INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”.
   If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in “sectioning” the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from “photographs” if essential to the understanding of the dissertation. Silver prints of “photographs” may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
NEELD, Ronald Louis, 1948-
GLOBAL CONSTRAINTS IN SYNTAX.
The Ohio State University, Ph.D., 1974
Language and Literature, linguistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan
GLOBAL CONSTRAINTS IN SYNTAX

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Ronald Louis Keeld

# * * * #

The Ohio State University
1974

Reading Committee:  Approved By

Michael L. Geis  Michael L. Geis
Arnold M. Zwicky  
David R. Dowty  Adviser

Department of Linguistics
...great scientists were never distracted by the fact-finding rage: they knew from the first what they were doing. Their task was always to relate the facts to each other, either as different cases of the same general fact, or as successive transformations of an initial fact according to some systematic principle, or (as an elementary stage of conception) as more and more exemplifications of "contingent laws," or generally observed uniformities.

The philosophical horizon widened in all directions at once, as horizons do with every upward step....Most new discoveries are suddenly-seen things that were always there.

Susanne K. Langer

*Philosophy in a New Key*
ACKNOWLEDGMENTS

I have had the honor of learning from so many people during the course of my studies that it is hard to thank them adequately. The following people stand out in my mind for their scholarship, teaching, and/or stimulating discussion of matters linguistic: David Dowty, Michael Geis, Georgia Green, Bob Jeffers, Mike Kenstowicz, Charles Kisseberth, Ilse Lehiste, Pat Miller, Jerry Morgan, Linda Norman, Haj Ross, David Stampe, and Arnold Zwicky. I owe special thanks to my adviser Mike Geis and to my reading committee David Dowty, David Stampe, and Arnold Zwicky. Their insightful comments have greatly improved this essay.

I would also like to thank the following people. Although they did not have an effect on this dissertation, they had an effect on its writer: Gary Neeld, Sharon Hess, Harry Hess, Simonne Crawford, Pete Crawford, Harry Lever, Beverly Lever, David Johnson, Jan Van Ort, Tracy La Tour, Greg Broadhurst, Bob Colborne, Thea Delisio, Jane Bannerman, Larry Schourup, Stephan Miller, Joan Taylor, and Steve Miller.

Above all, my love and gratitude goes to my guru, Shri Maharaj Ji.
VITA

January 1, 1948 . . . . . Born - Jacksonville, Florida

1970 . . . . . . . . . . . . . . B.A., cum laude, Louisiana State University in New Orleans

1971 . . . . . . . . . . . . . . M.A., University of Illinois

1971-1974 . . . . . . Teaching Associate, The Ohio State University

1972 . . . . . . . . . . . . . . Fellowship, Linguistic Institute of the Linguistic Society of America

PUBLICATIONS


FIELDS OF STUDY

Major Field: Linguistics

Studies in Syntax. Professors Arnold M. Zwicky, Michael L. Geis, Jerry Morgan, Georgia Green

Studies in Phonology. Professors Arnold M. Zwicky, David Stampe

Studies in General Linguistic Theory. Professor David L. Stampe
### TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>VITA</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>ONE. GLOBAL RULES AND ISLAND CONSTRAINTS</td>
<td>6</td>
</tr>
<tr>
<td>TWO. GENERAL CONSTRAINTS</td>
<td>141</td>
</tr>
<tr>
<td>THREE. RULE-SPECIFIC CONSTRAINTS</td>
<td>80</td>
</tr>
<tr>
<td>FOUR. TRANSFORMATIONS AND FILTERS</td>
<td>140</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>156</td>
</tr>
</tbody>
</table>
INTRODUCTION

Global derivational constraints have been introduced into syntactic theory in order to account for cases where reference must apparently be made to non-adjacent trees in a derivation. In the theory presented in Chomsky (1965) there is a list of transformations which apply in a fixed order, each one of which applies to the output of the immediately preceding transformation. Only information present in the phrase marker at the point at which a rule applies may be used to prevent or require the application of a rule. Information present at earlier or later stages of a derivation is not accessible.

However, several cases have been brought to light where it appears as if reference to non-adjacent phrase markers in a derivation is necessary. The most extensive theoretical interpretation of such phenomena has been made by George Lakoff, who argues for the introduction of derivational constraints into linguistic theory. These constraints, according to Lakoff's definition, 'state well-formedness conditions on configurations of corresponding nodes in nonadjacent trees in a derivation' (1970, 628). Transformations thus differ from derivational constraints in that the former are restricted to make reference only to adjacent trees in a derivation, while the latter are not.
The terms 'derivational constraint' and 'global rule' have been used interchangeably. Derivational constraints are not, however, rules in the sense that transformations are. They are filters that throw out certain derivations as being ill-formed. By now, a large number of cases have been found which require constraints that are more than transformational in nature. Yet most of the debate on this topic has been concerned not with explication and explanation of global constraints, but with ever more fine-spun arguments concerning whether it is necessary to have them. But even those who argue against the use of derivational constraints must add new descriptive devices to their theory, such as surface structure interpretation rules, doubly-filled nodes, and indexed nodes. It appears that natural language is the sort of thing which requires an explanation going beyond that of Chomsky's *Aspects of the Theory of Syntax*.

The purpose of this dissertation is to explore the nature of the concept 'global derivational constraint' in transformational syntax. An issue which has been repeatedly raised in the discussion of global rules is that of the relative power of theories. Detractors of global rules claim that a theory which incorporates global constraints describes a wider class of languages than one which does not. If it can be shown that there are natural-language phenomena which can be generated only by a grammar incorporating global constraints, then such constraints must be admitted as part of linguistic theory.
The difficulty comes in that a new descriptive device may allow a grammar to generate languages which are not, and could not be, natural human languages. A theory which allows for too wide a range of grammars must be restricted in some appropriate way: there must be some way to prevent the generation of languages outside the class to be described. All other things being equal, the theory which allows for the narrowest range of grammars is to be preferred. This is not necessarily because such a theory will be simpler, but because it gives a more explicit characterization of what a human language is.

But the matter does not end with the study of the power of purely formal devices. The remarks made by George Lakoff (1972) are very much to the point: Global grammar allows one to get rid of many of the arbitrary and unmotivated devices of classical transformational grammar as aux, modal, ADV, etc. Lakoff claims that the natural semantic categories of sentence, predicate, and argument can be used instead. A whole series of devices such as arbitrary categories, syntactic features, rule-ordering, interpretive rules, empty nodes, and doubly-filled nodes have been used in transformational grammar. A theory without these devices, and with a serious treatment of substantive language universals, would provide the needed narrowing of the excessive power of transformational grammar. As Lakoff says, 'The issue is not one of greater or less descriptive power, but of the ability of very different collections
of descriptive devices to state the general principles of language' (1972, 83).

There has been some confusion in terminology. Lakoff (1971) uses the term 'derivational constraint' to include almost any kind of constraint that blocks a derivation, to the point of using 'local derivational constraint' to refer to transformations. But calling a transformational rule a derivational constraint does not make it one. Transformational rules and global constraints have very different formal properties: the former are stated in terms of a structural description and a structural change, and map one phrase marker into another; while the latter are filters which throw out certain derivations as ill-formed. The issue I focus on here is 'global constraints on derivations', conditions that mark an entire derivation as ill-formed on the basis of information in nonadjacent phrase markers.

Restrictions of the form of global grammars can be seen to follow from an adequate explanation of what global constraints are. I therefore focus in the last chapter on explanation in syntax, by characterizing the function of global constraints. An outline of the thesis follows.

Chapter One begins with a discussion of rule ordering, wherein I accept the Unordered Rules Hypothesis. I also argue that no conclusive evidence for the use of a cycle in syntax has been proposed. Based on the URH, I show that the Sentential Subject Constraint is a global constraint.
In Chapter Two I present those global constraints which are general in their formulation. These fall into two classes: those which involve movement rules, and those which involve deletion/anaphora processes.

In Chapter Three I discuss those global constraints which mention specific rules or highly specific structures. I argue that the constraints involved are reduceable to more general and independently-needed grammatical phenomena, and that rule-specific constraints should be eliminated from the class of global constraints.

The focus of Chapter Four is on the concepts of description and explanation in syntactic theory. I discuss the general nature of global constraints and the distinction between transformations and filters in syntax.
CHAPTER ONE

GLOBAL RULES AND ISLAND CONSTRAINTS

1 Rule Ordering

The significance of island constraints for the study of global constraints is that some island constraints may be viewed as global in nature, rather than strictly transformational. In this chapter I will discuss the status of island constraints as derivational constraints, and present evidence that the Sentential Subject Constraint is global in nature.

A crucial part of the discussion will involve the concept of rule ordering. In this dissertation I accept the validity of the Unordered Rules Hypothesis. Ringen (1971) notes that there are three possible assumptions about rule ordering:

(1) a. All rules are extrinsically ordered.
   b. Some rules are extrinsically ordered.
   c. All rules are unordered — i.e. all constraints on the order of application are predicted by language-universal principles.

(1a) can be falsified (see Koutsoudas (1971)). (1b) cannot be proven either true or false. Therefore (1c) remains as the strongest empirical hypothesis and I will accent it until it is proven false.
As a preliminary to further discussion, I present a summary of the basic principles of the Unordered Rule Hypothesis. These appear in Koutsoudas (1973b).

(2) a. All restrictions on the application of rules are determined by universal principles (and hence there are no language-specific ordering restrictions between the rules of a grammar).

b. An obligatory rule must apply whenever its structural description is met, unless its application is precluded by some universal principle. It follows from this requirement that rules will apply simultaneously if possible; otherwise they will apply sequentially.

c. All rules are scanned for applicability to each new representation in a derivation.

d. A derivation is terminated when no obligatory rules are applicable.

The concept of rule ordering is closely tied up with the concept of the cycle. I shall discuss briefly the status of the cycle as a principle governing rule application before going on.

2 The Cycle: Nonevidence

Given the Unordered Rules Hypothesis, it is still an open question as to whether cyclic application of rules is to be allowed. We could apply all rules possible to the lowest domain of application, then move on to the next higher domain, etc. As was noted by Lakoff (1969, 138) many arguments for rule ordering involve cases where one rule is cyclic and the other is postcyclic, so that
the former must precede the latter. Lakoff speculates that the correct application of rules could be predicted by the principle of the cycle, such that 'the application of rules in their correct order would be given entirely by the principle of the cycle given an adequate account of underlying structures, and by the postcyclical grouping of other rules' (138). Other rule ordering constraints would be replaced by global derivational constraints. The proposal is attractive, for given a choice between cyclic application and extrinsic ordering, one would favor cyclic application. The principle of the cycle is a general (perhaps language-universal) feature of grammar, whereas each extrinsic ordering statement is an ad hoc language-particular restriction.

However, it is by no means clear that there is no third alternative. Given the Unordered Rules Hypothesis, there are language-universal constraints on the order of rule application, which may well eliminate the need for both extrinsic ordering and the cycle. It is indeed plausible that the elimination of extrinsic ordering should lead also to the elimination of the cycle, for many arguments for the cycle depend on extrinsic ordering. That is, it is shown that in some derivations, rule A must precede rule B, but that A must also apply after B has applied. The only way to avoid this paradox in a system which has extrinsic ordering is to use the cycle.

Some arguments have recently been given by Postal and Koutsoudas that the cycle may be used to eliminate extrinsic ordering. I shall now show that these examples do not provide evidence for the existence of a cycle. Koutsoudas (1973b) claims that the cycle
can explain that in the derivation of (3) from (4) Passive must precede Wh Rel Movement.

(3) The man who was pardoned by the governor breathed a sigh of relief.

(4)

```
S₁

NP₁

the man

NP₂

the governor

S₂

VP

breathed a sigh of relief
```

Given a cycle, Passive may apply on the S₂ cycle. Then on the S₁ cycle, Wh Rel Movement applies. Both rules could not apply simultaneously to (4), since both operate on NP₁, and this NP could not undergo both rules at once.³

However, Koutsoudas' argument depends on the assumption that there is no other principle that will predict the application of Passive before Wh Rel Movement. The putative distinction is that one rule is cyclic, the other noncyclic. (See Postal (1972d) for arguments that Wh Rel Movement is not cyclic.) It seems that the rules of grammar fall into two blocks, but it is not clear that cyclic vs. noncyclic is an adequate characterization of these blocks. Postal (1972c) has claimed that the distinction is just that between rules which change grammatical relations, and those which do not. Let us call rules in the former class relational. Examples are Passive, which makes a direct object into a subject, and Raising, which makes the subject of an embedded sentence into the direct
object of the immediately higher sentence. It turns out that relational rules have been termed cyclic, while non-relational rules have been termed postcyclic. Then a possible principle of rule ordering is:

(5) Relational rules precede nonrelational rules.

This predicts the precedence of Passive over Wh Rel Movement.\footnote{\textsuperscript{4}}

Passive affects grammatical relations, but Wh Rel Movement does not, for it moves something into leftmost position in a sentence. It does not make something into a subject, direct object, or indirect object. (5) is not just a restatement of the cyclic principle, with the word 'relational' substituted for 'cyclic'. The term 'cyclic' refers only to the fact that a rule may be reapplied after intervening application of another rule or rules. The term 'relational' has an independent use in that it refers to the syntactic function of a rule, i.e. that it changes grammatical relations.

Koutsoudas presents another argument for cyclic rule application, this one in regard to the rules Wh Fronting and Sluicing.

(6) He is writing something, but you can't imagine what.

(7) \begin{center}
\begin{tikzpicture}
  \node (S1) {$S_1$};
  \node (S2) at (0.5,0) {$S_2$};
  \node (S3) at (2,0) {$S_3$};
  \node (S4) at (3.5,0) {$S_4$};
  \node (S5) at (2,-1) {$Q$};
  \draw (S1) -- (S2);
  \draw (S1) -- (S3);
  \draw (S3) -- (S4);
  \draw (S4) -- (S5);
  \draw (S2) -- (\text{he is writing something});
  \draw (S3) -- (\text{you can't imagine $S_4$});
  \draw (S5) -- (\text{he is writing wh-something});
  \draw (S1) -- (\text{but});
\end{tikzpicture}
\end{center}

On the $S_3$ cycle or the $S_4$ cycle Wh Fronting will apply. It is only when we move to the $S_1$ cycle that Sluicing can apply, thereby in-
suring the operation of Wh Fronting before Sluicing. But the cycle is not necessary. When we examine Phrase Marker (7) we see that only the structural description for Wh Fronting is met. Then after this rule has applied, the conditions for Sluicing are met and this rule applies. The cyclic principle is not necessary here, for we need only the universal principle that rules apply whenever their structural description is met.

Postal (1972c) speaks favorably of the idea that extrinsic ordering can be eliminated (to be replaced in some cases by global and possibly transderivational rules, although he gives no examples of this). Postal also maintains that the cyclic principle is needed to constrain certain derivations. Upon closer examination, it turns out that his examples do not support the cycle. One example is based on the rules Raising and Reflexive. The relevant sentences are those of the class illustrated by (8).

(8) a. Melvin seems to have believed himself to be overly critical of himself.

b. Melvin considers himself to have proved himself to be innocent.

If we examine the deep structure of (8b), represented by (9), we see that Raising must apply to both occurrences of he.
If we use a cycle, NP_3 is raised on the S_2 cycle, and NP_2 and NP_3 become Clause Mates, allowing reflexivization of NP_3. Then on the S_1 cycle, NP_2 is raised, making NP_1 and NP_2 Clause Mates, so that NP_2 may then be reflexivized. Postal correctly points out that we cannot allow random application of rules here. Suppose Raising were to apply to the domain of S_1 first, raising NP_2. Reflexivization could not apply since the two NP would not be Clause Mates. But even after Raising applied to the domain of S_2, raising NP_3, the noun phrases still would not be Clause Mates, and the conditions for Reflexivization would never exist. Postal concludes: 'In fact, then, the ordinary cyclical mode is the only known way to derive all the correct structures and none of the wrong ones in cases like these'(251).

But this conclusion is not warranted. Let Raising apply (twice), once to the domain of S_1, and once to the domain of S_2. The result is (10).
At this point, $NP_1$ and $NP_2$ are Clause Mates, allowing Reflexivization. We would also like to simultaneously apply Reflexivization to $NP_3$, with $NP_2$ as antecedent. It would appear, however, that the clause-mate condition is not met, due to the presence of $S_2$. We could, of course, claim that Pruning has applied here, eliminating $S_2$ and allowing Reflexivization (cf. Ross 1967, 24-64). However, the status of Pruning as a principle of grammar is dubious (see Postal 1972c). But even without Pruning, Reflexivization could apply to $NP_3$ if we introduce the notion of Quasi-Clause into the theory of grammar. Postal (1972c, Part IV, section 0, 161 and Part VI, section E, 215-216) claims that a clause becomes a Quasi-Clause if the subject NP has been removed through the action of cyclic rules. Quasi-Clauses are not as strong a barrier to syntactic processes as are full clauses. We could allow Reflexivization to apply to Quasi-Clause Mates as well as to Clause Mates, in which case $NP_2$ and $NP_3$ would meet the conditions for forming an antecedent-reflexive pair. By using the alternative approach I have just sketched, we can avoid the use of the cycle.
Postal also gives an argument for the cycle based on Inclusion Constraint cases. The relevant examples (cited by Postal) are:

(11) a. *I believe us to know French.
    b. *Melvin believes me to be overly critical of us.
    c. *I seem to have believed us to know French.
    d. *Melvin considers me to have proved us to be innocent.

If there is a strict Clause Mate condition on the Inclusion Constraint, then (11c,d) can be ruled out only at an intermediate stage of the derivations, after Raising applied on the lower cycle, but before it applied on the higher cycle. Given that the deep structure of (11d), for example, is roughly as in (12), the underlined NP will be clause mates after Raising has applied on the $S_2$ cycle.

(12) $S_1[\underline{\text{Melvin considers }} S_2[\underline{\text{I have proved }} S_3[\underline{\text{we are innocent}}]]]$.

When we move on to the $S_1$ cycle, the subject of $S_2$ is lifted into $S_1$, and the underlined NP's are no longer Clause Mates.

