruning, see Ross (1967a), Chapter 3.

ADVERBS

1. The assumptions

In the study of generative grammar, perhaps the least studied and most maligned part of speech has been the adverb. In fact, most studies do not concede to adverbs the right to be a part of speech at all. Before the introduction of the semantic component, this tendency might have been understandable. At that time it was thought necessary to state co-occurrence restrictions at the syntactic level. To make co-occurrence relations general, it was important to reduce the number of deep grammatical relations in which a category took part. Adjectives submit fairly docilely to this reductionist tendency, since there is almost always a paraphrase for an Adj-N construction with a relative clause. Adverbs are more unruly, since constructions they occur in are less homogeneous, and since their paraphrase relations are much more widely varied as well.

In addition, there is another tendency that has hampered the study of adverbs. The lack of feature mechanisms forced early generative grammarians to subdivide major categories into a multitude of classes. For example, in Lees's The Grammar of English Nominalizations, verbs are divided by phrase structure rules into predicative verbs, activity verbs, transitive verbs,

intransitive verbs, and middle verbs; these classes are further
subdivided. In such a theory it is only natural to subdivide
adverbs into manner adverbs, locative, time adverbs, means ad-
verbs, degree adverbs, and so forth. However, the terminology
for adverbs has not kept up with the times; these categories
are still usually treated as entirely distinct in the phrase
structure.

The result of these two tendencies has been the ready
acceptance of quite different underlying structures for similar
adverbials, along with the implicit claim that <u>adverb</u> is only
a surface structure notion with no relevance to "deep" syntax.
In recent papers by Lakoff (1965,1967), Ruwet (1968), and Kuroda
(1968), deep-surface pairs as disparate as the following have
been proposed.

(1) Seymour used a knife to slice the salami $\Rightarrow$
Seymour sliced the salami with a knife. (Lakoff)

(2) It is certain that the <u>Pueblo</u> entered the terri-
torial waters of N. Korea $\Rightarrow$
The <u>Pueblo</u> certainly entered the territorial waters
of N. Korea. (Ruwet)

(3) John is careless at driving his car $\Rightarrow$
John drives his car carelessly. (Lakoff)

(4) The manner in which John disappeared was elegant $\Rightarrow$
John disappeared elegantly. (Kuroda)

(Cf. § 4.3 of Katz and Postal (1964) for several more examples.)
In addition, there are some adverbs for which no paraphrase
with an adjective seems appropriate.

(5) Albert is $\begin{Bmatrix} \text{truly} \\ \text{simply} \\ \text{merely} \end{Bmatrix}$ a fool.

It is $\begin{Bmatrix} \text{true(non-paraphrase)} \\ \text{*simple} \\ \text{*mere} \end{Bmatrix}$ that Albert is a fool.

In fact, for _merely_ it is obvious that no paraphrase involving a copula can work, since _mere_ never occurs in a copula.

An attempt will be made here to begin to give a somewhat more unified account of -_ly_ adverbs and some other adverbials that are not prepositional phrases. Hopefully the account can be extended to include prepositional phrases acting as adverbials.

First of all, given that lexical items have semantic interpretations, we can abandon the division of adverbs into multifarious syntactic categories. Adjectives are not divided into adjectives of color, size, quality, degree, frequency, and so forth. Rather it is taken for granted that their semantic representations will automatically account for these properties. We will assume that the same is true of adverbs. For example, the ungrammaticality of *John knew the answer terribly will require no special consideration in the syntax, any more than *the colorless idea does: both will be considered as simple violations of selectional restrictions.

How will these selectional restrictions be captured? A common argument is that adverbs must be reduced to adjectives so that shared selectional restrictions need not be stated twice. However, given the very plausible assumption that we can capture the notion "separate but related lexical items" by derivation rules in the lexicon, as was outlined in Chapter 1, §9 this argument does not hold. Rather the similarity of selectional restrictions will follow from the derivation rules that relate adjectives to -_ly_ adverbs. An -_ly_ adverb will be less

costly in the lexicon if there is a related adjective.

If, on the other hand, we accept the commonly held view that adverbs are derived from adjective paraphrases, the semi-productivity of the paraphrases will require a large number of different transformations and a liberal use of exception features; for each paraphrase relation such as (1)-(4) there will be a transformation and a rule feature. According to such an analysis, it is accidental that all the adverbs wind up in essentially the same positions in the sentence, and that -ly is the ending added to all adjectives to form adverbs. Why don't some sentence adverbs turn up between the direct and indirect objects, and why don't some adverbs end in, say, -guck instead of -ly? These generalizations are not expressed in the transformational approach.

We will assume therefore that the base mentions a category Adv. There will be no structural indication such as Adv Manner-- it is a matter of semantics whether something is an adverb of manner or some other kind. We will also assume that the existence or non-existence of a paraphrase with an adjective construction is more or less fortuitous. However, in case there is a paraphrase, we will assume that it indicates a lexical relationship and that the semantic structure of the paraphrase can tell us something about the semantic structure associated with the adverb.

2. Some data

There are three basic surface positions in a sentence in which an -ly adverb can occur: initial position, final position,

and as part of the auxiliary.[1] Various classes of adverbs can occupy different combination of these three.

One class can occupy all three positions, but changes meaning according to position, for example, in (6)-(8).

(6) John $\left\{\begin{array}{l}\text{cleverly}\\\text{clumsily}\end{array}\right\}$ dropped his cup of coffee.

(7) $\left.\begin{array}{l}\text{Cleverly}\\\text{Clumsily}\end{array}\right\}$ (,) John dropped his cup of coffee.

(8) John dropped his cup of coffee $\left\{\begin{array}{l}\text{cleverly}\\\text{clumsily}\end{array}\right\}$.

(6) is ambiguous, meaning either (7) or (8). These latter two are approximate paraphrases of (9) and (10) respectively.

(9) It was $\left\{\begin{array}{l}\text{clever}\\\text{clumsy}\end{array}\right\}$ of John to drop his cup of coffee.

(10) The manner in which John dropped his cup of coffee
      was $\left\{\begin{array}{l}\text{clever}\\\text{clumsy}\end{array}\right\}$.

Some other adverbs that behave this way are carefully, carelessly, happily, truthfully, specifically, and frankly.

For other adverbs which can occupy all three positions, there is no discernible change in meaning, for example, quickly, slowly, reluctantly, sadly, quietly, indolently, frequently, immediately, often, soon.

There are -ly adverbs which can only occur in initial and auxiliary position ((11)-(13)); they typically have paraphrases like (14).

(11) $\left.\begin{array}{l}\text{Evidently}\\\text{Probably}\end{array}\right\}$ Horatio has lost his mind.

(12) Horatio has $\left\{\begin{array}{l}\text{evidently}\\\text{probably}\end{array}\right\}$ lost his mind.

(13) *Horatio has lost his mind $\left\{\begin{array}{l}\text{evidently}\\\text{probably}\end{array}\right\}$.

(14) It is $\left\{\begin{array}{l}\text{evident}\\\text{probable}\end{array}\right\}$ that Horatio has lost his mind.

Often these adverbs will be acceptable in final position if separated from the rest of the sentence by a pause and accompanied with a drop in pitch, thus:

(15)   Horatio has lost his mind, $\begin{Bmatrix} \text{evidently} \\ \text{probably} \end{Bmatrix}$.

Some other members of this class are unbelievably, certainly, understandably, unfortunately, and apparently.

A fourth class can only occur in auxiliary and final position.

(16)   *$\begin{Bmatrix} \text{Completely} \\ \text{Easily} \end{Bmatrix}$ Stanley ate his Wheaties.

(17)   Stanley $\begin{Bmatrix} \text{completely} \\ \text{easily} \end{Bmatrix}$ ate his Wheaties.

(18)   Stanley ate his Wheaties $\begin{Bmatrix} \text{completely} \\ \text{easily} \end{Bmatrix}$.

Consistent paraphrase relationships for these adverbs with adjective constructions are not evident to me. Easily has the paraphrase (19), but completely has only the lame (20).

(19)   It was easy for Stanley to eat his Wheaties.

(20)   ??Stanley's eating (of) his Wheaties was complete.
       ??The degree to which Stanley ate his Wheaties was complete.

Other adverbs that have the paradigm (16)-(18) are purposefully, totally, altogether, handily, badly, mortally, tremendously.

Some adverbs, typically non -ly adverbs, only occur in final position.

(21)   *Hard John hit Bill.
       *Well Sam did his work.

(22)   *John hard hit Bill.
       *Sam well did his work.

(23)   John hit Bill hard.
       Sam did his work well.

Others like this are more, less, before, early, fast, home, slow,

terribly, lengthwise, indoors, and downstairs.

Finally, there is the class exemplified in (5), including at least merely, utterly, virtually, hardly, truly, and simply. This class can occur only in auxiliary position, and no adjective paraphrase is appropriate.

3. Some syntax

How can we distinguish all these classes of adverbs in a motivated way? The only attempt I know of to come to grips with this problem is Keyser (1968), which we will discuss in the next section. All other work I know of in generative grammar (references in §1) is concerned with finding ways of reducing small classes of adverbs to some other construction in the base, and no attention is paid to the problem of locating the appropriate positions for different kinds of adverbs, or to a general approach to -ly adverbs.

What we have to decide about the syntax of adverbs is which positions are generated in the base, which positions are transformationally derived, and what the transformations look like.

