NETWORKS OF ANAPHORA:

AN ESSAY

IN THE SYNTAX

OF PRONOMINALIZATION

by

GEORGE MIDDLETON WILLIAMS, JR.

B.A., Davidson College
(1966)

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

at the

MASSACHUSETTS INSTITUTE OF
TECHNOLOGY

September, 1971

Signature of Author .....
Department of Foreign Literatures
and Linguistics, August 16, 1971

Certified by .....
Thesis Supervisors

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

NETWORKS OF ANAPHORA: AN ESSAY IN THE SYNTAX OF PRONOMINALIZATION

by

George Middleton Williams, Jr.

Submitted to the Department of Foreign Literatures and Linguistics on August 16, 1971, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

After networks of anaphora have been defined, it is shown that every theory of pronominalization must contain a special device for filtering out sentences with ill-formed networks. In particular, it is shown that the rules of pronominalization, which apply to pairs of nodes, cannot prevent the generation of ungrammatical sentences if three or more anaphorically connected nodes are involved (coreference is one kind of anaphoric connection). The use of coreferentiality as the basic relation in a theory of anaphora is discussed, and rejected in favor of the use of antecedency. A Network Filter is formulated in terms of antecedency and tested on definite pronouns and Verb Phrase Deletion. Then it is demonstrated that the Filter also applies to reflexives and to nodes deleted by EQUI-NP Deletion. As a consequence, controlled Noun Phrases and reflexives must be subject to the normal rules of pronominalization throughout the derivation. Finally, a generalization of the network concept is proposed as an explanation for the inability of constituents out of which nodes have been chopped to serve as antecedents. This explanation may subsume Ross's Complex NP Constraint.

Thesis Supervisors: John Robert Ross and Paul Kiparsky
Titles: Associate Professors of Linguistics

This dissertation is dedicated to

H.I., G.M., and A.E.

ACKNOWLEDGEMENTS

The criticism and suggestions of John Ross and Paul Kiparsky contributed much to whatever clarity and insightfulness this dissertation possesses. Their willingness to read quickly under the pressure of deadlines made it possible for it to be submitted in the first place. Several conversations with Noam Chomsky helped me better understand the issues involved; and the comments of Bruce Fraser and Susumu Kuno have found their way into various formulations. Morris Halle deserves credit for keeping me working. What I have learned from Ken Hale has subtly influenced the thesis as a whole, as have a number of people who were in on it at the beginning: my wife, Helga, who suffered through hours of interrogation about the German pronoun dessen, whose peculiarities led me to many of the ideas expressed in this thesis; and the staff of the Forschungsinstitut für deutsche Sprache, Marburg, Germany, who submitted themselves to similar questioning. To all these people, and to the Danforth Foundation, which made everything possible, my deepest thanks.

TABLE OF CONTENTS

Abstract 2

Dedication 3

Acknowledgements 4

Introduction 6

Chapter One: The Need for Networks 12

Chapter Two: Coreference and Antecedency 40

Chapter Three: The Pairwise Antecedency Constraint 72

Chapter Four: A Constraint on the Internal Structure of Antecedents 120

Footnotes 166

Bibliography 176

Biography 178

INTRODUCTION

The purpose of this dissertation is to study somewhat more closely than has been done in the past the behavior of what can be called networks of anaphora. A network of anaphora will be said to exist if there are three or more nodes which are all coreferential with one another, stand in some sort of anaphoric relation to one another, or "refer to the same thing." In sentence (1), for example, James and the two occurrences of the pronoun he are coreferential with one another (refer to someone named James) and form a network:

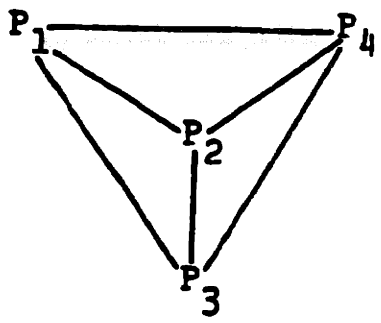
1) James said that he and the people he likes are all structuralists.

Sentence (2) also exemplifies a network, one in which the Verb Phrase go to Patagonia and learn Hittite and the two reduced Verb Phrases all stand in some sort of anaphoric relation to one another (refer to the same or equivalent acts):

2) Harry would go to Patagonia and learn Hittite if Sam also did and Jennifer wanted to.

The sentence is understood as if Sam also were followed by went to Patagonia and learned Hittite; and wanted to by go to Patagonia and learn Hittite. The reduced Verb Phrases function as "regular grammatical substitutes" for the full one and thus satisfy at least the definition of anaphora found in Webster's.

The term network was chosen because it is a central claim of this thesis that the members of a set of nodes $\{P_1, P_2, \ldots, P_n\}$, $n \geqslant 3$, cannot stand in an anaphoric relation to one another unless each P_i stands in an anaphoric relation to each of the members of the set $\{P_1, \ldots, P_{i-1}, P_{i+1}, \ldots, P_n\}$. Represented pictorially, with a line drawn between the nodes connected anaphorically in a grammatical anaphoric network, the term receives an interpretation which is in line with its other more established uses.

3) 

The above diagram could stand for the anaphoric connections between John and the three occurrences of the pronoun he in the following sentence:

4) John_{P_1} said that he_{P_2} would treat most favorably anyone he_{P_3} heard saying he_{P_4} should be governor.

In order for all four Noun Phrases (John and the three pronouns) to be coreferential or have the same referent, each one must, by hypothesis, have a direct anaphoric connection to each of the other three: $\underline{\text{he}}_{P_4}$ to $\underline{\text{John}}_{P_1}$, $\underline{\text{he}}_{P_2}$, and $\underline{\text{he}}_{P_3}$; $\underline{\text{John}}_{P_1}$ to $\underline{\text{he}}_{P_{1,2,3}}$; etc.

Since the pronouns in sentence (4) are definite pronouns, and since definite pronouns are understood to be coreferential with the nodes to which they are anaphorically connected, the multiple interconnections

shown in (3) would be the logical consequence of a somewhat smaller number of coreference relations: if P_3 is coreferential with P_2 and P_2 is coreferential with P_4, then by the transitivity of the coreference relation P_3 is coreferential with P_4. Similarly for some of the other connections in (3). The point of studying such an apparently unproblematic network of full Noun Phrases and definite pronouns is twofold: (a) to show that each of the connections among the nodes has purely syntactic, as well as logical, significance; (b) to determine whether some of the connections are obtained free, by deduction, or whether all of them must be established individually by the rule(s) of pronominalization.

The discussion of point (b) will require the study of the ways various potential and actually proposed theories of pronominalization account for networks. This discussion will be found in Chapter One. Using the example of a deletion theory of pronominalization, it will be shown that all theories must contain a device for forming or checking networks which is not identical to the devices or rules employed in determining whether a single pronoun can have a single NP, VP, or S node as its

antecedent.

In Chapter Two the nature of this device will be examined, and it will be argued that it does not involve the deduction of consequences from the coreference relation, or from any other equivalence relation. Instead, it will be proposed that a non-transitive relation called antecedency is established between anaphorically connected nodes and that a Network Interpretation Rule sorts out networks from the set of all connections (antecedency relations between nodes) provided by the rule(s) of pronominalization during a derivation.

The hypotheses proposed and defended in Chapter Two will be tested in Chapter Three, where the Reflexivization and EQUI-NP Deletion transformations will be analyzed. Reflexive pronouns and Noun Phrases deleted by EQUI will be shown to be connected anaphorically with Noun Phrases outside the range of the transformations or rules responsible for their generation (or interpretation). The demonstration of this fact involves using Postal's Crossover Principle as a diagnostic. Since all theories of Reflexivization and EQUI are affected by this demon-

stration, some attention will be paid to the evidence it provides concerning their formulation. A by-product of all this argumentation is proof of point (a) above, that all connections in a network are syntactically significant, i.e., that they affect the operation of syntactic rules.

In the course of the first three chapters emphasis will be placed on the surface-syntactic aspects of pronominalization, as opposed to the semantic. In line with this emphasis, Chapter Four will contain evidence for a previously unnoticed constraint on the internal structure of antecedents. A slight generalization of the network concept will be proposed as one possible explanation for the constraint, as will a (fragmentary) model of sentence perception. Finally, an attempt will be made to show that Ross's Complex NP Constraint is a special case of this Internal Structure Constraint.

Chapter One

THE NEED FOR NETWORKS

Most theories of pronominalization proposed within transformational grammar can be considered as the sum of decisions made concerning the following issues:

a) When the rules of pronominalization should apply: in deep structure, cyclically, last-cyclically, or post-cyclically in either shallow or surface structure;

b) Whether the rules are optional or obligatory;

c) What the rules do: whether they delete under identity, interpret readings, or filter out ungrammatical configurations which have been established by other rules earlier in the derivation;

d) Where relations such as coreference are marked: in deep structure by indices, or by the rules of pronominalization themselves, wherever they apply;

e) Which relations are primary:
coreference, some sort of equivalence,
or antecedency.

Two of the most widely known recent theories are those of Lakoff (1968) and Jackendoff (1969). In terms of the above scheme they can be characterized as follows:

Lakoff: Post-cyclic filter with coreference marked by indices in deep structure; certain kinds of deletion pronominalization are also permitted. The filters are probably obligatory.

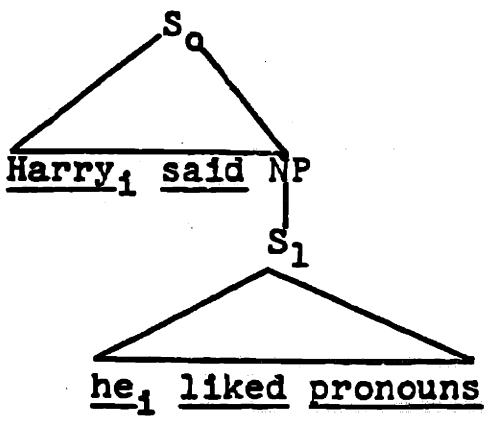
Jackendoff: Optional cyclic interpretive rules which themselves mark coreference.

An earlier theory argued for by Ross (1967) calls for the obligatory cyclic deletion of nodes under identity, where identity includes identity of coreference-marking indices, which are present in deep structure.

A sentence like (5) would have an underlying structure (very) roughly like (6) in Lakoff's theory.

5) Harry said that he liked pronouns.

6)



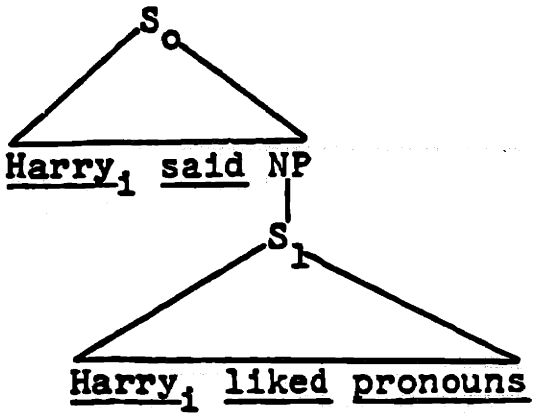
At the time post-cyclic rules apply, the structure of (5) would not differ greatly from the rough sketch in (6). One of several pronominal filters would be applied and would take into account that Harry$_1$ and he$_1$ have identical indices; that Harry$_1$ is a full Noun Phrase and he$_1$ a pronoun; and that Harry$_1$ is to the left of and commands he$_1$. Because all of this is as required by the filter, sentence (5) is declared grammatical (as far as pronouns are concerned).

In Jackendoff's theory, sentence (5) would again have an underlying structure like (6), except that Harry and he would not be subscripted. At the end of the S_1-cycle the interpretive rules of pronominalization

would apply, but nothing would happen, since there is no Noun Phrase in S_1 which could be the antecedent of he. Then at the end of the S_o-cycle the rules would be applied again; because Harry is to the left of and commands he, he would be marked as coreferential with Harry.

Ross's theory requires a slightly different underlying structure:

7)



At or near the end of the S_1-cycle the rule of deletion pronominalization would apply, but nothing would be deleted, since there are not two identical Noun Phrases in S_1. In S_o, however, the rule would have an effect: $\underline{\text{Harry}}_i$ in S_1 would obligatorily be deleted under identity with $\underline{\text{Harry}}_i$ in S_o, giving (eventually) the surface

structure which represents sentence (5).

It was within such a theory that Ross noticed and attempted to explain the most important fact about networks: that each member must be anaphorically connected with every other member. This formulation of Ross's insight is, however, somewhat anachronistic. In terms of his theory, what he observed was that the grammaticality of a particular anaphoric connection between two (Noun Phrase) nodes is dependent on the grammaticality of the connections between other nodes. He demonstrated this in his discussion (later amplified by Postal 1970) of the interaction between the Reflexivization, Pronominalization, and EQUI-NP Deletion transformations. In sentences like

8) Meeting him in the bus terminal
bothered Sam.

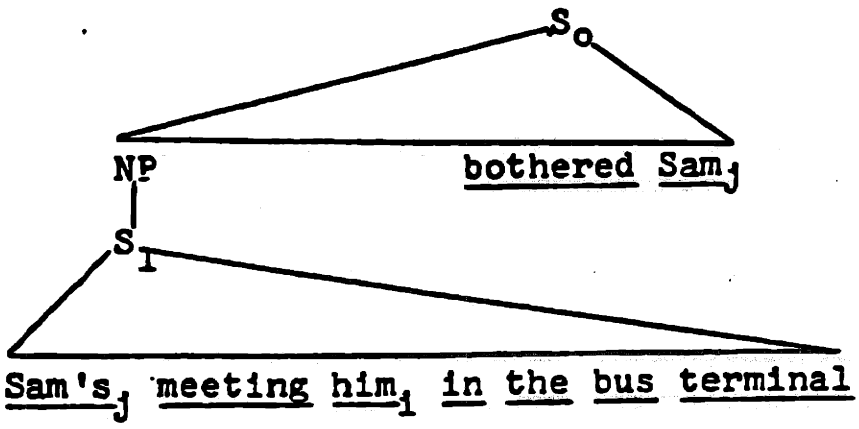
him cannot be coreferential with Sam, despite the fact that sentences with apparently identical structures (disregarding some lexical items), such as (9), do permit coreference:

9) Mary's meeting him in the bus terminal bothered Sam.

Ross accounted for the impossibility of coreference in sentences like (8) by claiming that their underlying structure was roughly the same as that of (10):

10) His$_j$ meeting him$_i$ in the bus terminal bothered Sam$_j$.

11)



In the course of deriving (7) from a structure like (11), the following transformations, among others, would apply: In the S_1-cycle, none. Since Reflexivization is obligatory in the framework of the deletion theory, it would have to apply if Sam$_j$ and him$_i$ were identical; its

non-application therefore guarantees that Sam (as the subject of meeting) has a different index from him. EQUI and pronominalization could both apply in the S_o-cycle, EQUI first. The Structural Description of EQUI is satisfied by the two occurrences of Sam_j; thus, the subject of meeting may optionally be deleted under identity with the object of bothered. Assume that deletion has taken place. Then at the time pronominalization could occur, the rules of Reflexivization and EQUI will have assured that him has a different index from the occurrence of Sam which is the subject of meeting, and that the subject of meeting has an index identical to that of the object of bothered. It follows that him cannot have the same index as the object of bothered, and therefore pronominalization cannot apply. Him can for this reason not be coreferential with Sam in sentence (8).

The same sort of logic prevents him and Sam from being coreferential in sentence (10). If Sam's has not been deleted by EQUI, but instead by the rule of pronominalization, it will still have to have the same index as the object of bothered. Once again this makes it impossible for him (the object of meeting) and the one

occurrence of Sam which reaches surface structure to have identical indices and be coreferential. The deductions which require that the objects (and the indices of the objects) of meeting and bothered be different become impossible under two conditions: if Sam's has an index not equal to j; or if some other Noun Phrase, like Mary in (9), is used as the subject of meeting. As the grammaticality of (9) shows, the impossibility of such deductions implies the possibility of pronominalization. Since this is what the deletion theory predicts, it receives important support.

Because the deletion theory marks indices in deep structure, no actual deductions are necessary to insure that pronominalization is properly applied in sentences like those discussed above. By being constrained to delete only under identity of index, the rules of pronominalization, Reflexivization, and EQUI take implicit advantage of the logic of identity. Their success in insuring only correct readings for the surface structures of sentences like (8) and (10) - that is, in making sure that the deleted subject, him, and Sam in (8), as well as his, him, and Sam in (10) do not form

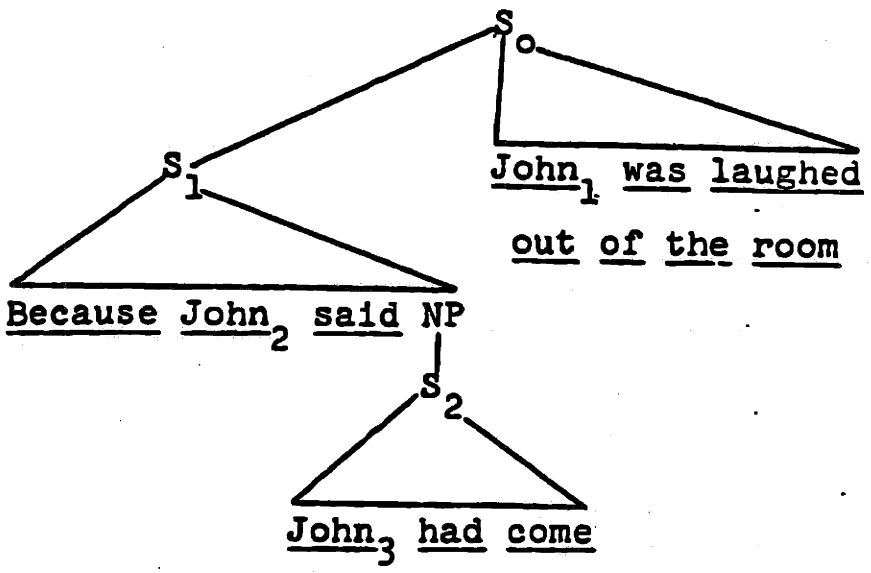
a network - thus makes it appear likely that a theory using obligatory cyclic deletion under identity would be maximally efficient. It would automatically generate networks (which contain three or more nodes) without recourse to any rules beyond the ones necessary to determine whether any two nodes are coreferential.

It will be the purpose of the remainder of this chapter to show that this impression of maximal efficiency is misleading. Beyond the rules for determining whether any pair of nodes (from now on called pronoun-antecedent or P-A pair) can be connected anaphorically, some additional device is necessary in a deletion theory in order to insure that the grammar generates only those networks which are grammatical. The arguments for this point can be carried over to other theories of pronominalization.

Cyclic pronominalization by obligatory deletion achieves its results by forcing a decision as to the possibility of coreference before the situation becomes ambiguous. Consider sentence (12), for example. If, because of optionality, pronominalization were not to

apply until the S_o-cycle, it would be possible by the normal constraints on backwards and forwards pronominalization for $\underline{John}_1$ to pronominalize $\underline{John}_2$ and then for $\underline{John}_3$ to pronominalize $\underline{John}_1$, giving sentence (13). Note that the numerical subscripts in (12) are used for easier identification, and not as indices marking reference.

12)



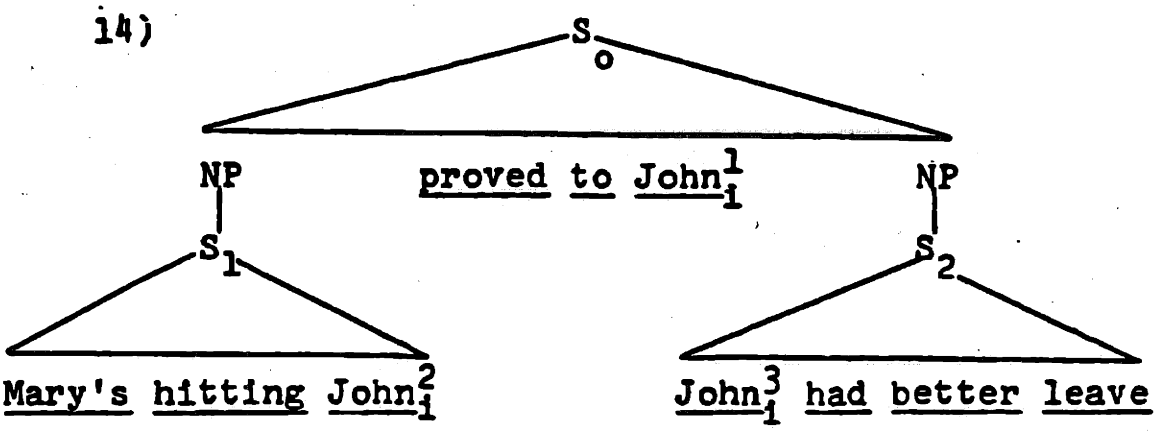
13) *Because he_i said that $John_i$ had come, he_i was laughed out of the room.

Clearly, however, the first occurrence of he in (13) cannot be considered coreferential with John. One of the major functions of obligatory cyclic pronominalization when applied to structures like (12) is therefore to

make sure that some action is taken in the S_1-cycle, depending on whether or not $\underline{John}_2$ and $\underline{John}_3$ have identical referential indices. If they do, $\underline{John}_3$ must be pronominalized, and sentences like (13) cannot result.