But here as above we can use the notion quasi-Clause Mate, and let the Inclusion Constraint apply to NP which are quasi-Clause Mates. Then in (11c,d) both applications of Raising can apply simultaneously, and the underlined NP become Quasi-Clause Mates, being subject to the Inclusion Constraint. Furthermore, in this case it can be shown that the constraint must apply to quasi-Clause Mates on independent grounds; for at no point in the derivation of (13a,b) are the underlined NP's strict Clause Mates:

(13) a. *I want to paint us.
    b. *I believe John to like us.

In (13a) the underlying subject of the complement sentence would,
before being deleted by Equi, be a strict Clause Mate of \textit{me}, meeting the conditions for the Inclusion Constraint. But this constraint cannot be stated at the level of deep structure, for the deep structure of (11a), wherein \textit{us} is the subject of an embedded sentence, does not meet the conditions for the Inclusion Constraint. And, indeed, if \textit{us} does not become a derived Clause Mate of \textit{I} in (11a) the sentence is acceptable:

\begin{enumerate}
\item[(14)] I believe that \textit{we} know French.
\end{enumerate}

Furthermore, in (13b) the underlined NP are not Clause Mates in deep structure, and so are not subject to the constraint. But if the subject of the embedded clause is raised, then the embedded clause becomes a quasi-clause, and the underlined NP's become quasi-Clause Mates, and so are subject to the constraint. If the subject of the complement sentence is not raised, the complement sentence remains a full clause, and the constraint is not applicable. Compare (15) and (16) where the underlined NP's in the latter are not quasi-Clause Mates, and the sentences are acceptable.

\begin{enumerate}
\item[(15)] a. *\textit{I} believe John to like \textit{us}.
\item b. *\textit{We} believe John to like \textit{me}.
\item c. *\textit{He} believes John to like \textit{them}.\j
\item d. *\textit{They} believe John to like \textit{him}.
\end{enumerate}

\begin{enumerate}
\item[(16)] a. \textit{I} believe that John likes \textit{us}.
\item b. \textit{We} believe that John likes \textit{me}.
\item c. \textit{He} believes that John likes \textit{them}.
\item d. \textit{They} believe that John likes \textit{him}.
\end{enumerate}

Since the Inclusion Constraint refers to quasi-Clause Mates, we do not
need a cycle to rule out the sentences in (11).

Another case of cyclic rule application that Postal considers involves the process of Answer Reduction, whose effects are illustrated in (17):

(17) a. Q: Who should the cop arrest?
   b. A: I think the cop should arrest Bob.

Significantly, Reflexivization never applies to the result of this process.

(18) a. Q: Who should Bob call?
   b. A: I think Bob should call me.
   c. A: I think me (*myself).

(19) a. Q: Who should Harry call?
   b. A: John\textsubscript{1} thinks Harry should call him\textsubscript{1}.
   c. A: John\textsubscript{1} thinks him\textsubscript{1} (*himself).

These results should follow from the cycle, since Reflexive can be shown to be cyclic (cf. Postal (forthcoming) part VI, section B). Assuming that Answer Reduction is postcyclic, the structure prior to the operation of Answer Reduction would not meet the conditions for Reflexivization. No cyclic rules could apply to the output of Answer Reduction, so that (18c) and (19c) could not be generated.

But given the Unordered Rules Hypothesis, there exist universal principles which constrain the order of application of certain rules. One such principle is that proposed in Ringen (1971):

(20) Obligatory-optional Precedence

If the same phrase marker meets the structural description of both an optional and an obligatory rule, the obligatory
rule must apply to this phrase marker — i.e. the obligatory rule may not apply after the optional rule has applied.\textsuperscript{10} This principle predicts the results shown in (18) and (19). The obligatory rule of Reflexivization may not apply to the structure derived by the optional rule of Answer Reduction, and of course prior to the operation of this latter rule the conditions for Reflexivization do not exist. Reflexivization could never apply in the derivation of (18c) or (19c).

The situation is much the same with Genitive Head Deletion and Reflexivization. Postal claims that the facts in (21) follow from the fact that Genitive Head Deletion is postcyclic.

\begin{enumerate}
\item[(21)] a. Martha's\textsubscript{i} dog didn't fall on her\textsubscript{i}, but Sarah's\textsubscript{j} fell on her\textsubscript{j}.
\item b. *Martha's\textsubscript{i} dog didn't fall on her\textsubscript{i}, but Sarah's\textsubscript{j} fell on herself\textsubscript{j}.
\item c. Martha's\textsubscript{i} dog didn't fall on her\textsubscript{i}, but Sarah's\textsubscript{j} dog fell on her\textsubscript{j}.
\end{enumerate}

I assume that Postal means that the genitive noun phrases are clausal in underlying structure, and lose their clausal status after Genitive Head Deletion. Then Reflexivization would have to apply when the genitive noun phrase still had clausal status, and the Clause Mate condition on Reflexivization would not be met. Clausal status would not be lost until after all cycles were completed, at which point it would be too late for Reflexivization to apply. But here again, Obligatory-optional Precedence is applicable, for the obligatory rule of Reflexivization could not apply after the optional rule of Genitive Head Deletion (cf. (21c)).\textsuperscript{11}

In summary, the evidence brought forth in favor of the cycle is
susceptible to reanalysis in which cyclic application is not necessary. The cycle would, if substantiated, be a strong universal principle of grammar of the type that generative grammarians are seeking. But since Koutsoudas and Postal both deny extrinsic ordering, they have in effect lost the basis for one class of arguments in favor of the cycle. The use of principles of rule application (needed in the Unordered Rules Hypothesis) eliminates the need for the cycle in still another class of cases. At this point, it seems to me that neither extrinsic ordering nor the cycle need be used in syntax. It is indeed the case that some rules are restricted to apply within a certain domain, but this is a different matter than ordering rules according to the cyclic principle.

3 The Complex Noun Phrase Constraint

I have argued above that if we accept the Unordered Rules Hypothesis, the cycle is not necessary. The acceptance of the Unordered Rules Hypothesis has consequences in other areas of syntax. Of particular concern here is the way in which the elimination of extrinsic ordering leads to the conclusion that (at least some of) the island constraint discussed in Chapter One are global derivational constraints. As an example, Koutsoudas (1973a) has shown that the Complex Noun Phrase Constraint (or CNPC) must be viewed as derivational in nature. His argument is as follows. Adopting the Unordered Rules Hypothesis, there is no extrinsic ordering imposed on the rules Extraposition from NP, Relative Clause Formation, Relative Clause Reduction, Question Formation, and Topicalization. Constituents may not be moved out of relative clauses, due to the Complex NP Constraint. Crucially, consti-
tuents also may not be moved out of reduced or extraposed relative clauses.

(22) a. The girl who was sewing the shirt was Jessica.
    b. *The shirt, the girl who was sewing was Jessica.
    c. The girl sewing the shirt was Jessica.
    d. *The shirt, the girl sewing was Jessica.

(23) a. The man who bought the car drove away.
    b. The man drove away who bought the car.
    c. *The car, the man drove away who bought.
    d. *What did the man drive away who bought?
    e. *There's the car which the man drove away who bought.

(22d) shows that Topicalization may not move an element from a reduced relative clause. Topicalization also may not apply to an extraposed relative clause, as shown by (33c). Wh Q Movement and Wh Rel Movement also may not extract elements from extraposed relative clauses, as shown by (23d,e). The significant fact here is that if a relative clause is reduced or extraposed, it is no longer part of a complex NP. This means that we cannot block movement from the relative clauses in (23c,d,e) and (22) by the use of the CNPC. Using the Unordered Rules Hypothesis, movement rules could not apply to the underlying structures of (22) and (23c,d,e) since the moving elements would be in complex noun phrases. But after the application of Relative Clause Reduction and Extraposition from NP, the relative clauses are not part of a complex noun phrase, and there is nothing to prevent movement rules from applying to these relative clauses.

Noticing these facts, Koutsoudas proposes that the CNPC is
derivational in nature, a position for which independent evidence is
given in Ross (1969, 276-277). The CNPC is reformulated by Koutsoudas
as follows:12

(24) The Complex Noun Phrase Constraint

If in a line in a derivation there is a complex noun phrase, 
NP₁, consisting of a lexical head NP₂ and a sentence, S₁, 
then there cannot be another line in this derivation in 
which an element which was contained in S₁ is not dom-
inated either by NP₁ or by S₁.

By viewing the CNPC as derivational in nature, (22d) and (23c,d,e)
can be blocked without reference to extrinsic ordering.

4 The Sentential Subject Constraint

Koutsoudas' analysis depends in an important way on how constraints 
on the order of rule application are to be treated. The relation 
between rule ordering and island constraints should be kept clearly 
in mind. Island constraints are defined on certain structural config-
urations, e.g. sentential subjects and complex noun phrases. There are 
some transformations which destroy the structures defining islands.
For example, Extrapolation from NP and Relative Clause Reduction destroy 
complex noun phrases. As mentioned above, when complex noun phrases 
are destroyed, they maintain their integrity as islands: no consti-
tuents may be moved out of them. This restriction may be expressed 
in global terms by saying that nothing may move out of a constituent 
if that constituent was under the domination of a complex NP node 
earlier in the derivation. In this case, once some node is an island, 
it remains an island.

Now let us consider the other logical possibility with respect 
to the destruction of islands. It could be the case that some node
is an island, that the island is destroyed by the application of some rule, and that elements may then be moved out of the former island. Using extrinsic ordering, we could order all movement rules to apply after the rule (or rules) which could destroy the island in question. But using the Unordered Rules Hypothesis, some other approach must be taken.

The Sentential Subject Constraint presents an example of just the sort that I have outlined in the paragraph immediately above. A sentential subject may be destroyed by the application of Extraposition, as shown in (25) and (26).

(25) That John read the book is \{obvious\} \{probable\}

(26) It is \{obvious\} that John read the book. \{probable\}

When a sentential subject is destroyed, elements may be extracted from it. Compare (27) and (28).

(27) a. *What is that John read \{probable\}? \{obvious\}

b. *There is the book which that John read is \{probable\}. \{obvious\}

(28) a. What is it \{probable\} that John read? \{obvious\}

b. There is the book which it is \{probable\} that John read. \{obvious\}

Even if we were to use extrinsic ordering, we would face a serious problem with respect to (25)-(28). Each extrinsic ordering statement is an ad hoc statement for a particular pair of rules, and does not generalize to other pairs. Given that we must order Wh Q Movement after Extraposition (to account for (26a)), there is no reason why
we should order Wh Rel Movement after Extrapolation (to account for (28b)). Any extrinsic ordering of one rule before another carries with it implicitly the possibility that the order could be opposite to that actually found. It would then be completely accidental that all movement rules followed the rules which could destroy a sentential subject. No explanation could be offered for the fact that all movement rules can operate on elements inside the former island.

Of course, adopting the Unordered Rules Hypothesis, we must find some other explanation for (25)-(28). I shall now sketch an alternative explanation wherein the Sentential Subject Constraint is viewed as a global constraint, stated at the level of surface structure, but making reference to an earlier stage in the derivation. I shall argue that the constraint is to be stated in the following way.

(29) The Sentential Subject Constraint

Given a phrase marker containing a clause $S'$ and a constituent $C'$, the derivation of the phrase marker is ill-formed if:

a. in surface structure $S'$ is a sentential subject,
b. $C'$ is not under the domination of $S'$ in surface structure,
c. in semantic structure, $C'$ is under the domination of $S'$.

The sentences in (27) illustrate the fact that the presence of a sentential subject in surface structure is sufficient to block movement out of the subject clause. The sentences in (28) show that the presence of a sentential subject in deep structure is not sufficient to block movement, for (28a,b) are derived from structures in which the constituents that have been moved were under the domination of a sentential subject. Furthermore, there are rules which can create
sentential subjects, such as Passive. The clauses which are moved into subject position by this rule are islands:

(30) a. Bob did not know that John had read the book.
   b. That John had read the book was not known by Bob.

(31) a. *What was that John had read not known by Bob?
   b. *That is the book which that John had read was not known by Bob.

The above facts indicate that the presence of a sentential subject in surface structure is crucial.

Given the Unordered Rules Hypothesis, there is no restriction against applying movement rules to a clause before that clause has become a sentential subject. For example, consider the deep structure of (31b), which is given in (32)

(32) \[ \begin{array}{c}
\text{NP} \\
\text{that} \\
\text{V} \\
\text{is} \\
\text{NP} \\
\text{the book} \\
\text{VP} \\
\text{Bob did not know} \\
\text{NP} \\
\text{that John had read the book} \\
\end{array} \]

There is nothing to block Wh Rel Movement from applying to move the noun phrase the book out of $S_3$. If we left things at this stage, the result would be:

(33) That is the book which Bob did not know that John had read.

But there is no restriction against going on to apply Passive to $S_2$
in (32), after Wh Rel Movement has applied, the result of which is (31b). There would be nothing to block (31b) and it would in fact be generable by the grammar if there were no global condition on movement out of a sentential subject. I claim that (31a,b) are ruled out because they violate the global constraint (29). I have already shown that the presence of a sentential subject in surface structure is crucial, so the constraint must be stated at that level. Furthermore, in the absence of extrinsic rule ordering constraints we have to know whether C' was under the domination of S' in semantic structure. Therefore, we have to make reference to an earlier stage of the derivations in question.

There is a set of sentences in which Passive applies, creating a sentential subject, and then Extraposition applies, destroying the sentential subject. Elements may move freely out of the extraposed clauses in such sentences, providing further support that reference to surface structure is necessary. An example of this situation is given in (34), where (34a) most closely resembles the deep structure of the sentences, and (34b) is blocked by (29).

(34) a. There is the dress which John expects Mary will buy.

b. *There is the dress which that Mary will buy is expected by John.

c. There is the dress which it is expected by John that Mary will buy.

An alternative to the global constraint analysis would be to claim that the object complement of expect in (34a) undergoes vacuous Extraposition and that Passive then applies, interchanging John and the morpheme it, the derived object of believe. Since the extraposed
clause was never a sentential subject (in this analysis), we expect constituents to move out of it freely. This type of derivation is sketched in (35).

(35) a. There is the dress \( S [ \text{John expects } N_P[S^1] \text{ that } \text{Mary will buy the dress} S, \] \)

\( \rightarrow \) b. There is the dress \( S [ \text{John expects } N_P[S^1] \text{ that } \text{Mary will buy the dress} S, \] \) (vacuous Extrapolation)

\( \rightarrow \) c. There is the dress \( S [ \text{it is expected by John } S^1, \text{ that } \text{Mary will buy the dress} S, ] \) (Passive)

\( \rightarrow \) d. There is the dress which it is expected by John that Mary will buy. (Relative Clause Formation)

Such an analysis holds little merit, for while the use of vacuous Extrapolation explains the lack of ill-formedness of (34c), it does not explain the lack of ill-formedness of (28), where the extrapolated clause must have been (on semantic grounds) a sentential subject in deep structure. Furthermore, it is not clear at the present time whether vacuous application of rules is to be allowed, where by vacuous application I mean that the structure of a phrase marker is changed, but there is no change in the linear order of morphemes. If there is a restriction against vacuous Extrapolation, there must be a sentential subject at an intermediate stage of the derivation of (34d). But since the sentential subject does not appear in surface structure, there is no restriction against moving constituents from the clause.

There are alternatives to the formulation of the constraint as presented in (29), but they are less than optimal. To begin with, (28a,b) and (35d) are acceptable because Extrapolation
has destroyed the sentential subject. We might try to make Extraposition obligatory, along the lines presented in (36).

(36) Extraposition is obligatory if:

a. there is a sentential subject, S', and some other constituent C' outside the domination of S',

b. C' was under the domination of S' earlier in the derivation.

The problems with (36) are manifold. First, there is the problem with making an optional rule obligatory, discussed in Chapter Three, section four with reference to the problem of Extraposition. Briefly, the problem is that if we make an optional rule obligatory under certain conditions, we rob the distinction between optional and obligatory of its validity. Second, I argue in Chapter Three that no global constraint may mention the name of a particular rule. And independently of the above considerations, there is damaging evidence in the fact that Raising can destroy sentential subjects, as demonstrated in (37).

(37) a. That John will read the book is certain.

b. John is certain to read the book.

If a sentential subject has undergone Raising, constituents may freely move out of it. Compare (37) with (38), where Raising has applied to the structures underlying the sentences in (39).

(38) a. *What is that John will read certain?
   (Wh Q Movement)

b. *That's the book which that John will read is certain.
   (Wh Rel Movement)

c. *It's the book that that John will read is certain.
   (Cleft Formation)
(38) d. *What that John will read is certain is the book.  
(Pseudo-Cleft Formation)

(39) a. What is Jon certain to read?
   b. That's the book which John is certain to read.
   c. It's the book that John is certain to read.
   d. What John is certain to read is the book.

If we chose (36), we would have to make Raising obligatory under the same conditions given for Extraposition. In fact, we would have to repeat the constraint expressed by (36) for any rule which could destroy a sentential subject. But this repetition leaves unexplained why the constraint should apply to rules which destroy sentential subjects and not to any arbitrarily selected set of rules. In fact, it is the obligatory triggering of just those rules which destroy sentential subjects, based on the conditions (36a,b), which will prevent the conditions expressed in (29) from existing. Repeating conditions (36a,b) for a set of rules does not express the connection which holds between the rules in the set. The crucial factor is whether or not a sentential subject exists in surface structure, something not expressed by (36), where it is derivative that the rule made obligatory happens to destroy a sentential subject.

Another alternative to (29) is the constraint given in (40):

(40) Given a clause S' and a constituent C', such that
   a. C' is under the domination of S',
   b. S' will appear as a sentential subject in surface structure,

then C' may not be moved out of S'.

The underlined restriction in (40b) is necessary because of cases
like (35d), where a sentential subject is created and then destroyed. It appears that (40) merely states constraint (29) in inverse terms. In (40) we are blocking a rule, but on the basis of 'future history'. But there is a conceptual flaw in (40). The deep structure of (41a,b) is as in (42).

[(41) a. That is the dress which that Alice would buy was suspected by her husband.

b. That is the dress which it was suspected by her husband that Alice would buy.