Let us start with auxiliary position. There is fairly good evidence that this should be an underlying position for -ly adverbs. All of the -ly adverbs can occur there, and there are some that can only occur there. But there is a more interesting reason than this.

Chomsky (1969) develops a generalized set of base rules that accounts for the fact that nouns and verbs take similar com-

plement structures. This, combined with a lexicon incorporating derivation rules, enables us to express the fact that verbs and their nominalizations behave alike. We are using similar derivation rules in the lexicon to relate adjectives and adverbs morphologically and semantically. Can we use the base rule schemata to relate them syntactically?

The answer is yes, if we are willing to give up the long-standing assumption that all adjectives in noun phrases are derived from reduced relative clauses. We will justify this move in a moment. But first let us see what there is to be gained.

It seems to me no accident that the surface position of adjectives in noun phrases is between the article and the head, a position exactly parallel to auxiliary position of adverbs in sentences. In particular, the parallelism between adjectives in derived nominals and adverbs in gerunds is striking.

(24) John's rapid reading of the letter
(25) John's rapidly reading the letter

In a grammar that derives adjectives and adverbs in these constructions from far-flung sources, there is no way to express this correspondence. However, a base rule schema which generates these positions directly could capture the generalization. Assume X is the set of features common to nouns and verbs, and that the feature $[\pm \text{Verb}]$ distinguishes nouns from verbs. Also, assume Y is the set of features common to adjectives and adverbs, and that the feature $[\pm \text{Adverb}]$ distinguishes them. Then the relevant base rule will look in part like (26).[2]

$$(26) \quad \overline{\begin{bmatrix} X \\ \alpha \text{Verb} \end{bmatrix}} \rightarrow \begin{bmatrix} Y \\ \alpha \text{Adverb} \end{bmatrix} - X - \text{Complement}$$
$$(\text{where } \overline{V} = VP)$$

Therefore, if we want to generate -ly adverbs at some position
in the base, the choice of auxiliary position lets us get adjec-
tives in at no extra cost to the syntax. Incidentally, this rule
will automatically generate the adjectives like mere and utter
which cannot occur in the copula, along with their adverbial
counterparts merely and utterly.

To get a reading for a base-generated adjective-noun
construction, we will need a projection rule[3] of attribution, more
or less like the one Katz and Fodor discuss for this construction.
Will this projection rule be needed for anything else? It is
likely that it generalizes, exactly as the base rule (26) does,
to handle the interpretation of adverbs in VPs as well (cf. §5),
so it would cost no extra. This is an intriguing proposal, but
I will not test it here. '

But more important is the question of what the projection
rule is that determines the meaning of relative clauses in noun
phrases. A relative clause can be thought of as a way of expres-
sing a complex property that cannot be expressed in a single
word. From this point of view we see that the projection rule
for relative clauses must in fact be the same projection rule of
attribution. In this case, extending the rule to adjectives costs
us virtually nothing in the grammar.

One thing we have to contend with in this account of ad-
jectives is that there are now two structural configurations with

identical co-occurrence relations: <u>Det-N is Adj</u> and <u>Det-Adj-N</u>.
The desire to reduce these to a single co-occurrence relation
is the primary reason for deriving <u>Adj</u> constructions from
<u>N-Relative Clause</u> constructions. Thus we seem to be playing
off a generalization in syntactic position of adjectives against
a corresponding loss of generality in co-occurrence relations.

A conceivable way out, though at present unformalizable,
might be to argue that the projection rules for these two con-
structions create very similar semantic readings. The relation
between the noun and the adjective is the same in both cases,
that is, the property denoted by the adjective is attributed to
the set denoted by the noun. But in the case of the copula, the
resulting expression is a proposition; in the case of $\bar{N}$, the re-
sult is a more restricted set. The difference is intuitively
illustrated by the contrast of (27) and (28).

(27) <u>N</u> has the property <u>Adj</u> (copula)
(28) <u>N</u> having the property <u>Adj</u> ($\bar{N}$)

That is, the only difference between the two projection rules
is some sort of formal operator. If this is the case, the simi-
larity of co-occurrence relations is more understandable, if we
maintain that co-occurrence relations are stated over readings.

To fully explore the consequences of the base rule (26),
one would have to go over all the data handled by the standard
theory of adjectives. In particular, it would be interesting
to see if the transformation deleting <u>wh- be</u> could be eliminated.
I will not follow up this possibility here. In Jackendoff (1968c)
I have pointed out some parallels between various kinds of rela-

tive clauses and sentence adverbials which may be pertinent to
the problem of adjectives and adverbs in the base rule schema.

To return to the main problem of this section, we have
given at least the main lines of an argument to show that adverbs
should be generated directly in auxiliary position by the base,
and that this will not add anything to the grammar that is not
independently motivated.

What about other positions? Initial position has no paral-
lels in adjective constructions. Furthermore, adverbs cannot
occupy initial position in subordinate clauses:

(29) ?George says that evidently Bob has disappeared.
(30) *For apparently Bob to be sick would worry Harriet.
(31) *Charley was scared by stupidly Violet's driving
      the car off the cliff.
(32) *Did you see that man who quickly Izzy ran away
      from?
(33) ?I won't come because probably my mother is sick.

Therefore it is feasible to move adverbs to initial position by
a transformation that optionally fronts them (presumably from
auxiliary position) in main clauses.

In final position, there is the problem of the non-_ly_
adverbs, such as _hard_, _before_, and _early_, which only occur there.
Klima (1965) analyzes these adverbs as intransitive prepositions,
which can be generated by the base in the same positions as
normal prepositional phrases. His evidence is essentially that
they often substitute semantically for prepositional phrases,
and that in addition, many of them can act as normal prepositions,
as shown in the following examples.

(34)

Johnny ran $\begin{cases} \text{home} \\ \text{downstairs} \\ \text{into the house} \\ \text{down the stairs} \end{cases}$

(35)

Tommy didn't do amusing things like that $\begin{cases} \text{alone} \\ \text{before} \\ \text{afterwards} \\ \text{after that} \\ \text{before his} \\ \quad \text{12th birthday} \\ \text{for many years} \\ \text{by himself} \end{cases}$

He suggests, therefore, that the base rule for prepositional phrases, like the base rules for noun phrases and verb phrases, contains an _optional_ NP following the head:

(36) PP → P - (NP)

This way, at no extra cost we can get these "adverbs" generated in the base, and use the same projection rules to account for their semantic relationships to the sentence as we use for prepositional phrases. It is not clear to me, however, what adjectives such as _hard_ and _fast_ are doing in this category.

For the -_ly_ adverbs in final position, there are now two solutions possible. We can either extend the base rule for prepositional phrases to include -_ly_ adverbs, or we can add a transformation to move them to the end from auxiliary position. I don't see any particular reason to prefer one analysis over the other, except that the transformational solution may generalize to adjective constructions like _a book yellow with age_.

To account for the three adverb positions, we therefore need the base rules (26) and (36), which cost us nothing, plus a preposing rule and either a postposing rule or an addition to

the complement of VP in (26). Now how do we prevent all -ly

adverbs from occupying all three positions? The worst possible

solution would be to use exception features to restrict the

applicability of the preposing and postposing rules. In §5 we

will see that the adoption of a derived structure semantic rule

permits a motivated division of the -ly adverbs into classes

based on their ranges of surface position.

4. Transportability--an aside

Actually, we have been somewhat oversimplifying the

account of -ly adverb positions up to this point. -Ly adverbs

can occur in the VP in positions other than at the end, and they

can occur between any two elements of the auxiliary.

    (37) John will send the money immediately back to
           the girl.
    (38) John immediately will send the money back to
           the girl.
    (39) John will have probably sent the money back to
           the girl.
    (40) John will have probably been getting upset by now.

Also, we have not yet accounted for final position with pause,

as in (15). It would not be nice if we had to state a special

transformation for each of these positions. Keyser (1968) pro-

poses an addition to the theory of grammar, called the Trans-

portability Convention, designed to handle these cases, in the

hope that it will prove useful generally in the description of

free word order phenomena.

Keyser observes that the positions in which adverbs occur

correspond to major syntatic breaks in the derived structure.

The transportability convention expresses this by permitting a

constituent marked specially, say [+transportable], to occupy
any position in a derived tree so long as the sister relation-
ships with all other nodes in the tree are maintained, that is,
as long as it is dominated by the same node. In English, -ly
adverbs are transportable; in other languages, such as Latin,
hopefully more different kinds of nodes will be characterized
as transportable.

By the transportability convention, we would expect ad-
verbs dominated by S to occur initially, before the auxiliary,
and finally: these are the three possible sister positions to
the subject and the VP. All three of these positions actually
occur. Presumably final position dominated by S is the position
with a pause seen in (15); at least this is consistent with the
analysis presented in the next section.

Adverbs dominated by VP should occur before the verb,
finally (this time without pause), and at various places in
between. Also, under the fairly reasonable assumption that the
aspectual markers have and be are daughters of the VP, transpor-
tability will account for (38)-(40). The one place where adverbs
sound particularly bad is between the verb and the following NP.

(41)   *?John sent immediately the money back to the
            girl.

To take care of this, Keyser appeals to a "surface-structure
tendency to prevent anything from intervening between a Verb and
the following Noun Phrase" (p.371). This constraint is only
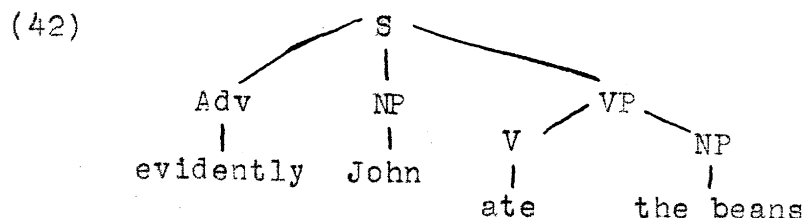scantily justified, but then it is not unreasonable either.