There are, however, structural configurations which do not permit cyclic deletion to take place until ambiguities are encountered which are exactly like those which would arise in (12) if pronominalization were optional. (14) shows such a configuration.

14)



Deletion is impossible in both the S_1 and S_2 cycles. When the S_o-cycle is reached, $\underline{John}_i^2$ could pronominalize $\underline{John}_i^1$; after that, $\underline{John}_i^3$ could pronominalize $\underline{John}_i^2$, since, as (15) shows, backwards pronominalization is possible from one S into another S which it both commands

and is commanded by.

15) Mary's kissing him_i proved that
$John_i$ was not her father.

But the application of pronominalization to (14) in the sequence just described does not produce a sentence in which all three relevant Noun Phrases can be understood as coreferential.

16) *Mary's hitting him^2_i proved to him^1_i
that $John^3_i$ had better leave.

Disregarding the subscripts, sentence (16) has only the following readings: $\underline{him}^1$ and $\underline{him}^2$ coreferential with each other but not with $\underline{John}^3$; or $\underline{him}^2$ coreferential with $\underline{John}^3$, but not with $\underline{him}^1$.

There is an apparently plausible way to avoid the attribution of the starred reading to (16) and sentences like it. Although nobody has proposed it and it does not work, it is still worth considering because of the insight it gives into networks. The downfall of

obligatory pronominalization in structures like (14) is its lack of instructions about what to do in case of ambiguity. If the application of the rule of pronominalization were properly constrained, it might be possible to generate only grammatical networks without having to introduce a special network-checking device. One such constraint might be: "First delete the left-most node under identity with the rightmost." The following constraint should be taken more seriously:

> Ambiguity Avoidance Rule: In a given cycle S_i take any possible antecedent (node with respect to which others can be deleted) NP_j; carry out all deletions which are possible with respect to NP_j; proceed to the next possible antecedent and do the same; continue until all possible deletions have been performed.

For structures like (14) the Ambiguity Avoidance Rule works. If $\underline{John}_i^2$ were chosen as NP_j it would have to pronominalize both $\underline{John}_i^1$ and $\underline{John}_i^3$, giving (17).

17) Mary's hitting $John_i^2$ proved to him_i^1 that he_i^3 had better leave.

If $\underline{John}_i^1$ happened to be NP_j it would be required to pronominalize both $\underline{John}_i^2$ and $\underline{John}_i^3$, which would result in (18).

18) Mary's hitting him_i^2 proved to $John_i^1$ that he_i^3 had better leave.

And finally, should $\underline{John}_i^3$ be NP_j only $\underline{John}_i^2$ could be pronominalized, leaving (19) as an intermediate stage.

19) Mary's hitting him_i^2 proved to $John_i^1$ that $John_i^3$ had better leave.

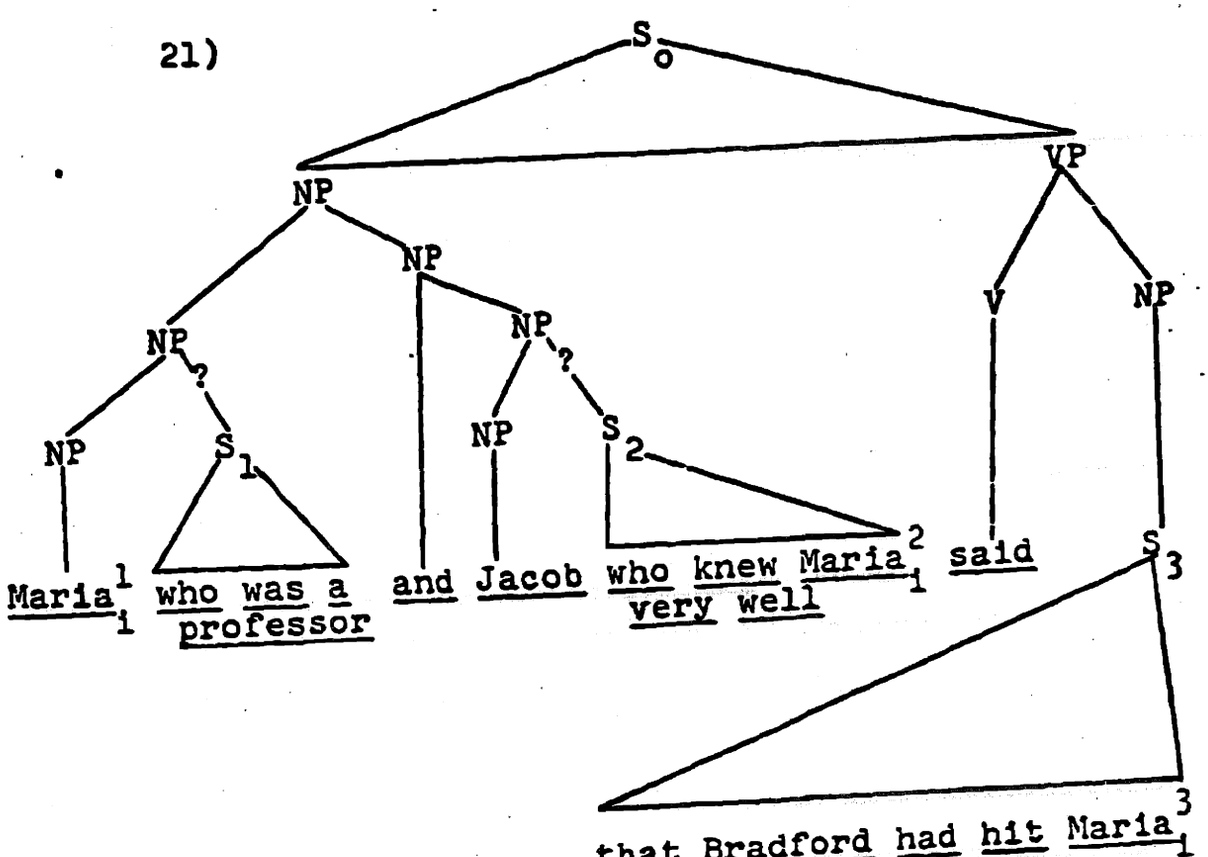
But then the next possible antecedent would be chosen. $\underline{John}_i^1$ is the only possibility; since it is to the left of and commands $\underline{John}_i^3$, deletion can occur, producing a surface structure identical to (18). The Ambiguity Avoidance Rule thus generates all and only the grammatical pronominalized surface structures from deep structures like (14).

Sentences exist, nevertheless, for which even the Ambiguity Avoidance Rule does not guarantee the correct assignment of coreference:

20) *Maria$_i^1$, who was a professor, and Jacob, who knew her$_i^2$ very well, said that Bradford had hit her$_i^3$.

Before pronominalization, (20) probably has a derived structure much like (21).

21)



All pronominalization in (21) must take place in the S_o-cycle, since there is only one occurrence of $\underline{\text{Maria}}_i$ in each of the other three S's. Furthermore, pronominalization probably must follow the rule which puts the appositive clauses (S_1 and S_2) in the positions shown above: while sentences like (22) are fully acceptable, the acceptability of their putative sources, sentences like (23), is questionable at best.

22) Jacob, who knew Maria_i very well, said that Bradford had hit her_i.

23) *Jacob said that Bradford had hit her_i, and he knew Maria_i very well.

It should perhaps be noted, however, that even if (23) were grammatical and the positioning of appositive clauses followed pronominalization, there would in this case (i.e., for sentences like 21) still exist an argument parallel to the one which follows that is capable of making the desired point.

Assume now that $\underline{\text{Maria}}_i^2$ in (21) is the NP_j mentioned in the Ambiguity Avoidance Rule. As (24)

shows, $\underline{\text{Maria}}_i^3$ can be deleted with respect to it.

24) Harry, who was a professor, and Jacob, who knew $\underline{\text{Maria}}_i^2$ very well, said that Bradford had hit $\underline{\text{her}}_i^3$.

$\underline{\text{Maria}}_i^1$, on the other hand, cannot, since $\underline{\text{Maria}}_i^2$ is to the right of and commanded by it.

25) *She_i^1 and Jacob, who knew Maria_i^2 very well, said that Bradford had left town.

If $\underline{\text{Maria}}_i^2$ has deleted $\underline{\text{Maria}}_i^3$, the only choice for the next possible antecedent is $\underline{\text{Maria}}_i^1$. The usual constraints allow $\underline{\text{Maria}}_i^2$ to become a pronoun in this environment.

26) $\underline{\text{Maria}}_i^1$, who was a professor, and Jacob, who knew $\underline{\text{her}}_i^2$ very well, decided to collaborate on a paper.

The sequence of pronominalizations just described, however, results in the ungrammatical sentence (20),

even though every step taken was in conformity with the Ambiguity Avoidance Rule.

The probable reason for the ungrammaticality of (20) can be found by examining sentence (27).

27) *Maria$_i^1$, who was a professor, and Jacob, who was a doctor, said that Bradford had hit her$_i^3$.

It is apparently impossible - for some speakers at any rate - for Maria$_i^1$ to pronominalize Maria$_i^3$. The ungrammaticality of the anaphoric connection between Maria$_i^1$ and her$_i^3$ in (20) must thus be understood as being capable of somehow rendering the anaphoric connections between the other nodes (Maria$_i^1$ and her$_i^2$; her$_i^2$ and her$_i^3$) null and void. This situation may be described as follows in terms of the definition of a network: her$_i^3$ is anaphorically connected in (20) to only one of the two other members of the putative network; by hypothesis, therefore, no network exists, and Maria$_i^1$, her$_i^2$, and her$_i^3$ cannot be coreferential.

If the Ambiguity Avoidance Rule is supplemented it will be able to prevent the generation of (26).

> Ambiguity Avoidance Supplement[1]: NP_j must be either the leftmost or the highest possible antecedent; the highest antecedent will be the one which commands all the others, and which is in turn commanded by none. If there is no highest possible antecedent, NP_j will be chosen from among those possible antecedents which command all other possible antecedents but are themselves also commanded.

This Supplement amounts to a definition of prominence and should therefore not be dismissed out of hand as irrelevant to the study of pronominalization. Because of the Supplement, $\underline{\text{Maria}}_i^1$ would have to be chosen as NP_j when pronominalization applies to (21): it is the sole occurrence of $\underline{\text{Maria}}_i$ which is not in an embedded sentence; it therefore commands all other occurrences. As (26) and (27) show, it would pronominalize $\underline{\text{Maria}}_i^2$,

but it is incapable of pronominalizing $\underline{\text{Maria}}_i^3$. Once $\underline{\text{Maria}}_i^2$ has been deleted, the only remaining unused possible antecedent is $\underline{\text{Maria}}_i^3$. But, as (28) shows, $\underline{\text{Maria}}_i^3$ cannot pronominalize $\underline{\text{Maria}}_i^1$.

28) *She$_i^1$, who was a professor, and
Jacob, who was a doctor, said that
Bradford had hit $\underline{\text{Maria}}_i^3$.

Therefore, the only surface structure which can result from (21) is (29).

29) *Maria$_i^1$, who was a professor, and
Jacob, who knew her$_i^2$ very well,
said that Bradford had hit Maria$_i^3$.

The ungrammatical sentence (20) cannot be derived. Within the deletion theory the ungrammaticality of (29) would in turn be accounted for by requiring that the Structural Description of pronominalization must be met if two identical Noun Phrases are found in a sentence; in the case of $\underline{\text{Maria}}_i^1$ and $\underline{\text{Maria}}_i^3$ it is not met, as has been shown (cf. sentences 27 and 28), so (29)

is starred.

At this point it might be asked if the Ambiguity Avoidance Supplement will work by itself, i.e., without the aid of the Ambiguity Avoidance Rule. That it will not is indicated by (16), which, although ungrammatical, can be derived if pronominalization is constrained only by the Supplement. On the other hand, certain sentences cause problems even when the Rule and the Supplement work together:

30) The first time I saw him_i was when $Aldous_i$ was three.

31) This picture of him_i proves that $Heinrich_i$ was once a satyr.

As the Supplement is formulated above, (30) and (31) should be <u>ungrammatical</u>. It must therefore be modified as follows:

> The Supplement does not apply if the NP_j it chooses is in an environment where it may be pronominalized by a

Noun Phrase which does not command
and is to the right of it.

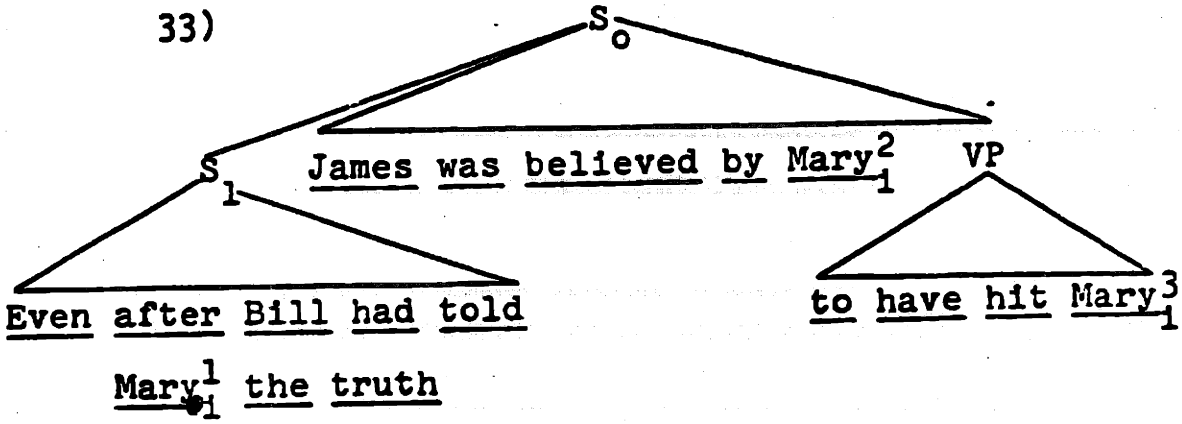
This modification of the Supplement succeeds in allowing (30) and (31) to be marked grammatical. Recall, however, the purpose of stating the Ambiguity Avoidance Rule and its Supplement: the deletion theory was to be constrained in such a way as to prevent certain ungrammatical derivations, but without modifying its most important characteristic, the ability to derive networks "blindly" by generating the connections between pairs. Once the above modification has been added to the Supplement, generating an anaphoric connection between the two nodes of one pair sometimes requires a look at the possible connections between the two nodes of another pair. The deletion theory is no longer blind, and it thereby loses much of its interest.

Not only does the modification in a sense destroy the deletion theory in order to save it; there is some indication that its inclusion in the Supplement will still not prevent the derivation of all ungrammatical sentences. Consider (32):

32) *Even after Bill had told Mary_i^1 the truth, James was believed by her_i^2 to have hit her_i^3.

If (32) has a derived structure like (33) at the time of pronominalization - in this case, the last cycle - then Noun Phrases could be deleted in the sequence described below.

33)



In (33) there is no environment where pronominalization of the sort exemplified in (30) and (31) can occur. It is therefore necessary to choose either the leftmost or the highest possible antecedent. If the leftmost possible antecedent, Mary_i^1, is taken as NP_j, a grammatical sentence should be the result of deleting Mary_i^2 and Mary_i^3,

assuming that pronominalization is otherwise acceptable in analogous environments. This assumption seems justified:

34) Even after Bill had told Mary_i^1 the truth, James was believed by her_i^2 to have destroyed the republic.

35) Even after Bill had told Mary_i^1 the truth, he was believed by most people to have told her_i^3 nothing but lies.

To explain the ungrammaticality (or extreme awkwardness) of (32) it is necessary to take into account the ungrammaticality or awkwardness of the anaphoric connection between $\underline{\text{Mary}}_i^2$ and $\underline{\text{Mary}}_i^3$ (that is, between the two occurrences of $\underline{\text{her}}_i$):

36) *James was believed by Mary_i^2 to have hit her_i^3.

37) *James was believed by her_i^2 to have hit her_i^3.

38) *James was believed by her_i^2 to have hit Mary_i^3.

The Ambiguity Avoidance Rule still fails to do that, despite its emendations.[2]

As far as I can see, the only additional revision of the rules governing deletion pronominalization which could handle cases like (32) would be one requiring a pre-deletion check of all possible applications of the pronominalization rule. Only those deletions would be permitted which involved nodes belonging to a set of nodes all of whose members had been determined by the check to be potentially connectable (anaphorically) to each other. This revision would be exactly the additional network-forming device whose use was to be avoided by subscribing to deletion pronominalization.

The structure of sentences like (13), (16), (20), and (32) is such that the kind of evidence they provide and the arguments they suggest apply to other theories of pronominalization besides cyclic deletion. Post-cyclic deletion, for example, would quite obviously suffer from all the defects discussed above. All the crucial examples presented in this chapter were cases where all pronominalization would have taken place in

the last cycle. With respect to pronominalization nothing basic would change in the interval between the last cycle and the post-cyclic part of the derivation.

Post-cyclic filters, whose use has been proposed by Lakoff, will accept sentences like (16) unless an extra device is added to check networks.

16) *Mary's hitting him_i^2 proved to him_i^1 that $John_i^3$ had better leave.

According to Lakoff's definition[3], $\underline{him}_i^2$ can be the antecedent of $\underline{him}_i^1$, since "NP_i (=$\underline{him}_i^2$) and NP_j (=$\underline{him}_i^1$) are identical in form and NP_i precedes NP_j." Furthermore, $\underline{John}_i^3$ can be the antecedent of $\underline{him}_i^2$; it cannot, however, be the antecedent of $\underline{him}_i^1$. Nothing in the filters themselves would reject (16) if it were claimed that the sentence should be understood as having only the reading associated with the two grammatical pronoun-antecedent pairs: $\underline{him}_i^1$ - $\underline{him}_i^2$ and $\underline{him}_i^2$ - $\underline{John}_i^3$. But there is no such reading. A provision might be added which states that if NP_i^1 - NP_i^2 and NP_i^2 - NP_i^3 have been

accepted as grammatical pronoun-antecedent pairs, then either NP_i^1 - NP_i^3 must also be acceptable, or the sentence is ungrammatical. This is a network-formation rule.

In Jackendoff's theory of cyclic interpretive coreference assignment, (16) causes the same kind of problems it causes post-cyclic filtering. The way Jackendoff supplements the ordinary rules of pronominalization to avoid these difficulties will be discussed in the next chapter. He adopts a particular kind of network-formation rule. Even if his rules applied post-cyclically they would have to be supplemented in a similar way.

If all theories of pronominalization require a special rule or device whose function is to check or form networks, a study of network formation is justified in paying somewhat less attention than is normal to the details of the rules which state whether or not NP_i can, in isolation, be the antecedent of NP_j. Knowledge of whether the anaphoric connection is grammatical is, of course, necessary, but this knowledge can be used with little regard for the way it is expressed in a

particular theory. The study of networks can be relatively self-contained, and it should not suffer from being isolated in this thesis from the many other issues in the theory of anaphora.

Chapter Two

COREFERENCE AND ANTECEDENCY

A discussion of the way Jackendoff explains the ungrammaticality of (16) will provide some insight into the use of coreferentiality as a basic concept in the theory of pronominalization. Arguments will then be presented against exploiting the logical properties of coreferentiality as a check in the formation of networks, and an alternative will be proposed.

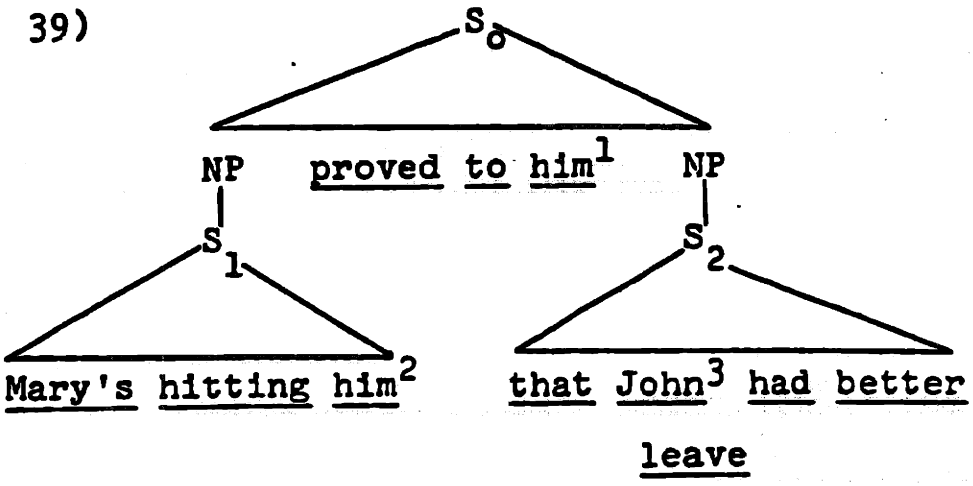
How networks are formed or checked has received detailed discussion only in Jackendoff (1969), where the concepts of coreferentiality and semantic interpretation are made responsible for insuring that all nodes which should be in a network are in one. The generation of sentence (16), for example, would be prevented as follows:

16) *Mary's hitting him_i^2 proved to him_i^1
that $John_i^3$ had better leave.

In Jackendoff's theory, the underlying structure of (16) is almost identical to that shown in (14), except that

in two places pronouns without indices are used instead of indexed full Noun Phrases.