(42) That is the dress which her husband suspected

NP [NP [NP [that Alice would buy which] ] ]

The constraint (40) claims that nothing may be moved out of a clause that will appear as a sentential subject in surface structure. (41a) is blocked by this constraint, while (41b) is not. But consider more closely the case in which Wh Rel Movement is blocked in (42) because Passive will later create a sentential subject. If we did not apply Passive, then the conditions for blocking Wh Rel Movement would not exist, and we should not have blocked Wh Rel Movement. In essence, Passive becomes obligatory on the basis of having blocked Wh Rel Movement. We would have to add another constraint making certain rules obligatory (just those which create sentential subjects) on the basis of constraint (40). In other words, we have to block Wh Rel Movement from applying to (42) because Passive must apply, where Passive must apply just because we blocked Wh Rel Movement from applying. The circularity, redundancy, and lack of insight of this analysis is obvious. I conclude that the proper way to express the
necessary constraint is as in (29).

I have shown above that the Sentential Subject Constraint is global in nature. It should be noted that there is a surface structure constraint similar in nature to the Sentential Subject Constraint. In (43) and (44), the violation could not result from moving something out of a sentential subject.

(43) a. *What is that John did surprising?

*I went out with a girl who it pleases that John showed up.

To account for the ungrammaticality of (43) and (44) Ross (1967, 251) proposed the following constraint:

(45) The S-Internal NP Clause Constraint

Grammatical sentences containing an internal NP which exhaustively dominates S are unacceptable, unless the main verb of that S is a gerund.

It might be claimed that all sentences which are putative violations of the SSC are actually violations of (45). But the following considerations indicate that the SSC is necessary. Constraint (45) refers to internal NP's: for example, Subject-Verb Inversion in (43) has caused the subject NP to become internal. Notice that in (46), the deep structure of (47), the sentential subject $S_3$ has material to both the left and right, but it is not clause-internal since it is the leftmost constituent of $S_2$. 
(47) We believe it is surprising that Harry read the book. Now notice that in (49), the deep structure of (48), NP' is not clause-internal.

(48) a. *Mary asked what John believes that Harry read is probable.

b. Mary asked what John believes it is probable that Harry read.

(49)

But in the structure derived from the application of Wh Q Movement (illustrated in (50)), NP' is not clause-internal either.
Since Wh Q Movement has an essential variable in its structural description it may move the wh-element to the left over any number of S boundaries. Therefore the wh-element can move over the S-node which dominates the NP complement that it originated in. By virtue of this property of Wh Q Movement, NP' in (50) does not appear as clause-internal in surface structure. It would appear that (48a) must be blocked because something has moved out of a sentential subject, and that the Sentential Subject Constraint is necessary.14

5 Coordinate Structures

We have seen that the Complex Noun Phrase Constraint and the Sentential Subject Constraint are global in nature. The question remains whether the constraints involving coordination, the Element Constraint and the Conjunct Constraint, are global as well.15 The answer to this question must be no. To begin with it is difficult even to find any evidence to test whether these constraints are
global. Recall that the demonstration that the CNPC and SSC are
global required the possibility that the structural configuration
defining an island could be created or destroyed by the application
of some transformational rule or rules. But coordinate structures
are not typically the sort of things that can be created or destroyed.
There seem to be no rules which destroy a coordinate structure
(with the possible exception of Conjunct Movement), and any rule
which appears to create a coordinate structure (like Coordination
Reduction) actually regroups constituents (after deletion) that were
originally part of a coordinate structure. This situation makes it
difficult to tell whether constraints on these structures are global.
There is one rule of English which might be viewed as destroying
a conjoined structure. This is the rule of Conjunct Movement, which
would derive (51b) from (51a).

(51) a. John and Mary went to the movies.

b. John went to the movies with Mary.

Notice that the NP Mary, which was formerly part of the conjoined
subject, can be moved.

(52) It's Mary who John went to the movies with.

This is the result we would expect if the Conjunct Constraint were
not global. Of course, if Conjunct Movement is not a rule (see
Andrews (1971) for evidence against the rule), then we have no evi-
dence at all on the question at hand. In any event, there is no
reason to suppose that the Conjunct Constraint is global.

The only other case that I can find that is relevant involves
the rule of Coordinate Deletion. This rule reduces (53a) to (53b).
(53) a. $s[\text{John cooked the eggplant}_S]$ and $s[\text{John ate the eggplant}_S]$

b. John cooked and ate the eggplant.

The NP the eggplant was part of a coordinate S node in deep structure, and is not part of a coordinate node in surface structure. It turns out that this NP can be moved after Coordinate Deletion.

(54) a. The eggplant is what John cooked and ate.

(Pseudo-Cleft Formation)

b. Here's some leftovers from the eggplant which John cooked and ate.

(Wh Rel Movement)

c. What did John cook and eat?

(Wh Q Movement)

These results ought to indicate that the Element Constraint is non-global, for only if we had to block movement by reference to the presence of an earlier conjoined structure would we need a global formulation. But there is a problem: when dealing with coordinate structures we can't factor out the effects of across-the-board application (cf. Ross (1968), section 4.2.4.1). An NP can move out of a coordinate structure in cases of across-the-board application. The deep structure of (54c) is (55).

(55) $s[\text{John cooked wh-something}_S]$ and $s[\text{John ate wh-something}_S]$

Application of Wh Q Movement (and other minor rules) yields (56).

(56) What did John cook and John eat?

Then Coordinate Deletion gives (54c). The possibility of across-the-board rule application, along with the difficulty of creating or destroying a conjoined structure, makes it virtually impossible to construct any argument which would show that the Conjunct Constraint
or the Element Constraint are global. We must therefore assume that these rules are not global.

2 Given extrinsic ordering, the order must be:
   (i) Passive
   (ii) Wh Rel Movement

for if the opposite order were to obtain, Wh Rel Movement would apply first, and then Passive could not apply. (3) could not be generated.

3 This can be seen more clearly in a case where the relativized NP is in subject position, such as (i) (derived from (ii)).

   (i) The man who the boy was hit by escaped.
   (ii) S1 [NP The man] S2 [who hit the boy] NP escaped S2

The structural change of Wh Rel Movement specifies that the relative pronoun be the leftmost constituent of S2 (or perhaps that it be Chomsky-adjoined to the left of S2), but the structural change of Passive specifies that the relative pronoun be in the postverbal by-phrase of S2. If Passive applied in the correct way, we would find:

   (iii) S1 [The man who the boy was hit by who] NP escaped S2

But Wh Rel Movement could not have applied here. The two rules are incompatible, because part of the structural description of both mentions the same NP (the relative pronoun) but the structural changes of the rules make different predictions as to where the mentioned NP ends up.

One might try to allow for the simultaneous application of the rules to (4) by using an analysis of the Passive along the lines of Hasegawa (1968). In this analysis, the superficial subject of the passive sentence is a deep structure subject. (4) could be replaced by (4')

(4')

```
 VP
 /   \  
 S1  
 /   \ 
 NP1  
 /   \ 
 S2  
 /   \ 
 NP2  
 /   \  
 the  
 man 
 /   \ 
 V    
 /   \ 
 S3  
 /   \ 
 was 
 /   \ 
 S1  
 /   \ 
 VP  
 /   \ 
 S3  
 /   \ 
 the  
 governor  
 pardoned  
 the  
 man 
```

This can be seen more clearly in a case where the relativized NP is in subject position, such as (i) (derived from (ii)).
The rules could apply simultaneously, since NP₂ would be subject to Wh Rel Movement, while Passive would affect only the remainder of S₂. But this proposal will not work for (i), derived from (ii).

(i) The man who the boy was hit by escaped.

(ii) NP
     /   
    NP   S₂
     |    /   
  the man₁ NP VP
     |    /   
  the boy V
      |    /   
     was S₃
           /   
          NP'
             /   
            VP
               /   
       hit the boy

This suffers the same problem as simultaneous application of the rules to (ii) does. NP', the subject of S₂, must be moved to the left by Wh Rel Movement, but must be moved to the right by Passive. Also, as R. Lakoff (1971) points out, Hasegawa's analysis of the passive is deficient on several grounds, and may not be appropriate for be-passives. Lakoff proposes an analysis in which Hasegawa's analysis is appropriate for get-passives. She proposes that be-passives have a higher predicate be. For example, the deep structure of (iii) is (iv).

(iii) The pie was eaten by Wilt.

(iv) S
     /   
    NP VP
     |    |
  be Wilt ate the pie

This analysis is interesting, but will not get around the problem of simultaneous application of Passive and Wh Rel Movement.

In his discussion of Quasi-Claususes, Postal suggests that constraints on multiple question word binding and double negatives are to be stated in terms of the notion quasi-Clause. As an example of this notion, the underlined NP in (i) are quasi-Clause Mates, while those in (ii) are not.

(i) a. I want Bob to help me.
b. John wants to kiss Mary.
c. I prevented Tom from calling Mary.

(ii)

a. I arranged for Bob to help me.
b. I would prefer for you to call me.
c. I would prefer that you called me.

6 Given a quasi-Clause condition on Reflexivization, we have to find some means of blocking

(i) *I want Bill to shoot myself.
(ii) *Betty believes us to have deceived herself.

But the presence of an intervening noncoreferential NP between an antecedent-reflexive couple would do the job, along lines reminiscent of Grinder's Intervention Constraint for deletion rules (Grinder (1970)).

An Intervention Constraint should apparently block (iii) also, because of the NP intervening between the antecedent-reflexive pair.

(iii) John promised Mary to wash himself.

But this problem arises only if Equi applies before Reflexive to (iv), the deep structure of (iii).

(iv) \[ S_1 \left\{ \begin{array}{c} \text{John} \\
\text{promised Mary} \end{array} \right\} \quad S_2 \left\{ \begin{array}{c} \text{John} \\
\text{wash John} \end{array} \right\} \]

In the absence of extrinsic ordering, Reflexive will apply in \( S_2 \) at the same time as Equi applies in \( S_1 \), and no problem arises.

In the derivation of (i) and (ii) Equi applies first only because that is the only rule applicable to the deep structures of these sentences.

Notice also that (vi) is ungrammatical, in contrast to (v).

(v) We believed Bill to have washed himself.
(vi) *I promised Bill to wash himself.

Apparently, Bill can serve as the antecedent for Reflexive in (v), but not in (vi) in spite of the fact that both sentences contain quasi-clauses. This is a superficial observation, however, for the problem with (vi) resides not in reflexivization at all. The subject of the verb promise, but not the direct object, can be the controller for Equi. The deep structure of (vi) is roughly as in (vii).

(vii) \[ S_1 \left\{ \begin{array}{c} \text{I} \\
\text{promised Bill} \end{array} \right\} \quad S_2 \left\{ \begin{array}{c} \text{Bill} \\
\text{wash Bill} \end{array} \right\} \]

Although Reflexive may apply within \( S_2 \), the subject of promise is not coreferential to the subject of \( S_2 \), so Equi may not apply. Therefore (vi) cannot be generated, no matter what the restrictions on Reflexive are.

7 The Inclusion Constraint prevents NP's under certain conditions
from overlapping in stipulated coreference:

(i) 
   a. *I like us.
   b. *We like me.
   c. *He praised them.
   d. *They, criticized him,.

The deep structure of (13b) is as in (i), the derived structure is as in (ii) where S' is a quasi-clause.

(i) I believe \[ \text{NP} \left[ S \left[ \text{NP} \left[ \overset{\text{John}}{\text{NP}} \right] \right] \right] \overset{\text{likes us}}{\text{NP}} \]

(ii) I believe \[ \text{NP} \left[ \overset{\text{John}}{\text{NP}} \left[ \overset{\text{to like}}{\text{NP}} \left[ \overset{\text{us}}{\text{NP}} \right] \right] \right] \]

Postal does not explicitly demonstrate that Answer Reduction is postcyclic, but I will assume that he can demonstrate this, presumably on the basis that the rule operates only on structures derived by cyclic rules.

As far as I can tell, this is the import of Ringen's proposal. What the paper states is that, given the conditions defined in the antecedent, the obligatory rule must apply first. There seems to me no reason that the obligatory and optional rules could not apply simultaneously, at least in certain cases. The point is that the obligatory rule must not apply to the structure derived by the optional rule, for in this case the obligatory rule would not have applied to the structure defining its structural condition.

It should be pointed out that even if genitive noun phrases lose clausal status before Genitive Head Deletion, the same results obtain, for the formation of a genitive NP from an underlying clause is optional:

(i) \[ \left\{ \overset{\text{Martha's bicycle}}{\text{NP}} \right\} \overset{\text{was stolen}}{\text{NP}} \]

The failure of Reflexivization from genitive NP's is in fact good evidence in favor of their underlying clausal status. If genitive NP's are clausal in underlying structure, the Clause-Mate condition on Reflexive will prevent this rule from applying.

This statement is not entirely complete, for the data in Ross (1969) show that if the element which is moved does not appear in surface structure, the violation is less severe. This matter is further discussed in Chapter Two, section 2.5, below.

If we were using the cyclic principle, we could claim that Passive applied first to the S, cycle. Then when we tried to apply Wh Rel Movement on the S, cycle, the relativized NP would be inside a sentential subject and we could not extract it. I have presented
arguments against the cyclic principle in section 2, and so shall reject this option.

14 Ross (1967, 134–138) gives two other pieces of evidence that the SSC should be included in the grammar of English.

15 The two coordinate structure constraints are:

(a) The Conjunct Constraint: no conjunct of a coordinate structure may be moved out of that structure.
(b) The Element Constraint: no element under the domination of one of the conjuncts of a coordinate structure may be moved out of that structure.

Grosu (1973) shows that certain transformations distinguish between these two constraints (and that therefore they are separate, not parts of a single coordinate structure constraint). For example, VP Deletion is not subject to the Element Constraint, but it is subject to the Conjunct Constraint.

(i) I can't play tennis but my son can and I am proud of him.
(ii) *I can't play tennis, but I know a man who can and run the marathon too.
The VP that is deleted in (ii) is one of the conjuncts of a coordinate structure, but the VP deleted in (i) is only part of one of the conjuncts of a coordinate structure.

16 The precise formulation of this rule is open to debate. A similar rule was proposed by Ross (1967, section 4.2.4.1) under the name of Conjunction Reduction. Hankamer (1971) provides an extensive discussion of rules which effect deletion in coordinate structures, with a revision of the earlier treatment by Ross. Hankamer combines Left Conjunction Reduction and forward Gapping into one rule of Coordinate Deletion, which deletes from left to right in conjoined structures. Likewise, he combines backward Gapping and Right Conjunction Reduction into a single rule which he calls Delay, which deletes from right to left.

17 In Hankamer's terms, the derivation is accomplished by deleting the rightmost occurrences of John and the eggplant, with subsequent application of a principle regrouping the VP constituents. There are other ways the derivation could be accomplished, for example by using both Right Conjunction Reduction and Left Conjunction Reduction à la Ross (1967), but such matters are irrelevant here. See Hankamer (1971, Chapter One) for more discussion.

18 It might be argued that Coordinate Deletion creates coordinate structures in examples like (ii) (derived from (i)).

(i) John plays the piano and Harry plays the piano.
(ii) John and Harry play the piano.
In (i) there is a coordinate S node, and (ii) there is a coordinate NP node. But we have not actually 'created' a coordinate de novo: we have only mapped one coordinate structure onto another coordinate structure. Any element which is part of the coordinate structure island in (ii) was also part of a coordinate structure island in (i). There is a difference: the NP's John and Harry are in (ii) conjuncts of a coordinate structure (subject to the Conjunct Constraint, as defined in note 15) while in (i) they are elements under the domination of one of the nodes of a coordinate structure (and therefore subject to the Element Constraint). But for tactical reasons it is impossible to construct any situation where this difference reveals anything about English syntax.
CHAPTER TWO

GENERAL CONSTRAINTS

In this chapter and the next I discuss the global derivational constraints that have been proposed in the literature. I have tried to be exhaustive: I treat all cases of derivational constraints that I have been able to find. I have put a rough grouping on the examples as follows: those in sections 3.1-3.5 involve movement rules, while those in 3.6-3.11 involve deletion rules or anaphora. The constraints discussed in Chapter Three mention specific rules. They clearly lack the generality in formulation which the rules in this chapter have. The distinction between general and rule-specific constraints will serve as the basis for a discussion of the general properties of global constraints. I shall argue in Chapter Three that a large number of proposed global constraints are not global at all, but that other principles of grammar are responsible for the data cited. I shall therefore indicate in the discussion of the general constraints why it is plausible that they should be global.

2.1 The Complex Noun Phrase Constraint. Koutsoudas (1973) presents evidence that the CNPC is a derivaitonal constraint. I summarized his arguments in Chapter One, section 3.

2.2 The Sentential Subject Constraint. In section 4 of Chapter One above, I argue for the global nature of the SSC.
There is strong evidence that the CNPC and the SSC are part of the grammar of English (and perhaps part of universal grammar). J. R. Ross wrote a seminal work in English syntax (1967) a large part of which is devoted to justification of these island constraints. Anyone working in syntax as the present time knows that the issue of rule ordering is being intensively investigated: new papers on the topic appear at a rate which is alarming to one trying to keep up with the literature. So far, the majority of evidence is on the side of the Unordered Rules Hypothesis (see Koutsoudas (1971), (1973a), (1973b), and the literature cited there). I showed in the last chapter that given the URH, it follows that the CNPC and SSC are global: there seems to be no way around this conclusion.

2.3 The Quantifier Constraint. Lakoff (1971) discusses constraints on quantifiers. A common analysis within Generative Semantics is that quantifiers originate in semantic structure outside of the noun phrases they bind. Sentences (1) and (2) are not synonymous for many speakers. The difference in meaning is due to the different heights of the quantifiers in deep structure. The deep structure of (1) and (2) are represented roughly as (3) and (4) respectively.

(1) Many men read few books.