Keyser does not discuss the formal nature of the trans-
portability convention, that is, whether it is a kind of
transformation or a condition on the base or possibly something
else. He does establish that it must operate before at least
one transformation, particle movement. If transportability were
stated as a new kind of transformation that said essentially
"transport transportable nodes," there is the possibility that
we could order it after adverb preposing. This way we could get
all positions with the two rules of preposing and transportability.
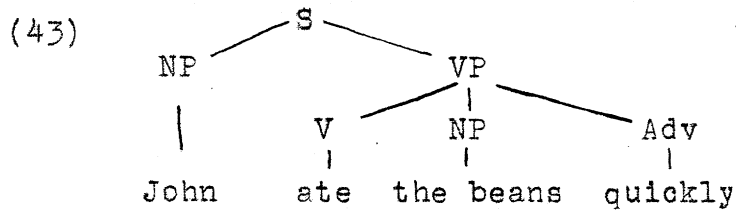
Nothing very crucial in my analysis of adverbs hangs on
the viability of Keyser's proposal. However, it is useful in
that it makes various things more convenient to state, so it
deserves mention here.

5. The Adverb Scope Rule

Now we turn to the problem of motivating the distributions
of the different classes of -ly adverbs. The approach will be
based on the rules of semantic interpretation needed for adverbs.

The derived structures for adverbs in initial and final
position are (42) and (43) respectively, according to fairly
conservative guesses.[4]

(42)

```
                         S
            ┌────────────┼────────────┐
          Adv           NP           VP
           │             │        ┌───┴───┐
       evidently       John       V      NP
                                  │        │
                                 ate    the beans
```

(43)

```
                    S
       NP         ╱    VP
        |        V    NP    Adv
        |        |     |     |
       John     ate the beans quickly
```
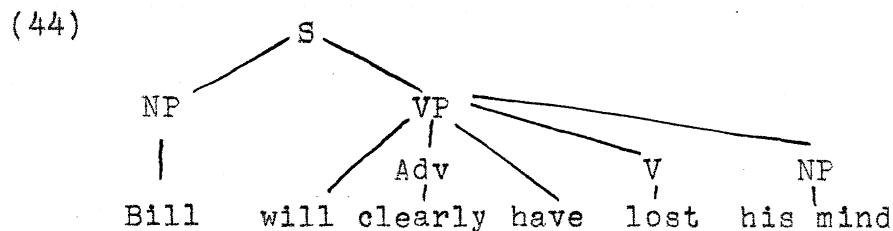
The structural difference between initial and final positions is
that in initial position, the adverb is a daughter of S, while
in final position, it is a daughter of VP.

Suppose then that adverbs are divided into S adverbs and
VP adverbs. Since the domination relations involved in S adverbs
and VP adverbs are different, there must be a separate semantic
projection rule associated with each of the structures (where I
use the term "projection rule" here to mean rule that combines
the readings of daughter nodes to form a reading for the mother
node, as in a Katz-Fodor P1 rule). Thus the classification of
adverbs into S and VP adverbs could conceivably be stated either
in terms of generation in the base, or in terms of which semantic
projection rule the adverb must undergo.

What about auxiliary position? Suppose the derived struc-
ture of adverbs in auxiliary position is something like (44).

(44)

```
                    S
       NP         ╱    VP
        |          Adv      V     NP
        |          |        |      |
       Bill  will clearly have  lost his mind
```

As in §4, I am assuming that the Aux is part of the VP in derived
structure, as indicated by the major constituent breaks. I am
also assuming, for the sake of simplicity, that <u>Aux</u> is just an
abbreviation for a certain string of elements preceding the verb

rather than a real node; thus the modal and aspect are actually

daughters of VP. Under these assumptions the adverb is also a

daughter of VP, so the projection rule applying to the configura-

tion $\begin{smallmatrix} VP \\ | \\ Adv \end{smallmatrix}$ will normally apply. This accounts for one of the

interpretations of the auxiliary adverb.

To account for the other interpretation, that of an S

adverb, I will propose a rule of semantic interpretation called

the Adverb Scope Rule. This rule will optionally change (seman-

tically) adverbs in aux position into sentence adverbs. This

rule will be similar to the scope rule for negation, which

changes the scope of a neg from a constituent to S negation.

Unlike the Neg Scope Rule, the Adverb Scope Rule operates only

on adverbs in aux position, so that they and not final adverbs

can be interpreted as S adverbs. If this rule is correct, it

implies that the division of adverbs into S adverbs and VP ad-

verbs is a semantic one, based on which projection rule the

adverb can undergo.

Let us work through the examples of §2, using the Adverb

Scope Rule, the S adverb projection rule (P_S), and the VP adverb

projection rule (P_{VP}). The adverbs in (11)-(13) (_evidently_,

probably, etc.) can only be S adverbs. If in initial position,

they can be so interpreted by P_S. If in final position, only

P_{VP} applies, so the reading is anomalous. In auxiliary position,

if the Scope Rule does not apply, P_{VP} is applicable, and we again

get a bad reading. If however the Scope Rule does apply, moving

the adverb (semantically) to S scope, P_S applies and a good

reading results. In final position with a pause (15), if the
derived structure has the adverb as a daughter of S (as I sup-
posed in §4), P_s applies to give a good reading.

The fact that these adverbs are S adverbs accounts for
the paraphrase many of them have with constructions of the form
it is Adj that S. Since an S adverb "modifies" (or perhaps
better, "modulates") the reading of an entire S, the reading
has the same sort of functional structure as an adjective pre-
dicated over the S.

The adverbs in (16)-(18) (completely, easily, etc.) are
VP adverbs only. In initial position only P_s applies, so a bad
reading results. In final position, P_{vp} is correctly applicable.
In aux position, we get an anomalous reading if the Scope Rule
applies, since then P_s is appropriate. But if the Scope Rule
does not apply, P_{vp} can apply to give a good reading.

The adverbs in (6)-(8) (cleverly, clumsily, etc.) can
function as either S or VP adverbs. When there is a change in
meaning, we can attribute it to some slightly different specifi-
cation for the two positions in the dictionary entry. The
ambiguity in aux position is of course due to the optional
operation of the Scope Rule.

The non-ly adverbs, which we have analyzed as intransitive
prepositions, are generally only VP adverbs. Some, such as the
locatives (inside, here) and time adverbials (then, afterwards)
are indifferent as to whether they are S or VP adverbs, and they
can prepose optionally, presumably by the same transformation that

fronts the corresponding prepositional phrases (<u>in the house</u>, <u>at 6:00</u>, etc.).

We have not yet accounted for the adverbs which only occur in auxiliary position (5). From what has been said so far, it should be the case that they are either S or VP adverbs and therefore capable of occupying in addition either initial or final position. We will return to these adverbs in §7.

There are some interesting complications in the Scope Rule when the auxiliary contains more than just Tense. Before aspect, a modal, or even emphatic <u>do</u>, only S adverbs are possible.

(45)

George $\begin{Bmatrix} \text{probably} \\ \text{*completely} \end{Bmatrix}$ $\begin{Bmatrix} \text{has read the book.} \\ \text{is finishing his carrots.} \\ \text{was ruined by the tornado.} \\ \text{will lose his mind.} \\ \underline{\text{did}} \text{ eat up the cabbage.} \end{Bmatrix}$

After any one of these auxiliaries alone, either kind of adverb is all right.

(46)

George $\begin{Bmatrix} \text{has} \\ \text{is} \\ \text{was} \\ \text{will} \\ \text{did} \end{Bmatrix}$ $\begin{Bmatrix} \text{probably} \\ \text{completely} \end{Bmatrix}$ $\begin{Bmatrix} \text{read the book.} \\ \text{finishing his carrots.} \\ \text{ruined by the tornado.} \\ \text{lose his mind.} \\ \text{?eat up the cabbage.} \end{Bmatrix}$

If there are two auxiliaries, and the adverb is between them, the S adverb is highly preferred.

(47)

George $\begin{Bmatrix} \text{will} \\ \\ \text{has} \\ \\ \text{is} \end{Bmatrix}$ $\begin{Bmatrix} \text{probably} \\ \text{?completely} \end{Bmatrix}$ $\begin{Bmatrix} \text{have read the book.} \\ \text{be finishing his carrots} \\ \quad \text{by now.} \\ \text{be ruined by the tornado.} \end{Bmatrix}$
$\begin{Bmatrix} \text{been finishing his carrots.} \\ \text{been ruined by the tornado.} \end{Bmatrix}$
being ruined by the tornado.

If the adverb is after two auxiliaries, only the VP adverb is
possible.