39)



In S_1 and S_2 no rules of pronoun interpretation would apply. At the end of the S_o-cycle, him^2 would be marked coreferential with him^1 and with John^3; him^1 cannot be marked coreferential with John^3, so nothing is done. It is also conceivable that him^1 might be (somewhat redundantly) marked coreferential in turn with him^2, since each of the two pronouns is in a position to be the antecedent of the other. The results of coreference marking are listed:

him^2 coref with him^1
him^2 coref with John^3
him^1 coref with him^2 (?)
him^1 non-coref with John^3

Since they could not be marked coreferential, him^1 and John^3 have by convention been marked non-coreferential. The semantic component interprets the list, that is, it carries out deductions on it, and it marks for grammaticality on the basis of the deductions. If him^1 is coreferential with him^2 (which it must be anyway, if him^2 is coreferential with him^1), and if him^2 is coreferential with John^3, then it follows from the transitivity of coreferentiality that him^1 is also coreferential with John^3. But the list shows that him^1 and John^3 are not coreferential. Because the assignment of coreference by the rules of pronominalization has allowed a contradiction to be deduced, the reading which led to the contradiction is ungrammatical.

Such an account of network checking or formation, perhaps better called network interpretation within a

theory like Jackendoff's, rests on the assumption that coreference is the relevant property exhibited by pronoun-antecedent pairs. It utilizes certain definitional characteristics of this property, namely that coreferentiality is a symmetrical, reflexive, and transitive relation. This makes sense within generative grammar, given the many past discussions whose purpose was to justify the notion of intended coreference as opposed to the notion of referential identity, on the grounds that even in cases where no referent exists in the real world the pronoun and its antecedent behave jointly with respect to some sort of designation. Despite the correctness of this objection to narrow views of reference, there are, I believe, some arguments against using coreferentiality as the fundamental concept in network interpretation.

The rules of pronominalization operate over the set of (NP, VP, and S) nodes in a given sentence, and they determine whether various pairs of nodes stand in a coreference relation[4] to one another. If coreferentiality were a relation which held over the set of nodes in a sentence, one would expect deductions based on the

properties of the relation not to lead to contradictions, just as the relation "greater than" does not lead to contradictory deductions when applied to pairs of real numbers. Such contradictions do, however, arise; and Jackendoff's network-interpreting device, a vital part of his theory, makes crucial use of them. One might therefore conclude that the assumption that coreferentiality holds over the set of nodes in a sentence is false. If it were false, deductions based on the transitivity of coreference (him^1 is coreferential with him^2; him^2 is coreferential with John^3; therefore him^1 is coreferential with John^3) would be illicit, and network interpretation would not work.

This argument is unfortunately only suggestive, since instead of rejecting the coreferentiality relation, it is also possible to reject the assumption that the set of NP nodes in an ungrammatical sentence is the kind of set over which coreferentiality holds. That is, it can be claimed that the nodes in such sentences are not analogous to the real numbers mentioned above. If they are not, there is no reason to expect that deductions based on the properties of coreferentiality will be

non-contradictory. The evidence in favor of this position is the perfect correspondence between ungrammatical sentences and sentences where contradictions of the form "A is coreferential with B and A is not coreferential with B" can be deduced. It is difficult to determine the strength of this evidence, since it would appear that the only thing responsible for the ungrammaticality of the sentences in question (such as 16) is the application of the rules of pronominalization, that is, of the rules assigning coreference. Readings of sentence (16) other than the one discussed are, for example, perfectly acceptable:

40) Mary's hitting him_i^2 proved to him_j^1 that $John_i^3$ had better leave.

When him^2 and $John^3$ are coreferential and him^1 refers to someone else, there appears to be nothing ungrammatical about the distribution of NP nodes in sentences like (16/40). If there is no characterization of the ungrammatical reading of (16) which is independent of the assignment of coreference, then the correspondence mentioned above becomes more like a stipulation than like evidence;

rejecting the assumption that coreferentiality holds over sets of nodes would thus be slightly more justified than denying the analogy between the nodes and real numbers.

The following argument provides a further reason for believing that coreferentiality should not be the basis for network interpretation[5]: Network interpretation in Jackendoff's theory relies on producing contradictory readings. One would expect sentences whose contradictory readings have been derived in the course of network interpretation to behave like sentences which have obtained their contradictory readings in other ways. Sentences (41) and (42) exemplify the latter sort of contradiction:

41) It is raining here and now and it is not raining here and now.

42) These married men are not married.

Sentences like (41) and (42) are not meaningless; they have a definite truth value, namely False. They can also be used legitimately - if uninterestingly - in arguments; with a contradictory sentence as a premise, any and

all conclusions will follow. Sentences like (43), however, have none of these properties.

43) *Because he^1_i said that $Amos_i$ was a linguist, he^2_i was laughed out of the room.

In terms of Jackendoff's theory, (43) would embody the following contradiction: by the rules of pronominalization, $\underline{he}^1$ is coreferential with $\underline{he}^2$, and $\underline{he}^2$ is coreferential with John. Therefore, $\underline{he}^1$ is coreferential with John. By the same rules of pronominalization, $\underline{he}^1$ is non-coreferential with John. (43) thus supposedly asserts that $\underline{he}^1$ and John both are and are not coreferential. It is not possible, however, for one to use (43) as a contradictory premise and to draw true conclusions from it, as one can with (41) and (42). Furthermore, (43) does not seem to have a definite truth value, either True or False. Finally, the supposed contradictory reading of (43) cannot be perceived at all. The reading gives every appearance of not existing at all, just as no reading exists for a sentence like (44), which has violated rules having no obvious

connection with the semantics of English, that is, which has violated only the syntactic norms.

44) *She$_i$ said that he$_j$ would be in Tokyo, my friend$_j$, Elaine's governess$_i$.

(44) is an example of the impossibility of applying the rule of Right Dislocation twice. Single applications produce grammatical results, as (45) and (46) indicate.

45) Elaine's governess said that he$_j$ would be in Tokyo, my friend$_j$.

46) She$_i$ said that my friend would be in Tokyo, Elaine's governess$_i$.

The difference in the behavior of (43) and (41)-(42) suggests very strongly that it is incorrect to prevent the generation of (43) and sentences like it by allowing the semantic component to produce or interpret a contradiction. And the similarity between (43) and sentences like (44) which are ungrammatical for purely syntactical reasons hints that setting up a non-semantic device for network formation is the correct way to explain

why (43) is out.

Some further evidence that network interpretation must be syntactic rather than semantic comes from the behavior of anaphoric elements which are not coreferential with their antecedents. For this kind of anaphora, coreferentiality is not available as a means for deducing contradictions, and some other solution must be found. Verb Phrase Deletion, as exemplified in (47), establishes an anaphoric connection between two Verb Phrases; this connection cannot be identified with coreferentiality in its normal sense, since the two Verb Phrases do not refer to the same "action" the way John and he in (48) can refer to the same person.

47) Sam will go to the store, and Mary probably would like to (go to the store).

48) John said that he would read it soon.

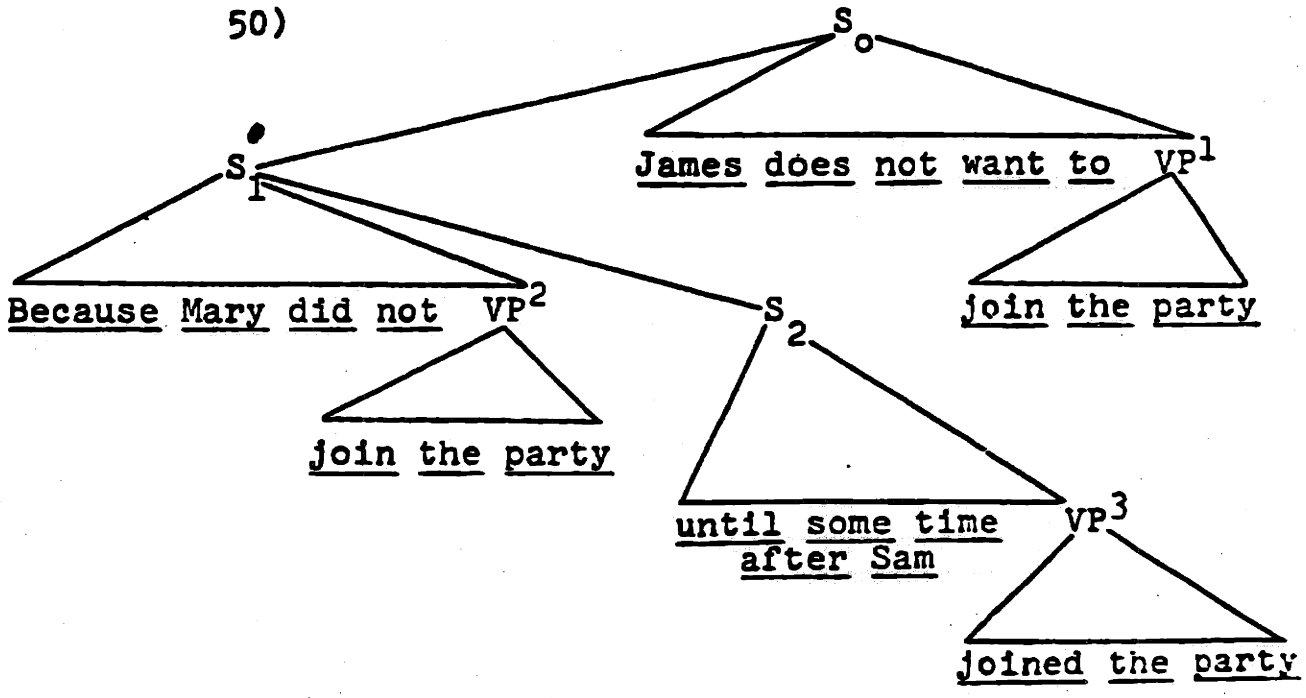
Certain sentences which have undergone multiple applications of the Verb Phrase Deletion transformation are ungrammatical on readings which should be possible,

given the normal constraints on that transformation.

49) *Because Mary didn't until some time after Sam joined the party, James doesn't want to.
(= Because Mary didn't join the party until some time after Sam joined the party, James doesn't want to join the party.)

(49) has the following structure at one point in its derivation:

50)



Sentence (51) shows that Verb Phrase Deletion can go from left to right and delete a Verb Phrase which commands the "controller" VP; i.e., that a Verb Phrase in the position of VP^1 above can be deleted by a Verb Phrase in the position of VP^3.

51) Because Mary complained after Sam joined the party, James didn't want to (join the party) anymore.

Sentence (52) indicates that right-to-left deletion is possible from a Verb Phrase in the position of VP^1 to a Verb Phrase in the position of VP^2.

52) Because Mary didn't until some time after Sam had given her money, James doesn't want to join the party.

Since Verb Phrase Deletion is not obligatory, as (53) shows, and since it ought to be able to occur in any cycle (even though post-cyclic application would not make any difference here), the sequence of deletions described below ought to generate a sentence (49) which is

grammatical.

53) Because Sam said that Mary had joined the party, James now also wants to join the party.

In the S_2-cycle, Verb Phrase Deletion cannot apply; assume that in the S_1-cycle it does not apply, because optional. Then, in the S_o-cycle, VP^1 deletes VP^2 on the pattern of sentence (52), after which VP^3 deletes VP^1 on the pattern of sentence (51). This leaves a structure which results in the ungrammatical sentence (49).

The probable reason for the ungrammaticality of (49) lies in the ungrammaticality of (54) and (55).

54) *Mary did not until some time after Sam joined the party.

55) *Because Mary did not until some time after Sam joined the party, James was upset.

Compare these two sentences with (56):

56) Because Mary did not join the party until some time after Sam did, James was upset.

Sentences (54) through (56) show that VP^2 cannot be anaphorically connected with VP^3 unless it is VP^3 which is the deleted node. This is basically the same pattern of grammaticality and ungrammaticality as is revealed by sentences (57) - (59).

57) *He_i did not come until some time after Sam_i had finished the story.

58) *Because he_i did not come until some time after Sam_i had finished the story, $James_j$ was upset.

59) Because Sam_i did not come until some time after he_i had finished the story, $James_j$ was upset.

Network interpretation uses the evidence of sentences like (57) - (59) to explain the ungrammaticality of (60).

60) *Because he_i^1 did not come until some time after Sam_i had finished, he_i^2 was upset.

As (57) and (58) indicate, $\underline{he_i^1}$ cannot be coreferential with $\underline{Sam_i}$. Thus, although $\underline{he_i^1}$ can be coreferential with $\underline{he_i^2}$, and $\underline{he_i^2}$ with $\underline{Sam_i}$, (60) is starred. The best solution to the problem of explaining the ungrammaticality of (49) would be to handle it in the same way as (60). (60), however, is rejected by deducing a contradiction from coreferentiality, and even two Verb Phrases which can be connected anaphorically are not coreferential with one another.

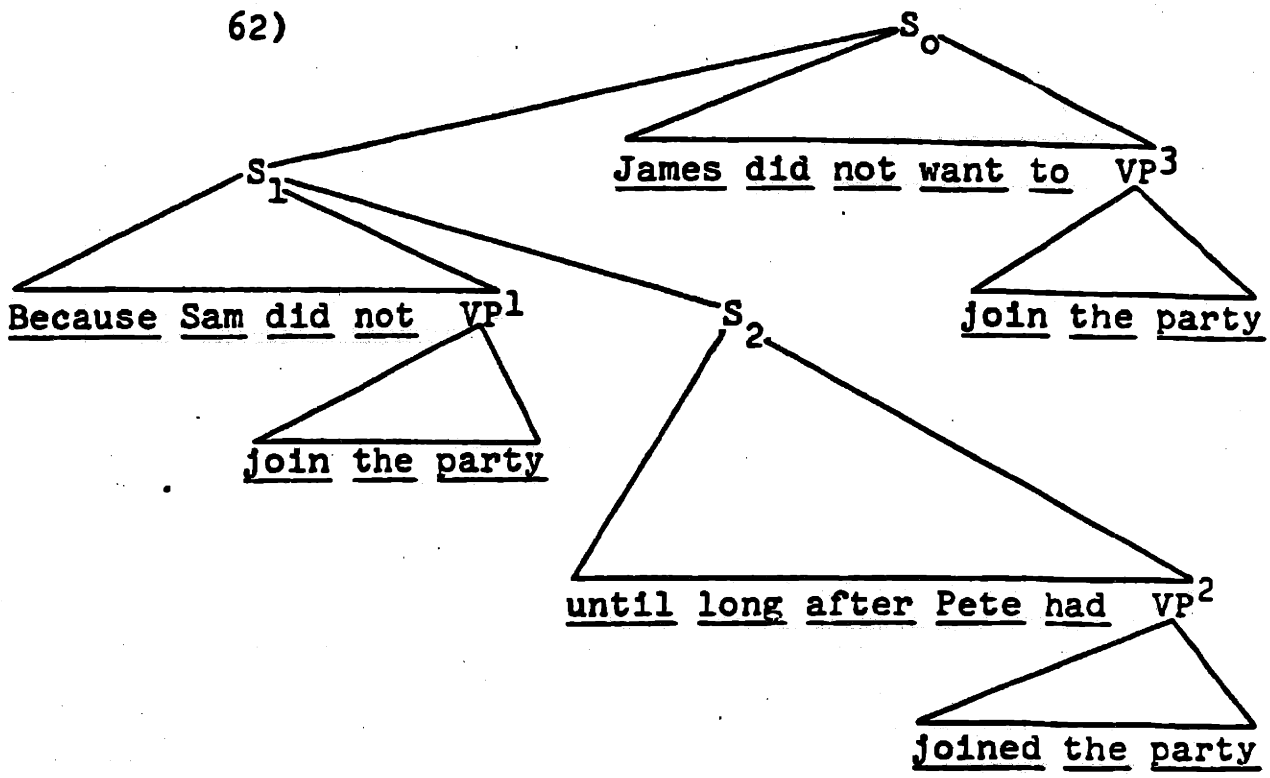
Perhaps some equivalence relation other than coreference is necessary. Since equivalence relations are all transitive, some relation such as "is an equivalent act or state to" would allow the deduction of contradictions in sentences like (49): VP^2 is an equivalent act to VP^1, and VP^1 is equivalent to VP^3; therefore VP^2 is equivalent to VP^3. But, by constraints on Verb Phrase Deletion, VP^2 is not equivalent to VP^3. This gives the desired contradiction.

There are two problems with such a solution. First, Verb Phrase Deletion has never applied between VP^2 and VP^3, so network interpretation does not have access to the information that they are not equivalent Verb Phrases. That lack of access to the information does not automatically mean that the sentence is marked ungrammatical is shown by sentences in which two equivalent deleted nodes do not have to have been deleted by the same (third) node.

61) Because Sam didn't until long after Pete had, James didn't want to join the party.

If (61) has a structure like (62) at some point in its derivation, VP^1 could delete VP^2 and could in turn be deleted by VP^3, giving (61).

62)



Even though VP^2 has not been deleted by VP^3, they are equivalent. A way must thus be found to distinguish sentences like (61) from sentences like (49), given that Verb Phrase Deletion is optional. Making sure that network interpretation knows whether two Verb Phrases can be connected anaphorically is also a problem, however, for solutions other than those using a special equivalence relation. It will discussed again later.

The second problem with the proposed equivalence relation is the following: Although (49) is ungrammatical, the series of deletions which produced it

guarantees that VP^2 and VP^3 actually are equivalent. Verb Phrase Deletion can take place only if the constituent to be deleted has the same semantic content (with the exception of pronouns) and syntactic structure as the constituent which is the controller (the node with respect to which deletion under identity occurs). Since VP^1 has deleted VP^2, they have the same semantic content and syntactic structure; and since VP^3 has deleted VP^2, they do, too. It follows that VP^2 and VP^3 must also have identical semantic content and syntactic structure, and that they therefore must be equivalent. To say that (49) is ungrammatical because VP^2 and VP^3 are not equivalent is consequently incorrect, and a rule which declares them not to be equivalent is asserting something false. It would be closer to the truth to say that they cannot, because of the rules of anaphora, be <u>understood</u> to be equivalent. That is, the anaphoric (in this case deleted) element cannot be to the left of and command its antecedent, regardless of whether other applications of any rules of pronominalization assure the equivalence of the two nodes. The conceivably perceptual rules of anaphora itself, and not the semantic notions associated with a particular kind of anaphoric connection, would

appear to be what counts in the interpretation of networks.

In a solution which bases network interpretation on the syntax of anaphora, the most obvious alternative to relations like coreferentiality and "equivalent act to", which are interpreted in the semantic component, is the relation of antecedency, which probably is not. Use of the antecedency relation was proposed most recently by Lakoff (1968), but his use of the term "antecedent" is slightly different from the one which will be adhered to in this dissertation. For Lakoff, X is the antecedent of Y if X stands higher in a certain Noun Phrase hierarchy than Y, or if X and Y occupy the same hierarchical rung and X is to the left of Y. For example, Bob is the antecedent of he because Bob is a full Noun Phrase, and Noun Phrases of that sort stand above pronouns in the hierarchy. An anaphoric relation may hold if the antecedent stands in a certain structural position with respect to the node which is not the antecedent. In this dissertation, however, X will be said to be the antecedent of (to antecede) Y only if X and Y occupy the proper positions relative to one another in

the Noun Phrase hierarchy and if an anaphoric relation also exists between them. This would seem to be more in accord with normal usage. The phrase "X antecedes Y" will be abbreviated Ant(X,Y).

Lakoff justifies the use of the antecedency relation by pointing out that certain words with anaphoric function, such as one in sentence (63), cannot be said to be coreferential with their antecedents.

63) Sam chased a butterfly, and Roberta chased one, too.

In order to have as general a theory of pronominalization as possible, one which will work for both "coreferential" and "non-coreferential" pronouns, antecedency should be taken as the basic notion, Lakoff argues. A pronoun like one would then receive a particular kind of semantic interpretation after an antecedency relation has been established between it and an indefinite Noun Phrase. Similarly, coreferentiality would be attributed to a Noun Phrase and a definite pronoun if an antecedency relation held between them. Lakoff further argues that a certain

difficulty in the formulation of Postal's Crossover Constraint could be overcome by utilizing antecedency rather than coreferentiality. My reasons for rejecting this particular argument will be presented in Chapter Three. To Lakoff's arguments can be added the arguments against coreferentiality presented so far in this chapter.

It is now necessary to discuss how antecedency can be used in a theory of network interpretation. On the basis of what the word "antecedent" means it can be concluded that the relation "is the antecedent of (antecedes)" is irreflexive. X cannot be the antecedent of X, even though it certainly is coreferential with itself. It is also non-symmetrical; it does not necessarily follow that, if X is the antecedent of Y, then Y is the antecedent of X, although it does follow that, if X is coreferential with Y, then Y is coreferential with X. Furthermore, antecedency is not transitive. Examples used throughout this dissertation (cf. sentences 16, 20, and 49) show that even if X is the antecedent of Y and Y is the antecedent of Z, nothing follows about whether X is the antecedent of Z.