(2) Few books are read by many men.
'many are the men who read few books'

A rule of Quantifier Lowering lowers a quantifier onto the NP which it binds. The problem is that if Passive applies to $S_3$ in (3), and Q-Lowering applies, we derive (2), which does not have the meaning represented by (3). Also, if Passive does not apply to $S_3$ in (4), the
application of Q-Lowering gives (1), again a wrong result. In order to prevent cases of derivations like the above (as well as others) Lakoff proposes essentially the following constraint:

(5) If there is an underlying structure $P$, in which there are two quantifiers $Q_1$ and $Q_2$, such that $Q_1$ commands $Q_2$, then if $Q_1$ and $Q_2$ command each other in surface structure $P_n$, $Q_1$ must precede $Q_2$ in $P_n$.

Lakoff also speculates that this constraint holds for any stage in a derivation, but this stronger formulation remains an open question.

Constraint (5) will prevent the derivation of (2) from (3).

In deep structure many commands few. Therefore, since in surface structure many and few command each other, many must precede few. Likewise, (1) cannot be derived from (4). Since few commands many in (4), few must precede many in a surface structure where both the quantifiers command each other. In other words, if the command relation is neutralized, the precede relation takes over.

2. Extraction from Complex NP. Postal (forthcoming) presents a global constraint (part V, section A) stemming from observations that NP cannot be extracted from a complex NP, $N$, where $N$ is a subject.

(6) a. Pictures of Bob lay on the table.

b. *Who did pictures of lay on the table?

c. *The one who pictures of lay on the table was Bob.

(7) a. Bill thought pictures of Sally lay on the table.

b. *Who did Bill think pictures of lay on the table?

There is no restriction where $N$ is not a subject:
(8) a. Who did you buy pictures of?
   b. The only person who I heard rumors about was Ted.

Now notice that the restriction applies where N has been raised:

(9) a. *Who did you believe pictures of to be hidden here?
   b. *Who did you prove stories about to be false?
   c. *Who do you expect stories about to terrify John?
   d. *Who do you believe a picture of to be on the table?

Since the relevant NP have been raised, they are no longer subjects.

Postal formulates the constraint in terms of the notion 'cyclic subject'. An NP is a cyclic subject if it has a corresponding constituent NP which is the subject of a clause at the end of some cycle. The restriction is then as follows:

(10) NP cannot be extracted from a complex NP, N, where N is (i.e. corresponds to) a cyclic subject. 2

2.5 Island-forming Nodes. Ross (1969) points out that island constraints hold if the island-forming node is present in the surface structure, but if some transformation has deleted that node then the constraint does not hold. For example, notice (11), claimed by Ross to be derived from (12).

(11) a. ?Irv and someone were dancing, but I don't know who.
   b. ?Mary met a man who had worked for someone famous, but she wouldn't tell me who.

(12) a. *Irv and someone were dancing, but I don't know who
   b. *Mary met a man who had worked for someone famous, but she wouldn't tell me who she met a man who had worked for.

The sentences in (11) are better than those in (12) because the island-
forming node has been deleted by Sluicing. It is therefore necessary to make reference to surface structure, as well as to the point at which the movement rule which is subject to an island constraint applies. Since we need to refer to nonadjacent phrase-markers, the constraint must be global.

Baker and Brame (1972) claim that there is no global constraint here, because (11a) is ungrammatical and while (11b) is grammatical, it is derived from (13).

(13) Mary met a man who had worked for someone famous, but she wouldn't tell me who the man had worked for.

The problem is that for many speakers, including myself, both (11a) and (11b) are partially ill-formed, but much better than their counterparts in (12). For such speakers, (13) could not be the source for (11b), for (13) is fully grammatical and the derived structure should be also. The island constraints have lessened force if the island-forming node is deleted, so clearly a global condition is operative.

Baker and Brame note that the condition quite likely is not limited to island constraints. They give some examples which indicate that '... the effects of any type of node violation are reduced if the node in question is deleted under identity'(62). Although this condition is more general than Lakoff's, it is still global.

2.6 Controller Cross-Over. Grinder (1971b, 97-131) discusses a constraint which he terms Controller Cross-Over. He argues that...
Lakoff (1970) mentions a constraint which prevents application of Passive to a for-to complement sentence to which Equi also applies. Thus, we block:

(14) *To leave was desired by John.

But Lakoff's constraint is actually an instance of Grinder's more general constraint, which applies to whole classes of rules. The application of Passive to (14) causes the underlying subject to move to the right over an NP which is identical in reference to that subject. Yet the NP that the subject sweeps over is deleted by Equi. It is Grinder's claim that the crossover of two NP's which take part in an identity deletion rule is banned.

The generality of Grinder's constraint is shown by appeal to cases where deletion transformations other than Equi, and permutation rules other than Passive, apply.

To begin with deletion rules, Grinder offers four cases where Passive may not apply if one NP moves over another NP, where the former NP serves as controller for deletion of the latter.

(1) By Subject Deletion relates (15a) and (15b).

(15) a. ?Mary irritated John by her ignoring him.

b. Mary irritated John by ignoring him.

Yet while the relative grammaticality of the a version remains the same when it is passivized, the passive of the b version is uninterpretable.

(16) a. *John was irritated by Mary by her ignoring him.

b. #John was irritated by Mary by ignoring him.
In the b version, By Subject Deletion has applied, and Passive causes the controlling NP to move along a path which crosses the position of the deleted NP. Lakoff's Passive/Equi constraint could not block this derivation, for Equi is not involved.

(ii) Pronoun Apparel Elision derives the b from the a sentence below.

(17) a. Sam took Pete's hat off him.
   b. Sam took Pete's hat off.

(17) is ambiguous, since in the structure it is derived from, the pronoun could have either of two NP's as antecedent. Controller Cross-Over predicts that when (17b) is passivized, it will only have one reading. The reading where the subject is the coreferent of the deleted pronoun is not possible, for the subject would cross over the deleted pronoun. This is just what we find:

(18) a. Sam took Pete's hat off.
   b. Pete's hat was taken off by Sam.

In (18a), either Sam or Pete might have been sporting a headpiece, but in (18b), the hat must have been removed from Pete's head.

(iii) Since-When-While Deletion is responsible for the synonymy of (19a) and (19b).

(19) a. Louise saw Don [while] [was leaving home].
   b. Louise saw Don [while] [leaving home].
   c. Don was seen by Louise [while] [leaving home].

(19b) is ambiguous as to whether Louise or Don was leaving home.
Controller Cross-Over predicts that the passive of (19b) will be unambiguous, and this is just what we find. (19c) cannot mean that Louise saw Don while she was leaving home, for in that event the subject would move over the NP for which it is the controller in an identity deletion transformation.

(iv) In Order to Deletion relates (20a) and (20b).

(20) a. Liz called Tom in order that she could get his attention.
   b. Liz called Tom in order to get his attention.

Controller Cross-Over predicts that if the subject of (20) is moved over the NP which was deleted under identity to that subject, the result will be ill-formed.

(21) *Tom was touched by Liz in order to get his attention.

Notice that the passive of (20a) is well-formed, for here there has been no deletion:

(22) Tom was called by Liz in order that she could get his attention.

The data above indicate that Controller Cross-Over is general across the set of coreferential deletion rules. The constraint is also general across the set of movement rules: no matter what the rule is which crosses a controller NP over its coreferent, a violation ensues. Grinder cites ten additional cases of supporting evidence attesting to the generality of the phenomenon. I shall summarize two of these in order to give a flavor of what is involved. The additional evidence may be found in Grinder (1971b, 97-131).

(v) About Movement applies in the derivation of (23b) from (23a).
(23) a. David talked to Mildred about Tom.
    b. David talked about Tom to Mildred.

Now consider the following examples.

(24) David talked to Mildred about perjuring himself;

(25) David talked to Mildred about perjuring herself.

The embedded reflexive enables us to tell what the deleted subject of perjure must have been in each case. Controller Cross-Over predicts that when About Movement applies to (24), the result will be well-formed, as there is no crossover of controller and deleted node:

(26) David talked about perjuring himself to Mildred.

But when About Movement applies to (26), the controller Mildred crosses over the position of the deleted subject of perjure, so the result is ill-formed:

(27) *David talked about perjuring herself to Mildred.

(vi) The rule Psych Movement has operated in the derivation of (28a,b).

(28) a. Sam is annoying to me.
    b. Leroy is amusing to Maxine.

If Psych Movement had not applied to the structures underlying (28a,b) the resultant surface structures would have been (29a) and (29b) respectively.

(29) a. I am annoyed at Sam.
    b. Maxine is amused at Leroy.

Controller Cross-Over predicts that if Psych Movement causes the two
NP's mentioned in the structural description of Equi to cross, the result is ill-formed:

(30) *Washing herself$_i$ is annoying to Marge$_i$.

(30) is derived from a more remote structure something like (31).

(31) Marge$_i$ is annoyed at $\text{\textgreek{s}Marge$_i$ wash herself$_i$}$

Notice that the surface structure is better if Equi does not apply.

(32) ?Her$_i$ washing herself$_i$ is annoying to Marge$_i$.

(32) is not fully acceptable, due to the failure to apply Equi, which is obligatory. But (30) is worse than (32), so there must be an additional restriction at work in (30). Note also the similar sentence (33), where Equi applies, but not Psych Movement.

(33) Washing herself$_i$ annoyed Marge$_i$.

These results are consistent with Controller Cross-Over.

The above results show that it doesn't matter whether Passive, About Movement, or Psych Movement is responsible for movement. The constraint is thus general for the class of identity deletion rules and the class of movement rules. The Controller Cross-Over constraint could not be stated at any single level in a derivation, for it involves the interaction of two sets of transformations, and the members of one set could easily apply at a different point than the members of the other set.

2.7 The Chaining Constraint. Grinder (1971a) discusses what he calls the Chaining Constraint. A set of terms in a pronominalization relationship is said to constitute a chain of coreference. The constraint prevents the deletion of a term in a chain of coreference if that term is an anaphor or if the chain includes an anaphor of
variable reference. Thus, notice that (34b) and (34c) are not synonymous.

(34) a. Laura asked (someone) what time it was.
   b. Laura asked someone_i when he_i was planning on leaving.
   ≠ c. Laura asked when he_i was planning on leaving.

The optional deletion of someone cannot take place if this term is in a chain of coreference. This constraint is global in nature: we must identify the set of all coreferential terms, T_i, in a specified underlying structure. No rule later in the derivation can delete a term in T_i if it does not preserve coreference relations.

2.8 The Deletion Constraint. Hankamer (1971, chapter 3, section 2, 237-252) discusses a derivational constraint which he calls the Deletion Constraint:

There is an output constraint on the transformational component to the effect that any sentence containing a verb must contain overtly all of the complements for which it is strictly subcategorized. (245)

Thus if the object of a transitive verb is deleted, so must the verb be. Some examples illustrating this constraint follow. Hankamer proposes that the rule of Gapping be combined with the rule of Conjunction Reduction into one rule of Coordinate Deletion which deletes from left to right in conjoined structures. If this be the case, then direct objects cannot be deleted by Coordinate Deletion unless the verb is also deleted.

   b. *Harry bought the book, and George bought.
c. Harry bought the book and George read \( \{ \text{the book} \} \).

There are similar results with respect to the rule of Stripping, which deletes all parts of a sentence which is an answer to a question, save for the questioned element. Thus (37) may, when it is the answer to (36), be reduced to (38).

(36) What did Harry do to the eggplant?
(37) Harry boiled the eggplant.
(38) Boiled it.

This rule can delete the object of a transitive verb only if the verb is also deleted:

(39) What did Harvey do to the eggplant?
(40) a. Boiled it.
    b. *Boiled.

Hankamer discusses a rule which he calls Wrong, which deletes from corrective answers to questions everything but the corrected material. 5

(41) Harvey ate the eggplant this morning.
(42) a. No, Albert.
    b. No, at noon.

This rule never leaves any verb stripped of any NP for which it is strictly subcategorized.

(43) Q: Did Albert mop the floor?
    b. *Swept.
(45) Q: Did John lend Harriet the can opener?
Sweep is strictly subcategorized for a direct object; give is strictly subcategorized for an indirect as well.

The above three examples can be accounted for by Hankamer's Deletion Constraint. This constraint is derivational:

... the information required to state this constraint at the level of surface structure is the set of subcategorization features of the verb, which is not available at that level, unless it has been carried along all through the derivation as a set of features on the verb itself. If we presume this to be the case, such a preservation of information which originates at the level of deep structure is clearly equivalent to a derivational constraint. (248)

Hankamer points out (243-44) that this constraint refers only to sentences which undergo identity deletion rules, not to those having free deletion. Notice:

(47) Ivan was reading.

(48) Susan kept on eating.

Grinder (1971b, 23-35) argues that the lack of an overt direct object in (47) and (48) is not due to deletion, but rather to failure to apply the process of 'optional lexicalization' to the unspecified semantic material underlying the direct objects in (47) and (48).

If Grinder is right, we can easily explain the failure of Hankamer's Deletion Constraint to apply to (48), because deletion is not involved.

Also, Hankamer notes that the constraint applies only to topmost sentences: 'every topmost sentence of English which contains a verb also contains all the complements for which that verb is strictly subcategorized (though they may be moved around by movement rules)'(244).
2.9 Anaphoric Islands. Postal (1969) claims that the pseudo-adjectives in (59) are derived from underlying noun phrases like those in (50).

(59) a. the **Chinese** attempt to invade **Panama**.
   b. the **governmental** claim that the men were spies
   c. the **imperial** attack on the outlying islands

(50) a. China's attempt to invade Panama
   b. the government's claim that the men were spies
   c. the empire's attack on the outlying islands

The contrasts in (51)-(53) show that pseudo-adjectives may not take part (on the surface) in a coreference relation.

(51) a. China's attempt to destroy itself
   b. *the Chinese attempt to destroy itself

(52) a. the government's justification of itself
   b. *the governmental justification of itself

(53) a. the empire's claim that the democracies were trying to destroy it
   b. *the imperial claim that the democracies were trying to destroy it

In (59) there is a coreference relation, but due to the application of Equi there is no **surface** pronoun coreferential to the NP forming a pseudo-adjective. Since there must be a pronoun present in order for blockage to manifest itself, Postal states the constraint as:

(54) Throw all derivations in which both:
   
   (i) the semantic representation contains a pair of NPs, $N^a_\text{P}$ and $N^b_\text{P}$, which are stipulated coreferents; and
(ii) a. the surface-structure correspondent of NP<sub>a</sub>
is a pseudo-adjective; and

b. the surface-structure correspondent of NP<sub>b</sub>
is a pronominal NP.

2.10 Wh Crossover. There is an argument for a global constraint on pronominalization in Postal (1972b). This constraint is stated by Postal as follows:

(55) Mark as ill-formed any derivation in which:

a. there are two nominal constituents A and B, in the output structure of a Wh Movement rule, where:
   (i) A is a pronoun
   (ii) B is a wh form
   (iii) A is to the left of B, and

b. the corresponding constituents of A and B in the output structure of the Wh Movement rule, call them A' and B' respectively, are aligned such that B' is to the left of A'; and

c. in the semantic representation, A and B (or, more precisely, their corresponding elements) are marked as stipulated coreferents.

This constraint will block the derivation of the sentences in (56).

(56) a. *Who<sub>i</sub> did Mary say his<sub>i</sub> mother yelled at?

b. *What<sub>i</sub> did the fact that it<sub>i</sub> got rusty prove was no good?

c. *What official<sub>i</sub> did the policeman who arrested him<sub>i</sub> claim was drunk?

d. *Which tigers<sub>i</sub> were there<sub>i</sub> keepers eaten by?

In the structures before the application of the movement rules in these sentences, indicated schematically in (57), the wh form is to the right of the pronoun with which coreference is banned in the surface form.

(57) a. [Mary said his<sub>i</sub> mother yelled at who<sub>i</sub>]

b. [the fact that it<sub>i</sub> got rusty proved what<sub>i</sub> was no good]
c. [the policemen who arrested him\textsubscript{1} claimed what official\textsubscript{1} was drunk]\\
d. [their\textsubscript{1} keepers were eaten by which tigers\textsubscript{1}]

However, the sentences in (58) are not blocked

(58) a. Who\textsubscript{1} do you think said Mary yelled at his\textsubscript{1} mother?

b. The car\textsubscript{1} which\textsubscript{1} Bill said got rusty later developed a hole in its\textsubscript{1} exhaust pipe.

c. What official\textsubscript{1} claimed he\textsubscript{1} had been roughed by the police?

d. Which tigers\textsubscript{1} did you claim indicated that they\textsubscript{1} wanted a better grade of keeper?

But in the pre-Wh Movement stage of these sentences the wh form is to the left of the pronoun, so the Wh-Constraint does not apply.

(59) a. [b. the car\textsubscript{1} Bill said which\textsubscript{1} got\textsubscript{1} rusty later developed a hole in its\textsubscript{1} exhaust pipe]\\
b. [the car, Bill said which\textsubscript{1} got rusty later developed a hole in its\textsubscript{1} exhaust pipe]\\
c. [what official\textsubscript{1} claimed he\textsubscript{1} had been roughed up by the police]\\
d. [you claimed which tigers\textsubscript{1} indicated that they\textsubscript{1} wanted a better grade of keeper]

It is of interest here that the two levels of structure referred to here are contiguous, while typically derivational constraints are defined on noncontiguous trees. This casts some suspicion on the constraint. Furthermore, I argue in Chapter Three that global constraints may not be rule-specific. This constraint is not quite rule-specific, but it applies only to cases where either Wh Q Movement or Wh Rel Movement applies. These rules do not constitute a large-scale general class in the way that deletion rules, for example, do.
In fact, Cole (1972, 1973a) provides evidence that we do not have a global constraint here. Cole argues that there is a constraint against backward pronominalization which blocks the relevant sentences. Postal's claim that movement is involved stems from two facts:

(a) Incredulity questions, which apparently have the same underlying form as the cases in which movement applies, are grammatical:

(60) a. Mary said his\textsubscript{1} mother yelled at which\textsubscript{1}?
   b. The fact that it\textsubscript{1} got rusty proved what\textsubscript{1} was no good?
   c. The policemen who arrested him\textsubscript{1} claimed what official\textsubscript{1} was drunk?
   d. Their\textsubscript{1} keepers were eaten by which\textsubscript{1} tigers?