(48)

$$\text{George} \begin{cases} \text{will have} \\ \text{will be} \\ \text{will be} \\ \text{has been} \\ \text{has been} \\ \text{is being} \end{cases} \begin{cases} \text{*probably} \\ \text{completely} \end{cases} \begin{array}{l} \text{read the book.} \\ \text{finishing his carrots.} \\ \text{ruined by the tornado.} \\ \text{finishing his carrots.} \\ \text{ruined by the tornado.} \\ \text{ruined by the tornado.} \end{array}$$

I can see no principled semantic factor on which these
differences can be based, particularly since the results are in-
dependent of the auxiliary elements chosen. It seems likely,
therefore, that these paradigms are due to the way the structural
description of the Scope Rule is stated. Here is one way to
state the rule, which makes the proper predictions:

(49) (Adverb Scope Rule)
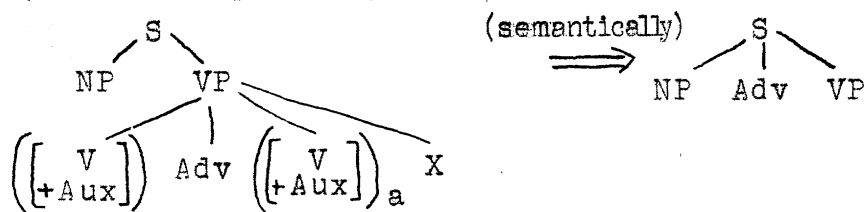
Conditions: If <u>a</u> present, obligatory
otherwise optional

There may of course be notational conventions that permit an
easier statement of the rule.

Since the rule may depend on the presence of the <u>surface
structure</u> auxiliary elements <u>be-en</u> and emphatic <u>do</u>, the scope
rule must be a surface structure rule of semantic interpretation.
Thus it is formally very similar to the Scope Rule for negation
and quantifiers discussed in Chapter 4; both are surface struc-
ture rules which move semantic elements up to an S node.

A couple of other miscellaneous aberrations deserve
mention here. I have no explanation for either of them. The
first is that S adverbs do not feel comfortable in questions.

(50) *Did Frank probably beat all his opponents?
      *Who certainly finished eating dinner?
      ?What has Charley evidently discovered?

VP adverbs in these sentences are often better.

(51) Did Frank easily beat all his opponents?
      Who completely finished eating dinner?
      What has Charley suddenly discovered?

Note, by the way, that (50) constitutes evidence against deriving
S adverbs from their paraphrases it is Adj that S. According to
that hypothesis, (50) should be synonymous with the acceptable (52).

(52) Is it probable that Frank beat all his opponents?
      Who is it certain finished eating dinner?
          (If this sounds funny, it is because who has
          been moved from an embedded subject.)
      What is it evident that Charley has discovered?

The other odd fact was pointed out to me by Jay Keyser.
It seems that adverbs that can function as either S or VP adverbs
sound funny when they precede a modal, either in aux position or
in initial position; they are much better after the modal.

(53) ?Slowly John will open the door.
      ?Dave quietly may leave the room.
      ?Stealthily you should climb the wall.
      ?Willy cleverly can touch his tongue to his nose.

(54) John will slowly open the door.
      Dave may leave the room quietly.
      You should stealthily climb the wall.
      Willy can touch his tongue to his nose cleverly.

According to the Scope Rule, the adverbs in (53) should all be
S adverbs. There does not seem to be any particular semantic

reason for all of these to sound funny; for example,

> (55) It is clever of Willy to be able to touch his
>   tongue to his nose

is not odd in the same way that the last sentence of (53) is.

6. The Orientation of S Adverbs

Some S adverbs are understood as relating the speaker's
attitude towards the event, and some somehow comment on the sub-
ject of the sentence. All the adverbs that are only S adverbs,
such as evidently, unfortunately, etc., seem to be speaker-oriented.
Thus they often have a paraphrase of the form it is Adj that S,
or even sometimes it is Adj to me that S (as with evidently,
apparently, but not unfortunately). We have noted already that
some adverbs that can be either S or VP adverbs shift in meaning
with shift in scope. In VP scope they are all "manner" adverbs.
In S scope, some are speaker-oriented:

> (56) Happily (,) John won the game.
>   Truthfully (,) ⎫
>   Frankly (,)   ⎬ John lied to Bill.

Note that the last two as VP adverbs are anomalous in this par-
ticular sentence, though they are not generally bad as VP adverbs.

> (57) *John lied to Bill ⎰truthfully.
>   ⎱frankly.
>   John told the story to Bill ⎰truthfully.
>   ⎱frankly.

Others in the class of adverbs that can be S or VP are
subject-oriented--they express some additional information about
the subject, as can be seen by the kinds of paraphrases they have.

> (58) Carefully ⎫
>   Cleverly ⎬ (,) John spilled the beans.
>   Clumsily ⎭

(59)
$$John\ was \begin{Bmatrix} ?clumsy \\ careful \\ clever \end{Bmatrix} to\ spill\ the\ beans.$$

$$It\ was \begin{Bmatrix} clever \\ clumsy \\ *careful \end{Bmatrix} of\ John\ to\ spill\ the\ beans.$$

Is there any reason to believe that this difference in orientation is anything more than a matter of semantic interpretation? In the papers I have cited, it is more or less taken for granted that the difference in orientation is attributable to deep structure, involving derivations from adjectival paraphrases.

But the paraphrases are hopelessly varied. It is Adj that S, which works for many of the speaker-oriented adverbs, will not work for those in (56). Happily might come from I am happy that S in some cases; nothing at all reasonable can be constructed to come close for truthfully and frankly. (59) shows that a consistent deep structure for subject-oriented adverbs is out of the question too.

A stronger argument for considering orientation a purely semantic fact comes from clauses and prepositional phrases which function as sentence adverbials. There is an incredibly wide variety of these.

(60) In all probability,
In my opinion,
According to Albert,
In spite of his mother's admonitions,
In order to kill his mother,
By going to Cincinnati,
(Being) sick at heart,
Having lost the game,
Now that he is married to Sally,
To tell the truth,
Taking all things into consideration,
} Bill has ruined his chances of an inheritance.

Notice now that some of these are speaker-oriented and some are subject-oriented. In particular, the orientation is revealed by the reference of deleted subjects. In order to clauses, by (means) clauses, and nominative absolutes (-ing clauses) all have deleted subjects understood to be subjects of the main clause. However, to tell the truth and taking all things into consideration certainly do not have subjects co-referential with the subject of the main clause: the understood subject is I or one or something more obscure--these are speaker-oriented adverbials. There is no question of deriving the adverbials in (60) from adjectival paraphrases; no one would suggest such a possibility. Yet they exhibit the same semantic range as -ly adverbs in the same position.

It therefore seems much more reasonable to attribute the range of meanings of sentence adverbs to the property of being a sentence adverb, rather than to a corresponding range of deep structures. The range of adjective paraphrases, even where possible, does not correspond in any nice way; and in the case of clauses, adjective paraphrases cannot exist. Instead, I suggest that the orientation of an adverb is an idiosyncratic semantic feature of the adverb, provided for in the projection rule for sentence adverbs. If the morphology of a particular adverb happens to relate to an adjective with appropriate syntatic properties, some paraphrase or another may exist. In other cases, where there is no such adjective, it will be impossible to find a paraphrase. Such a solution seems much more

in the haphazard spirit of the data.

7. Attraction to focus

In "Deep Structure, Surface Structure, and Semantic Interpretation" Chomsky discusses the division of the reading of a sentence into focus and presupposition. Intuitively, the presupposition is the material assumed by the speaker to be common to both speaker and hearer, and the focus is more or less new material (see examples in Chapter 1, §5). Chomsky shows that the focus is always a constituent containing the stress maximum of the sentence, where the stress maximum is determined by regular, obligatory rules. If there is emphatic (contrastive) stress, the focus is associated with a constituent containing the emphatic stress.

Since the rules for stress contours occur late in the derivation of a sentence, even after surface structure, the determination of focus must be a semantic process occurring at an equally late point in the derivation. I will assume that Chomsky's analysis is essentially correct and that there is a semantic rule called Focus Assignment which marks a certain constituent as focus and the rest of the sentence as presupposition, connected with the position of maximum stress.

It has been pointed out by various people, including Lakoff (1965) that yes-no questions and sentence negation do not necessarily question or negate the entire sentence. For example, the sentences in (61) are at least three ways ambiguous, the readings having paraphrases such (62), (63) and (64). Assume

normal, non-emphatic stress throughout.

(61) Did Marcel beat his wife with a hammer?
Marcel didn't beat his wife with a hammer.

(62) Was it with a hammer that Marcel beat his wife?
It wasn't with a hammer that Marcel beat his wife.

(63) Is what Marcel did beat his wife with a hammer?
One thing Marcel didn't do was beat his wife with
a hammer.

(64) Is what happened that Marcel beat his wife with
a hammer?
What didn't happen
Something that didn't happen } is that Marcel beat
his wife with a hammer.

Note that the constituents that can be negated or questioned are

exactly those that can be assigned focus under Chomsky's rules.

His wife cannot be negated or questioned unless it carries em-

phatic stress: (65) but not (61) can be paraphrased by (66).

Stress is indicated by underlining.

(65) Did Marcel beat his wife with a hammer?
Marcel didn't beat his wife with a hammer.

(66) Was it his wife that Marcel beat with a hammer?
It wasn't his wife that Marcel beat with a hammer.

Likewise, the subject can only be negated or questioned if it

has emphatic stress. (68) is only a paraphrase of (67), not of

(61).

(67) Did Marcel beat his wife with a hammer?
Marcel didn't beat his wife with a hammer.

(68) Was it Marcel who beat his wife with a hammer?
It wasn't Marcel who beat his wife with a hammer.
It was Marcel who didn't beat his wife with a hammer.
(depending on a subtle difference in intonation)

Thus the ambiguity in (61) is attributable to an ambiguity in

focus, reflected in different clefting possibilities. Where em-

phatic stress shifts focus, different cleft sentences are para-

phrases, and the negation and questioning are associated with the focus.