Since a non-transitive relation like antecedency does not allow deductions the way coreferentiality does; and since a rule which filters out ungrammatical sentences by deriving in any way contradictions like $\sim Ant(X,Y) \wedge Ant(X,Y)$ would appear to be making claims much like the rejected coreferentiality solution; network interpretation should be set up as follows:

First networks are defined:

> Network: If $P_1, P_2, \ldots, P_n$ are nodes, $n \geq 2$, the set $\{P_1, \ldots, P_n\}$ forms a network if and only if for each P_i and all x either $Ant(P_i,P_x)$ or $Ant(P_x,P_i)$ holds, where $i = 1 \ldots n$ and $x \neq i$. Note that $n \geq 2$ above, rather than 3. 3 was used earlier for heuristic purposes.

Then a condition on sentence grammaticality is stated:

> Network Filter: For each node P_q in a sentence S_j and for all nodes X_i, $i = 1 \ldots n$,

such that either $Ant(P_q,X_i)$ or $Ant(X_i,P_q)$, S_j is grammatical if and only if the set $\{P_q, X_1,\ldots, X_i,\ldots, X_n\}$ forms a network.

An example will illustrate how the Network Filter operates.

64) *Because he_i^1 claimed that Sam_i had finished the work, he_i^2 had to leave.

Let $\underline{he}_i^1$, $\underline{he}_i^2$, and $\underline{Sam}_i$ be P_1, P_2, and P_3, respectively. P_1 (=$\underline{he}_i^1$) can have P_2 (=$\underline{he}_i^2$) as its antecedent, but neither of P_1 and P_3 (=$\underline{Sam}_i$) can have the other as an antecedent. P_2 (=$\underline{he}_i^2$) can have both P_1 and P_3 as antecedents. For P_2, therefore, both $Ant(P_1,P_2)$ and $Ant(P_3,P_2)$ hold. Sentence (64) (=S_j) will thus be grammatical only if $\{P_2, P_1, P_3\}$ forms a network. It does not, since for those three nodes to form a network, either $Ant(P_1,P_3)$ or $Ant(P_3,P_1)$ would, by definition, have to hold. Neither does, and (64) is not grammatical. Notice in particular that in order to decide on the ungrammaticality of (64) it is nowhere necessary to deduce a contradiction, which in this case would be

that $\underline{he}_i^1$ both does and does not have $\underline{Sam}_i$ as its antecedent.

Sentence (65), on the other hand, is grammatical, even though its surface structure is identical to that of (64).

65) Because he_j^1 claimed that Sam_i had finished the work, he_i^2 had to leave.

If $\underline{he}_j^1$ equals P_1, $\underline{Sam}_i$ equals P_2, and $\underline{he}_i^2$ equals P_3, the Network Filter will function as follows: For P_1 none of $Ant(P_1,P_2)$, $Ant(P_1,P_3)$, $Ant(P_2,P_1)$, $Ant(P_3,P_1)$ holds. For P_2 only $Ant(P_2,P_3)$ holds, while the only antecedency relation that P_3 participates in is $Ant(P_2,P_3)$. For P_1, there is no node X_i such that either $Ant(P_1,X_i)$ or $Ant(X_i,P_1)$ holds; the Network Filter thus lets it pass. P_2 and P_3 each have one node X_i for which an antecedency relation holds. The sets $\{P_2, P_3\}$ and $\{P_3, P_2\}$, which are of course identical, form networks. Since it is the case that in (65) networks exist wherever they are required by the Network Filter, the sentence is grammatical.

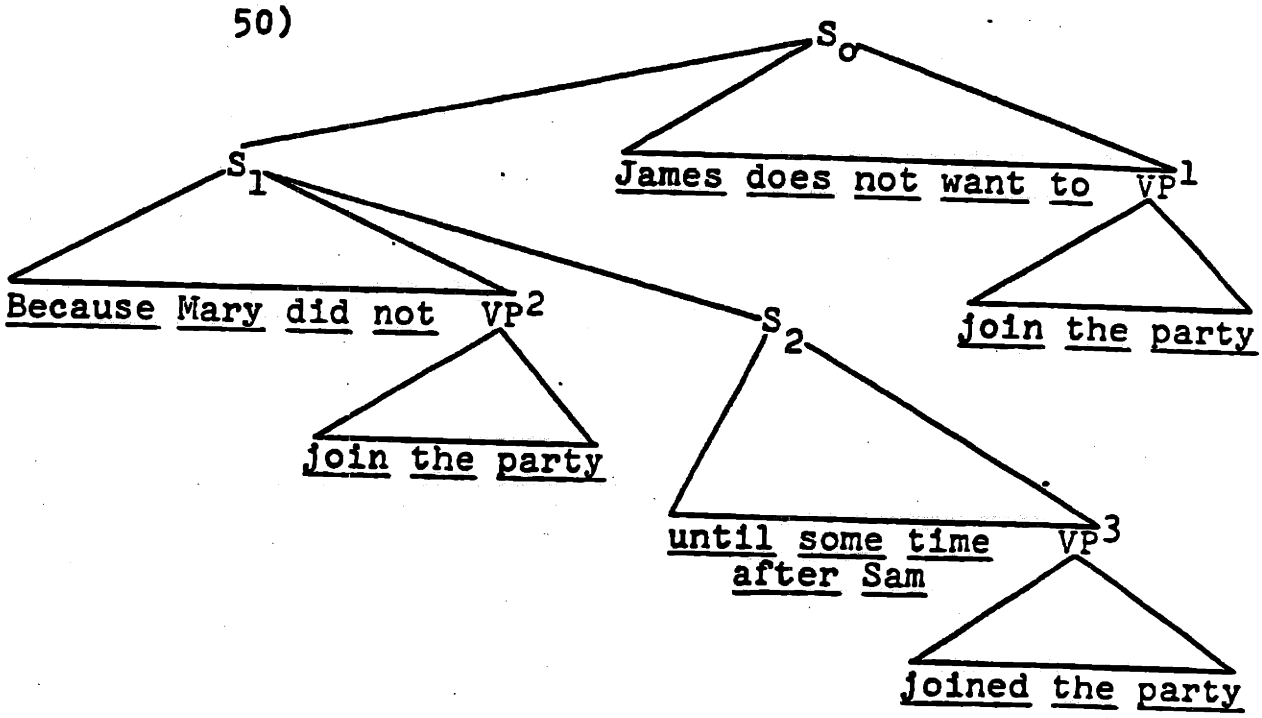
Attempting to account for the patterns of grammaticality and ungrammaticality exhibited by Verb Phrase Deletion brings up the problem mentioned earlier: how access is obtained to information about whether a node which has been deleted can have as its antecedent a node other than the one which deleted it. There are at least two ways of acquiring such access. The first way is to mark whether the Structural Description of Verb Phrase Deletion was met in any environment, regardless of whether deletion is then carried out. Instead of being deleted upon satisfaction of the SD, Verb Phrases are then marked with DOOM. In this way the structure of even those nodes which are to be deleted is available in later cycles for a check as to whether they and still higher nodes satisfy the SD. The second way to find out whether antecedency can hold is actually a variant of the first. This device has been proposed by Lakoff (1971) as one which might have many uses: it should, since it is extremely powerful. The Structural Description and the Structural Change of Verb Phrase Deletion would apply at widely separated stages of the derivation, the SD perhaps cyclically, and the SC post-cyclically, somewhere near shallow or surface structure. All situations where

deletion could apply (even those where it need not) will thus be marked before any deletions are carried out. Both ways of evading the access problem require one further condition: Assume that two nodes, VP^1 and VP^2, are marked as satisfying the SD of Verb Phrase Deletion, and that VP^2 is the antecedent or controller, while VP^1 is the node which could be deleted. If VP^1 is eventually deleted by <u>any</u> Verb Phrase, VP^3, which itself gets deleted by VP^2, then the anaphoric connection $Ant(VP^2,VP^1)$ will be considered to hold in addition to the anaphoric connections established by deletion: $Ant(VP^3,VP^1)$ and $Ant(VP^2,VP^3)$. Call the method of obtaining $Ant(VP^2,VP^1)$ indirect deletion. Only those anaphoric connections derived by either direct or indirect deletion will be said to hold.

Given such devices, it is easy to see how the Network Filter would decide on the grammaticality of sentences derived (partially) by means of Verb Phrase Deletion. For example, sentence (49), which has (50) as a stage in its derivation, would be treated as follows:

49) *Because Mary didn't until some time after Sam joined the party, James doesn't want to.

50)



In the S_2-cycle nothing happens; in the S_1-cycle it is determined that VP^2 and VP^3 satisfy the Structural Description of Verb Phrase Deletion, and that it is VP^3 which would be deleted. The SD check in the S_o-cycle determines that VP^1 satisfies the SD of Verb Phrase Deletion with respect to both VP^2 and VP^3: it could delete or be deleted by either one of them. When the deletions which lead to (49) are carried out, VP^2 is

deleted by VP^1, and VP^1 is in turn deleted by VP^3. Since VP^2 was not the node of the VP^2-VP^3 pair which would or could have been deleted upon satisfaction of the SD, the sequence of deletions which is assumed to have derived (49) did not generate an indirect deletion of VP^2 with respect to VP^3. $Ant(VP^3,VP^2)$ therefore does not hold. For the operation of the Network Filter, VP^1 is the most significant node in this example. There are two nodes X_i with which VP^1 $(=P_q)$ stands in an antecedency relation: VP^2 $(=X_1)$, for which $Ant(VP^1,VP^2)$ holds; and VP^3 $(=X_2)$, for which $Ant(VP^3,VP^1)$ holds. For (49) to be grammatical the Network Filter thus requires that the set $\{VP^1, VP^2, VP^3\}$ $(= \{P_q, X_1, X_2\})$ form a network. It is clear, however, that it does not, since the above discussion has shown that neither $Ant(VP^2,VP^3)$ - because the SD was not satisfied - nor $Ant(VP^3,VP^2)$ - because that is an impossible indirect deletion - holds. (49) is therefore ungrammatical, as desired. How a grammatical sentence is filtered should be apparent.

One could argue that the treatment of Verb Phrase Deletion offered here is unnecessarily complicated, and that this complication, namely the solution

to the access problem, results from assuming that there is such a transformation as Verb Phrase Deletion. What really happens, so the argument would run, is that the surface structures of sentences in which Verb Phrases have apparently been deleted are generated with empty Verb Phrase nodes in the base. These empty nodes are then treated just like pronouns, and the Network Filter would apply to them in the normal way.

Bresnan (1971) has provided evidence, however, which supports the claim that deletion actually occurs. Her arguments will be summarized here: Sentences like (66) are peculiar in that, although they are grammatical, the pronoun it is missing an antecedent.

66) My uncle didn't buy anything for Christmas, but my aunt did, and it was bright red. (Bresnan 5)

Anything cannot be the antecedent, as (67) shows, and nothing else is available on the surface.

67) *My uncle didn't buy anything for Christmas, and it was bright red. (Bresnan 6)

The obvious way out is to assume that (66) comes from (68) and that pronominalization is so ordered that the antecedent is still there when it applies.

68) My uncle didn't buy anything for Christmas, but my aunt did buy something for Christmas, and it was bright red. (Bresnan 7)

A solution which assumed a null pronoun in the Verb Phrase position would have to allow another pronoun to have as its antecedent a Noun Phrase which would be where the null Verb Phrase stands if that Verb Phrase were fully expanded to correspond to its antecedent. Thus, in (66) above, it would have as its antecedent the Noun Phrase something (out of which anything has resulted in the first Verb Phrase because of negation) because that Noun Phrase is part of the interpretation of did. This kind of solution comes close to what Lakoff has called a

transderivational constraint. Sentences like (69), however, appear to render such a solution untenable.

69) *My uncle didn't buy anything for Christmas, so my aunt did it for him, and it was bright red. (Bresnan 9)

If part of the content of a pronoun's interpretation could provide a missing antecedent, one would expect overt and null pronouns to behave identically. The ungrammaticality of (69), in which the it immediately following did is probably not derived by deletion, shows that they do not (and is in fact independent evidence for deleting Verb Phrases but interpreting it). If Verb Phrases are required to be spelled out in deep structure in order to explain the divergent behavior of overt and null pronouns, then the problem of access discussed above is not a pseudoproblem.

Now that the application of the Network Filter has been discussed and exemplified, it is appropriate to consider some further consequences of its use. In particular, anaphoric elements which are normally assumed

to be much more restricted in their behavior than normal pronouns, namely reflexives and the Noun Phrases deleted by the EQUI-NP Deletion transformation, will be shown to behave exactly as normal pronouns do with respect to the Network Filter. This will provide an interesting test of the Filter's generality and prove that even anaphoric rules with a tightly circumscribed range (Reflexivization and EQUI) operate within the general system of anaphora, rather than outside of and parallel to it. It is the essence of the Network Filter whose generality will be tested; this essence may be informally restated as the Pairwise Antecedency Constraint:

> In order for an anaphoric element (pronoun, etc.) to belong to a network and for a sentence to be grammatical, the element must participate in an antecedency relation with every other node in the network.

The ramifications of the Pairwise Antecedency Constraint are the subject of the next chapter.

Chapter Three

THE PAIRWISE ANTECEDENCY CONSTRAINT

Before the evidence concerning Reflexivization and EQUI is presented, the Pairwise Antecedency Constraint must be defended against the criticism found in Witten (1971). Witten denies that the constraint exists. He presents his arguments while discussing Jackendoff's coreference-based theory of pronominalization, but they will be seen to apply also to a theory of pronominalization which takes antecedency to be fundamental.

Witten finds sentences like (70) acceptable, and sentences like (71) and (72) unacceptable.

70) Jane beat Sam_i up, according to his_i father.

71) *According to his_i father, Jane beat Sam_i up.

72) *Jane, according to his_i father, beat Sam_i up.

He then notes that (73) seems perfectly acceptable to him.

73) According to his_i^1 father, his_i^2 friends beat Sam_i up.

If pairwise coreference (or antecedency) were necessary, he argues, (73) should also be out, since its structure is analogous to that of (71). $\underline{His}_i^1$ in (73) corresponds to $\underline{his}_i$ in (71); if $\underline{his}_i$ cannot have $\underline{Sam}_i$ as its antecedent in (71), then $\underline{his}_i^1$ should not be able to have $\underline{Sam}_i$ as its antecedent in (73). Because $\underline{him}_i^1$, $\underline{him}_i^2$, and $\underline{Sam}_i$ would not form a network if $\underline{Sam}_i$ did not antecede $\underline{him}_i^1$, (73) would not pass the Network Filter and would be starred. Since it is perfectly grammatical, Witten concludes that the requirement of pairwise coreference (the Pairwise Antecedency Constraint) should be dropped. I do not agree that (71) and (72) are unacceptable. More importantly, however, something special seems to be going on in (73).

Switches in grammaticality or acceptability similar to that claimed by Witten for (71) and (73) also occur when the potential antecedent is one of two or more coordinate nodes. Thus, (74) is unacceptable for many people if $\underline{he}_i$ is unstressed, but sentences (75) and

(76) - and the stressed version of (74) - are totally acceptable for the same people.

74) *Heinrich$_i$ and Sybille said that he$_i$ was going away.

75) Both Heinrich$_i$ and Sybille said that he$_i$ was going away.

76) Heinrich and his^1_i wife said that he^2_i was going away.

Apparently, there are both anaphoric (as in 76) and non-anaphoric (as in 75) ways of making a previously inaccessible node, such as Heinrich in (74), accessible to anaphora. Sentence (75) is particularly important, since it demonstrates that an additional (accessory) node is not necessary in order to overcome the rules of pronominalization, despite what his^2_i in (73) would lead one to believe. Arguing on the evidence of (75), one could say that for people disinclined to accept (71) and (72), the additional pronoun his^2_i in (73) in some way deflects attention away from the Noun Phrase immediately adjacent to the according to phrase, and that this deflection is necessary because according to phrases are

usually in some way very closely connected to the nearest Noun Phrase.

In support of this hypothesis, note that (77) and (78) have some of the (mild) awkwardness found in (71) and (72).

77) Jane_i beat Sam up, according to her_i father.

78) Jane_i heard Sam, according to her_i father, at three o'clock.

Sentences (79) and (80), in which the <u>according to</u> phrase is closely connected to the adjacent Noun Phrase, sound more natural.

79) Jane_i, according to her_i father, beat Sam up.

80) According to his_i father, Sam_i heard Jane leave.

(81) and (82), which employ accessory pronouns, also sound somewhat less awkward than (77) and (78).

81) Jane$_i$ beat her$_i$ brother up, according to her$_i$ father.

82) Jane$_i$ heard her$_i$ brother, according to her$_i$ father, at three o'clock.

If the existence of some sort of connection to a given Noun Phrase actually is what is important in the use of <u>according to</u>, (81)-(82) might sound as natural as (79)-(80) because the accessory pronoun establishes exactly such a connection (a substitute for adjacency). In this case, the "deflection" mentioned above would then have a simple syntactic explanation.

Furthermore, even when <u>according to</u> has no connection to the adjacent Noun Phrase some non-anaphoric devices can prevent the sentences from sounding awkward.

83) According to his$_i$ father at any rate, Jane beat Sam$_i$ up.

84) Jane$_i$ beat Sam up, according to her$_i$ father at any rate.

Thus, whatever is involved in the supposedly greater

naturalness of (73) in comparison to (71) and (72), forces are apparently at work which have no bearing on the Pairwise Antecedency Constraint. An explanation must rather be sought in terms of stress, adjacency, length, etc.

Notice that even if Witten's arguments were granted, more damage would be suffered by a theory based on coreference than by one based on antecedency. The anaphoric connections whose existence in (71) and (72) Witten denies can be deduced by transitivity from the (coreferential) anaphoric connections he assumes in (73). Thus, Witten's evidence, if accepted, would undermine the fundamental notion of Jackendoff's theory: the theory would have to allow a contradictory sentence (contradictory in terms of network interpretation) to be grammatical. The same evidence would force a theory based on antecedency only to add an exception clause to the Network Filter, since no contradiction could be deduced.

Assuming now that the Pairwise Antecedency Constraint has been successfully defended against Witten's

criticisms, it remains for this chapter to show what that constraint reveals about Reflexivization and EQUI-NP Deletion. Although both these rules appear to establish anaphoric connections between nodes, they differ from regular pronominalization in that they have a much more restricted range of application. EQUI insures that the subject (the controlled NP) of one sentence is identical to a Noun Phrase (the controller) in some higher sentence, and that the controlled Noun Phrase does not appear in surface structure. Sentence (85) is a typical example.

85) Being in the army bored Sam.

The sentential subject of (85), <u>being in the army</u>, has as its (understood) subject <u>Sam</u>, which has been deleted by the overt occurrence of <u>Sam</u> in the matrix sentence. Not every overt Noun Phrase in a higher sentence can serve as the controller.

86) *John$_i$ said that ($\emptyset_i$) being in the army bored Sam$_j$. ($\emptyset_i$ indicates that the understood subject of <u>being in the army</u> is supposed to be <u>John$_i$</u>)

Pronouns, however, can often have as antecedents the Noun Phrases inaccessible to the controlled Noun Phrase.

87) $John_i$ said that his_i being in the army annoyed Sam_j.

Reflexivization is similarly restricted. For regular (non-emphatic, non-picture-noun) Reflexivization to take place, the reflexive form (or the node to become a reflexive) must be in the same clause as its antecedent.

88) $John_i$ hit $himself_i$.

89) *$John_i$ said that Mary had hit $himself_i$.

The constraints on picture-noun reflexives are looser,

90) $John_i$ said that a picture of $himself_i$ had been stolen.

but they are still stricter than those on regular pronominalization.

91) *$John_i$ said that $Bill_j$ had stolen a picture of $himself_i$.

92) $John_i$ said that $Bill_j$ had stolen a picture of him_i.

Given the restricted range of EQUI and Reflexivization, what should be expected in those cases where the controlled Noun Phrase or the reflexive form has as its antecedent a pronoun which itself has as an antecedent a Noun Phrase which is outside that range? The Pairwise Antecedency Constraint requires that the controlled Noun Phrase and the reflexive form must also have the out-of-range Noun Phrase as an antecedent if a network is to be formed and the sentence is to be grammatical. Rules of anaphora beyond EQUI and Reflexivization must apply to them, and they must be treated <u>throughout the derivation as if they were regular pronouns</u>.

Trivial cases of this have been known for a long time in connection with EQUI. Sentence (93) is such a case, one almost like sentence (8) above.

93) Discovering that Mary knew Sam was in the room surprised him.

Him is the controller Noun Phrase which has caused the deletion of the understood subject of discovering. Normally, him could stand in an antecedency relation with Sam, but in (93) it cannot, since the rules of pronominalization would prevent the subject of discovering from having Sam as an antecedent. The set of nodes composed of the controlled Noun Phrase, Sam, and him would not form a network, and the sentence would be filtered out. On the basis of (93), however, it is not necessary to conclude that the controlled Noun Phrase must be checked throughout the derivation for participation in antecedency relations. It might, for example, appear reasonable to claim that it is subject to the normal rules of pronominalization up to, but not including, the cycle in which EQUI applies, from which point on it enters into only those antecedency relations entered into by the controller. This is certainly what the original conception of EQUI as a deletion transformation would predict.

Postal and Jackendoff have both argued against the hypothesis that EQUI deletes complement subjects under identity. Postal (1970) proposed that the controlled

Noun Phrase be a real pronoun, since like real pronouns, it cannot be to the left of a [-Specific] indefinite Noun Phrase.

94) *Going to the movies during the day excited sm (=some, reduced stress) old ladies.