(b) Multiple \textit{wh} questions are grammatical:

(61) What tyrant tortured her\textsubscript{1} mother in front of which helpless princess\textsubscript{1}?

(60) and (61) have not undergone movement, while the sentences in (56) have. But Cole claims that there is not a restriction on movement, rather that ungrammaticality is due to properties of the structure of (56) prior to \textit{wh} movement. Cole presents evidence that \textit{wh} movement does not have to do with the ungrammaticality of these sentences. In particular, it is Cole's contention that (56) and (60) differ in underlying structure. The differences in underlying structure could be responsible for the difference in grammaticality of (56) and (60), and a constraint on \textit{wh}-movement would not be necessary. Cole's argument is as follows. First, he states 'I shall assume that sentences differ in meaning when they can be used in different situations. Questions differ in meaning when they have different sets of appropriate replies'(1972, 5).
Cole goes on to show that different replies are appropriate to incredulity and multiple wh questions on the one hand, as opposed to questions like (56a-d) on the other hand. The data cited by Cole are the following (p. 6):

(62) a. Speaker 1: Did you hear about it? They arrested the president of Eastern Alaska State University for possession of over a kilo of grated banana peel!

b. Speaker 2: They arrested who for possession of over a kilo of grated banana peel?!

b'. Who did they they arrest for possession of over a kilo of grated banana peel?

c. Who did they arrest for possession of over a kilo of grated banana peel?!!

(63) a. Speaker 1: Say, Max! Did you hear about it? Pat Nixon ran off with a Red Guard and was inducted into the Red Army!

b. Speaker 2: Pat Nixon ran off with who?!!

c. Speaker 2: Who did Pat Nixon run off with?!!

(64) Speaker 2: Oh yeah? That's nice.

(65) a. Who did you say Pat Nixon ran off with?!!

b. You said Pat Nixon ran off with who?!!

(62) shows that incredulity questions are used in different circumstances than genuine requests for information. In this case, we may or may not have wh movement associated with the incredulity question. But (63) shows a differentiation of the movement and non-movement cases. Immediately after hearing (63), Max might respond with either (63b) or (63c). But suppose Max is absorbed in his work and responds with (64). If Max later meets speaker 1, he could ask (65a) but not (65b). Increduility questions with wh movement are appropriate
both immediately and at a later time, while those without wh movement are only appropriate immediately. Cole concludes:

I have shown that there are differences in meaning which must reflect different underlying structures for 1) genuine requests for information, 2) immediate incredulity questions, and 3) incredulity questions not asked immediately. These differences in semantic structure correlate with which sorts of questions permit backward pronominalization.

Since incredulity questions and genuine requests for information have different underlying structures, backwards pronominalization can be blocked for genuine questions on the basis of semantic structure, and no crossover constraint is needed. Cole does not give in detail what the different semantic structures might be, but what is important is the correlation he points out. The constraint on backwards pronominalization applies to some semantic structures but not to others. We do not need a constraint on wh movement.

Postal's constraint can be maintained only in the absence of any plausible nonglobal alternative. In this case we have an alternative which is more than merely plausible, for Cole presents strong evidence in favor of his analysis. This stems from dialect/idiol ect variation. An inherent weakness of a formal descriptive measure such as (55) is that it cannot adequately account for linguistic variation. Cole presented informants with (66) and asked them to compare acceptability of the sentences. (66a) and (66b) are incredulity questions. (66c) is a genuine request for information. The results are given in (67).

(66) a. The claim that he₁ was drunk bothered who₁?!!
   b. Who₁ did the claim that he₁ was drunk bother?!!
c. Who did the claim that he was drunk bother?

<table>
<thead>
<tr>
<th></th>
<th>No wh movement (a)</th>
<th>Incredulity Question with wh movement (b)</th>
<th>Incredulity Question without wh movement (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Group II</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Group III</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Group IV</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Group V</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

The largest group was Group I. There was no informant who accepted (66c) and rejected (66a) or (66b). Postal's idiolect falls into Group II.

This variation presents serious problems for Postal's position. For Postal must claim that a certain dialect either has the constraint or is does not, and the prediction is that only Groups II and V exist. But this prediction is empirically disconfirmed, drastically under-cutting the validity of Postal's analysis. Note also the implicational relationship. Speakers reject incredulity questions like (66a,b) only if they also reject questions like (66c) (although there is variation as to which type of incredulity question they reject). This indicates that the constraint is one blocking backward pronominalization with antecedents that are genuinely questioned, and that it has been extended to other sorts of wh antecedents. Postal would have to claim that the rejection of a sentence in Groups II and III is on some other basis than the wh constraint, for not all cases involving wh movement are blocked. We would need
an additional principle to account for the ungrammaticality of
(66a) for Group IV. It should be clear that a unified analysis
is preferable, for there is no evidence that separate constraints
are involved.

Cole's position can account for the variation involved. He
claims that speakers fall into two groups: those who accept backward
pronominalization with genuinely questioned antecedents, and those who
do not. For some people in the second group, the constraint is
extended analogically to other wh antecedents. To summarize,
Postal's constraint is not the only way to account for the difference
in grammaticality of (56) and (60), for these must differ in underlying
structure. Furthermore, Postal's constraint cannot account for
dialect variation in a unified way.

Cole then presents a similar argument concerning questions like
(61) having multiple wh nominals. Cole questions whether sentences
like (61) are acceptable as genuine requests for information. For
most people (61), like incredulity questions, is acceptable if the
speaker knows the answer to the question. (61) is acceptable as a
quiz show question, but not as a genuine request for information.
Thus (61) would be grammatical because backward pronominalization
is not banned where there is not a genuine request for information.
Cole gives some evidence that even if (61) were a genuine request,
there are special properties of being the second nominal which would
allow backward pronominalization to take place.
The force of Cole's work is to indicate that there is a constraint against backwards pronominalization with genuinely questioned nominals, and that this constraint has been extended to: incredulity question nominals and wh relativized nominals. Genuinely questioned nominals are indefinite, so this case may be viewed as part of a wider prohibition against backward pronominalization with indefinite antecedents. For further details concerning backwards anaphora with indefinite noun phrases, see Cole (1973b). The analogical extension of the constraint is further explored in Cole (1973a). Cole shows how semantically motivated regularities become generalized, and lose their semantic motivation. Cole presents considerable cross-linguistic evidence in favor of analogical extension of constraints on wh forms in questions to wh forms in relative clauses, based on their morphological similarity. In languages where question formation and relativization are closely related processes, the constraint is extended, but where these are not closely related processes, the constraint is not extended. Postal's treatment cannot predict in which languages the backwards pronominalization constraint will be extended to relativized antecedents.

Cole's treatment offers a natural basis for the constraint: speakers understand personal pronouns to be definite when they appear in isolation, or more precisely, when they do not follow an antecedent. When a pronoun precedes its antecedent, speakers suppose that a definite referent exists. But then it could not have an indefinite antecedent. Since Cole's treatment offers an explanation, it is preferable to Postal's purely descriptive treatment.
2.11 Equi/Pronominalization Confluence. Postal (1970) presents a case which has the character of a derivational constraint, although it is not explicitly so-called. Postal first shows that Equi is cyclical and that Pronominalization (if indeed it is a rule) is post-cyclical. He then shows that Equi is dependent on Pronominalization in the sense that certain restrictions on Equi are the same as restrictions independently needed on Pronominalization. The constraints on Equi can be explained if this rule applies to structures which have already undergone Pronominalization. As a way out of the inconsistency, Postal proposes that NP's to be deleted are marked with the ad hoc syntactic feature [+Doom]. Then there is a postcyclic rule of Doom Erasure which deletes NP's marked [+Doom] if they have undergone Pronominalization. We are thus able to unify the constraints on Pronominalization and Equi. Postal's solution is an ad hoc notional device to express the presence of a global constraint. We have to refer to the point before Equi applies, and to the point after constraints on Pronominalization apply.

Postal presents several arguments that restrictions are shared between what I shall call overt surface anaphors and null surface anaphors (e.g. the output of Equi). I use these terms for convenience, and ascribe no theoretical significance to them: in particular I do not claim that there is a null surface anaphor \( \emptyset \) to which rules may make reference. I merely use this term to abbreviate the output of Equi.

(a) The Indefinite Argument. There is a constraint blocking backwards anaphora with indefinite nongeneric NP's as antecedent.
(68) a. If the man_{i} comes, arrest him.
   b. If he_{i} comes, arrest the man_{i}
   c. If a man_{i} comes, arrest him.
   d. *If he_{i} comes, arrest a man_{i}

(69) *The man who lost it_{i} needs to find something_{i}.

Notice that Equi does not apply if the complement subject would refer to an indefinite NP.

(70) *Finding out the answer pleased some students.

(71) *Winning is difficult for somebody.

(b) The Blocked Forward Pronominalization Argument. There are cases where forward anaphora is blocked:

(72) Realizing that he_{i} was going to be fired bothered Bill_{i}.

(73) *Realizing that Bill_{i} was going to be fired bothered him_{i}.

Not only is forward anaphora blocked in cases like these, but Equi does not occur. Notice:

(74) a. *His_{i} realization that Bill_{i} was going to be fired didn't worry him_{i}.
   b. *The realization that Bill_{i} was going to be fired didn't worry him_{i}.

(75) a. His_{i} realization that he_{i} was going to be fired didn't worry Bill_{i}.
   b. The realization that he_{i} was going to be fired didn't worry Bill_{i}.

In (75) Pronominalization can occur and Equi can occur. But in (74) Pronominalization cannot occur. The ill-formedness of (74b) would follow from this restriction on anaphora if Equi applied to a structure like (74a). (But Postal argues it can't, because of ordering restrictions.)
(c) Modal Constraint Arguments. In (76), he may refer to either Harry or Max.

(76) Harry told Max that he was ill.

But in (77), he may refer only to Max:

(77) Harry told Max that he ought to visit Greta.

Postal proposes an Ought Modal Constraint which prevents a pronoun from referring to subjects of higher sentences when the pronouns are subjects of complements with ought or should, and the complements are embedded with certain verbs. Now if (79) is derived from (78),

(78) Harry told Max that he should bring the tickets.

(79) Harry told Max to bring the tickets.

then the fact that the deleted complement subject in (79) refers to Max is explained. For he is prevented from referring to Harry; the latter could not be the controller NP for deletion. The constraint on deletion automatically follows from the Ought Modal Constraint on coreference.

(d) Arguments from Plural Coordinate Constraints. In (80) the deleted NP is either (a) you or (b) John and Bill jointly.

(80) John and Bill spoke to you about going to New York.

On the latter reading, the sentence is derived from (81):

(81) John, and Bill spoke to you about their going to New York.

Why could not the deleted subject in (80) refer to either John or Bill singly? Because with a conjoined subject, a following anaphor cannot refer to either conjunct singly.
(82) *John\textsubscript{j} and Bill\textsubscript{j} spoke to you about \(\{\text{his}_i\} \) going to New York.

Since Pronominalization cannot occur in (82), Equi cannot apply. This explains why the deleted NP in (80) could not refer to John or Bill singly. Again, a constraint on coreference is reflected in the operation of Equi.

Certain restrictions on Equi exactly parallel restrictions on Pronominalization. By having prior pronominalization, we can avoid stating the restrictions twice in the grammar. In Postal's summary:

Certain otherwise ad hoc restrictions on Equi are predictable from independently necessary constraints on Pronominalization, if Equi applies to structures only when they have already undergone Pronominalization.

But because of extrinsic ordering problems, Postal is driven to split deletion into two operations, applying at separate points in the grammar: one, a rule which marks segments for \([+\text{Doom}]\); and the other a rule actually deleting segments marked \([+\text{Doom}]\). If elements are only marked for deletion, rather than actually deleted, they will still be present later in the derivation for pronominalization constraints to apply to them.

Postal's analysis is based on two assumptions:

(a) there is a transformation of Pronominalization,

(b) there is extrinsic (and cyclic) rule ordering.

The problems with the use of a transformation of Pronominalization are given in McCawley (1972) and Fauconnier (1971). I have noted in several places in this essay that extrinsic ordering has very shaky foundations, and I argued in Chapter One that there is little evidence for a cycle.\textsuperscript{6}
Any argument can only be as good as the assumptions it is built upon. The shakiness of Postal's assumptions casts some doubt on his analysis. Furthermore, the resulting analysis differs from other cases of global constraints. It is a constraint that 'looks ahead' in a derivation, and it is actually transderivational. If we wish to block Equi by reference to constraints on Pronominalization, we would have to formalize the constraint in something like the following manner: At the point in derivation $D_1$ where Equi applies, check another derivation, $D_2$, which is identical to $D_1$ except that Equi does not apply. If at a later point in derivation $D_2$ Pronominalization of the NP corresponding to the NP in $D_1$ which Equi would delete is blocked, then Equi may not apply in $D_1$. This seems like too much machinery to add to the grammar in order to preserve extrinsic ordering. Little is known about transderivational constraints, but we would certainly want extremely convincing evidence to add them to the theory of grammar.  

My feeling is that with the elimination of extrinsic ordering, Postal's problem does not arise. We may let the process which introduces pronouns apply before Equi. Then whatever the restrictions are which exist on the occurrence of pronouns, they will be reflected in the operation of Equi.

It is my feeling that pronouns do not result from a rule of Pronominalization which adds the feature [+Pro] to a node. The approach using a rule of Pronominalization was taken by Lees and Klima (1963), who present an analysis of anaphora in which the rule of Pronominalization converts a full noun phrase to a pronoun under
certain conditions. An alternative approach has been developed with the advent of a more semantically-based syntax. Within the framework of Generative Semantics, the semantic representation of a sentence must include, at least in part, its logical structure.

The Lees and Klima analysis is deficient in that it does not properly explain the relation between the logical forms and surface forms of sentences containing pronouns. I believe that a proper explanation of this relation must take a form roughly along the lines sketched by McCawley (1970). McCawley's analysis maintains that noun phrases are represented as variables in a predicate calculus, and that there are rules which specify that some variables are replaced by a full noun phrase while the other occurrences of the variable are replaced by pronouns. This proposal is superior to that presented by Lees and Klima on syntactic as well as semantic grounds, for it avoids certain syntactic difficulties inherent in the transformational rule treatment of pronominalization.

Given this proposal, the constraints on pronominalization must be reformulated as constraints upon which variable can be filled in by a full noun phrase. If the replacement of variables by pronouns takes place before deletion, then we can quite naturally explain why the rule Equi reflects the constraints on anaphora. Such an explanation is not available to Postal, for his analysis implies that any given language might either have or not have the rule Doom-marking. It is then an accidental fact about English that Equi reflects constraints on anaphora.
The insertion of pronouns prior to the application of Equi can be insured be a general principle postulated by Postal: that any deletion of an NP under conditions of coreference can take place only when that NP is pronominal. Or as Postal (1970, 489) puts it:

(83) Universal Deletion Constraint

If a transformation T deletes an NP, subject to the existence of a coreferent NP, NP, in the same structure, then at the point where T applies, NP must be pronominal.

The plausibility, and possible greater generality, of this condition is indicated by Perlmutter and Orešnik (1973), who show that in Slovenian identity of sense deletion goes through a pronominal stage. They show that many apparently idiosyncratic facts about adjective agreement in Slovenian can be given a general explanation if a pronominal form is present at a stage prior to deletion. They suggest that 'the universal proposed by Postal may be but a special case of a more general principle: any deletion of a noun phrase under identity to another noun phrase (whether identity of sense or identity of reference) passes through a pronominal stage' (442).

Given the Universal Deletion Constraint (UDC), Equi could not delete an element prior to insertion of pronouns, for application of Equi would depend on prior insertion of pronouns. The UDC will insure that the insertion of pronouns precedes deletion, and make the presence of intermediate pronouns available in grammatical explanation. A substantive claim about Universal Grammar is embodied in the proposal.

If the UDC is correct, there are some interesting implications for the Doom-marking analysis. These implications will hold also
for a global rule analysis, which is descriptively equivalent to the
Doom-marking proposal. It is only on abstract theoretical grounds
that they differ. I will refer to Postal's analysis, and its
equivalent(s), as the Global Deletion Analysis. Consider what the
Global Deletion Analysis predicts about Universal Grammar. In any
particular language either deletion precedes pronoun-insertion, or
pronoun-insertion precedes deletion. If pronoun-insertion precedes
deletion, then constraints on overt and null surface anaphors will
be shared. Alternatively, deletion may precede pronoun-insertion
(which Postal claims to be true of English). Then for any such
language, Global Deletion (i.e. Doom-marking) may or may not exist.
If Global Deletion does exist, then overt and null surface anaphors
will share restrictions, just as for languages wherein pronoun in-
sertion precedes deletion. But if Global Deletion (i.e. Doom-marking)
does not exist then overt and null surface anaphors will not share
restrictions. Remember, Global Deletion was introduced in order to
account for the sharing of restrictions on null and overt surface
anaphors. The Global Deletion theory makes the prediction that in
some languages restrictions on null and overt anaphors will be shared,
and in other languages they won't be. Now if restrictions are
always shared, the Global Deletion theory (or the descriptively
equivalent Doom-marking theory) is disconfirmed. If it turns out
that restrictions are always shared, the analysis without Global
Deletion is preferred, for the universal sharing of constraints by
different types of anaphors follows from the UDC in conjunction with
the lack of extrinsic ordering. For practical reasons, we cannot at
the present time examine all the languages of the world in order to
determine whether or not in all of them restrictions on null and overt
anaphors are shared. But since the analysis with the UDC and with-
out Global Deletion makes a stronger empirical claim about universal
grammar, I prefer it to the Global Deletion theory.

In fact if the Universal Deletion Constraint is correct, the
Global Deletion analysis makes an implausible claim about universal
grammar. Suppose that in a given language deletion precedes pronoun-
insertion and that the language lacks Global Deletion. Given the
UDC, deletion could not apply to an underlying form, for pronouns
would not be present until later in a derivation than the point
at which deletion applies. But Doom-marking can't apply either,
for the language lacks the rule. Then later in the derivation
pronoun insertion takes place, but deletion cannot now occur,
because of the ordering restriction. We thus predict that some lan-
guages lack identity deletion rules. But the existence of such rules
is highly natural, for they convert unwieldy semantic structures
into more compact surface structures, and avoid the needless repiti-
tion (in the surface form) of much material.