Before we formalize this notion, it should be pointed out that the use of focus in this connection explains various examples of Lakoff's without resorting to any higher Ss in deep structure. Locatives and manner adverbs work exactly the same way, merely by virtue of their occurring at the end of the sentence and thus normally receiving maximum stress. The sentences in (69) have ambiguities parallel to (62)-(64), and the sentences with emphatic stress, (70) and (71), are just like (65) and (67).

(69)   Does Klaus study German in the garden?
       Does Frank play his etudes carefully?
       Klaus doesn't study German in the garden.
       Frank doesn't play his etudes carefully.

(70)   Does Klaus study _German_ in the garden?
       Does Frank play his _etudes_ carefully?
       Klaus doesn't study _German_ in the garden.
       Frank doesn't play his _etudes_ carefully.

(71)   Does _Klaus_ study German in the garden?
       Does _Frank_ play his etudes carefully?
       Klaus doesn't study German in the garden.
       Frank doesn't play his etudes carefully.

Also the ambiguity of (72) is explained simply by the fact that either the subordinate clause or the whole sentence can be focus.

(72)   Moe doesn't beat his wife because he loves her.

Notice that emphatic stress on _loves_ makes the subordinate clause (or perhaps just the verb) obligatory focus; then the reading must include as part of the presupposition the fact that Moe beats his wife.

The elegance of this solution involving focus turns on the

fact that it refers to a single property, the stress center, which is independent of any considerations about the meaning of the items receiving it. Since focus is a semantic property orthogonal to the so-called "logical" reading of the sentence (the properties involved in finding truth-value), the independence of stress from lexical meaning makes a nice correlation. A dependence of focus on deep structure, on the other hand, does not capture this independence immediately, and it is not clear to me that it ever could. Certainly the variety of deep structures Lakoff has proposed for different cases of focus do not even seem to allow one to treat it as a unitary phenomenon.

To formalize the interpretation of these sentences, I will propose two ordered rules of semantic interpretation. The first is Chomsky's Focus Rule, which labels one constituent of the sentence _focus_ and substitutes variables for the focus to form the _presupposition_. The second rule will be a rule of Attraction to Focus, which so far will apply to neg and Q of yes-no questions, but will be extended in a moment. This rule links the interpretation of neg and Q to the focus, so only the focus is understood as being negated or questioned.

Whether this rule represents an actual shifting of scope or some more subtle kind of attraction is unclear to me. If it is an actual shifting of scope, it will have to take place after the application of the some-any selection restriction, in order to get (73), for example, rather than (74).

(73)   Bill didn't send a _book_ to any of the men.
(74)   *Bill didn't send a _book_ to some of the men.

The some-any rule must apply which the neg has S as its scope, in order for _any_ to be correct. If Attraction to Focus reduced the scope before the application of some-any, we would expect (74) instead.

The Focus Rule and Attraction to Focus together account for the readings of all the examples discussed so far but one. The only exception is the second reading of the negated sentence in (67), "It was Marcel who didn't beat his wife with a hammer." This reading of (67)-neg is distinguished from the other by intonation curve; the first reading has higher pitch on the last stressed syllable, or rising intonation, whereas this one has level or falling. Actually, this same difference in intonation produces a difference in readings of (65)-neg, the reading with level intonation being an answer to "Who didn't Marcel beat with a hammer?"

Somehow the intonation curve has to affect Attraction to Focus. Perhaps the rule deriving these special readings is a rule of Attraction to Presupposition, but I suspect it is simply an automatic consequence of not doing Attraction to Focus. If there were also a rule of Attraction to Presupposition, we might expect a third reading, in which neither Attraction Rule had applied.

Some similar phenomena were noticed by Susan Fischer (1968), dealing with the use of _even_ and _only_, and by Francois Dell (1969) for _first of all_. The same facts also seem to hold for _just_. These words can occur before NPs and in the aux. If in an NP, no

ambiguity results. But if <u>even</u> is in the aux, a number of possible readings present themselves. (Subsequent remarks referring to <u>even</u> apply identically to <u>only</u> and <u>just</u>, though the examples have to be changed for appropriateness.)

(75) Even John gave his daughter a new bicycle.
(76) John gave even his daughter a new bicycle.
(77) John gave his daughter even a new bicycle.
(78) John even gave his daughter a new bicycle.

In (78), the readings seem to be associated with different choices of focus: <u>even</u> can go with <u>a new bicycle</u>, the VP, and perhaps the entire S. In (78) S focus is not so clear, but in (79) (due to Steve Anderson) it is.

(79) The results of today's games will be remarkable:
      Harvard will even defeat Loyola.

Note however that emphatic stress in (78) results in lack of ambiguity. <u>Even</u> must go with the stressed constituent.

(80) <u>John</u> even gave his daughter a new bicycle.
      John even gave his <u>daughter</u> a new bicycle.
      John even gave <u>his</u> daughter a new bicycle.
      John even gave <u>his</u> daughter a <u>new</u> bicycle.
      John even gave his daughter a new <u>bicycle</u>.
      John even <u>gave</u> his daughter a new <u>bicycle</u>.

Furthermore, stress in (75), (76), and (77) renders the sentence unacceptable unless the stress is in the NP preceded by <u>even</u>.

(81)   Even <u>John</u> gave his daughter a new bicycle.
      *Even John <u>gave</u>
      *   <u>his</u>
      *   <u>daughter</u>
      *   <u>new</u>
      *   <u>bicycle</u>.

(82) *<u>John</u> gave even his daughter a new bicycle.
      *   <u>gave</u>
      ok   <u>his</u>
      ok   <u>daughter</u>
      *   <u>new</u>
      *   <u>bicycle</u>.

(83)   *<u>John</u> gave his daughter even a new bicycle.
       *   <u>gave</u>
       *   <u>his</u>
       *   <u>daughter</u>
       ok   <u>new</u>
       ok   <u>bicycle</u>.

To account for these facts, let us discuss for a moment the meaning of <u>even</u> (cf. Anderson (1969)). The association of <u>even</u> with a constituent implies that there is something special, unusual, or unexpected about the connection of that constituent with the event. If there is something unexpected about the constituent, it must be new to the hearer, and hence by definition part of the focus.

How does <u>even</u> get associated with the focus? Obviously, by an extension of the rule Attraction to Focus. We have just shown that it is in the semantic nature of <u>even</u> that it must be attracted to focus to produce a good reading (this is not true of neg). The failure of the starred sentences in (81)-(83) must be due to the inability of <u>even</u> to attract to the focus because of the structure of the sentence.

Let us assume that <u>even</u> is subject to the usual Adverb Scope Rule. Since this rule applies to give sentence scope to adverbs in aux position, <u>even</u> will be an S adverb in (78). But in (75)-(78) it must have only the NP it precedes as scope, since the Scope Rule can only raise an adverb in aux position.

If this is the case, we can produce the correct results by adding a simple condition to Attraction to Focus: the focus must be contained in the scope of the element being attracted.

In (80), where <u>even</u> is an S adverb, everything in the sentence is contained in its scope, and so focus can go anywhere. In (81)-(83), <u>even</u> only takes as scope the NP it precedes, so focus can only be within that NP, although it can be any element within the NP.

As a confirmation of this analysis, observe that <u>even</u> cannot attract to a stressed subject if it follows two auxiliary elements. This is exactly the situation in which the Scope Rule cannot operate to produce sentence scope.

(84)   *<u>John</u> will have even given his daughter a new bicycle.
        *<u>John</u> has been even giving his daughter French lessons.

But any other focus is acceptable in these sentences.

(85) John will have even given his <u>daughter</u> a new bicycle.
                    etc.

If focus must be within the scope of the adverb that is undergoing attraction, (84) will be out because the scope of <u>even</u> can only be the VP, and the subject is the focus. On the other hand, in (85), the focus is within the VP, so the sentences are all right.

This condition on Attraction to Focus makes it easy to show the difference between the Scope Rules for negation and ad-verbs. Contrast (82) to (86), where <u>even</u> has been replaced by neg. (Read all of these with rising intonation.)

(86) <u>John</u> gave nobody a new bicycle.
        <u>gave</u>
            <u>nobody</u>
                  <u>new</u>
                      <u>bicycle</u>.

All of these sentences are acceptable, and in each case the neg

attracts to the focus. These sentences work because the Scope
Rule for negation can move a neg up from any position in the
sentence to sentence scope. With the neg of nobody interpreted
as S negation, all possible foci fall within the scope of neg,
so Attraction to Focus can work in every case. The adverb even,
on the other hand, can only assume S scope from aux position, so
in (82), if the focus is outside the indirect object, even cannot
attract and a bad reading results. The difference between (82)
and (86), then, is a difference in the applicable scope rules,
not in the use of Attraction to Focus or the selection of focus
itself.

As Stephen Anderson has pointed out, this analysis of the
semantics of even eliminates the need to syntactically restrict
the occurrences of even to one per sentence. Such a need prompted
Kuroda's proposal (1965) for handling even by means of an attach-
ment transformation that changed meaning. Since there is only
one focus per sentence, and since one can (presumably) only men-
tion once that the focus is remarkable, this leaves semantic
room for only one even per sentence. In the syntax, then, we
can generate even freely in all constituents it occurs with.
There is no need for transformations to move even around, and no
need to make any syntatic restrictions on its occurrence.