EQUI would apply cyclically and mark the pronoun subject (of going to the movies...) with [+DOOM], and at or near the end of the derivation the pronoun would be deleted. Deletion would not occur earlier, since in Postal's theory the rules governing pronominalization also do not apply until late in the derivation. Jackendoff (1969) suggested that the use of a null pronoun (Δ) would be more appropriate; he argued that DOOM-marking is too powerful a device. Sentences would be marked grammatical if all Δ's have received an interpretation by the end of the derivation; otherwise they will be rejected.

Although they represent some progress over earlier conceptions of EQUI, the theories of both Postal and Jackendoff are vague as to which Noun Phrases, if

any, can serve as antecedents of the controlled Noun Phrase both in the cycle in which a given application of EQUI takes place (to be called from now on "the EQUI cycle") and later. As far as I can tell, however, on an uncharitable interpretation Postal's theory is consistent with the view stated above: that after the cycle has been reached in which EQUI will apply, the controlled Noun Phrase has only indirect (non-pairwise) anaphoric connections with other Noun Phrases, namely those established by the controller.[6] Jackendoff's theory is not consistent with the view stated above, since in it, pronominalization is cyclic and precedes EQUI. Jackendoff's theory is therefore totally vague only about what happens after the EQUI cycle.

That the controlled Noun Phrase must, however, be marked for participation in antecedency relations at least through the EQUI cycle is indicated by sentences like (95).

95) *In John's$_i$ house, ($\emptyset_i$) getting up early annoys him$_i$.

Him$_i$ cannot have John$_i$ as its antecedent, even though sentences like (96) are grammatical.

96) In John's$_i$ house, my getting up early annoys him$_i$.

Given (96), the ungrammaticality of (95) can best be accounted for by assuming that, for some reason, the understood subject of getting up cannot have John$_i$ as its antecedent. But by EQUI him$_i$ is obligatorily an antecedent of the subject of getting up (i.e., of the controlled Noun Phrase); if it itself has John$_i$ as an antecedent, as in (95), the Network Filter will throw the sentence out, for the following reason: The set of nodes made up of John$_i$, him$_i$, and the controlled Noun Phrase contains one member, him$_i$, which stands in an antecedency relation to each of the other two. Both Ant(him$_i$,controlled NP) and Ant(John$_i$,him$_i$) hold. Therefore, by definition, the other two members of the set, John$_i$ and the controlled Noun Phrase, must stand in an antecedency relation to one another if the sentence is to be grammatical. By hypothesis (above), they do not; (95) is starred.

An example of a similar sort is provided by (97).

97) *In John's$_i$ house, by (Ø$_i$) closing the windows it is possible for him$_i$ to sleep.

Some variety of EQUI has applied in (97), deleting the subject of closing under identity with him$_i$. As in (95), him$_i$ cannot have John$_i$ as its antecedent, despite the grammaticality of closely related sentences like (98) and (99).

98) In John's$_i$ house it is possible for him$_i$ to sleep.

99) In John's$_i$ house, my opening the windows makes it possible for him$_i$ to sleep.

Direct parallels to (97), with only a change in the subject of the by-phrase, are unavailable, since EQUI apparently must apply.

100) *By my opening the windows it is possible for him to sleep.

That constructions with a preposed by-phrase are possible at all is shown by (101) and (102).

101) By ($\emptyset_i$) opening the windows it is possible for him$_i$ to sleep.

102) In Mildred's house, by ($\emptyset_i$) opening the windows it is possible for him$_i$ to sleep.

An explanation for the ungrammaticality of (97) which is parallel to that suggested for (95) would seem to be highly plausible: him$_i$ antecedes ($\emptyset_i$), by EQUI, and is itself anteceded by John's$_i$; the Network Filter thus requires that ($\emptyset_i$) and John's$_i$ stand in an antecedency relation to one another. The evidence of sentences (98) through (102) suggests that such a relation cannot hold; the Network Filter thus must throw (97) out.

If the above explanations for the ungrammaticality

of (95) and (97) are accepted, then for the following reason the controlled Noun Phrase must enter normally into antecedency relations at least through the EQUI cycle. Adverb Preposing, which is responsible for positioning the phrase in John's house at the front of both sentences, must precede pronominalization. This is true even in the theory propounded in Lakoff (1968), according to which pronominalization is a post-cyclic filter. The filter must apply after all movement rules to obtain the correct results. If pronominalization did not follow Adverb Preposing, sentences like (103) could not be derived.

103) In the house which Sam_i owns in the country he_i likes to catch centipedes.

104) *He_i likes to catch centipedes in the house which Sam_i owns in the country.

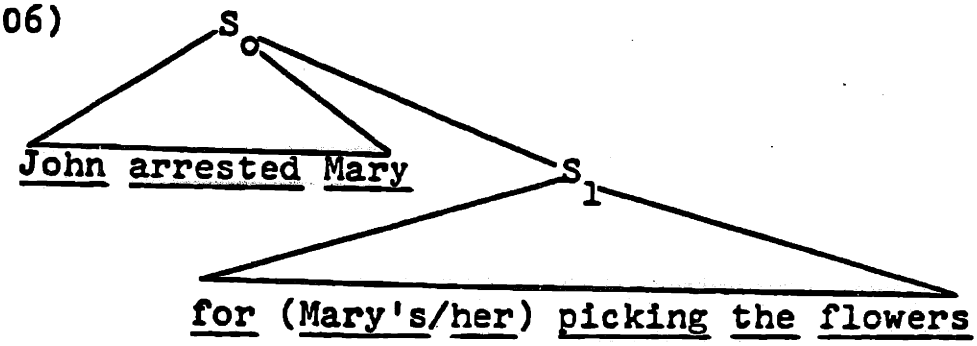
Adverb Preposing cannot be earlier than cyclic; thus, in (95) it cannot apply earlier than the cycle containing the verb annoys, which in John's house modifies.

In (97), it cannot apply earlier than the cycle containing the adjective possible. In both cases the cycle of application is the EQUI cycle or later. Given that the rules of pronominalization apply to preposed phrases after they have been preposed, it then follows that the determination of whether the controlled Noun Phrases in (95) and (97) can be anteceded by a Noun Phrase in the preposed adverb must also take place in the EQUI cycle or later. The determination that this particular antecedency relation does not hold in (95) and (97) is an essential part of the explanations provided above for their ungrammaticality. If the controlled Noun Phrases had all and only the antecedents of the controller Noun Phrases during and after the EQUI cycle, (95) and (97) would be grammatical, since in both sentences the controller can be anteceded by the relevant Noun Phrase in the preposed adverb. The rules of pronominalization must therefore apply to the controlled Noun Phrases at least during the EQUI cycle, and perhaps later; the view that controlled Noun Phrases are marked for antecedency relations independently of controller Noun Phrases only for the cycles preceding the EQUI cycle must be dismissed.

It is more difficult to prove convincingly that what happens in even later cycles affects the controlled Noun Phrase directly. The structure of English is such that, if something is done to a node on a higher clause, it will almost always have the same effect (in terms of antecedency relations) on both the controller and the controlled Noun Phrases somewhere down in lower clauses. But of course the crucial case is one in which only the controlled Noun Phrase is (negatively) affected. Certain not uncontroversial sentences may, however, provide the desired evidence. Consider (105), which has (106) as an underlying structure.

105) John arrested Mary for picking the flowers.

106)



S_1 is connected to S_o by EQUI. The verb arrest requires that its object be the controller if the subject of a for-

phrase is to be deleted by EQUI. If the subject could be the controller, (107) should be grammatical (with Reflexivization applying in the S_1-cycle under identity with the subject of hitting: John), but it is not.

107) *John$_i$ arrested Mary for ($\emptyset_i$) hitting himself$_i$.

For many people, sentence (105) is also grammatical if the adverbial phrase (for...) is preposed.

108) For picking the flowers John arrested Mary.

Even after fronting, the understood subject of picking the flowers remains Mary.

109) *For ($\emptyset_i$) hitting himself$_i$ John$_i$ arrested Mary.

Moving the for-phrase to a position immediately after John seems, on the other hand, to be impossible, or at least extremely awkward, despite the fact that other

adverbials, such as because-clauses, may be placed there.

110) *John - for picking the flowers - arrested Mary.

111) ?John - because she had annoyed him - arrested Mary.

A few people find relative clauses based on sentences like (108) to be acceptable for purposes of emphasis.

112) I met the man who for picking the flowers arrested Mary.

Because (110) is out, I assume that (112) could only come from a sentence in which the for-phrase had been moved to the front of the relative-clause-to-be before Relative Clause Formation applied.

Observe now what happens to (108) if it is attached as a relative clause to a Noun Phrase whose head is Mary.

113) *I met Mary, who for picking the flowers John (had) arrested.

Contrast this with (112) and (114).

114) I met Mary, who John (had) arrested for picking flowers.

Apparently, what is responsible for the ungrammaticality of (113) is the fact that the controller (the antecedent) has crossed over the controlled Noun Phrase (the pronoun). For those people who find (112) unacceptable, (113) is even worse, almost uninterpretable in fact. The assumption of a crossover effect in (113) is therefore necessary to explain extreme, clear, and consistently reported degrees of grammaticality or acceptability, as well as to account for the distinction between certain grammatical and certain ungrammatical sentences. The existence of a crossover effect in sentences related to (113) will prove crucial; (113) shows that the existence of the effect in such sentences is not accidental.

There is one other conceivable explanation for the ungrammaticality of (113). John Ross has suggested to me that (113) is out because it is impossible to have preposed adverbs in relative clauses. That this is

incorrect, at least for some speakers, is shown by the acceptability of the sentences immediately below.

115) I met Mary, who because it was raining John had arrested.

116) I met Sam, who without any warning John had decided to arrest.

117) I spoke with my lawyer, who in the middle of the day the police had decided to arrest.

(115) through (117) are all better than (113), even if they are not especially graceful. The assumption of a crossover effect remains the best explanation for the ungrammaticality of (113).

Reflexives are also possible as the object of arrested in sentences like (105);

118) John arrested himself for picking the flowers.

and parallels to (108) are grammatical for some people.

119) For picking the flowers John arrested himself.

120) For causing great harm to the nation John turned himself in.

If (119) is used as a relative clause, an ungrammatical sentence results.

121) *I met the man who for picking the flowers arrested himself.

For those who find sentences like (115)-(117) acceptable, the grammaticality of (119) should imply the grammaticality of (121). Crossover accounts for the failure of these expectations. The only way a crossover effect could arise is if who has been moved over a pronoun for which it is or could be the antecedent, that is, if there is something for it to cross over, as was the case in (113). Since it obviously could not have crossed over himself, who must be considered to have crossed over the understood subject of picking.

That the understood subject has actually been

crossed over is guaranteed by the ungrammaticality of (110). As further evidence for this point, consider (122).

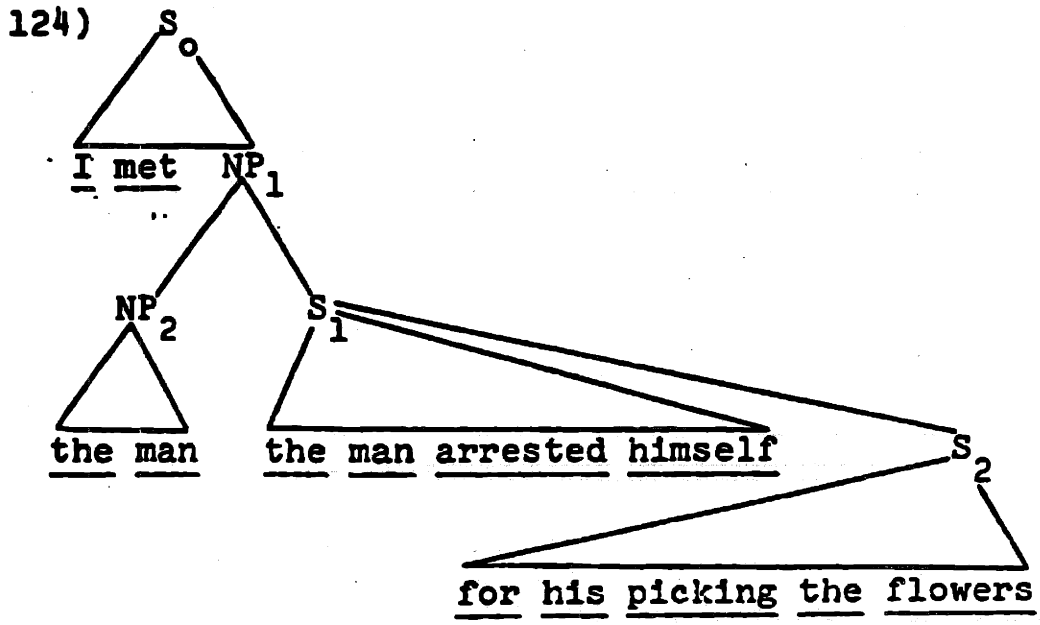
122) I met the man who because he had done great harm to the nation turned himself in.

Sentence (111) showed that a <u>because</u>-clause could be positioned after the subject of <u>arrested</u>. If the relative clause in (122) is derived from something like (111) rather than from something like (123), the pronoun subject of <u>had done great harm</u> will not have been crossed; in that case, (122) should be grammatical, as it is.

123) Because he_i had done great harm to the nation the man_i turned $himself_i$ in.

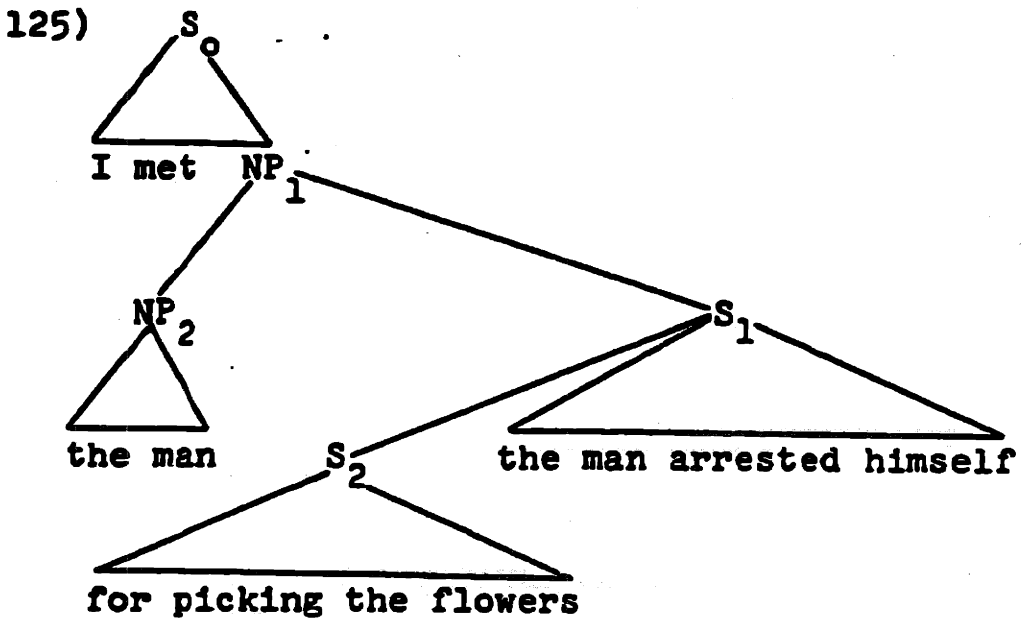
Sentences (119) and (121) have the following significance: It has already been shown that the understood subject of the <u>for</u>-phrase can be connected by EQUI

only to the object of arrest. The object of arrested in (119) is a reflexive and therefore (in this case) obligatorily has the subject of arrested as its antecedent. In (121) it is the subject, and not the object, of arrested which crosses over the understood subject of the for-phrase; that is, the Noun Phrase which is moved is not the one which served as the controller when EQUI applied. Relative Clause Formation applies in (121) in the cycle following the EQUI cycle, if it is cyclic; this can be seen from (124). If it is post-cyclic, it applies even later; this would only strengthen the argument being presented here.



In the S_1-cycle, EQUI and Reflexivization apply and S_2 is preposed, giving (125) as an intermediate structure.

125)



The Structural Description of Relative Clause Formation is not satisfied until the cycle is reached which contains the head noun, that is, until the S_0-cycle (or perhaps the NP_1-cycle). Since Relative Clause Formation moves <u>the man</u> over <u>for picking the flowers</u>, crossover also occurs in the cycle after the one in which EQUI has applied. Note here the importance of having Adverb Preposing apply before Relative Clause Formation; this ordering has been justified above in the discussion of sentences (110)-(113) and (115)-(117), and of (122)-(123) as well.

Crossover can have an effect only on those pronouns which could conceivably be marked for participation in an antecedency relation with the Noun Phrase being moved. Since it is not the controller which is moved in the derivation of (121), the controlled Noun Phrase must be markable with respect to Noun Phrases other than the controller; and since the crossover occurs in a cycle later than the EQUI cycle, the determination of whether the controlled Noun Phrase is anteceded by the moved Noun Phrase must also be made in a cycle later than the EQUI cycle. This means that the controlled Noun Phrase must be treated like a regular pronoun in whatever cycle the later cycle might be, as was to be shown. It was demonstrated earlier that it must be treated as a regular pronoun both before and during the EQUI cycle; it therefore follows that it must be so treated throughout the derivation. The Pairwise Antecedency Constraint, it will be remembered, predicted exactly this result.

The Network Filter characterizes (121) as ungrammatical in the normal way. Himself stands as controller in an antecedency relation with the controlled

Noun Phrase, and as anaphoric element in an antecedency relation with the subject of arrested. For (121) to be grammatical, then, the set composed of the three nodes himself, the controlled Noun Phrase, and the man (who) must form a network. It does not, since for the controlled Noun Phrase and the man (who) neither Ant(controlled NP,the man) nor Ant(the man,controlled NP) holds. (121) is therefore rejected.[7]

In addition to supporting the Pairwise Antecedency Constraint, the result just obtained has certain consequences for the formulation of EQUI. As was mentioned earlier, Jackendoff utilizes a null pronoun in his theory of how EQUI operates. Sentences are rejected if the null pronoun has not received an interpretation by the end of the derivation. If it were true that the controlled Noun Phrase participated in the normal rules of pronominalization only up to the cycle in which EQUI applies, the null pronoun would probably receive an interpretation only by virtue of a permissible application of EQUI, for the following reasons: It makes sense to say that a pronoun which is an antecedent does not receive an interpretation from the pronoun which refers

to it. If EQUI is to be carried out, nothing in the clause containing the controlled Noun Phrase can be to the left of it.

126) *At two o'clock going downtown bothered John. (on the reading "Going downtown at two o'clock bothered John.")

There can thus be no antecedent to the left of the controlled Noun Phrase and in the same clause; that is, left-to-right pronominalization in that clause with the controlled Noun Phrase as anaphoric element is impossible. Furthermore, because the controlled Noun Phrase is always the subject, and all cycles preceding the EQUI cycle are lower in the tree, right-to-left pronominalization also cannot apply. Only left-to-right pronominalization with the controlled Noun Phrase (the Δ) as antecedent will produce grammatical results, as in (127).

127) Δ_i knowing that he_i would be deleted amused Sam_i.

Only the clause Δ_i knowing that he_i would be deleted is is relevant in (127), since it is the only part of the sentence which would be treated on a cycle lower than the EQUI cycle. By hypothesis, since the null pronoun will be serving only as an antecedent before the EQUI cycle, it will not receive an interpretation before that cycle - in the case of (127), before the cycle containing amused Sam. EQUI would then apply, providing the desired interpretation.

By otherwise neglecting what happens during and after the EQUI cycle, as has been done in the literature, one could conclude that EQUI also provides the only interpretation. But it has just been shown that the controlled Noun Phrase must be subject to the normal rules of pronominalization throughout the derivation. At some point during any of an infinite number of derivations, a full Noun Phrase will be in the proper structural position to serve as the antecedent of the null pronoun, which will then receive an interpretation. Even if EQUI never applied, the fact that the null pronoun is always subject to the normal rules of pronominalization would insure it an interpretation in an infinite number of sentences.

Because of this, ungrammatical sentences like (128) and (129) would not be thrown out by Jackendoff's constraint on the use of the Delta.

128) *Sam$_i$ said that Δ_i went to town.

129) *James$_i$ said that Mary had kissed Δ_i.

To rescue a theory using null pronouns it thus becomes necessary to stipulate that a sentence with a Delta is acceptable only if the Delta has acquired an interpretation by way of EQUI. This, of course, is equivalent to marking a regular pronoun with DOOM. The supposed advantages of the null-pronoun theory over DOOM-marking disappear.

In order to decide whether there is any reason to distinguish between the theories of Jackendoff and Postal, some way must be found to show that one of the following two possibilities is in better accord with the facts:

a) Permitting only non-null pronouns and allowing their deletion only under certain

general conditions;

b) Enlarging the stock of pronouns and greatly restricting the range of interpretation of some of them.

I cannot resolve these issues. Note, however, that the putative restrictions on the interpretation of null pronouns would not receive support from the fact that (130), with an overt pronoun, is ambiguous, while (131), with a null pronoun, is not.[8]

130) Sam_i disliked $his_{i,j}$ being elected janitor for the party.

131) Sam_i disliked Δ_i being elected janitor for the party.