Although the considerations discussed in the last few para-
graphs is speculative due to the lack of our substantive knowledge
about universal grammar, I believe they indicate a certain insuffi-
ciency in the Global Deletion analysis. Furthermore, given the UDC
and the lack of extrinsic ordering, such an analysis is not necessary.
2.12 Summary. In this chapter I have presented those global constraints that are general in their formulation: they do not mention the name of any particular rule. It seems that two of the constraints, the Wh Crossover Constraint (section 2.10) and the Doom-marking Constraint (section 2.11) are not actually global constraints. This foreshadows the task of the next chapter: to show that many apparently global constraints are only apparently global.

It might be argued that there are no global constraints in syntax, either general or rule-specific. I think this is too strong a claim. The way to show that a constraint is not global is to show that a constraint is not global is to show that it can be stated at deep structure, surface structure, or after the application of some specified rule. In the last case mentioned, we are really talking about the way some rule operates; i.e. the restriction is on the operation of that rule.

But the constraints presented in this chapter could not be reformulated in any of these ways. It should be clear that they could not be stated at some level of derivation defined by a rule R_i. To state the constraint at this level is to claim that something about the way R_i operates (but not the way other rules operate) is crucial to the constraint. But this is just to claim that the rule R_i is crucial in the formulation of the constraint, and in this event the constraint is not general but actually rule-specific. The definition of general constraint precludes reformulation in the above manner.
To state that a constraint is a deep structure is to claim that some aspect of the meaning of the sentence is ill-formed: that it is contradictory, violates selection restrictions, etc. In a framework where transformations preserve meaning, the resulting surface structure will be unacceptable no matter what rule applies. But the data motivating the constraints presented in this chapter show in each instance that the application of a member of some class of rules is necessary for the instigation of deviance. This means that in cases where a member of that class of rules does not apply, the result is acceptable. For example, the derivation including the deep structure underlying (84) is unacceptable only if About Movement results in violation of Controller Cross-Over (section 2.6).

(84) a. David\textsubscript{i} talked to Mildred\textsubscript{j} about perjuring herself\textsubscript{j}.  
    b. *David\textsubscript{i} talked about perjuring herself\textsubscript{j} to Mildred.

Because of the necessity of referring to a derived structure, and presence of alternative well-formed surface structures, the constraints could not be reformulated at deep structure.

On the other hand, it might be possible to reformulate a given rule-specific constraint in terms of deep structure, because a specific rule can only apply to certain types of deep structure (i.e. those which meet the structural description of the rule).\textsuperscript{10} In certain cases, then, it might well be the case that some aspect of certain of the types of structure to which a rule applies is ill-formed, making the constraint a deep-structure constraint. An example of this type is presented in section 3.8 below, where a subclass of the type of structures to which a rule can apply is ill-formed.
Such a reanalysis would not be possible for general global constraints, for since they do not involve specific rules, they are not defined over a specific class of structures.

To state that a constraint is a surface structure constraint is to say that there is some specific configuration of surface structure which is banned. A general global constraint covers a large variety of rules. In the case of some constraints we have to mention two classes of rules. For example, Controller-Cross-Over mentions the class of identity deletion rules and the class of permutation rules. Different rules (even different rules of the same class) have different kinds of effects on what the surface structure will be. Thus it is unlikely that we could find any specific surface structure that is in common to all the banned sentences. For example, what is in common in the surface structures of these sentences which are blocked by Controller Cross-Over?11

(85) a. *To leave was desired by John.
   b. *John was irritated by Mary by ignoring him.
   c. *David talked about perjuring herself to Mildred.

Or what is in common to the surface structures of the following sentences blocked by the Comples Noun Phrase Constraint?

(86) a. *What did the man who bought left the store?
   b. *There's the girl who I know the man who kissed.
   c. *It's Alice who the man who married is Swedish.

Similar examples showing the diversity of possible banned surface structures can be constructed for each general constraint.
It might be possible, given sufficient time and ingenuity, to find for each case a list of possible types of blocked surface structure: e.g. we could block surface structures of the type represented by (85a), of the type represented by (85b), of the type represented by (85c), etc. But obviously we would only have a list of unrelated constraints and we could not explain the unity inherent in the whole class of types of constraints. A single constraint blocks each of (85a-d), not four different constraints. It is thus difficult to see how one could possibly reformulate any of the general constraints in this chapter as surface structure constraints.

Furthermore, there are cases where it is literally impossible to reformulate a general global constraint as a surface structure constraint. These are the cases where one of the readings of a potentially ambiguous sentence is blocked. For example, the Quantifier Constraint (section 2.3) blocks the reading of (87) where many has wider scope than few.

(87) Few books are read by many men.
Controller Cross-Over blocks the reading of (88) where Sam removes a hat from his own head (cf. (17)-(18) above).

(88) Pete's hat was taken off by Sam.
Likewise, this constraint blocks the reading of (89) where Louise was leaving home (cf. (19)above).

(89) a. Don was seen by Louise while leaving home.
   b. Louise saw Don while leaving home.
(89), to which the constraint does not apply, is ambiguous. In the above cases we obviously could not use a surface structure constraint, for the surface structures are perfectly well-formed, although they may not be related to certain semantic structures. General constraints which have the property of filtering out one of the readings of a sentence can clearly not be reformulated as surface structure constraints.

I have sketched above the main reasons why it is unlikely that the general global constraints could be reduced to nonglobal accounts. Therefore, we must conclude that the general constraints presented in this chapter in sections 2.1-2.9 can only be stated in global terms.
The formulation presented here is due to Schiebe (1970). Lakoff splits the constraint into two parts, but there seems to be no difference in the effects of the different formalizations.

Postal notes that, as Kuno, McCawley, and Ross have pointed out, the constraint may be more general than one which refers only to subjects:

(i) a. I consider a picture of Betty terrible.
   b. *Who do you consider a picture of terrible?

(ii) a. I gave a disciple of the swami a lot of money.
   b. *The scholar who I gave a disciple of a lot of money is in jail.

An anaphor of variable reference is one that cannot represent a bound variable unless the antecedent appears in the surface structure. These are himself, herself, itself, etc. (ii) contains an anaphor of variable reference, while (i) does not.

(i) It was pleasant to buy myself some new clothes.
(ii) *It was pleasant to buy himself some new clothes.

Gapping applies not just to verbs, but in Hankamer's terms to the segment \( V = \text{(aux)} \, \text{(adv)} \, V \, (\text{ind.} \, \text{obj.}) \). This is in order to account for cases like (i), where more than just the verb has been deleted.

(i) John quickly gave Alice the ring, and George the key.

But the category \( V \) is not a natural category, but one created just for the purposes of stating the rule efficiently. For a critical look at such arbitrary categories, see Zwicky (1968).

Some speakers require a no before the response, some do not.

Note that Postal does not show that Equi must precede Pronominalization. He argues that Equi is cyclical and Pronominalization is postcyclic, and that for this reason Equi must apply first.

It may well be that the only bona fide cases of this phenomenon are those which have the character of Hankamer's No-Ambiguity Condition (discussed below in Chapter Three, section 4.12): some derivation is blocked if the output of that derivation is identical to the output of some other derivation. That is in certain cases, ambiguity is not allowed.

See Bach (1970), where it is shown that if there is a rule of
Pronominalization, certain sentences cannot be given a finite underlying structure.

9 There is also the possibility that the constraint could be stated at shallow structure. But at the present time the notion 'shallow structure' is ill-defined, and may not even exist as a single well-defined level of structure.

10 Of course, a specific rule might apply to derived structures, but it will always be possible to find instances where a given rule can apply directly to some deep structure.

11 The structures prior to the operation of the relevant rules are respectively:

(i) a. [the man who bought wh-something] left the store
    b. [there's the girl I know the man who kissed wh-the girl]
    c. [the man who married Alice] is Swedish
CHAPTER THREE
RULE-SPECIFIC CONSTRAINTS

Fauconnier (1971) noted that two types of global constraints had been proposed. On the one hand, there are particular constraints that mention specific rules in specific languages. On the other hand, there are general constraints (not necessarily universal: quantifier constraints do not hold for all dialects of English) which do not refer to specific rules of grammar. Fauconnier speculates: "... it may be possible to dispense altogether with language-specific global constraints" (255). To be sure, the issue of generality of constraints and the issue of universality are partially separate. But only partially: showing the generality of constraints is a preliminary to showing their universality.

A constraint that mentions a specific rule in a particular language is not universal. The issue of the generality of constraints is interesting enough to worth pursuing further. Of the constraints presented in the literature some are general and some are particular. But are the particular constraints really global? In this chapter I examine this question.

Suppose we wish to examine the general properties of dollar bills, but some are counterfit. Obviously, we first have to elimin-
ate the bogus bills. Likewise, if we wish to examine the general properties of global constraints, we must eliminate the counterfit constraints. There is a tendency for proponents of global rules to justify everything that has been proposed as a global constraint. But certainly no one who accepts transformations believes that everything that has been proposed as a transformation is in fact a transformation. I now turn my attention to those constraints that are highly specific in character in order to show that they may be removed from the class of global constraints.

3.1 Passive/Equi. It was noticed by Robin Lakoff (reported in G. Lakoff (1970)) that no single lexical item may take a for-to complementizer and undergo both Passive and Equi:

(1) a. Sam expected to leave last night.

   b. *To leave last night was expected by Sam.

This restriction requires use of a global constraint. However, Grinder (1971b, 97-131) shows that the examples given by Lakoff are blocked by a far more general constraint, Controller Cross-Over, discussed in Chapter Two, section 2.6 above. This more general constraint refers to classes of rules, rather than to specific rules. We do not have a rule-specific global constraint, for the data cited follow from the rule-general global constraint Controller Cross-Over.
3.2 Greek Case-Agreement. Andrews (1971) argues that in Greek 'a predicate modifier agrees with that NP which was its subject at the end of the first cycle applying to that predicate modifier' (147). Notice the examples (2)-(6).

(2) Tauta dikia estin.
   (nom.) (nom.)
   'these things just be'
   'These things are just.'

(3) Tauta legetai dikai einai.
   these things-are said-just-be
   'These things are said to be just.'

(4) Ismen tauta legemena dikai einai.
   (acc.) (acc.) (acc.)
   we know-these things-being said-just-be
   'We know these things to be said to be just.'

(5) Emmenomen toutois has Ismen legomena.
   (dat.) (acc.) (acc.)
   we abide-by-those things-which-we know-being said
   dikai einai.
   (acc.)
   just-be
   'We abide by those things which we know are said to be just.'

(6) Emmenomen hois Ismen legomenois dikafois einai.
   (dat.) (acc.) (acc.)
   'We abide by what we know is said to be just.'

In (3) tauta and dikai agree although tauta has undergone Raising and Passive. In (4) tauta and legomena agree where tauta has undergone Raising, Passive, and Raising again. This indicates agreement takes place at a late level. But in (6) toutois is optionally deleted from the structure that underlies (5), and as a result the relative pronoun becomes dative (hois). The adjectives and participles switch to dative, although they are not in the right configuration for the agreement rule which gives a predicate modifier
the same case as the NP it modifies. Case marking takes place at a late level, but an earlier level must be examined in order to tell what NP a modifier is to agree with. (This example is also discussed in Lakoff (1970), (1972), Smonds (1973); Baker and Brame (1972), and Perry (1973).)

3.3 French Adjective/Participle Agreement. Casagrande (1970) discusses global .rules in regard to grammatical agreement. He discusses Greek case agreement (see 3.2 above) and proposes a similar solution to the problem of agreement of adjectives and past participles in French. An example of agreement of adjectives is given in (7), of past participles in (8). The underlined elements agree in gender and number.

(7) Marie est jeune. 'Marie is young.'
(8) Noun les avons toujours admirées. 'We have always admired them.'

Casagrande argues that there is a single agreement rule for adjectives and past participles. This rule must occur after any object placement rule, for a past participle agrees with its direct object only if the direct object precedes the past participle. There are two object placement rules, a syntactic rule which moves pronominal NP, and a stylistic rule which moves full NP. The agreement rule then is as follows: The adjectival element of a deep être-verb agrees (in gender and number) with its deep subject NP and the adjectival element of a deep avoir-verb agrees with its deep direct object if that object is to the left of the adjectival element in question. Agreement applies after rules which permute
direct objects, but must make reference to the notions deep subject and deep object. The rule must therefore be global in nature.

It appears as if the two constraints just discussed mention specific rules. I suspect, however, that no global constraints in syntax are rule-specific. This means that an apparent rule-specific global constraint is either an instance of a general constraint or it is not global in nature.

If the Greek example is considered in isolation, it appears to require a rule-specific constraint. However, when the Greek and French examples are considered together, it is obvious that the same phenomenon is exhibited in both cases: some node is stipulated as agreeing with some other node at an early level, but the actual assignment of those features for which agreement is marked must take place at a later level. The rule of case agreement in Greek and the rule of gender/number agreement in French are different rules, but both exhibit the same global phenomenon. We need to view case agreement and gender number agreement as separate rules, for if they were the same rule, we would predict that a language either has that rule or not. Yet there are some languages, such as French, which have gender/number agreement but do not have case agreement.

Notice, for example, (9) and (10).

(9) Je crois qu'il est fameux.
    'I believe that he is famous.'

(10) Je le crois etre fameux.
    'I believe him to be famous.'

The adjective fameux agrees in gender and number with its antecedent,
but not in case. There simply are no varying case forms for adjectives in French.

Also, in languages that have syntactic gender, like French, the case form of a pronoun is governed by its grammatical relation to the verb, but the gender/number form of the pronoun is governed by its antecedent. Notice (11) and (12).

(11) J'ai trouvé le crayon, et il était rouge. 'I found the pencil, and it was red.'

(12) Marie a perdu le crayon, et je l'ai trouvé. 'Marie lost the pencil, and I found it.'

In (11) il is masculine singular, to agree with its underlined antecedent. But although the antecedent is a direct object, the pronoun is in the nominative form. In (12) le (contracted) is masculine singular to agree with its antecedent. Le is accusative because it is itself a direct object rather than because the antecedent is a direct object. Assignment of case and agreement of gender and number must be separate.

We can therefore view the French and Greek cases as instances of a general global condition which refers to a class of agreement processes. This condition would specify that if two nodes are part of some agreement process which must be stated at an early level, the features that the controlling element bears in surface structure will be assigned to the element that agrees with it. Thus, we do not have here an instance of a rule-specific global constraint.

There is some debate as to whether the relevant constraint is actually global. Fauconnier (1971) presents an analysis where no global constraint is required. There has been objection to Fauconnier's
proposal on the grounds that it requires use of indexing, which extends the power of the theory. However, because the matter is still in debate, I summarize Fauconnier's proposal. Whether or not Fauconnier's analysis can be maintained, we certainly do not have a rulespecific global constraint.

Fauconnier (1971) argues that there are unexpanded indexed nodes in deep structure and that adjectives may become part of a network of coreference by virtue of an agreement rule which copies the index of a noun phrase onto an adjective. He provides argumentation that the use of unexpanded NP nodes can handle the problems involving definite descriptions discussed in McCawley (1970), Karttunen (1969, 1971), and Kuroda (1971).

Fauconnier (1971, 1973) argues that facts concerning Agreement, Quantifier Floating, and Pronominalization in French can be accounted for only by use of unexpanded indexed NP nodes, along with a process of index-copying. This approach will also account for the facts motivating the two global constraints presented.

The most important parts of Fauconnier's proposal are as follows:

(a) There is an adjective agreement rule of roughly the form:

S.D.: $\text{NP}_x - \text{V}' - A$

S.C.: $\text{NP}_x - \text{V}' - A_x$

Here $V'$ is a copulative predicate, $A$ is an adjective or participle. The rule copies the index of the NP onto the adjective. This is a formal way of specifying that the adjective is under the control of that NP.
(b) There is a rule Feature-copying of the form:

Copy the features of a noun phrase \( NP \) onto all nodes \( W \) that are coreferential with \( NP_X \) and not already marked for those features.

(c) There is a Closeness Constraint on feature-copying.

First, we need the definition: 'Node B is closer to node C than node A is, if the lowest S dominating A, B, and C dominates the lowest S dominating B and C' (Fauconnier 1971, 144). Then the constraint is defined as:

**Closeness Constraint:** Given two coreferential NP's, \( NP^1_X \) and \( NP^2_X \), both marker for the feature \( F \), and an unmarked node \( V_x \) with the same index, if \( NP^2_X \) is closer to \( V_x \) than \( NP^1_X \) is, then feature copying of a feature specification for \( F \) cannot operate between \( NP^1_X \) and \( V_x \).

Fauconnier motivates the Closeness Constraint on the basis of anaphoric definite descriptions and epithets in English and French (146-9), the 'accusativus cum infinitivo' construction in Latin (149-54), and relative clause reduction in Latin (154-60).

Fauconnier summarizes the Lakoff-Andrews data concerning Greek and proposes that agreement does not actually copy any features, it 'only establishes control relations between noun phrases, adjectives and participles' (161). Case-marking is a late rule and therefore all nodes bearing the same referential index will share the surface features of the antecedent node.

Fauconnier analyzes the examples concerning Greek case agreement in the following way. The common deep Structure of (5) and (6) is (13) (= Fauconnier's (46), p. 162).
The A node in S₃ receives the index of NPₓ², and NPₓ² is then raised in S₂ and passivized. At this point leimen is indexed to agree with NPₓ². Again NPₓ² is raised, and finally comes to rest in S₁. The result in (14) (=147), p. 163.
The rule which distinguishes (5) from (6) replaces $\text{tauta}$ by $\text{NP}_x^2$.

Suppose this rule does not apply. Then $\text{NP}_x^1$ is marked dative since $\text{emmenomen}$ governs this case. $\text{NP}_x^2$, being in object position, is marked accusative. The case of $\text{NP}_x^1$ shows up on its head ($\text{tauta}$) and case of $\text{NP}_x^2$ on the relative pronoun. Feature-copying can now apply, and since $\text{NP}_x^2$ is closer than $\text{NP}_x^1$ to $P_x$ and $A_x$ these latter two will receive the case feature of $\text{NP}_x^2$ (namely $[^{\text{accusative}}]$) by the Closeness Constraint. The result is (5).