Also in the class of adverbs that undergo Attraction to
Focus are the -ly adverbs merely, truly, simply, and hardly.
(87) is ambiguous in the usual way for sentences with such adverbs,
depending on choice of focus; stress automatically attracts the

adverb.

(87)
Tom $\begin{Bmatrix} \text{merely} \\ \text{simply} \\ \text{truly} \\ \text{hardly} \end{Bmatrix}$ put the book on the table.

(88) Tom $\begin{Bmatrix} \text{merely} \\ \text{simply} \\ \text{truly} \\ \text{hardly} \end{Bmatrix}$ put $\begin{Bmatrix} \text{the \underline{book} on the table.} \\ \text{the book on the \underline{table}.} \end{Bmatrix}$

These adverbs do seem to be more restricted in use than the _even_ group. Like most of the _even_ group, they cannot occur preposed without being automatically associated with the subject, and they cannot occur postposed. Their occurrence in front of NPs seems to be restricted larely to indefinite NPs.

(89) Donald brought along $\begin{Bmatrix} \text{simply} \\ \text{merely} \end{Bmatrix}$ $\begin{Bmatrix} \text{a piece of cheese} \\ \text{*the first thing he found} \end{Bmatrix}$
for lunch.

(90) This is truly $\begin{Bmatrix} \text{a fantastic result.} \\ \text{*the fantastic result you've been} \\ \text{expecting.} \end{Bmatrix}$

(91) Hardly $\begin{Bmatrix} \text{a mouse was} \\ \text{*the mice were} \end{Bmatrix}$ stirring.
(note however: Hardly even the mice were stirring.)

Syntactically, these constructions seem to be related to _many a_, _such a_, and _not a_, which are only indefinite. In particular, the appearance of the quantifiers _many_, _hardly_, and _not_ seems no coincidence. There is undoubtedly also a relation between these constructions and things like _a mere boy_, _a true believer_, and _a simple contradiction_. Under our assumptions this cannot be a transformational relationship. Rather these are the syntactic analogues of (87) generated by the base rule schema (26) when X is a noun instead of a verb. It may be that in these NP constructions we can find evidence for an NP analogue of the Adverb Scope

Rule.

At least two of the adverbs that occur in aux position only, _utterly_ and _virtually_, are still more restricted than the ones we have discussed so far. These probably also attract to focus, but they only allow the focus to be the verb. Contrastive stress is bad elsewhere in examples like (92).

(92)   The enemy $\begin{Bmatrix} \text{utterly} \\ \text{virtually} \end{Bmatrix}$ly destroyed the city.

   *   _enemy_

ok   _destroyed_

   *   _city_.

Also, they do not occur in front of indefinite NPs.

(93)   *This is virtually a fantastic result.
         *Utterly a mouse was stirring.

The same is probably true of _really_, at least when it occurs in aux position.

Thus, by means of the two rules, Adverb Scope and Attraction to Focus, we can distinguish all the classes of -_ly_ adverbs mentioned in §2 on the basis of their semantic behavior. This being the case, there is no need to place any restrictions on their syntactic behavior. They can be permitted to move about freely in the sentence under the preposing and transportability rules, without exception. Then the surface structure rules of interpretation will filter out the bad sentences. Note that since the semantic rules depend on surface structure, they do not depend on the correctness of the highly speculative analysis of §§ 3-4.

Incidentally, there is at least one other semantic rule that is similar to Attraction to Focus. Bowers (1969) points out that in the preposed _of_-phrase connected with superlatives,

the reference shifts with contrastive stress.

(94)   Of the three men, $\begin{cases} \text{John hates } \underline{\text{Bill}} \text{ the most.} \\ \text{*John hates } \underline{\text{Mary}} \text{ the most.} \\ \underline{\text{John}} \text{ hates Bill the most.} \\ \text{*}\underline{\text{Mary}} \text{ hates Bill the most.} \end{cases}$

This cannot be a syntactic linkage originating in deep structure.

For one thing, *John of the three men is impossible. For another,

the of-phrase really goes syntactically with the superlative.

(95)   John hates Bill the most of the three men.

With the of-phrase here, however, stress only sounds good on Bill.

In the passive, stress only sounds good on John.

(96)   Bill is hated by John the most of the three.

This may be another scope phenomenon. Note that when the of-

phrase is in the VP, only the noun phrase within the VP can be

stressed. But when the of-phrase is attached to S, as in (94),

either NP can be stressed. However, I would prefer not to

speculate further.

8. Until and some other odd things

It was noticed at least as early as Klima (1964) that until

has a strange set of conditions on its occurrence. In sentences

with durative verbs, until can always occur, but with point-action

verbs, the sentence must be negative.

(97)   John slept until 6:00.
(98)   John didn't sleep until 6:00.
(99)   *John arrived until 6:00.
(100)   John didn't arrive until 6:00.

This restriction was just taken for granted without any analysis.

Let's try to find some sort of explanation for it.

Notice first that there is more to the restriction than just the facts of (97)-(100). The preposing properties with durative and point-action verbs differ.

(101)   Not until 6 did John $\begin{cases} \text{arrive} \\ \text{*sleep} \end{cases}$ 5

(102)   Until 6 (,) John didn't $\begin{cases} \text{arrive} \\ \text{*sleep} \end{cases}$

(103)   Until 6, John $\begin{cases} \text{*arrived} \\ \text{slept} \end{cases}$

In particular, neither of the possible preposed forms of (98), the negative durative case, are grammatical. There is more than meets the eye here.

Let's analyze the meaning of until. In X happened until Y, it is implied that X is a continuous state of affairs that underwent a change at time Y, into a different or opposite continuous state of affairs. To see that until implies a change at time Y, compare (104), with until, to (105), with before, where there is no such implication.

(104)   *John was sleeping until 6. In fact, he slept
        straight through supper too.
(105)   John was sleeping before 6. In fact, he slept
        straight through supper too.

The fact that until implies a change in a continuous state accounts for the fact that it cannot occur with point-action verbs. However, if the sentence has a plural sense, a point-action verb may be acceptable, with the sense keep Xing:

(106)   Guests arrived until 6, when the last straggler
        finally showed up.

(106) shows that the restriction is not on the verb, but rather on the capability of the whole sentence to be interpreted duratively.

Next look at <u>until</u> in negative sentences. (98) seems to
be a straightforward case of sentence negation; it can be para-
phrased by <u>it is not the case that John slept until 6; he woke
up before</u>. But (100) has no such paraphrase. How can we account
for this?

As a hypothesis, suppose that <u>not until 6</u> in (100) and
(101) functions as a semantic unit which acts differently than
S negation and <u>until 6</u> do separately. We can make them a semantic
unit by applying Attraction to Focus in the usual way. What does
<u>X didn't happen until 6</u> mean? It implies that X is a point-action
which happened <u>at</u> time Y, but not <u>before</u>. As in (104)-(105), we
can contrast <u>until</u> with <u>before</u> to show that <u>until</u> implies that X
actually happened at Y, but <u>before</u> has no such implication.

(107)   *John didn't arrive until suppertime; he finally
           did turn up the next morning.
(108)   John didn't arrive before suppertime; he finally
           did turn up the next morning.

Let us suppose then that <u>until</u> with point-action sentences
must attract the sentence negation in order for the sentence to
be good. To show that the Attraction to Focus rule is what is
essential, let us move the focus away from <u>until</u>: the sentences
immediately sound strange as point-action (With the reading "keep
hitting," they are all right, particularly with rising intonation,
which strongly attracts neg to focus.).[6]

(109)   ?<u>John</u> didn't hit Bill until 6.
(110)   *John did <u>not</u> hit Bill until 6.
(111)   ?John didn't hit <u>Bill</u> until 6.
(112)   *John didn't <u>hit</u> <u>Bill</u> until 6.

It must be emphasized, however, that <u>not until 6</u> is only a

semantic unit, not a syntactic one. If it were a syntactic unit,
we would expect (113) to correspond to the preposed form (101),
instead of (100) doing so.

(113) *John arrived not until 6.

It is possible to find sentences where the neg associated
with a preposed adverbial is syntactically connected with the
adverbial, and then we get forms like (113), and no inversion in
the preposed forms.

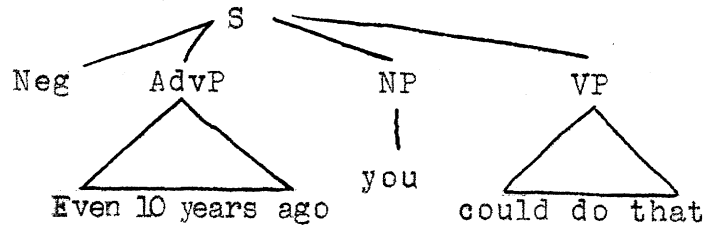(114) Not even 10 years ago you could do that.
(115) You could do that not even 10 years ago.

In this particular case, there is also a similar pair of sentences,
where the neg is S negation and does not go syntactically with
the adverbial.

(116) Not even 10 years ago could you do that.
(117) You couldn't do that even 10 years ago.

These sentences are discussed in Klima's Negation paper, but they
are, unaccountably, not distinguishable by his rules. His rule
absorbing neg into NPs and adverbials must occur before subject-
aux inversion, so sentence negation (114) and constituent negation
(116) cannot be told apart when it comes time to invert. However,
if in the case of sentence negation the absorption is a _semantic_
one, we can give (114) and (116) different structures at the time
of inversion and still account for the meanings. (114) will have
the derived structure (118), and (116) will have (119).