EQUI simply establishes an obligatory pronoun-antecedent pair, so sentence (131) should be unambiguous regardless of which solution is ultimately chosen. In terms of the Network Filter, establishing an obligatory antecedency relation may, in fact, be the major or even sole function of EQUI. Regardless of what the other antecedency relations might be, Ant(controller,controlled NP) will

hold, and any network which contains one of those two nodes must contain the other, plus the additional antecedency relations in which the other participates. This significantly reduces the number of possible networks in comparison to the situation where no node is obligatorily bound to another and where a node brings only its own load of antecedency relations along. Besides reducing the number of nodes in surface structure by deletion, EQUI may thus also make sentences easier to understand by limiting the number of possible readings, which is based on the number of possible networks.

It now remains to be shown that reflexives also satisfy the Pairwise Antecedency Constraint. I have not been able to find a structural configuration containing regular reflexives which would allow a test of the Constraint, but one containing picture-noun reflexives is available. In all relevant details it is exactly what the proper kind of example with regular reflexives would be.

For many people both (132) and (133) are grammatical.

132) A picture of him_i could get $Sam's_i$ parents out of jail for him_i.

133) A picture of $himself_i$ could get $Sam's_i$ parents out of jail for him_i.

Since even the rules governing picture-noun reflexivization do not permit a reflexive to have as its antecedent a genitive modifying a head noun (see 134)[9],

134) *A picture of $himself_i$ could get $Sam's_i$ parents out of jail.

the antecedent of $\underline{himself_i}$ in (133) must be $\underline{him_i}$, and not $\underline{Sam's_i}$. That $\underline{him_i}$ can serve by itself as the antecedent of a picture-noun reflexive is shown by (135).

135) A picture of $himself_i$ could get Bernice's parents out of jail for him_i.

A theory of pronominalization based on (non-transitive) antecedency, but without the Pairwise Antecedency Constraint, might describe (133) by stating: With $\underline{Sam's_i}$

outside the range of Reflexivization, himself$_i$ has him$_i$ as its sole antecedent; this could be expressed by Reflex(him$_i$,himself$_i$). Him$_i$ participates normally in pronominalization and has Sam's$_i$ as its antecedent, giving Ant(Sam's$_i$,him$_i$). Assume a pronoun receives an interpretation from the constituent which is its antecedent. Then him$_i$ will receive an interpretation from Sam's$_i$. Because Sam's$_i$ is not the antecedent of himself$_i$, himself$_i$ will have to acquire its interpretation mediately, through him$_i$ by way of the reflexivization relation (Reflex). From this indirect manner of providing himself$_i$ with an interpretation it would seem to follow that, whatever interpretation him$_i$ can have, himself$_i$ will have it also.

Sentence (136) indicates that this is not the case.

136) *Yesterday James told the man$_i$
whose$_i$ parents I had said a picture
of himself$_i$ would get out of jail
for him$_i$ to go to a photographer.

(136) should not be confused with (137), which is a

grammatical sentence with the same surface structure as (136), but with a different reading.

137) Yesterday James told the man$_j$ whose$_j$ parents I had said a picture of himself$_i$ would get out of jail for him$_i$ to go to a photographer.

(138) makes it clear that, in sentences like (136), him$_i$ can normally be anteceded by whose$_i$.

138) Yesterday James told the man$_i$ whose$_i$ parents I had said a picture of Alice would get out of jail for him$_i$ to go to a photographer.

Since him$_i$ can have whose$_i$ as its antecedent in (136), and since himself$_i$ can normally be anteceded by him$_i$ (see 133), the hypothetical theory of pronominalization sketched on the preceding page would predict that himself$_i$ could receive the same interpretation as him$_i$ and whose$_i$. If that were so, (136) would be grammatical. Because (136) obviously is not grammatical, that make-believe

theory, which tried to do without the Pairwise Antecedency Constraint, cannot be accepted.

The ungrammaticality of (136) is probably due to a crossover effect. This hypothesis receives some support from the ungrammaticality of (139), which is identical to (136), except that $\underline{\text{himself}}_i$ has been replaced by $\underline{\text{him}}_i$, and $\underline{\text{the man}}$ by $\underline{\text{Sam}}_i$.

139) *Yesterday James told Sam_i, whose_i parents I had said a picture of him_i^1 would get out of jail for him_i^2, to go to a photographer.

Since $\underline{\text{him}}_i^1$ is expected as a regular pronoun to obey the Pairwise Antecedency Constraint, it must participate in antecedency relations with every Noun Phrase that stands in an antecedency relation to $\underline{\text{him}}_i^2$. This follows from the assumption, expressed by the identical indices, that $\underline{\text{him}}_i^1$ and $\underline{\text{him}}_i^2$ are themselves anaphorically connected; that they can be is shown by (132). If $\underline{\text{whose}}_i$ is to be the antecedent of $\underline{\text{him}}_i^2$ (and it can, since it has not crossed over $\underline{\text{him}}_i^2$), it must therefore also stand in either of the

relations Ant(whose$_i$,him^1_i), Ant(him^1_i,whose$_i$) to him^1_i. Only the first of these two relations could hold, but because whose$_i$ has crossed over him^1_i even it does not; the Network Filter thus rejects (139).

For an explanation in terms of crossover to apply equally well to a sentence containing a reflexive (sentence 136) and a sentence containing a regular pronoun (sentence 139), the relevant properties of the regular pronoun and the reflexive must be assumed to be the same. In other words, the normal rules of pronominalization should also apply to reflexive forms. Proof that this is true would show that the Pairwise Antecèdency Constraint generalizes to reflexives, as well as to the controlled Noun Phrase of EQUI; showing this was given earlier as one of the major aims of this chapter.

A slight change of perspective will bring the desired proof somewhat closer. Crossover makes certain anaphoric connections impossible. Suppose that reflexives were affected only by the rules of Reflexivization. There would thus be no direct anaphoric connection between himself$_i$ and Sam's$_i$ in (133), or between himself$_i$ and

whose$_i$ in the structure underlying (136) before Relative Clause Formation has applied. After Relative Clause Formation has caused whose$_i$ to cross over himself$_i$, the Crossover Principle will declare that no direct anaphoric connection can exist between whose$_i$ and himself$_i$. But this situation - the nonexistence of an anaphoric connection - is no different from the situation before crossover, and there is thus no way to distinguish the grammatical (133) from the ungrammatical (136). To avoid this difficulty, it is necessary either to reject the supposition that reflexives are affected only by the rules of Reflexivization or to add an exception clause to the Network Filter. This clause would state that no direct anaphoric connection need exist between a reflexive and the other members of the network (beyond the one Noun Phrase picked as an antecedent by Reflexivization) except when the reflexive form has been crossed over by a member of the putative network, in which case the absence of a direct connection will lead to the network's being characterized as ungrammatical. Since this is just a verbal trick, rejecting the supposition mentioned above is the preferable solution; it is also a solution which preserves the generality in the application of the Pairwise Antecedency Constraint. In addition, it is in

accord with the results obtained during the discussion of (139).

It might be objected that the generality of the Pairwise Antecedency Constraint is not at all what has been demonstrated. What has been revealed, so the argument runs, is that the use of antecedency rather than coreferentiality is misguided. If the reflexive ($\underline{\text{himself}}_i$) is coreferential with its antecedent (in the case of (136) a pronoun), and if that pronoun is coreferential with its own antecedent ($\underline{\text{whose}}_i$), then the reflexive must also be coreferential with $\underline{\text{whose}}_i$. The Crossover Principle would stipulate that an anaphoric element cannot be coreferential with a node which has crossed over it. $\underline{\text{Himself}}_i$ would then be both coreferential and non-coreferential with $\underline{\text{whose}}_i$, and the sentence would be declared ungrammatical because of the contradiction.

This objection suffers from a modified version of the problem discussed on the preceding page. The theory based on coreferentiality also requires that checks be made on the basis of the Pairwise Antecedency Constraint, or Pairwise Coreferentiality Constraint. In (133), $\underline{\text{himself}}_i$ would be declared coreferential with $\underline{\text{him}}_i$, by

Reflexivization; and $\underline{him}_i$ would be declared by pronominalization to be coreferential with $\underline{Sam's}_i$. But if reflexives are marked directly for coreferentiality only by Reflexivization, $\underline{himself}_i$ and $\underline{Sam's}_i$ would not be so marked with respect to one another. The convention mentioned earlier that nodes unmarked for direct (non-deduced) coreferentiality with respect to one another are automatically interpreted as [-coreferential] would then proclaim $\underline{Sam's}_i$ and $\underline{himself}_i$ to be non-coreferential. Remember, however, that $\underline{himself}_i$ is coreferential with $\underline{him}_i$, and $\underline{him}_i$ with $\underline{Sam's}_i$; it follows that $\underline{himself}_i$ and $\underline{Sam's}_i$ must also be coreferential, giving a contradiction. The use of coreferentiality instead of antecedency would thus cause the grammatical (133), as well as the ungrammatical (136), to be thrown out. The choices available are the same as those discussed two pages above: (133) can be passed while (136) is rejected if either the supposition is dropped that reflexives are marked for coreferentiality only by Reflexivization, or an exception clause is added to the Network Filter. This clause would exempt reflexives from the convention which automatically interprets two nodes as [-coreferential] if they are not directly marked as [+coreferential]. Thus, $\underline{himself}_i$

need not be marked by convention as non-coreferential with Sam's$_i$ in (133), but it may be marked directly as non-coreferential with whose$_i$ in (136) by the Crossover Principle. That is, a distinction will be made between marking non-coreferentiality directly (as by Crossover) and marking it by convention.

Despite the introduction of this distinction by adding the exception clause to the Network Filter, the following sentence would be declared grammatical, even though it is ungrammatical: the Network Filter in a coreference-based theory still makes false predictions.

140) *A picture of himself$_i$ could get someone's$_i$ parents out of jail for him$_i$. (where someone is [-specific] and [-generic])

The ungrammaticality of (140) contrasts with the grammaticality of (141).

141) Your picture could get someone's$_i$ parents out of jail for him$_i$.

Him$_i$ is therefore not responsible for the rejection of (140). Sentence (142) shows that himself$_i$ probably is.

142) *A picture of him$_i$ could get someone's$_i$ parents out of jail for him$_i$.

(142) is out for the same reason that (143) is.

143) *A picture of him$_i$ could get someone$_i$ ([-specific]) out of jail.

Definite pronouns cannot be to the left of a [-definite, -specific] antecedent. (141) is grammatical because there the antecedent is to the left of the pronoun. If himself$_i$ has no direct anaphoric connection with someone's$_i$, because someone's$_i$, like Sam's$_i$ and whose$_i$, is outside the range of Reflexivization, there is no reason for the ungrammaticality of (140). Himself$_i$ can have him$_i$ as its antecedent, as (135) has shown, and someone's$_i$ can be the antecedent of him$_i$, on the evidence of (141). The exception clause on the Network Filter does not require himself$_i$ and someone's$_i$ to be marked non-coreferential;

therefore no contradiction arises when the coreferentiality of himself$_i$ and someone's$_i$ is deduced from the coreferentiality of himself$_i$ with him$_i$, and of him$_i$ with someone's$_i$. Since each of the three nodes in question (himself$_i$, him$_i$, and someone's$_i$) is by symmetry and transitivity coreferential with the other two, they should all form a network. If they did, (140) would be grammatical. But it is not; nothing other than a violation of the constraints on anaphora seems to be responsible for the ungrammaticality of (140). Therefore, the three nodes do not form a network. If no network has been formed, the questions arises: Why not?

In the case of (140), the Crossover Principle is not available to declare himself$_i$ and someone's$_i$ noncoreferential, since nothing has been moved. It must be assumed that (140) is out for the same reason that (142) is: a definite pronoun cannot be to the left of its antecedent if the antecedent is [-definite, -specific]. Thus, the regular rules of pronominalization must also apply to reflexives; given that they do, the Pairwise Antecedency Constraint can apply without an exception clause and determine in the normal manner whether a network has been

formed. The generalizability of the Pairwise Antecedency Constraint to reflexives has thus been demonstrated. One way of accounting for the fact that this generalization is possible is to assume that reflexives are analyzed as pronoun+self. That reflexives undergo the normal rules of pronominalization would then be no surprise; that their distribution is also much more restricted than that of normal pronouns would simply be the result of Reflexivization, a special rule triggered by self.[10] If the Pairwise Antecedency Constraint applies to reflexives in languages which give no reason to believe that pronoun+some morpheme is the correct analysis, however, then it need only be assumed that reflexives are pronouns, regardless of their morphology. The demonstration that the Constraint generalizes depended in no way on the form of the words involved.

The results obtained above for Reflexivization provide the grounds, mentioned in Chapter Two, for my disagreement with one of Lakoff's arguments in favor of antecedency. Lakoff wants to explain the following pattern of grammaticality.

144) Sam talked to Mary_i about herself_i.

145) *Sam talked about Mary_i to herself_i.

146) Sam_i talked to himself^1_i about himself^2_i.

147) Sam_i talked about himself^2_i to himself^1_i.

When a prepositional object is the antecedent of the reflexive, only the order to...about... is possible. If, however, the subject of the sentence is the antecedent of the two reflexives, both to...about... and about...to... are grammatical. Lakoff proposes that use of the notion "antecedent" could account for this. Assume to...about... to be the underlying order. (145) involves the movement of what is to become the antecedent of the reflexive over that very reflexive; Lakoff suggests that the Crossover Principle accounts for its ungrammaticality. In (147), he argues, the antecedent of himself^1_i has not crossed over it, since its antecedent is Sam_i. Crossover therefore does not apply, and the sentence is grammatical. The Pairwise Antecedency Constraint, however, requires that himself^2_i either antecede himself^1_i or be anteceded by it, and that Sam_i antecede them both, since otherwise a network would not be formed and the sentence would be ungrammatical. Although the structure of sentences like (144)-(147) is

such that the Constraint cannot be demonstrated to hold, the evidence that has been presented so far in its favor seems to warrant the assumption that it does. Antecedency therefore cannot explain the difference in grammaticality between (145) and (147). What constrains the movement of prepositional phrases must remain an open question; I have no evidence for an alternative solution. Perhaps in some cases one antecedent is more important than the others.

The arguments about Reflexivization and the evidence for them, when combined with the arguments and evidence concerning EQUI, make one additional point which was mentioned earlier, this one in the Introduction. All of the connections among the nodes of a network are syntactically significant. The interaction of the Pairwise Antecedency Constraint, the Crossover Principle, the constraints on pronominalization, and the operation of EQUI and Reflexivization shows that. Networks and the Network Filter are more than an expression of a vague feeling that anaphorically connected nodes all somehow stand for the same thing or belong together. If one connection of the (at least, depending on the situation

and on how one counts) n(n-1)/2 in a network is missing, the entire sentence is ungrammatical, regardless of how obvious it is what the nature of the missing connection must be. The arguments about coreferentiality in Chapter Two bolster the claim that this ungrammaticality is syntactic in origin, rather than semantic. The next chapter gets away from the total concern with networks which has characterized this study up until now and discusses another way anaphoric connections affect and are affected by the rest of the syntactic component.

Chapter Four

A CONSTRAINT ON THE INTERNAL STRUCTURE OF ANTECEDENTS

A concern with the syntactic, rather than the semantic, aspects of pronominalization is the natural outcome of the use of antecedency as the fundamental relation in a theory of anaphora. More emphasis is placed on the position of antecedent and pronoun with respect to one another in the sentence and on the actual presence of pronouns and antecedents, and less on expanding the power of the semantics to interpret certain kinds of anaphoric processes, such as EQUI. The emphasis is obviously a matter of degree, since semantics plays an important role in determining what the content of a given anaphoric connection might be, as in McCawley's infamous sentence:

148) I dreamt that I was Brigitte Bardot
and that I kissed me.

Even in Lakoff's counterpart sentences, however, there are interesting examples of the importance of simple pronominal syntax.[11]

149) In my mind Nixon and Humphrey became one person, and he said things I could not understand.

In some way he refers to both Nixon and Humphrey, when this reference is considered semantically. But syntactically there is no way of referring to this strange Nixon/Humphrey entity without using a Noun Phrase like person as an intermediary. Sentences like (150), which one would expect to be grammatical if semantic interpretation by itself could create counterparts, are not.

150) *In my mind Nixon and Humphrey merged, and he said things I could not understand.

The counterpart must be concretized before it can be referred to. What the semantics does in this case is assign the properties of the hybrid counterpart to the intermediary Noun Phrase, where they then become accessible to anaphora.

On the other hand, the emphasis on syntax

certainly does not necessarily imply that everything contained in the constituent structure of the antecedent must in all cases be part of the interpretation or meaning of the pronoun. In particular, it does not follow from this emphasis that a surface pronoun referring to a sentence must have exactly the same interpretation as that sentence. (151) shows that surface pronouns indeed do not need to have an interpretation identical to that of their antecedents.

151) I might frolic and jump when
I eat a garlic at the carnival,
but it would seem out of place
in a church.

(151) can mean "I might frolic and jump when I eat a garlic at the carnival, but frolicking and jumping when I eat a garlic would seem out of place in a church." In order for (151) to have this meaning, it would have to have as its antecedent almost everything before the comma. If it has at least frolic and jump when I eat a garlic as its antecedent, then the prepositional phrase at the carnival must also be a part of that antecedent,

on the assumption that pronouns have constituents as antecedents. The phrase is part of the when-clause, and it modifies eat. In order for it not to be part of the constituent frolic...garlic, it would have to be attached to a node higher than the one immediately dominating frolic...garlic; that is, it would have to be outside the when-clause, since the when-clause, as a member of the constituent frolic...garlic is obviously not attached to a node higher than the one immediately dominating that constituent. There is no evidence that the phrase at the carnival is ever at any stage of the derivation that far away in the tree from its surface position. But if at the carnival is part of the antecedent of it, and if (151) has the meaning attributed to it above, then it follows that surface pronouns need not have exactly the same interpretations as the sentences (or perhaps Verb Phrases?) which antecede them.[12]

Within the generally syntactic orientation sketched above it is important to ask whether there are parts of phrase structures which by their very construction are disqualified from serving as antecedents (apart from such factors as proper node labeling).

It is the purpose of this chapter to point out a previously unnoticed constraint on the internal (i.e., constituent) structure of antecedents, where the word internal is to be understood in contrast to the usual kinds of external constraints on pronominalization, constraints governing the position of pronoun and antecedent with respect to one another. After the basic nature of the constraint has been made clear, different ways of accounting for it will be discussed. One of these ways will involve a generalization of the network concept. Finally, an attempt will be made to utilize the constraint as a more general replacement for Ross's Complex NP Constraint.

Sentences like (152) and (153) are perfectly grammatical.

152) I met the mother of those children, and she seemed to be perfectly normal.

153) Jeff believes that Heinrich bought a helicopter, and Sandra believes it, too.

In addition, object Noun Phrases like the mother of those children and object complements like that Heinrich bought a helicopter may have nodes taken out of them by movement transformations.

154) Which children did you meet the mother of?

155) What does Jeff believe that Heinrich bought?

But sentences like (156) and (157) are ungrammatical.

156) *The children that I met the mother of were all very charming, and I liked her, too. (cf.: The children whose mother I met were all very charming, and I liked her, too.)

157) *The helicopter which Jeff believes that Heinrich bought turned out not to exist, but Sandra believes it anyhow.

The following is a list of similar examples. Where possible, for each transformation one example each of a

sentence with a partially moved Noun Phrase (as in 156) and of a sentence with a partially moved sentential complement (as in 157) will be given.

Relative Clause Formation (see 156 and 157)

158) *The book that I was picking up the pages of turned out to be deadly boring, but I kept on hunting for them.

159) *A girl that Bob believed that he had seen has strangely disappeared, so now he doubts it.

Question Formation

160) *Which book did Tom think Bill had collated the pages of, and why did he let them fall on the floor?

161) *Which book did Tom believe that Bill had collated, and why did he think so.

As Beautiful As...

162) *As beautiful as Tom thinks Sandra is, I don't think so.

Beautiful Though...

163) *Beautiful though I think Sandra is, Emil doesn't believe it.

Topicalization

164) *My friend Sam said (that) he had seen the children of, but I have not been able to find them.

165) *My friend Sam believes (that) he has outfoxed, but my brother doesn't believe it. (cf.: Sam believes that he has outfoxed my friend, but my brother doesn't believe it. And: My friend Sam told that he had outfoxed him, but my brother doesn't believe it.)

In all of the example sentences above, the intended antecedent has had a part of itself removed by a transformation. That the mere absence of a part is not sufficient to produce ungrammaticality is shown by (166).

166) After John said that he would go to Bamberg today, Emil said he might tomorrow, but I don't believe it (that Emil will go to Bamberg tomorrow).

A portion of the Verb Phrase after _might_ has been deleted by the Verb Phrase Deletion transformation, but _it_ can still have the (sentential) object of _Emil said_ as its antecedent. It cannot be claimed that the _it_ is somehow interpreted on the basis of the information in the object of _John said_ rather than on the basis of what is in its apparent antecedent, since the _it_ can certainly be understood as meaning "that Emil will go to Bamberg tomorrow," and _tomorrow_ is contained only in the sentential object of _Emil said_.

Also, as one would expect, supposed copying transformations, such as Left and Right Dislocation (see _167_ and _168_), do not render a constituent unacceptable as an antecedent (see _169_ and _170_).