But suppose the replacement rule substitutes $\text{NP}_x^2$ for $\text{tauta}$. The result is (15) (=48), p. 164).
In this tree, $NP^1_x$ is marked [dative] and this feature appears on its head $NP^2_x$. In his event, Feature-copying marks $P_x$ and $A_x$ [dative], and the result is (6).

It seems to me that Fauconnier is not arbitrarily assigning indices to certain nodes in order to keep track of them, as Baker and Brame (1972) do. The point is that during a derivation a network of coreference is established between certain nodes, and Feature-marking takes place on the basis of this network of coreference. According to Fauconnier no rule-specific global constraint is needed.

Fauconnier's analysis can also be applied to French agreement. In French a participle agrees in gender and number with an object if and only if the object precedes the participle. But the structural description for agreement is always met before and not after the object moves, which led Casagrande to propose a global constraint. Fauconnier (1971, 1973) shows that adjective agreement must be an
indexing transformation. The situation with past participle is parallel, so he formulates Object-Participle Agreement as an indexing rule:

\[(16) S.D.: \text{Aux} - \boxed{\text{P}} - NP_x \]

\[S.C.: \text{P becomes } P_x \]

Feature-copying (FC) is responsible not only for the features of adjectives and participles, but for the lexical forms of pronouns. Certain constraints on anaphora, as well as other facts, can be explained if FC obeys the Ross-Langacker constraint.²

(17a) is derived from (17b) as follows.

\[(17) a. \text{Je les } \text{ai prises.} \]

\[
\begin{array}{ll}
\text{fem.} & \text{fem.} \\
\text{plur.} & \text{plur.}
\end{array}
\]

b. 

\[
\begin{array}{l}
S_1 \\
\text{Je} \\
V \\
\text{ai} \text{P} \\
\text{pris} \\
\text{NP}_i \\
\text{fem.} \\
\text{plur.}
\end{array}
\]

Indexing changes P to P_i. NP_i is moved to precede P_i. Then FC copies the features of NP_i onto the participle P_i. Notice that (18) has an underlying structure similar to (17b).

\[(18) \text{J'ai pris les sacoches. 'I took the bags.'} \]

\[
\begin{array}{ll}
\text{no features} & \text{fem.}
\end{array}
\]

But no rule moves NP_i (it is not a pronoun), and FC cannot apply, since P_i precedes and commands NP_i.
Fauconnier offers several other considerations (119-122) which indicate that his proposal is superior to that offered by Casagrande. I shall not repeat the details, for what concerns me here is that no rule-specific global constraint is needed.

By using rule-specific global rules we claim that each constraint is an idiosyncratic fact about some particular language. Fauconnier claims that a number of facts about separate languages can be explained by the general processes that languages may draw on in constructing their grammars. This solution, in view of its widespread motivation in different languages, is superior to the Baker and Brame (1972) indexing proposal. There they assign arbitrary indices only in order to mark nodes as being within the same simple S. But Fauconnier uses referential indices which are needed anyway in accounting for definite descriptions. Such naturally limited use of indices does not intolerably extend the power of grammars, and the advantage is that the two examples discussed here automatically follow from a more general analysis.

The above discussion opens up a number of problems due to the debatable status of indexing in grammar. A more insightful analysis may well show that indexing is not necessary, but in any case the examples of agreement do not require use of rule-specific global constraints.
3.4 'Obligatory' Extrapolation. Verbs like seem, appear, happen, strike, etc. (which I will refer to as seem-class verbs) cannot appear in certain grammatical structures which have not undergone Extrapolation.

(19) a. It seems to me that Louise is a good cook.
    b. *That Louise is a good cook seems to me.

(20) a. It appears to me that Harry will win.
    b. *That Harry will win appears to me.

Extrapolation is ordinarily optional:

(21) That Jean-Pierre is a revolutionary is well-known.

We cannot account for (19) and (21) by making Extrapolation obligatory for seem-class verbs. For the sentences in (22) are grammatical, where Raising has applied (to the structures underlying (23)) but not Extrapolation.

(22) a. John seems to please you.
    b. Watson happened to be in London.

(23) a. [John please you] seems
    b. [Watson be in London] happened

Furthermore, we cannot say that Extrapolation must apply if applicable, for there are cases like (24b) and (25b) where Extrapolation has not applied.

(24) a. It seems strange that Betty can't tapdance.
    b. That Betty can't tapdance seems strange.

(25) a. It appears (to be) true that Winchell cheats at tic-tac-toe.
    b. That Winchell cheats at tic-tac-toe appears to be true.
Postal (1972a) discusses the above problems, and tries to find a solution using rule-features. Assuming that the rule-feature assigned to the verb by Extraposition is [Extra] and the one assigned by the rule marking a complement as a that-clause (as opposed to an infinitive or gerund) is [That], the constraint is:

\[(26) \text{ Throw out all derivations in which the verbs seem, appear, happen, etc. occur with the feature markings } [-\text{Extra}, +\text{That}].\]

This rules out (19b) and (20b) but allows (22), (24) and (25). Of course this proposal won't work, as Postal himself points out, because Extraposition is obligatory even when a seem-class verb is embedded and has its subject NP raised, as in (27).

\[(27) \text{ a. } \text{*That Harry threw the game is likely to seem.} \]
\[(27) \text{ b. It is likely to seem that Harry threw the game.} \]

Since Extraposition applies to the clause containing the verb likely, rather than the clause containing the verb seem, the verb seem is marked [-Extra] and (26) wrongly predicts that both (27a) and (27b) are ungrammatical. Postal therefore has no solution.

The problem is that Postal is trying to find conditions that make Extraposition obligatory. This does violence to the distinction between optional and obligatory. An obligatory rule is one in which no derivation is acceptable where that rule has not applied (if its structural description was met at some point in the derivation). An optional rule is one in which any derivation is acceptable in which that rule has not applied. Now if some derivation is blocked
because because an optional rule has not applied, then we violate the definition above of optional and meet the definition above of obligatory. We would have to create more complicated definitions of the notions 'optional' and 'obligatory' which would rob them of their generality and simplicity. Furthermore, suppose we adopted more complicated redefinitions which would allow a rule to be either optional or obligatory under certain circumstances. It then becomes impossible to state that any given rule either is or is not obligatory. If some rule appears to be optional, that may be because no one has found conditions where it is obligatory. And if some rule appears to be obligatory, it may really be optional, with the conditions engendering obligatoryness so numerous that no one has found examples of optional application. The distinction optional/obligatory becomes useless. We might just as easily speak of the conditions under which a rule can apply. The way Postal uses the term 'obligatory' it becomes devoid of meaning.

In addition, Postal's use of ad hoc feature markings like [Extra] provides no way to explain why the features are used that are used. Any arbitrarily selected set of features should be available. Postal's solution carries with it the claim that the constraint might in other dialects of English mention any two randomly selected rules, a highly unnatural claim.

Postal notes that both Raising and Extraposition give acceptable sentences, but fails to draw the proper conclusion from this fact. The real question is: What is it that Raising has done to make the surface forms acceptable that Extraposition also does?
Once the question is phrased this way, the solution is obvious. They place an NP or VP after the matrix verb. **Seem-class predicates** always have a surface form in which either an NP, VP, or adjective follows the verb. **Seem-class predicates** do not appear as the rightmost element of a sentence in surface structure. One does find sentences like *John is* except under the existential reading of the predicate *be*.

I propose that there is a surface structure constraint which throws out structures not of the form:

\[(28) \quad \text{NP} - V' - X \left\{ \begin{array}{c} \text{VP} \\ \text{NP} \end{array} \right\} \]

where \(V'\) is a seem-class predicate. This constraint blocks (19b), (20b), and (27a). But it allows (19a), (20a), (22), (24), (25), and (27b).

There are sentences like *John seems rich* and *John appears happy* where an adjective follows the verb. But these are derived from *John seems to be rich* and *John appears to be happy*, which indicates that the constraint holds at 'shallow structure' before the minor rule which deletes *to be*. If this is so, we can eliminate the curly brackets in (28) by specifying that there must be some constituent following the matrix verb which itself contains a verb. Or perhaps it is best to say that a major constituent must follow the verb, where a major constituent is one which results from the first expansion of the S-node by the phrase-structure rules. In any event, no global constraint is needed.
3.5 *Say.* Lakoff (1970) claims that if the verb *say* takes a *for-to* complementizer, it must undergo Raising, Passive, and Agent Deletion.

(29) a. *Sam said for John to be tall.*
   b. *Sam said John to be tall.*
   c. *John is said to be tall by Sam.*
   d. John is said to be tall.

Lakoff claims that a global constraint is needed here, but it would have many complications in its statement. A global constraint throws out certain derivations as ill-formed, so in order to mark (29a,b,c) as ungrammatical, the constraint would have to be stated as: Throw out any derivation in which (a) the verb *say* appears in the matrix sentence, (b) the complement sentence takes a *for-to* complementizer, and either (c) Raising applies, or (d) both Raising and Passive apply, but not Agent Deletion. Such a formulation fails, of course, to state what conditions (c) and (d) have in common.

A similar approach is taken by Stockwell, Schacter and Partee (1973) (cf. pp. 530-1, 560-1), who present a derivation of (29d) which is the same as Lakoff's. They claim that (29b) can be blocked by making Passive obligatory with Raising. There are serious theoretical problems with such a position, namely the problems of making an optional rule obligatory (discussed in section 3.4) and the use of a rule-specific global constraint. Furthermore, I think the Lakoff/Stockwell approach is based on a false analogy with sentences like (30) which undergo Raising and Passive in their
derivation which follows.

(30) a. One believes \((\text{Lloyd is the fastest gun in the West})\).

\[ \rightarrow \]

b. One believes Lloyd to be the fastest gun in the West.

\[ \rightarrow \]

c. Lloyd is believed by one to be the fastest gun in the West.

\[ \rightarrow \]

d. Lloyd is believed to be the fastest gun in the West.

What is ignored here is the grouping of verbs into lexical classes. Baker and Brame insightfully point out that \text{say} can be broken up into two lexical items, \text{say}_1 and \text{say}_2. \text{Say}_1 occurs in (31), and \text{say}_2 in (29d) and (32).

(31) Hermione said that Marcelle tickled her fancy.

(32) It is said that John is tall.

Baker and Brame then claim (incorrectly, I think) that Extraposition is obligatory for \text{say}_2:

(33) "That John is tall is said."

They propose that \text{be said} is generated as a Passive by the base rules, and adopt Emond's analysis of complementation, so that \text{be said} has an empty subject NP and Intraposition (the reverse of Extraposition) is blocked. But Postal (1972c) provides extensive argumentation against replacing Extraposition by Intraposition, and Lakoff (1972) discusses the arbitrary nature of empty nodes. Lakoff's remarks in this regard are much to the point, but both parties in the dispute focus on formal devices and not insight into grammatical phenomena.

The clue to what is going on appears when Baker and Brame say:

"There are are a number of respects in which \text{be said} and \text{rumored}
behave like predicates such as *seem* and *appear* (67). It is plausible, then, that *is said* is a predicate which takes a sentential subject, as in (34).

(34)

```
S
 -> NP
  S
  is said

John BE tall
```

If Raising applies, (29) results, if Extraposition applies, (32) results. If neither applies, the surface structure constraint (28) discussed in section 3.4 throws out (35).

(35) *That John is tall is said.

(29a,b,c) are blocked because while *say₁* takes an object complement, that complement cannot have a stative predicate if it bears a for-to complementizer:

(36) a. *Sam said for John to know the answer³, be tall³.
    b. Sam said for John to open the door.
    c. Sam said that John knew the answer³, was tall³.
(36b) is all right because *open* is a non-stative predicate. The verb in (29a-c) could not be *say₂*, because that predicate does not take an object complement. We now have more evidence against Lakoff's proposal: since the deep structure of (29) is ill-formed due to the the constraint on stative predicates demonstrated in (36), how could the application of any transformations (much less a list of specific transformations) make the surface sentence acceptable?.
I conclude that there is no global constraint here, rather that what we have is simple a case of the more general restriction for seem-class predicates developed in 3.4.

3.6 Double-Ing. Ross (1972) argues that there is a derivational constraint which rules out certain sequences of present participles. Notice that (37d) is unacceptable.

(37) a. It continued to rain.
   b. It continued raining.
   c. It is continuing to rain.
   d. *It is continuing raining.

Ross first shows that the Double-Ing constraint must be an output condition: there are some intermediate stages where such sequences must be allowed. But the constraint also has to refer to earlier levels of structure. To begin with, the second ing form has to be a verb in the complement of the first verb. There is no violation in (38) because the second ing form is a noun, and no violation in (39) because the second ing is not in the complement of the first.

(38) The police are stopping drinking on campus.

(39) I saw the man who had been drinking opening up the cash register.

Furthermore, notice that (41), derived from (40) by Raising, is acceptable.

(40) His is expecting that breathing deeply will benefit us is naive.

(41) His expecting breathing deeply to benefit us is naive.
(41) is acceptable because the constraint blocks only contiguous verbs that were in immediately adjacent clauses in remote structure. The deep structure of the subject of (41) is as in (42), where expecting (=\text{\textsc{\text{v}}}_1) and breathing (=\text{\textsc{\text{v}}}_3) are not in adjacent clauses.

(42)

\begin{center}
\begin{tikzpicture}
  \node (s1) {S_1};
  \node (v1) [below left of=s1] {V_1};
  \node (np1) [left of=v1] {NP};
  \node (he) [below of=v1] {he expects};
  \node (s2) [below of=he] {S_2};
  \node (v2) [below of=s2] {V_2};
  \node (np2) [left of=v2] {NP};
  \node (will) [below of=v2] {will};
  \node (s3) [below of=will] {S_3};
  \node (v3) [below of=s3] {V_3};
  \node (deeply) [below of=v3] {deeply};
  \node (we) [below of=deeply] {we breathe};
  \node (benefit) [below of=deeply] {benefit};
  \node (us) [below of=benefit] {us};

  \draw (np1) -- (v1);
  \draw (v1) -- (he);
  \draw (he) -- (s2);
  \draw (np2) -- (v2);
  \draw (v2) -- (will);
  \draw (will) -- (s3);
  \draw (v3) -- (deeply);
  \draw (deeply) -- (we);
  \draw (benefit) -- (us);
\end{tikzpicture}
\end{center}

Ross formalizes the Double-Ing constraint as follows:

(43) All surface structures containing a subtree of the form,

\begin{center}
\begin{tikzpicture}
  \node (s) {S};
  \node (x) [left of=s] {X};
  \node (v) [below of=x] {V};
  \node (v_a) [below of=v] {V_a};
  \node (v_b) [below of=v_a] {V_b};
  \node (v_\text{ing}) [below of=v_b] {V_\text{ing}};
  \node (y) [below of=v_\text{ing}] {Y};
  \node (np) [left of=s] {NP};
  \node (z) [right of=s] {Z};

  \draw (x) -- (v);
  \draw (v) -- (v_a);
  \draw (v_a) -- (v_b);
  \draw (v_b) -- (v_\text{ing});
  \draw (v_\text{ing}) -- (y);
  \draw (y) -- (np);
  \draw (np) -- (s);
  \draw (s) -- (z);
\end{tikzpicture}
\end{center}

in which the node corresponding to \text{V}_a in remote structure was immediately dominated by \text{S}_1, and the node corresponding to \text{V}_b in remote structure was immediately dominated by \text{S}_j, and in which no \text{S} node intervened in remote structure between \text{S}_1 and \text{S}_j, are ungrammatical.
This formulation of the Double-Ing Constraint has been amply criticized by Pullum (ms.) who offers a solution which does not require a global constraint. Pullum first summarizes the proposals by Emonds (1973) and Milsark (1972) that the constraint may be stated without reference to derivational history. In essence, they both propose that surface sequences of V-ing's are prohibited unless an NP boundary intervenes. They give evidence that there is no NP node in surface structure over the complement of a verb of temporal aspect, so that (44) is blocked.

(44) #John is continuing singing soprano arias.
But there is an NP node over the ing forms that have traditionally been called gerunds, so (45) is not blocked.

(45) John was considering getting into college.
Emond's constraint will allow (38), (39), and (41) because an NP dominates the second ing form that does not also dominate the first. Pullum then presents counterexamples to Emond's constraint (sentences from Pullum's ms., p. 7).

(46) a. Three policemen dragged the screaming, struggling girl away.

b. I was sitting thinking about my troubles when there was a knock at the door.
In these sentences there is no NP node over one of the ing forms that is not also over the other, so (46a,b) should be blocked.
The problem is that Emonds tries to state the constraint in terms of surface sequences of categories, rather than in terms of a phrase marker.
Pullum presents strong evidence against a global statement of the Double-Ing constraint. The remote structure of (47) must include something like the substructure (48).

(47) He was charged with numerous offences, including stealing apples and assaulting the Queen.

(48) 

The surface structure of (47) must include a substructure of the form given in (43), with including as $V_a$ and stealing as $V_b$.

Ross' constraint wrongly predicts that (47) is blocked.

Emonds (45) points out that it may be a general property of surface structure constraints that they do not prohibit a sequence of items when an NP boundary occurs between them. Pullum suggests that if this is the case, the Double-Ing Constraint may be stated as:

(49) Any sequence $v_i [V + \text{ing}] v_i - v_j [V + \text{ing}] v_j$ in surface structure is ungrammatical if $v_j$ in in the complement of $v_i$. 
We may conclude that this constraint is not a global constraint.

3.7 One-Pronominalization. Lakoff (1970) discusses a constraint which blocks structures roughly of the form one of NP under certain conditions. But the constraint does not hold if ones is spelled those in surface structure (by application of an optional morphophonemic rule which converts the ones into those). Thus, we have the contrast between:

(50) a. Max had known the kings of England and I had known the ones of Spain.
   b. Max had known the kings of England and I had known those of Spain.

Lakoff claims that the constraint must refer to an intermediate level of structure, as well as to surface structure, because the constraint must precede the rule of One(s)-deletion, which derives (51a) from (51b).