- 236 -

(118)

```
                          S
          Neg    AdvP        NP        VP
                  /\          |        /\
                 /  \         you     /  \
                /____\               /____\
           Even 10 years ago      could do that
```

(119)

```
                       S
             AdvP         NP        VP
             /\           |         /\
            /  \          you      /  \
           /____\                 /____\
     Not even 10 years ago    could do that
```
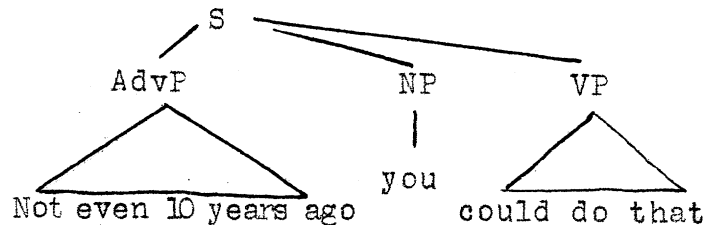
The difference also shows up in the some-any rule. Only in the case where there is inversion can <u>any</u> occur, showing that it is S negation and the other is not.

  (120) Not even 10 years ago could you do anything about that.
  (121) *Not even 10 years ago you could do anything about that.

There will of course have to be some transformation that moves neg into initial position. This rule will be conditioned by the presence of a preposed AdvP. It is optional: in addition to (116) and (117) there is the intermediate synonymous sentence (122).

  (122) Even 10 years ago you couldn't do that.

A few more examples of this difference between S negation and constituent negation with preposed adverbials:

  (123) Not for more than 5 minutes did he hesitate.
    For more than 5 minutes he didn't {hesitate / move (more plausible meaning)}
    He didn't hesitate for more than 5 minutes.
    *Not for 5 minutes he hesitated.
    *He hesitated not for 5 minutes.
     (S negation only; difference in meaning due to order of quantifiers?

(124) Not long ago there was a rainstorm.
There was a rainstorm not long ago.
*Not long ago was there a rainstorm.
*There wasn't a rainstorm long ago.
*Long ago there wasn't a rainstorm.
(only constituent negation)

(125) In not many years will Christmas fall on Sunday.
In many years Christmas won't fall on Sunday.
Christmas won't fall on Sunday in many years.
In not many years Christmas will fall on Sunday.
Christmas will fall on Sunday in not many years.

(these examples and example types from Klima (1964, pp.305-7)

These examples show that Attraction to Focus does not involve an actual shifting of the scope of negation, but rather some sort of semantic linking that leaves scope intact. If it did shift the scope, (114) and (116) would have the same semantic representation, i.e. be synonymous. It is only if we assume that Attraction to Focus does something other than shift scope that (114) and (116) will be distinct. Since scope remains intact, the selection restriction for some-any can be ordered after Attraction to Focus. This means it can maintain the status of a normal selection restriction, that is, a condition on readings.

Our conclusion about until, then, is that it cannot occur syntactically with negation. If it does not absorb neg by Attraction to Focus, it can only occur in a durative sentence. If it does absorb neg semantically, it can only occur in a point-action sentence. Let us review how this restriction operates in the examples under discussion. (97) is a durative sentence, so until cannot attract neg. (98) seemingly is in violation of this condition. However, note that the focus of (98) is at least sleep until 6:00, not until 6:00. To see this, put emphatic stress

on 6:00 in (98), so that focus must be just the PP and not the whole VP. The sentence becomes strange. If stress (and thus focus) is put anywhere else in (98), the sentence is good. In (98), since neg is attracted to at least the whole VP, the semantic unit not until 6:00 does not form, so the non-negative restriction to a durative sentence holds on until.

(99) is in straightforward violation of the restriction of until to durative sentences. However, if the verb stays point-action but a plural subject is used, a durative sense is possible, as we saw in (106). In (100), until 6:00 is the focus, so neg attracts to it. (109)-(112) show that if until 6 is not the focus and thus does not attract the neg, the sentence is out.

To deal with the preposed cases (101)-(103), it seems safe to assume that a preposed adverbial is focus. Certainly putting emphatic stress anywhere else in the sentence disturbs the grammaticality. This being the case, the neg must attract to until 6 in (101) and (102), making only the point-action reading possible. It cannot attract to the VP or a stressed element as it could in (98), which would make a durative reading acceptable as well. In (103), there is no neg, so of course only the durative reading is grammatical.

NOTES

1. But see § 4 for more positions.

2. (26) still leaves us with the problem of explaining why there can only be one adverb in auxiliary position, but numerous adjectives. I have no explanation for this. However, the usual solution for multiple adjectives, embedding one relative clause in another to the desired depth, is not the only possible solution. The evidence that successive adjectives are embedded is from indefinite pronominalization: <u>John bought a new red house and I bought an old one</u>. But <u>one</u> could be considered to represent the presupposed parts of the N̄, much as the pro-VP <u>do so</u> represents presupposed parts of the VP (cf. footnote 4). If this is the case, no embedding of adjectives need be supposed.

3. I.e., a deep structure semantic rule that interprets a structural configuration.

4. The arguments about final adverbs and prepositional phrases being inside or outside the verb phrase, in connection with discussions of <u>do so</u> (as in Lakoff and Ross (1966) and Anderson (1968)) seem to me totally irrelevant. Under more traditional analysis of the verb phrase, the distinction turns out to be simply that <u>do so</u> obligatorily replaces the verb and all parts of the verbal complement strictly subcategorized by the verb. These are the elements which Ross and Lakoff regard as

"inside the verb phrase." In addition, <u>do so</u> may also replace elements which are optional in the verbal complement, but it need not. These are the elements which they regard as "outside the verb phrase."

5.   Note that there is a point-action reading of <u>sleep</u>, "fall asleep, begin to sleep," that is good exactly where <u>arrive</u> is good. This kind of reading seems possible with all (most?) durative verbs. In further discussion in this section the * in a sentence with a durative verb refers to the durative reading only.

6.   (110), which was the germ of much of this discussion, was pointed out to me by Peter Culicover.

INDEX OF RULES

(*denotes a rule not in the framework of this thesis)

Base rules

$$\begin{bmatrix} X \\ \propto Verb \end{bmatrix} \rightarrow \begin{bmatrix} Y \\ \propto Adverb \end{bmatrix} - X - Comp \qquad \text{Chapter 5, § 3}$$

$$PP \rightarrow P - (NP) \qquad\qquad 5.3$$

Transformations

Focus <u>rules</u> (depending on stress contours)

    Attraction to focus   footnote to 4.2, 5.7 5.8

    Focus assignment   footnote to 3.8, 5.7

<u>Semantic rules that interpret structure</u>
    <u>(i.e. Katz-Fodor projection rules)</u>

    Adj-N (probably derived structure)   1.2, 5.3

    Copula (deep structure)   5.3

    Generic (probably derived structure)   1.5

    S adverbs (derived structure)   5.5, 5.6

    Verb, subject, and object
        (deep structure)   1.2

    VP adverbs (derived structure)   5.3, 5.5

<u>Other semantic rules</u>

    Comparative (maybe)   3.2

    Gapping (maybe)   3.2

    Rules establishing parallel thematic
        relations between Ss and NPs   3.10, 3.11

    Rule that optionally duplicates subject
        of S in determiner of definite NP
        (this may be same as above rules)   2.9

<u>Well-formedness conditions on semantic</u>
    <u>interpretations</u>

    Convention that rules out uninterpreted
        reflexives   2.3

    Convention that rules out uninter-
        preted Δ 3.2

    Networks of coreference   3.8, 3.9

    Some-any selection restriction   4.4, 4.5, 5.7, 5.8

    Thematic hierarchy condition   2.12, 3.4

BIBLIOGRAPHY

Akmajian, Adrian, and Ray Jackendoff, "Anaphora and stress," to
 appear in Linguistic Inquiry (1969).

Anderson, Stephen R., "Pro-sentential Forms and Their Implications
 for English Sentence Structure," in Kuno (1968).

Anderson, Stephen R., "On How to Get Even," M.I.T., unpublished
 (1969).

Bach, Emmon, "Problominalization," University of Texas, unpublished
 ditto (1967).

Bach, Emmon, "Nouns and Noun Phrases," in Bach and Harms (1968).

Bach, Emmon, and Robert T. Harms, eds., Universals in Linguistic
 Theory, New York: Holt, Rinehart, and Winston (1968).

Baker, C. Leroy, "Definiteness and Indefiniteness in English,"
 University of Illinois Master's dissertation (1966)

Bierwisch, Manfred, "Semantic Universals of German Adjectivals,
 Foundations of Language 3.1-36 (1967).

Bolinger, Dwight, "The Atomization of Meaning," Language 41.555-73
 (1965).

Bowers, John S., "Surface Structure Interpretation in English
 Superlatives," M.I.T., unpublished (1969).

Carden, Guy, "English Quantifiers," Harvard University Master's
 dissertation (1967).

Carden, Guy, "English Quantifiers," in Kuno (1968).

Chafe, W. L., "Language as symbolization," Language 43.57-91 (1967).

Chapin, Paul, "On the Syntax of Word-Derivation in English," M.I.T.
 doctoral dissertation (1967).

Chomsky, Noam A., Syntactic Structures, The Hague: Mouton (1957).