167) My brother, you will probably like him. (Left Dislocation)

168) He will probably cause you a lot of trouble, our pet dragon. (Right Dislocation)

169) My peregrine falcon, John said once that he would like to see it, and I expect him to say it again soon.

170) Heinrich told me once that she scared him to death, our lioness, and no doubt he will say it again when he house-sits.

Furthermore, movement transformations which do not involve essential variables, such as Subject Raising and Passive, do not make it impossible for the constituents from which they have removed pieces to function as antecedents.

171) His friend may have been believed by Maria to have caused some trouble, but I would never have believed it (of him).

All the transformations which produced ungrammatical sentences in the examples above are chopping transformations (Ross 1967a): transformations which move a node across an essential variable and do not leave an overt trace behind (besides a gap). A preliminary statement of the constraint on the internal structure of antecedents may thus read:

I) Internal Structure Constraint:
A constituent may not serve as an antecedent if any node properly contained within it has been moved outside of it by a chopping transformation.

The phrase "properly contained within it" is necessary, since the Noun Phrases moved by a chopping transformation may themselves still serve as antecedents, even though the constituents from which they have been extracted cannot.

172) Bob_i I met the mother of, and I later told him_i what I had done.

173) *Bob I met the mother$_i$ of, and I later met her$_i$ landlord.

It is also crucial that the node be moved outside of what is to serve as an antecedent. Within the sentential part of a relative clause construction, the Relative Clause Formation transformation may front a node from indefinitely far down, but the entire [NP S]$_{NP}$ structure of the relative may, of course, still be the antecedent of a properly located pronoun.

174) The man who Bill said that Jeff believes he met the mother of told us that he intended to cause Mary to become not single.

Similarly, if Topicalization applies within a sentential complement, the complement can still become an antecedent.

175) Sam said that Bill he had met yesterday, and if I know him he will say it again to everyone he can get his hands on.

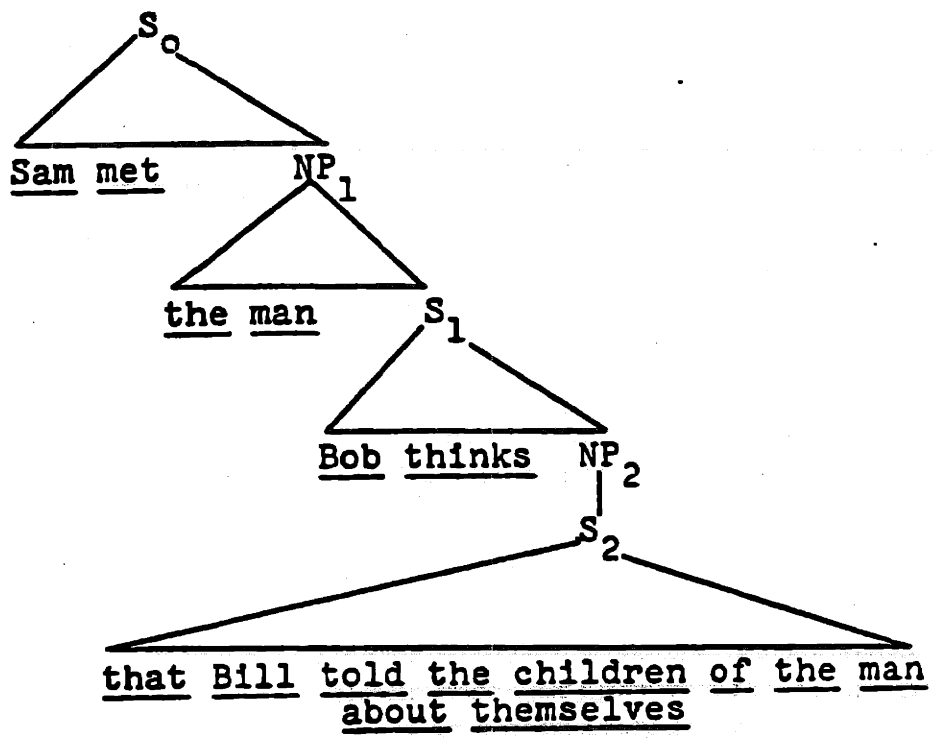
In an attempt to understand why the Internal Structure Constraint exists, one of the first possibilities that comes to mind is that it might be the result of the otherwise unconstrained interaction of correctly ordered rules. It is conceivable, for example, that by ordering all chopping rules before the determination of antecedency, or vice versa, the Internal Structure Constraint could be obtained "free." Whether this possibility is a real one will be examined below.

What happens in sentences where antecedency definitely must be determined before the application of a chopping transformation can be seen from (176).

176) *Sam met the man who Bob thinks
that Bill told the children of
about themselves.

(176) has approximately the underlying structure shown in (177).

177)



Since Reflexivization is cyclic, themselves will be marked as a permissible reflexive with respect to the children of the man in the S_2-cycle (even if the reflexive is derived by deletion instead of being lexically present in underlying structure). Relative Clause Formation, however, cannot apply until the S_o-cycle at the earliest - or perhaps in the NP_1-cycle, depending on how the issue of the NP-cycle is resolved. In either case, it will apply at some time after the application of Reflexivization. Unless Noun Phrases are marked when they become antecedents, there is no way, in sentences like (176), that the Internal

Structure Constraint can follow automatically from the ordering of Reflexivization and Relative Clause Formation. On the other hand, if it were to be decided that antecedents must be marked, (176) would provide evidence that the Internal Structure Constraint constrains chopping rules, and not the rules determining antecedency - unless, of course, there is a second antecedency check somewhere near the end of the derivation.

Sentences like (178) show that some sort of pronominalization takes place after at least one chopping transformation, Question Formation.

178) Which of the four auks that Sam_i saw did he_i like?

Before the fronting of <u>which of the four auks that Sam saw</u>, <u>he</u> could not have had <u>Sam</u> as its antecedent:

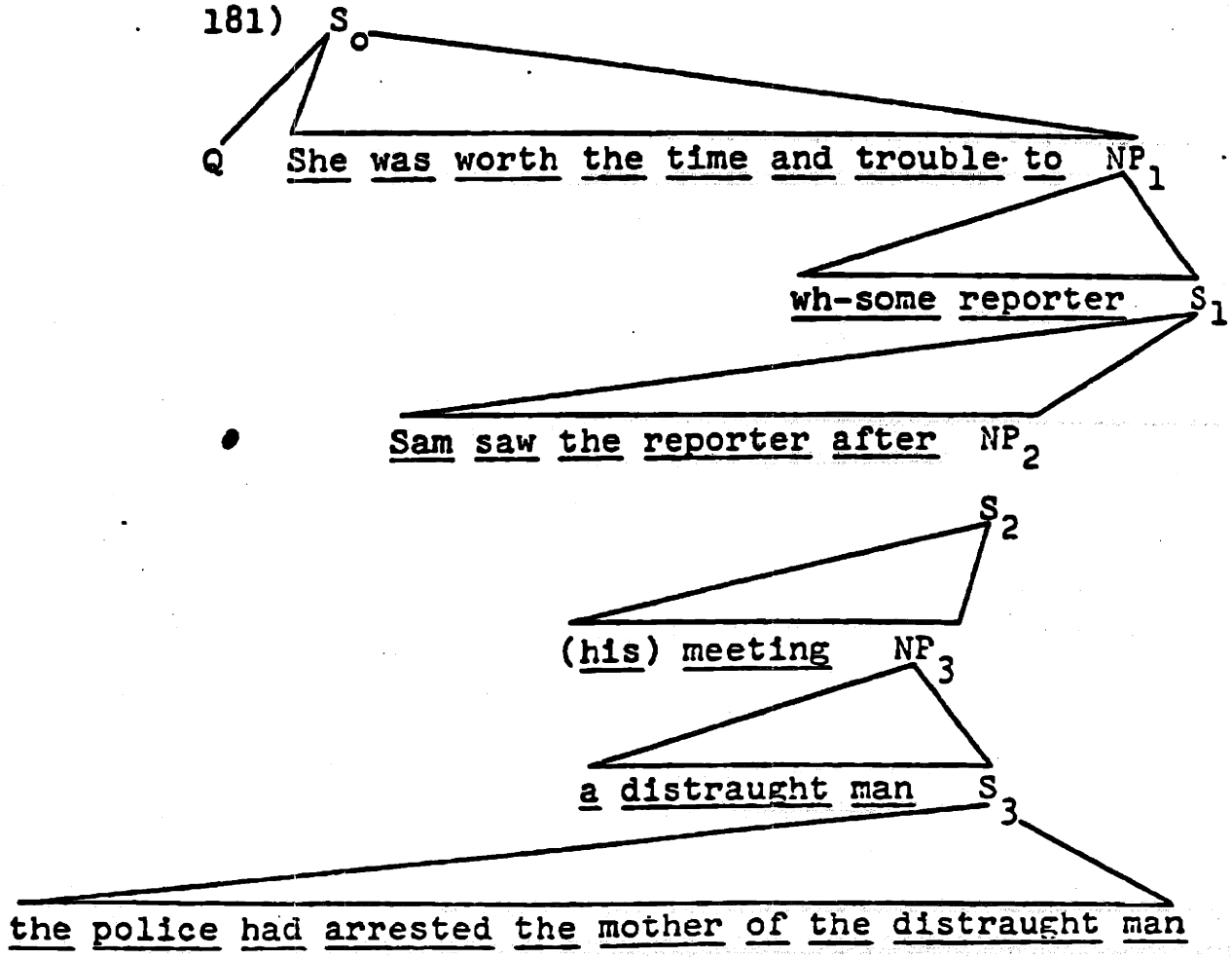
179) *He_i liked the second of the four auks that Sam_i saw.

Consider, then, sentence (180).

180) *Which reporter that Sam saw after meeting a distraught man that the police had arrested the mother of was she worth the time and trouble to?

(180) is derived from an underlying structure very roughly like (181).

181)



That (180) is out because of its derivation rather than because of its reading is shown by the grammaticality of (182) and (183).

182) Which reporter that Sam saw after meeting a distraught man whose mother the police had arrested was she worth the time and trouble to?

183) Which reporter that Sam saw after meeting a distraught man whose mother had been arrested by the police was she worth the time and trouble to?

A discussion of how (180) is derived from (181) will reveal the reason for its ungrammaticality. Note first that, as in (178)-(179), the relevant pronoun, she, cannot be marked as having the mother of the distraught man as its antecedent until that phrase has been fronted along with the questioned Noun Phrase:

184) *She$_1$ was worth the time and trouble to some reporter that Sam saw after meeting a distraught man whose mother$_1$ the police had arrested.

(182) and (183) have already shown that the relevant Noun Phrase can antecede she after fronting. The most reasonable assumption is then that (180) is out because the Internal Structure Constraint has been violated. The question arises as to how the violation occurred.

Since the fronting of the mother of the distraught man occurs when NP_1 is questioned, and since the Question Formation transformation is cyclic at the earliest, it must have been fronted in the S_o-cycle or later: NP_1 is a node in S_o, and the Q-marker is also located there. In addition, there is reason to believe that Relative Clause Formation is cyclic, since such an ordering of the rule would guarantee the correct generation of sentences like (185), where whose itself comes from a relative clause.

185) I met a man whose bowling shoes Harry had disposed of.

Assuming the cyclicity of Relative Clause Formation, then in the derivation of (180) from (181) the Noun Phrase the distraught man (in S_3) will be broken off from the Noun Phrase the mother of the distraught man during the S_2-cycle (or the NP_3-cycle), when the head noun of the relative construction is in the domain of Relative Clause Formation. This means that when Question Formation applies in the S_0-cycle or later, the phrase the mother of, and not the mother of the distraught man will be fronted along with NP_1. As a result, when an antecedent is being sought for she only the mother of will be available. Since the mother of violates the Internal Structure Constraint, (180) is ungrammatical. Because the rule which produced the mother of (Relative Clause Formation) must have proceeded Question Formation in the derivation of (180), and because Question Formation must itself have proceeded pronominalization, the Internal Structure Constraint cannot be a constraint on chopping rules: Relative Clause Formation, the relevant chopping rule, must have applied before the mother of (the distraught man) could have been made the antecedent of she. It therefore has no way of knowing it is applying to a (future) antecedent.

Given the result just obtained, it will have been shown that the ordering chopping rules/pronominalization will not give us the Internal Structure Constraint free. Sentence (176) has shown that the reverse order also does not lead to the desired result. It therefore appears that the Internal Structure Constraint must be understood as a derivational constraint on pronominalization. Any reading of a sentence is rejected if in its derivation a node is at any time marked as the antecedent of a pronoun or other anaphoric element (e.g., the null result of Verb Phrase deletion), and if that node at any time has had a proper part of itself moved by a chopping transformation.

Although a derivational constraint would certainly separate the grammatical sentences from the ungrammatical ones, there is a more elegant way of accomplishing the same thing which comes close to working. Consider three nodes: A, B, C. Suppose that B is the antecedent of C and that A is the antecedent of some part of B; call this a chain of anaphora.[13] A chain is identical to a network, and subject to the Network Filter,

if the part of B which A antecedes is not a proper part, but rather the entire node. On the basis of the above definition of a chain, the Internal Structure Constraint can be reformulated as follows:

> II) Revised Internal Structure Constraint: If any of the nodes within B and C have been deleted or chopped during the derivation, the chain and the sentence containing it are grammatical if and only if it is in principle possible to recover the missing nodes by looking at the antecedents of each node, and the antecedents of the antecedents, etc., on up the chain, in surface structure.

For example: If B has had a proper part of itself deleted, it can serve as the antecedent of C only if the missing part can be supplied by looking at the surface form of A. Sentence (166) provides an instance where exactly this happens.

166) After John said that he would go to Bamberg today, Emil said he might tomorrow, but I don't believe it (that Emil will go to Bamberg tomorrow).

A = go to Bamberg; B = Emil said he might tomorrow; C = it.. Between might and tomorrow a portion of the Verb Phrase is missing in B; if that were all, B would not be acceptable as the antecedent of C. But since go to Bamberg can be supplied in B by looking at the surface form of A (the antecedent of the deleted proper part of B), B can serve as C's antecedent after all.

It may be asked what would happen if A were not present in (166) and how the sentence resulting from its absence would be interpreted.

186) Emil said he might tomorrow, but I don't believe it.

This is the same as asking about the origin of he might tomorrow, which apparently is the product of an

application of Verb Phrase Deletion, except that nothing is visible with respect to which anything could have been deleted. In order for the Internal Structure Constraint (Revised) to succeed, it must be assumed that something actually has been deleted out of he might tomorrow with respect to an identical antecedent, and that the it is interpretable only insofar as the content of might tomorrow is interpretable with respect to its putative antecedent. It does, in fact, seem to be the case that (186) is acceptable only if an antecedent for the missing part of might tomcrrow can be found in the discourse.

Sentence (158) provides a contrasting example.

158) *The book that I was picking up the pages of turned out to be deadly boring, but I kept on hunting for them.

Pages of has no Noun Phrase as its antecedent, whether that Noun Phrase be understood or overt. Even though it is totally obvious that the pages of some book are

meant, the information that it is the book that they are pages of cannot be obtained by consulting an antecedent. The book certainly cannot be the antecedent of pages of, since a book (in the non-physical sense) and its pages are neither coreferential nor equivalent in any way.

Sentence (158) also suggests something which would follow directly from the Revised Constraint and other facts, but which would be an extra appendage to the unrevised formulation. Not only does pages of not have a real or imagined antecedent; no node like it can ever function as the "pronoun" portion of a pronoun-antecedent pair. Consequently, no such node could be the antecedent of yet a third node. Consider a sentence like (187).

187) He counted the pages of a book, even though the book he was counting the pages of was boring.

(187) is definitely grammatical, and it is clear that only one book is being talked about; but the words the pages of in the part of (187) after the comma are not understood

as an anaphoric element with respect to the pages of a book in the way that the second occurrence of Melvin in (188), for example, is understood as anaphoric with respect to the first.

188) I would take either Harry or Melvin, but Melvin is better.

Rather, it is the entire Noun Phrase the book he was counting the pages of which functions as an anaphoric element with respect to the first occurrence of book.

The weaknesses of the Revised Internal Structure Constraint lie in the exceptions which have to be admitted. The sentences generated by Subject Raising and Passive would be one class of such exceptions.

189) Marie is thought to have been believed by Harry to have visited the Black Hole of Calcutta, but I find it (that Marie has visited...) highly unlikely.

The sentences generated by Tough-Movement are another.

190) Marie is tough for people to convince, but John wants to try to.

These two classes of exceptions have in common that: (a) the Noun Phrase is moved out of a constituent which is or will become an infinitive; (b) it is moved from a definite grammatical position (subject position in the case of Subject Raising, object position - including object of preposition - for Tough-Movement) to a definite grammatical position (either subject or object position for Subject Raising, depending on the matrix verb, and subject position for Tough-Movement).

The fact that the movement is out of an infinitive suggests the following possibility: if a node is moved from a sentence and that sentence is at any time during the derivation changed into an infinitive, then the moved node may be "retrieved" and the original sentence may serve as an antecedent - provided no sentence boundary intervenes between the moved Noun Phrase and its original location. That this will not work, however, is shown by (191).

191) *The things that Harry promised Helena to buy are so ugly I don't know why he promised it.

Contrast (192):

192) The man who promised Helena to buy those things is so poor I don't know why he promised it.

In (191) a relative pronoun has been moved from object position in an infinitive (just as happens in Tough-Movement), and no S-boundary intervenes between the pronoun and its original position. In spite of that, (191) is ungrammatical. In support of the statement that no S-boundary intervenes in (191), consider (193).

193) The man who is believed by Harry to have swum the Hudson is such a hypochondriac that I can't bring myself to believe it, no matter what.

If an S-boundary intervened in (191), one would have to

intervene in (193). Either none does, or the S-boundary condition is irrelevant. Either way, the matter need not be pursued further.

Another class of apparent exceptions is provided by relative clause constructions and other complex Noun Phrases. Even when the sentential portion of such Noun Phrases has been extraposed, they may still serve as antecedents.

194) The man came in who owns the bank, and he asked to borrow some money.

195) The claim is amusing that we have done the enemy a favor, and I will not pursue it.

Such cases may be dealt with in a variety of ways; they will be discussed later.

The problems posed by Subject Raising, Tough-Movement, and Extraposition may be handled in either of two ways: (a) the derivational constraint formulated before the discussion of chaining may be revived and

amended to insure that it applies to anaphoric elements as well as antecedents; (b) chaining can be retained, and the exceptions can be made as palatable as possible. Since the first alternative is always available, an attempt will be made below to make the second one plausible. The attempt must be considered speculative, however.

Suppose that in the process of speech perception, sentences are reconstructed in stages, and suppose that among these stages are those which:

a) find the underlying location of all nodes moved by non-chopping transformations;
b) assign the semantic contents of antecedents to the pronouns which refer to them;
c) find the underlying location of chopped nodes.

Assume these stages apply in the order listed. Stage (a) would not require knowledge of which transformations have applied in order to separate chopped from unchopped nodes. As was noted above, Subject Raising and Tough-Movement

take nodes to prescribed grammatical positions. Chopping rules of the sort exemplified throughout this chapter (Relative Clause Formation, Question Formation, etc.) do not: they move nodes to the front of sentences and adjoin them in such a way that they do not take part in any of the traditional grammatical relations, such as subject of the sentence. Nodes moved by Subject Raising and Tough-Movement can thus be kept separate from chopped nodes which have moved to the left simply by observing their location in surface structure.

Nodes which have been moved to the right can easily be characterized in a perceptual model, since they never leave their sentences. Interestingly, the movement of such nodes does not render the nodes to which they were originally attached unable to be antecedents. Extraposition from NP, as in (194) and (195), is an example of a rule which moves nodes to the right. The rule which shifts prepositional phrases is another.

196) A story was in the magazine about the war, and it was horrifying.
(= A story about the war was in the magazine, and it (the story) was horrifying.)

Some minor support for a perceptual distinction (in terms of surface structure) between chopped and non-chopped nodes comes from the following observation. For some people, (197) is more acceptable than (198).

197) Mary, who is believed by many people to have probably robbed the bank, looks so innocent that I can't believe it of her.

198) ??Mary, who many people believe to have probably robbed the bank, looks so innocent that I can't believe it of her.

If there is a difference between these two sentences as to their degree of acceptability it might be explained as the result of a general constraint on derived constituent structure. Conceivably, who in (197) has not been moved by Relative Clause Formation, since it was already as far left as it needed to go even before Relative Clause Formation applied. In (198), on the other hand, who had to be chopped out of object position and moved left in order to derive a grammatical relative clause. The

general constraint which would explain why who has not been moved in (197) is a prohibition against vacuous movement. In his thesis (Ross 1967a), Ross has argued for prohibiting vacuous rightward movement, and against prohibiting vacuous leftward movement. Vacuous movement to the left is necessary if the ungrammaticality of sentences like (199) and (200) is to be explained as resulting from a violation of the Coordinate Structure Constraint.

199) *The man who and Bill visited me was an expert on Dutch syntax.

200) *Who and Sam called her up?

The Coordinate Structure Constraint disallows the chopping of a proper part of a coordinate node. If no chopping and leftward movement has occurred in (199) and (200), the CSC obviously cannot explain why they are out. If Ross were correct in positing vacuous leftward movement (and Chomsky adjunction) for rules like Relative Clause Formation and Question Formation, however, he would be unable to account for the grammaticality of (201).