Chomsky, Noam A., Aspects of the Theory of Syntax, Cambridge, Mass.:
 M.I.T. Press (1965)

Chomsky, Noam A., "Remarks on Nominalizations," in R. Jacobs and P. Rosenbaum, Readings in Transformational Grammar, Waltham, Mass.: Blaisdell (1969).

Chomsky, Noam A., "Deep Structure, Surface Structure, and Semantic Interpretation," in Jakobson and Kawamoto, eds., Studies in General and Oriental Linguistics: Commemorative Volume for Dr. Shiro Hattori, Tokyo: T.E.C. Corporation (to appear).

Dean, Janet, "Nonspecific Noun Phrases in English," in Kuno (1968).

Dell, François, "On the Scope of some French Adverbs," M.I.T., unpublished (1969).

Dougherty, Ray C., "A Transformational Grammar of Coordinate Conjoined Structures," M.I.T. doctoral dissertation (1968a).

Dougherty, Ray C., "A comparison of two theories of pronominalization and reference," M.I.T., unpublished mimeo (1968b).

Emonds, Joseph E., "Root and Structure-Preserving Transformations," M.I.T. doctoral dissertation (1969).

Fillmore, Charles, "The Case for Case," in Bach and Harms (1968).

Fischer, Susan, "Cleft Sentences and Contrastive Stress," M.I.T., unpublished (1968).

Fischer, Susan, and Byron Marshall, "The Examination and Abandonment of the Theory of Begin of D. M. Perlmutter as Carried Out by Two of the Inmates of Room Twenty-E-Two-Fifteen, under the Direction of Divine Providence," M.I.T., unpublished ditto (1969).

Gruber, Jeffrey, "Studies in Lexical Relations," M.I.T. doctoral dissertation (1965).

Gruber, Jeffrey, "Look and see," Language 43.937-47 (1967a).

Gruber, Jeffrey, Functions of the Lexicon in Formal Descriptive Grammar, Santa Monica, Calif.: Systems Development Corporation, TM-3770/000/00 (1967b).

Jackendoff, Ray S., "Quantifiers in English," Foundations of Language 4.422-42 (1968a).

Jackendoff, Ray S., "On Some Incorrect Notions about Quantifiers and Negation," in Anderson, Jackendoff, and Keyser, Studies in Transformational Grammar and Related Topics, Brandeis University, Department of English (1968b).

Jackendoff, Ray S., "Speculations on Presentences and Determiners," Bloomington, Ind.: Indiana University Linguistics Club (1968c).

Jackendoff, Ray S., "An Interpretive Theory of Negation," *Foundations of Language* 5.66-89 (1969).

Jackendoff, Ray S., "Morphological and semantic regularities in the lexicon," (forthcoming).

Katz, J. J., *The Philosophy of Language*, N.Y.: Harper and Row, (1966).

Katz, J. J., "Recent Issues in Semantic Theory," *Foundations of Language* 3 (1967).

Katz, J. J., and Jerry Fodor, "The Structure of a Semantic Theory," *Language* 39 (1963). Also in Fodor and Katz, *The Structure of Language*, Englewood Cliffs, N.J.: Prentice Hall (1964).

Katz, J. J., and Paul M. Postal, *An Integrated Theory of Linguistic Descriptions*, Cambridge, Mass.: M.I.T. Press (1964).

Keyser, S. J., Review of Sven Jacobson, *Adverbial Positions in English*, *Language* 44.357-74 (1968).

Kiparsky, Paul, "How Abstract is Phonology?", M.I.T., unpublished ditto (1968).

Kiparsky, Paul, and Carol Kiparsky, "Fact," in Bierwisch and Heidolph, eds., *Recent Advances in Linguistics*, The Hague: Mouton & Co. (1969).

Klima, Edward S., "Negation in English," in Fodor and Katz, *The Structure of Language*, Englewood Cliffs, N.J.: Prentice Hall (1964).

Klima, Edward S., "Studies in Diachronic Syntax," Harvard University doctoral dissertation (1965).

Kraak, A., "Presuppositions and the Analysis of Adverbs," M.I.T., unpublished (1967).

Kuno, Susumo, principal investigator, *Mathematical Linguistics and Automatic Translation*, Report NSF-20, Aiken Computation Laboratory of Harvard University (1968).

Kuroda, S.-Y., "Attachment transformations," M.I.T., unpublished (1965), (an enlarged portion of his M.I.T. doctoral dissertation, "Generative Grammatical Studies in the Japanese Language" (1965))

Kuroda, S.-Y., "Remarks on English Manner Adverbials," in Kuno (1968).

Lakoff, George, On the Nature of Syntactic Irregularity, Mathematical Linguistics and Automatic Translation, Report NSF-16, Aiken Computation Laboratory of Harvard University (1965).

Lakoff, George, "Deep and surface grammar," Harvard University, mimeographed (1966).

Lakoff, George, "Pronominalization and the Analysis of Adverbs," Harvard University, unpublished (1967).

Lakoff, George, "On Instrumental Adverbs and the Concept of Deep Structure," Foundations of Language 4.4-29 (1968a).

Lakoff, George, "Pronouns and Reference," Bloomington, Ind.: Indiana University Linguistics Club (1968b).

Lakoff, George, "Repartee," Foundations of Language 5 (1969).

Lakoff, George, and John R. Ross, "A Criterion for Verb Phrase Constituency," in Mathematical Linguistics and Automatic Translation, Report NSF-17. Aiken Computation Laboratory of Harvard University (1966).

Lakoff, Robin, "Some reasons why there can't be any Some-Any Rule," Cambridge, Mass.: Language Research Foundation (1968).

Langacker, Ronald, "Pronominalization and the Chain of Command," to appear in Schane and Reibel, eds., Modern Studies in English, Englewood Cliffs, N.J.: Prentice Hall (1969).

Lees, Robert B., The Grammar of English Nominalizations, The Hague: Mouton (1960).

Lees, Robert B., and Edward S. Klima, "Rules for English Pronominalization," Language 39.17-28 (1963).

McCawley, James D. "The Role of Semantics in a Grammar," in Bach and Harms (1968a).

McCawley, James D., "Lexical insertion in a transformational grammar without deep structure," in Proceedings of the Fourth Regional Meeting, Chicago Linguistic Society, April 1968. Dept. of Linguistics, University of Chicago (1968b).

McCawley, James D., "Where do NPs come from?", in R. Jacobs and P. Rosenbaum, Readings in Transformational Grammar, Waltham, Mass.: Blaisdell (1969).

Perlmutter, David, "The two verbs begin," M.I.T., unpublished ditto (1967).

Perlmutter, David, "Deep and Surface Structure Constraints in Syntax," M.I.T. doctoral dissertation (1968).

Perlmutter, David, "On the article in English," in Bierwisch and Heidolph, eds., Recent Advances in Linguistics, The Hague: Mouton & Co. (1969).

Postal, Paul M., "Restrictive Relatives and other matters and now less than 42.5% believed," I.B.M. Corporation, unpublished ditto (1967).

Postal, Paul M., Cross-Over Phenomena: A Study in the Grammar of Coreference, Yorktown Heights, N.Y.: I.B.M. Corporation, Thomas J. Watson Research Center (1968a).

Postal, Paul M., On Coreferential Complement Subject Deletion, Yorktown Heights, N.Y.: I.B.M. Corporation, Thomas J. Watson Research Center (1968b).

Rosenbaum, Peter, The Grammar of English Predicate Complement Constructions, Cambridge, Mass.: M.I.T. Press (1967).

Ross, John R., "Constraints on Variables in Syntax," M.I.T. doctoral dissertation (1967a).

Ross, John R., "On the cyclic nature of English pronominalization," in To Honor Roman Jakobson, The Hague: Mouton (1967b).

Ross, John R., "Gapping and the order of constituents," M.I.T., unpublished ditto (1967c).

Ruwet, Nicolas, "Adverbs: A Note on the Question: Where Do They Come From?", M.I.T., unpublished (1968).

Vetter, David, "Abstractness in Phonological Representations," M.I.T., unpublished ditto (1968).

Warshawsky, Florence, "Reflexivization I, II," M.I.T., unpublished (1965).

Weinreich, Uriel, "Explorations in Semantic Theory," in Current Trends in Linguistics, ed. by Thomas A. Sebeok, vol. III, The Hague: Mouton (1966).

BIOGRAPHICAL NOTE

The author was born on 23 January 1945 in Chicago. After a few years of moving around, his family settled in Glenside, Pennsylvania, a suburb of Philadelphia. Here the author underwent primary and secondary education, graduating from Abington High School in 1961 and receiving a National Merit Scholarship to Swarthmore College. At Swarthmore he majored in mathematics, minored in physics and philosophy, and discovered linguistics. Upon graduation in 1965, at which time he was elected to Phi Beta Kappa and Sigma Xi, he decided not to go to the Curtis Institute of Music, but rather to study linguistics at M.I.T. During the summer of 1966 he worked on a linguistics research project at Technical Operations, Inc.; in the summer of 1967 he did research under S. J. Keyser of Brandeis University. He was a participant in the 1968 and 1969 La Jolla Syntax Conferences. During the summer of 1969 he will be teaching at the Linguistic Institute at the University of Illinois.

The author's publications include "Quantifiers in English," in Foundations of Language (November 1968), "An Interpretive Theory of Negation," in Foundations of Language (May 1969), "Les constructions possessives en anglais," in Langages (June 1969), and with Adrian Akmajian, "Anaphora and Stress," to appear in Linguistic Inquiry. He is a member of the Linguistic Society of America.