201) Bill, whose uncle and Lyndon Johnson have been in the television business together since 1934, came to visit me yesterday.

The existence of sentences like (201) casts doubt on the kind of solution Ross proposes, even though the phenomenon manifested by (201) is not very general.

202) *Whose uncle and Lyndon Johnson have been in the television business since 1934?

203) *Bill, Lyndon Johnson and whose uncle have been in the television business together since 1934, came to visit me yesterday.

The status of vacuous leftward movement is at least unsettled.

(204) and (205) parallel (197) and (198), except that they are derived by Tough-Movement and Relative

Clause Formation, rather than Relative Clause Formation and Subject Raising.

204) Bill, who is tough for people to interview, is important enough to make it interesting for me to try to (interview him).

205) ??Bill, who Sam claims is tough for people to interview, is important enough to make it interesting for me to try to.

As was the case with (197) and (198), the sentence whose derivation involves obvious movement of the relative pronoun appears to be less acceptable than the one whose derivation does not. Furthermore, some people find the difference between (204) and (205) to be greater than that between (197) and (198).

Assume for the moment that (205) and (198) differ from (204) and (197) in acceptabilty. If this difference could not be explained in terms of whether or not movement has actually taken place, it would at least

be necessary to distinguish between real movement which involves crossing nodes and real movement which does not, which involves only the change in tree structure that is associated with Chomsky adjunction.[14] The perceptual model being discussed here would no longer function, since it would not be able to discriminate between the two types of movement on the basis of surface structure. The Internal Structure Constraint would then have to be formulated as a derivational constraint, and it would have to include the following condition: A constituent is incapable of serving as an antecedent not only when a node has been chopped out of it and moved left, but also when a node has been chopped and moved subsequent to its being moved out of the constituent by a non-chopping transformation, as in (205) and (198). Moving without crossing a node will not count. When formulated this way, the Internal Structure Constraint is at least as speculative as it is in its perceptual formulation. Furthermore, the perceptual model has the advantage of predicting that sentences could differ the way (197)-(198) and (204)-(205) perhaps do, given a prohibition against vacuous movement. It makes this prediction because it hypothesizes that only surface structure position is

relevant to deciding what can be an antecedent. Because it makes more predictions than does the derivational constraint, and because these predictions do not appear definitely to be incorrect, the perceptual model will be retained.

The above discussion concerns only stage (a) of the perceptual model. In stage (b), pronouns will be assigned the syntactic structure of their antecedents; this structure will later provide the basis for the pronouns' semantic interpretation. If not all of the missing nodes have been moved back into the antecedent during stage (a), the pronouns anaphorically connected to it will receive an incomplete interpretation and be declared ungrammatical. Anaphoric elements still missing nodes after stage (a) will also receive incomplete interpretations, with one exception: structures derived by Verb Phrase Deletion will be expanded according to the chaining relations which hold. Strictly speaking, however, this is not an exception. By hypothesis, in stage (b) all matters are taken care of which concern the assignment of syntactic content (what will later become semantic readings) to nodes participating in antecedency

relations of all sorts.[15]

At the end of stage (b), the only anaphoric elements with incomplete interpretations are those whose antecedents have lost nodes by way of chopping transformations, or which themselves have lost nodes in that manner. In stage (c) the missing non-vacuously moved chopped nodes are returned to their underlying positions. If its underlying position happens to be in an antecedent or anaphoric element, the return of a moved node will not prevent the sentence from being declared ungrammatical, since stage (b), where interpretations are fixed, does not apply again.

Besides the kind of argument advanced in the discussion of sentences (197), (198), (204), and (205), there is very little of a precise nature that can be said for or against the fragmentary perceptual model just sketched. It is speculative and without precedent, but it handles the data. As was mentioned above, it also provides a rationale for the differing effects of chopping and non-chopping rules on pronominalization; and it predicts, perhaps correctly, the distinction between

vacuous and non-vacuous movement. The derivational constraint which was considered does neither. In addition, it may or may not be significant that the use of ordered stages in sentence perception assigns negation a definite place in the scheme of things, whereas the derivational constraint provides no framework at all for the integration of the facts concerning the way negation interacts with movement rules.

206) Who did Sam claim that Hugo did not like?

207) Mary is easy for me not to like.

(206) and (207) show that nodes moved by either Question Formation or Tough-Movement, which represent the two basic kinds of rules under consideration, are understood to be under the scope of negation, even though they are above and to the left of the sentence containing the negative. In the perceptual model this would follow if the scope of negation were not delineated syntactically until after all nodes had been returned to their underlying positions, i.e., after stage (c).

Regardless of whether the Internal Structure Constraint is expressed in a perceptual model or as a derivational constraint, it is applicable to a somewhat wider range of cases than has been considered so far. Extraposition from NP was said above not to violate the Internal Structure Constraint because it moves nodes to the right. There may be yet another reason why it does not cause violations: relative clauses and the sentential complements of complex Noun Phrases may be anaphorically connected to the head nouns of their constructions. Just like Verb Phrases which have been deleted, complex Noun Phrases which have undergone Extraposition from NP could thus be reconstituted by utilizing chaining or something similar.

More importantly, if an anaphoric connection does exist between the head Noun Phrase and the S of complex Noun Phrases, which have the structure $[NP\ S]_{NP}$, the Internal Structure Constraint will also account for the facts which Ross adduces in support of his Complex NP Constraint (Ross 1967a). This constraint declares that a sentence is ungrammatical if a node has been chopped out of the S of an $[NP\ S]_{NP}$ structure and if the head Noun

Phrase is lexical, rather than the dummy it. That some sort of constraint is necessary to prevent movement out of relative clauses and complement sentences can be divined by noting that sentences like (208) and (209) are ungrammatical.

208) *Who did he meet the man who saw?

209) *What did he challenge the presupposition that John did?

Suppose there is an anaphoric connection between the man and who saw (who) in (208) and between the presupposition and that John did (what) in (209); then both sentences would be characterized as ungrammatical by the Internal Structure Constraint because a node has been chopped out of an anaphoric element. Admittedly, the kind of anaphoric connection being postulated is unusual, but in some sense the relative clause in sentences like (208) and the sentential complement in sentences like (209) do seem to refer to or "stand for" the head Noun Phrase. Like the second occurrence of the pages of (the book) in (187), they cannot participate in antecedency relations if they have lost a node to chopping.

Support for the hypothesis that there is an anaphoric connection between the head Noun Phrase and the S of complex Noun Phrases comes from sentences (210) and (211).

210) *The book which Sam collected the pages of which still remained on the floor later fell apart again.

211) *The man I met the girlfriend of who was beautiful was upset.

If the complement sentence is an anaphoric element, the head Noun Phrase must be an antecedent. If it is an antecedent nothing should be able to move out of it by way of a chopping rule. (210) and (211) show that, in fact, nothing can be chopped out. This fact, not noted by Ross, could be accounted for in his theory by saying that nothing can be chopped out of any part of a complex Noun Phrase. This extends the Complex NP Constraint from the S to the entire [NP S] structure of complex Noun Phrases. Such an extension is certainly legitimate, and it is motivated by the facts. Use of the Internal Structure Constraint, however, also motivates the

extension theoretically; it predicts that the extension is necessary. For this reason, a generalization of the Internal Structure Constraint is preferable to a reformulation of the Complex NP Constraint, all other things being equal.

Ross mentioned lexical head Noun Phrases in his statement of the Complex NP Constraint because nodes can be chopped out of sentences extraposed from it.

212) What did it seem that James needed?

The Internal Structure Constraint is also capable of accounting for the grammaticality of such sentences. The complement sentence may be considered to have no antecedent. If it is not an anaphoric element, the Internal Structure Constraint will permit chopping. The extraposed subject of seem appears to have it as an antecedent, but this it should rather be understood as a place-marker. The place-marker disappears when the subject is not extraposed, whereas legitimate antecedents do not.

213) That James needed something seemed unlikely.

214) *It that James needed something seemed unlikely.

215) A man came in who was carrying a big bag of brushes.

216) A man who was carrying a big bag of brushes came in.

Within a theory employing the Complex NP Constraint it is peculiar, although not inexplicable, that nodes cannot be chopped out of reduced relatives, such as swimming in the pool in (217).[16]

217) Sam spoke to the girl swimming in the pool. (= the girl that was swimming in the pool)

218) *What did Sam speak to the girl swimming in?

Presumably the sentence node of such a reduced relative clause has been pruned; if it has, the conditions on the

Complex NP Constraint are no longer satisfied. The conditions on the Internal Structure Constraint, however, still are. The relative clause remains an anaphoric element even when reduced, and nothing can be chopped out of it.

Similarly, it comes as no surprise that extraposed relatives and complements do not permit chopping, even though they are no longer dominated by an NP node with a lexical head.

219) *Who did a man come in who saw?

220) *What did the plan bother Bill that James should buy?

Extraposed sentences remain anaphoric elements. Accommodating the Complex NP Constraint to these facts is done by ordering Extraposition from NP after all applications of all chopping rules.

The Internal Structure Constraint has been shown above to handle all the facts concerning the movement of

nodes out of complex Noun Phrases, including the previously unnoticed prohibition against chopping out of the head Noun Phrase. In fact, it predicts this prohibition. It is thus reasonable to conclude that it should replace the Complex NP Constraint, because it is more general. One objection to this conclusion is that the complex NP cases seem much more severely ungrammatical than the cases used at the beginning of this chapter to justify the Internal Structure Constraint. This, I think, is a serious objection. The difference in the degree of ungrammaticality may, however, have something to do with the fact that the anaphoric connection in complex Noun Phrases is obligatory; there are no alternative, non-anaphoric readings that one could imagine. Regular definite pronouns almost always permit alternative readings, both anaphoric and non-anaphoric. Reflexives, on the other hand, do not, and sentences where a node has been chopped out of the antecedent of a reflexive sound the worst of all the sentences listed near the beginning of the chapter.

221) *Who did Sam talk with the girlfriend of about herself?

The objection is, therefore, not deadly, and the Internal Structure Constraint (Revised or Unrevised) can subsume the Complex NP Constraint.

One of the original constraints on movement rules thereby comes to respond not so much to the bracketing as to the function of certain nodes - antecedents and anaphoric elements - with respect to one another. What "with respect to one another" means for such nodes has been the overarching theme of this study. The concepts of network and chain, which explicate the meaning of this phrase, have been shown to have interesting consequences for the analysis of pronominalization, Reflexivization, EQUI, and now, movement rules. It would be interesting to discover the applicability of what has been learned to languages other than English.

FOOTNOTES

1. Bruce Fraser pointed out to me that something like the Supplement should at least be considered.

2. There is a discussion of sentences like (36)-(38) in Postal (1970a). Postal employs the notion of coreferentiality and claims that Noun Phrases like $\underline{\text{Mary}}_i^2$ and $\underline{\text{her}}_i^3$ cannot be coreferential with one another. If this is interpreted as a constraint on something other than pronominalization, (32) would be ungrammatical regardless of how any theory of pronominalization derives it. But there is no reason to accept such an interpretation. The use of coreferentiality will be analyzed in Chapter Two.

3. Lakoff (1968), p. 19.

4. Pronouns referring to VP's and S's are not actually coreferential with them. See the discussion of (47) and (48) below.

5. This argument was suggested to me by Noam Chomsky.

6. It would, of course, be difficult to state this as a restriction on the pronominal filters: they would have to recognize after the fact which cycles EQUI had applied in. The presence of DOOM should make it possible, however.

7. Noam Chomsky has pointed out to me that the treatment of controlled Noun Phrases proposed above could have undesirable consequences. Consider a sentence like (a):

a) John criticized Mary for failing to win.

EQUI is not obligatory in such sentences, as (b) shows:

b) John criticized Mary for her failing to win.

Furthermore, the object of criticized may be a reflexive.

c) John criticized himself for failing to win.

Why, then, should (d) sound peculiar?

d) ?John criticized Mary for his failing to win.

It is not the case that the subject of failing cannot

refer to the subject of criticized:

e) John criticized himself for his failing to win.

By the Pairwise Antecedency Constraint, his must be anteceded by both John and himself for the sentence to be grammatical. This means that his may have John as its antecedent independently of whether himself is also its antecedent. But the object of criticize controls the subject of failing when EQUI applies; it might also control any pronominal subject of failing. If it did, a pronominal subject which did not refer to the object of criticize would be ungrammatical. (d) contains such a pronominal subject, and it is peculiar. Therefore, control also affects overt pronouns. A controlled node cannot be anteceded by anything but the controller, so the argument goes, and because of this it must be concluded that his in (e), contrary to the Pairwise Antecedency Constraint, has John as its antecedent only indirectly: himself antecedes his, and John antecedes himself. Either that, or the concept of control must be rejected; rejecting control would be undesirable.

It is not necessary, however, to draw such extreme conclusions from the above facts. His behaves in (d) and (e) much the way a node deleted by EQUI would. It has been shown several times in this chapter that a deleted controlled Noun Phrase must be marked for antecedency relations independently of the controller. The Network Filter then checks the antecedents of both the controller and the controlled Noun Phrase. Only if they have the same antecedents is a network formed and the sentence declared grammatical. Because of this, in grammatical sentences it will always appear to be the case that indirect antecedency holds: the controlled Noun Phrase will have all and only the antecedents of the controller. But this is just an illusion created by the successful application of the Network Filter.

The difference in acceptability between (d) and (e) can be accounted for by adding a condition like the following to the rules of pronominalization: control may also apply to overt pronouns; antecedents which are also controllers may be obligatory, depending on the verb. The subject of promise, for example, is not

an obligatory antecedent of properly located overt pronouns, even though it is a controller.

f) Bill promised Sam to be there.
(the subject of be = Bill)

g) Bill promised Sam that he would be in Havana in three hours.
(He may be anteceded by either Bill or Sam.)

Given the above condition, the Network Filter will apply to (e) by noting that his is obligatorily anteceded by the controller himself. In turn, himself is obligatorily anteceded by John. Because himself stands in an obligatory antecedency relation to both his and John, these two nodes must be able to stand in a direct antecedency relation to one another if a network is to be formed. They can, and one is.

In (d), Mary must obligatorily antecede the subject of failing, but it cannot, since his does not agree with it in gender. (d) is therefore out because an obligatory antecedency relation cannot be properly constituted. Whether his has John as its antecedent

is irrelevant; however, it probably does. Saying that control is obligatory may be too strong, however, even in the case of (d). It might be more accurate to say that the control reading is the preferred one for some verbs.

8. Note the following additional facts in connection with sentences (130) and (131):

   a) Sam disliked his performing in the play. (his ≠ Sam's)

   b) Sam disliked Ø performing in the play. (Ø = Sam's)

   c) Sam disliked his/Ø having to perform in the play. (his/Ø = Sam's)

   d) Sam disliked his/Ø being forced to perform in the play. (his/Ø = Sam's)

   e) Sam disliked his/Ø being ordered, requested, asked to perform in the play. (his/Ø = Sam's)

   f) Sam disliked his/Ø being expected to perform in the play. (his/Ø = Sam's)

   g) Sam disliked his wanting to perform in the play. (???)

h) Sam disliked his knowing the answer.
(his ≠ Sam's ?)

i) Sam disliked his seeing the play.
(his ≠ Sam's)

j) Sam disliked his/Ø having been similar to his uncle in early childhood.
(his/Ø = Sam's)

Apparently, EQUI is obligatory if the controlled Noun Phrase is in command of the situation or perceiving it, and optional otherwise (when the person it designates is commanded to do something, etc.). Thus, EQUI depends on the thematic relations in the complement S as well as on those in the matrix S, which determine which Noun Phrase will be the controller.

9. Except in sentences like the following:

a) A picture of himself is Sam's favorite possession.

b) Sam's favorite possession is a picture of himself.

c) Sam's best friend is a picture of himself.

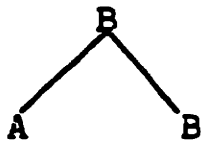
10. See Helke (1971) for a detailed discussion of the analysis of reflexives as pronoun + self.

11. See Lakoff (1970).

12. For two different analyses of how such surface pronouns actually receive an interpretation, see Akmajian (1970) and Ross (1969).

13. This is very close to the definition of chaining found in Grinder (1971), which did not appear until some time after I wrote this chapter.

14. If node A is Chomsky-adjoined to node B, the following structure results:



15. John Ross has pointed out to me that sentences like (a) and (b) below might cause problems for stage (b), and for the claim that nothing can be chopped out of anaphoric elements.

a) You can eat all the candy you want to.

b) Hurl insults at anyone you care to.

There is something missing after the to in each sentence: a verb or verb + object (eat in (a) and hurl insults in (b)); and a Noun Phrase (candy in (a) and anyone in (b)). It looks as if the Noun Phrases might have been moved out of an anaphoric element by Relative Clause Formation. It is also possible, however, that the Noun Phrases have been moved first, and that the sequences of verb or verb + object are then eliminated. If this were so, nothing would be moved out of an anaphoric element. In support of the latter hypothesis, it may be noted that (c) and (d) are alternates of (a) and (b), respectively.

c) You can eat all the candy you want to eat.

d) Hurl insults at anyone you care to hurl insults at.

16. The following two explanations are possible: (a) no

chopping is permitted out of the structure $[_{NP}\ X]_{NP}$, which is a generalization of the complex NP structure; or (b), all chopping rules are ordered before all instances of Relative Clause Reduction.

BIBLIOGRAPHY

Akmajian, Adrian (1970). Aspects of the Grammar of Focus in English. Unpublished MIT Ph.D. dissertation.

Bresnan, Joan (1971). "A Note on the Notion 'Identity of Sense Anaphora'." Unpublished paper, MIT.

Grinder, John (1971). "Chains of Coreference." Linguistic Inquiry 2 (1971), 183-202.

Helke, Michael (1971). The Grammar of English Reflexives. Unpublished MIT Ph.D. dissertation.

Jackendoff, Ray (1969). Some Rules of Semantic Interpretation for English. Unpublished MIT Ph.D. dissertation.

Lakoff, George (1968). "Pronouns and Reference." Indiana University Linguistics Club mimeograph.

Lakoff, George (1970). "Counterparts, or the Problem of Reference in Transformational Grammar." Mathematical Linguistics and Automatic Translation, Report No. NSF-24 to The National Science Foundation, Harvard University, The Aiken Computation Laboratory.

Lakoff, George (1971). "The Global Nature of the Nuclear Stress Rule: A Rejoander." University of Michigan Phonetics Laboratory mimeograph.

Postal, Paul (1970). "On Coreferential Complement Subject Deletion." Linguistic Inquiry 1 (1970), 439-500.

Postal, Paul (1970a). "On the Surface Verb 'Remind'." Linguistic Inquiry 1 (1970), 37-120.

Ross, John Robert (1967). "On the Cyclic Nature of English Pronominalization." To Honor Roman Jakobson, Vol. III, pp. 1669-1682. The Hague: Mouton & Co.

Ross, John Robert (1967a). Constraints on Variables in Syntax. Unpublished MIT Ph.D. dissertation.

Ross, John Robert (1969). "Guess Who?" Papers from the Fifth Regional Meeting, Chicago Linguistic Society, pp. 252-286.

Witten, Edward (1971). "Pronominalization: A Handbook for Secret Agents." Unpublished paper, Brandeis University.

BIOGRAPHY

I was born on March 29, 1944, in Newberry, South Carolina. After graduating from Francis C. Hammond High School in Alexandria, Virginia, I attended Davidson College in Davidson, North Carolina, from 1962 - 1966, with a year off to study German literature at the University of Munich, Germany (1964-1965). In my senior year I was named to Phi Beta Kappa. In June, 1966, Davidson granted me a B.A. summa cum laude in English and German, and the following September I began studying German literature at Princeton with the aid of Danforth and Woodrow Wilson Fellowships. After half a semester, and with the kind help of William G. Moulton, I switched to the Program in Linguistics.

Despite a hasty and late application, MIT admitted me to its Linguistics Department in September, 1967. During the summers of 1968 and 1969, I had the pleasure of working for Susumu Kuno at the Aiken Computation Laboratory. In the fall and winter of 1969-1970, I did research in historical syntax at the

Forschungsinstitut für deutsche Sprache, Marburg, Germany. From June, 1970, through May, 1971, I was a part-time employee of the Language Research Foundation, Cambridge, Massachusetts, where I studied the phonology and syntax of the English spoken by young Puerto Rican children in Boston. During the summer of 1971, I taught syntax at the Linguistic Institute of the Linguistic Society of America, in Buffalo, New York. In the fall of 1971 I began teaching linguistics at Florida State University, Tallahassee, Florida.

Publications:

"Be as a Main Verb." Mathematical Linguistics and Automatic Translation, Report No. NSF-24 to The National Science Foundation, Harvard University, The Aiken Computation Laboratory (February, 1970), pp. 273-286.

"What Seem's to Be." Ibid., pp. 287-303.

"Germanisches AI und AU im Altsächsischen und Althochdeutschen." Zeitschrift für Dialektologie und Linguistik 37 (1970), 44-57.

"Puerto Rican English: A Discussion of Eight Major Works Relevant to its Linguistic Description." Language Research Report No. 3 (1971), pp. 75-123.