(51) a. *I knew six girls from England and Irv knew five ones from Spain.
   b. I knew six girls from England and Irv knew five from Spain.

Lakoff maintains that the constraint must apply before one(s)-deletion in the derivation of (52b), for only at the point where (52b) is identical to (52a) is the structure defining the constraint present.

(52) a. *I knew six kings of England and Irv knew five ones of Spain.
   b. *I knew six kings of England and Irv knew five of Spain.

Lakoff also claims that the constraint must precede the rule of Pseudo-Adjective Formation (which converts, for example, king
of Spain to Spanish king) in order to rule out (53).

(53) *I met the English king and Sam met the Spanish one.

The evidence for referring to surface structure comes from (50). Now if we did not need to refer to intermediate stages, the constraint would be a surface structure constraint. But do we really need to refer to earlier structures? The evidence that the constraint must hold before certain rules comes from examples (51), (52), and (53). But I think all these can be blocked by other restrictions than Lakoff's. (51a) and (52a) can be blocked by an independently-needed restriction which prohibits structures in which ones is preceded by a quantifier of the class including numerals and items such as these, those, some, many. Notice the examples in (54).

(54) a. *I kissed five girls and George kissed six ones.

b. *I kissed many girls and George kissed \{some \} \{many \} \{a whole \} \{roomful of \} \{ones.\}

(52b) and (53) can be blocked in the following way. Baker and Brame give some indication (55) that there is a structural difference between NP like king of England and NP like a picture of Mary. It seems that king of England is dominated by the node NP. Notice that there are sentences like (55), where one of NP occurs, indicating that an N under the domination of NP has been reduced to one.

(55) John took a picture of Mary, and I took one of Alice.

But phrases like king of England and Man of La Mancha are single
nouns in some sense, as indicated by the difference between (55) and (56).

(56) I met the Man of La Mancha and George met
    \[
    \begin{cases}
        \text{the man of the year.} \\
        \text{"the one of the year."}
    \end{cases}
    \]

Now suppose that the NP's in (52b) and (53) have the structures shown in (52b') and (53').

(52b')

```
S
  .
  .
   NP1
     Adj
     six
     N1
     kings of England
   NP2
     Adj
     five
     N2
     kings of Spain
```

(53')

```
S
  .
  .
   NP1
     Det
     the
     N1
     the English king
   NP2
     Det
     the
     N2
     Spanish king
```

The rule of One-Pronominalization can apparently apply to both NP and N nodes:

(57) Alice bought a blue dress, and Harriet bought
    \[
    \begin{cases}
        \text{one, too.} \\
        \text{a green one}
    \end{cases}
    \]

But the rule cannot apply to (53'), because NP_2 and NP_1 do not meet any identity condition, nor do N_2 and N_1, so (53) cannot be generated. The situation is the same with NP_2 and N_2 in (52b').
Since One-Pronominalization can't apply, neither can One-Deletion, so (52b) can't be generated. The structural distinction between NP's like the English king and the cheerful lady is supported by the following considerations. Notice that One-Pronominalization can apply to (58).

(58) a. I met the generous king and Sam met the mean king.
   b. I met the generous king and Sam met the mean one.

Now suppose the NP's in (59) are as in (60).

(59) I met the English king and Sam met the mean king.

(60) S
     NP₁         NP₂
     \     / \     /
    Det N₁  Det Adj N₂
     \     \     /
         the English king the mean king

Now since N₂ and N₁ do not meet an identity condition, we should expect that N₂ cannot pronominalize to one, and this is just what we find:

(61) *I met the English king and Sam met the mean one.

This sentence can not be blocked by Lakoff's constraint, for while it has a superficial structural similarity to (53), the NP the mean one cannot be derived from *one of mean.

In fact, (50a) can be blocked by the same means used to block (52b) and (53). Perhaps no one of NP constraint is necessary, only a restriction against combining ones with an immediately preceding quantifier (illustrated by sentences (51a), (52a), and
(54)). Baker and Brame (1972, 54-5) provide evidence against the rule changing the ones to those, further vitiating Lakoff's proposal. This example therefore cannot be maintained as a global constraint.

3.8 *Numerous such ones. Postal (1972a) proposes the filter:

(62) Throw out all derivations in which the substructure

\[ \text{NP}[\text{numerous such ones}]_{\text{NP}} \]

occurs.

Postal derives one forms from such forms. Thus the \textbf{b} sentences in (63) and (64) are derived from the respective \textbf{a} sentences.

(63) a. John was looking for a yellow robin, but he couldn't find such a one.

\[ \rightarrow \]

b. John was looking for a yellow robin but he couldn't find one.

(64) a. Smith was searching for non-returnable boomerangs, but he couldn't find any such ones.

\[ \rightarrow \]

b. Smith was searching for non-returnable boomerangs, but he couldn't find any such.

The filter blocks sentences like:

(65) a. *Harry needed atomic ray guns and George sold him numerous.

b. *Harry needed atomic ray guns and George sold him numerous such \( \leftarrow \) (*ones).

The sentences in (65) are acceptable if we replace numerous with many. (62) refers only to a single tree, but this cannot be a surface structure treesince no such appears in (65b). The filter is unusual in that it is a single tree filter that is not stated at either deep or surface structure.

Postal notes, correctly, that the constraint cannot be a
surface structure constraint, since in some forms ones does not appear, and in others such does not appear. But from this one cannot logically conclude, as Postal does, that the constraint must be stated at some level or levels of intermediate structure. Postal dismisses the possibility of a deep structure constraint with the comment

'No non-ad hoc way of preventing generation of the underlying structures appears to exist, especially in view of the normal distribution of the closely related form many'(149). But an ad hoc deep structure constraint is surely preferable to an ad hoc constraint of a new type. (Remember, this constraint would be odd as a global rule, for it does not relate corresponding structures at nonadjacent points in a derivation.) And in all of the blocked sentences, the banned structure numerous such ones appears in something very much like deep structure. Furthermore, it has by no means been demonstrated that the deep structure blocking does not have to do with the semantics of numerous as opposed to many. This filter may be removed from the class of global constraints.3

4.9 Each Shift. Postal (1972c) presents a derivational constraint on the rule of Each Shift, which 'has the effect of moving the quantifier each out of the NP corresponding to the variable which each binds and attaching it to (or after) the end of a (usually) numerically quantified NP whose numerical quantifier is under the scope of each'(189). Examples of the rule are:

(66) a. Each of the boys kissed her three times.

b. The boys kissed her three times each.
(67) a. I gave each of them five dollars.
   b. I gave them five dollars each.

One NP (the each-Source) is moved to the end of another (the each-Target (at the point of application).

(68) a. He sent three men to each of the stores.
   b. *He sent three men each to the stores.

Furthermore, there is a clause-mate condition on Each Shift. In the sentences in (69), there is a clause boundary after about, and each may not hop over this boundary.

(69) a. I talked to each of the senators about (my) blocking three bills.
   b. *I talked to the senators about (my) blocking three bills each.
   c. I talked to each of the advisers about displaying three pictures of myself.
   d. *I talked to the advisers about displaying three pictures of myself each.

There are two significant restrictions here: (a) the requirement that each-Source precede each-Target, (b) the Clause Mate condition. Now notice that unbounded leftward movement rules (like Topicalization, Adverb Preposing, Wh Rel Movement) can move NP's so that Each Shift can apply, even when it could not apply in the structures which were input to those rules.

(70) a. Harry bought three diamonds for each of those girls.
   b. *Harry bought three diamonds each for those girls.
   c. For those girls, Harry bought three diamonds each.
   d. The girls, for whom Harry bought three diamonds each, are happy.
The relevant movement rules are unbounded. Notice, for example,

Adverb Preposing:

(71) a. For those girls, I am sure Harry bought three diamonds each.

b. For those girls, it was later learned that Mary claimed that he bought three diamonds each.

Bringing the facts about Each Shift together, we notice that the condition that each-Source precede each-Target is defined on the output of unbounded movement rules (as shown in (70) and (71)), but the clause-mate condition could only be met before the application of these rules. Postal's account is that Each Shift is postcyclic, subject to a left-right condition. The Clause Mate requirement is a global condition referring to earlier stages of a derivation. That is, the correspondents of each-Source and each-Target must be Clause Mates at the end of the lowest cycle covering both of them.

But I think the logic of this argument is faulty, for Postal is trying to build two restrictions into one rule. Such a move would perhaps be necessary if (a) the condition that each-Source precede each-Target is applicable only to the rule of Each Shift, (b) this condition is defined at the point of application of the rule Each Shift. I think that both of these assumptions should be called into question, thereby vitiating an analysis based upon them.

There is some interesting data given in Fauconnier (1971, 7-10, 171-95) which bears on the first assumption. There is a rule of Quantifier Floating (QF) in French which derives (73) from (72).

(72) Chacun des hommes a vu l'auto.

(73) Les hommes ont chacun vu l'auto.
Both (72) and (73) mean 'each of the men saw the car'.

QF says in brief: In a clause containing (Prep) $\{\text{chacun} \}$ $\{\text{NP}, \text{tous} \}$ move (Prep) $\{\text{chacun} \}$ into post-auxiliary or postverbal position.

Fauconnier notes (p. 10) that the remaining NP must precede the detached chacun:

(74) a. *J'ai mangé chacun des gâteaux.
'I ate each of the cakes'.

$\rightarrow$ b. *J'ai chacun mangé les gâteaux.

The same condition holds on the parallel rule of Quantifier Floating in English:

(75) a. I have seen each of those movies.

$\rightarrow$ b. *I have each seen those movies.

The conditions here seem to be the same as the condition on Each Shift: the moved quantifier must follow the NP it moved off of. Since the same restriction applies to separate rules in separate languages, it is unlikely that it should be built into the rule of Each Shift in English.

Furthermore, Fauconnier gives interesting data which indicate that the constraint could be viewed as a surface structure constraint.

In French chacun or tous can be moved from an NP in object position only if the NP from which it moves is a pronoun which ends up as a clitic or relative pronoun. Thru compare (74b) with (76).

(76) a. Je les ai chacun mangés.
'I ate each of them'.

b. Les gâteaux que j'ai chacun mangés étaient bons.
'the cakes which I ate were good'.

But the application of the separate rules of Clitic Movement and
Wh Rel Movement will in (76) insure that the NP precedes its quantifier in surface structure. If the pronoun cannot be moved to the left of the quantifier, the surface structure is banned. Notice (77), where the pronoun eux cannot be cliticized:

(77) a. J'ai mangé chacun d'eux.
   'I ate each of them'

b. *J'ai chacun mangé eux.

c. *Je eux ai chacun mangé.

The fact that separate rules create the acceptable sentences indicates that the constraint is stated at surface structure, for otherwise we could not explain why these separate rules both engender surface acceptability. We could, then, formulate a surface structure constraint which insures that a detached quantifier will follow the NP it is detached from. Roughly, the constraint would state:

Block structures of the form Q - X - NP where Q is not immediately dominated by the node NP. The data motivating the left-right condition on Each Shift would automatically follow from the more general condition. Now since this condition is separate from the clause-mate condition on Each Shift, no global constraint is necessary.

We are still missing something, it seems to me. This surface structure constraint is puzzling in view of the fact that non-detached quantifier precedes its NP. Perhaps the structures which we explained by use of a surface structure constraint actually result merely from the way the rules operate. We then would not need a surface structure constraint. In this case, perhaps the reason all detached quantifiers follow their NP's is that the quantifier
detachment rules are rightward movement rules. \(^7\) (70b) is ruled out because the rule can't move each to the left. (70c,d) and (71a) are acceptable because the rules of Adverb Preposing and Wh Rel Movement carry the whole NP each of those girls to the left. After that, Each Shift moves each to the right. The derivation of (71a) is:

\[(78)\]
\[
\begin{align*}
&\text{(a. I am sure Harry bought three diamonds for each of those girls.)} \\
\rightarrow &\text{b. For each of those girls, I am sure Harry bought three diamonds. (Adverb Preposing)} \\
\rightarrow &\text{c. For those girls, I am sure Harry bought three diamonds each. (Each Shift)} 
\end{align*}
\]

According to this derivation, we violate the Clause Mate condition on Each Shift. But the Clause Mate condition may well be illusory. Postal cites (69) as evidence for a clause-mate condition. But Each Shift is independently blocked from moving each into the about-phrase after the verb talk:

\[(79)\]
\[
\begin{align*}
&\text{(a. I talked to each of the senators about five bills.)} \\
\rightarrow &\text{b. *I talked to the senators about five bills each.} \\
\end{align*}
\]

\[(80)\]
\[
\begin{align*}
&\text{(a. Each of the authors talked to the editor about five books.)} \\
\rightarrow &\text{b. *The authors talked to the editor about five books each.} \\
\end{align*}
\]

No clause-mate condition can be invoked here. J. Geis argues that adverbial prepositional phrases originate as 'higher predicates' which take sentential subjects. If about \(X\) is such an adverbial, then the deep structure of (79) is (81):
(69) has a similar deep structure. But now we can easily see why
Each Shift is blocked, for rightward movement rules are upward-
bounded. (They may not move an element into a clause higher that
the one they originated in. See Ross (1967), 146-84.) Therefore
each cannot move out of $S_2$ in (81).

The only other evidence which could motivate a clause-mate
condition would be a case where one $S$ is embedded within another,
and the each in some NP of the matrix sentence could move onto an
each-Target in the lower sentence. If the movement is blocked, we
could claim that there is a clause-mate condition. An example of this
situation is (82).

(82) a. Each of the farmers thinks that Zebe owns five acres.

b. *The farmers think that Zebe owns five acres each.

But notice that in the examples (66) and (67) motivating the rule
of Each Shift, the each-Source and each-Target command each other.
In (82a), however, each asymmetrically commands five and precedes it.
In (82b) each and five command each other, but five precedes each.
It is just such a situation which is blocked by Lakoff's Quantifier
Constraint, which specifies that if $Q_1$ asymmetrically commands $Q_2$
in deep structure, then if $Q_1$ and $Q_2$ command each other in surface
structure, \( \alpha_1 \) must precede \( \alpha_2 \)

Since we cannot motivate a clause-mate condition, we cannot assume there is one, and derivations of the type presented in (78) cannot be blocked. In such an event, Postal' data can be accounted for by the simple restriction that Each Shift is a rightward movement rule, along with other independently-needed restrictions in the grammar. By focusing on the nature of the rules involved, we can explain what is going here, and we do not need to use a global constraint or a surface structure constraint.

4.7 Indirect Object Movement. Postal (1972a) points out that, as Fillmore (1965) first noticed, for-prepositional phrases behave differently with respect to Passive than do to-prepositional phrases.

\[(83)\]
\begin{align*}
\text{a. Marsha gave a rose to Emily.} \\
\text{b. Marsha gave Emily a rose.} \\
\text{c. A rose was given to Marsha by Emily.} \\
\text{d. Emily was given a rose by Marsha.}
\end{align*}

\[(84)\]
\begin{align*}
\text{a. Emily bought a rose for Marsha.} \\
\text{b. Emily bought Marsha a rose.} \\
\text{c. A rose was bought for Marsha by Emily.} \\
\text{d. *Marsha was bought a rose by Emily.}
\end{align*}

Indirect Object Movement derives the b sentences form the a sentences in (83) and (84). Postal (1972a) claims that application of the rules Passive and IO Movement in a clause with a main verb \( V_i \)
leads to assignment of the features [+Passive] and [+IO Movement] to 

\( V_i \). leads to assignment of the features [+Passive] and [+IO Movement] to 

(85) Throw out all derivations in which a single verb \( V_a \) both:

(i) occurs in an underlying structure with a for-indirect

object; and

(ii) occurs in a derived structure marked 

This is merely a restatement of the fact that no verb which has a

for-indirect object can undergo both Passive and IO Movement.

Constraint (85) mentions two specific rules, as well as the presence

of a for-indirect object as opposed to a to-indirect object. There

is no difference in the structural configuration of these indirect

objects: the difference is indicated by the preposition to as op­

posed to for.

The trouble with Postal's constraint is that the distinction

between to-indirect objects and for-indirect objects is not suffi­

cient. Notice, for example (86a-d) where a to-indirect object may

not be moved.

(86) a. John sang a new song to Mary.

b. John sang Mary a new song.

c. A new song was sung to Mary by John.

d. *Mary was sung a new song by John

The generalization seems to be that both Passive and IO Movement

occur with three-place predicates but not with two place predicates.

This is true of necessity because one argument is the subject of the

predicate, and IO Movement can take place only when there are two
arguments in the VP to be interchanged. A passivized indirect object does not appear with verbs that are two-place predicates, like buy and sing. To see the distinction between two and three place predicates, notice (87)-(89).

(87) *John gave a book.
(88) John bought an apple.
(89) John sang 'Keep on Truckin'.

A sentence like (87) can occur only where the third argument has been deleted by some rule. Now if buy is a two-place predicate, why is it that (84a) appears to have an indirect object, and why does it, look as if IO Movement has applied in (84c)? One possibility would be to claim that buy is a three-place predicate. Then the third argument in (88) is deleted by a rule of Indefinite Dative Deletion, similar to the way in which the indefinite direct object of the two-place predicate read is deleted by Indefinite Object Deletion. 8

(90) a. John was reading something.
   b. John was reading.

The trouble is that while (90b) has the meaning of (90a), (88) does not have the meaning of the supposed source (91):

(91) John bought an apple for someone.

Another possibility is that verbs like buy and sing may be optionally either two- or three-place predicates. But this does violence to the notion of saying that some verb is an n-place predicate. If there is a verb which apparently differs in the number of arguments it can take, I claim that we actually have two homophonous lexical items. Notice the verb rent is apparently either a two- or
three-place predicate:

(92) Albert rented a cabin.

(93) Albert rented a cabin to the Quigleys.

There is a difference in the meaning of the verb rent in (92) and the verb rent in (93). In (92) the subject of the sentence is paying money, but in (93) the subject is receiving money. (92) has another reading in which an indefinite dative to someone has been deleted, but this other reading is irrelevant. A promising way to account for this difference is to claim that it is based on the existence of two verbs rent: rent₁, which takes two arguments, and rent₂, which takes three arguments. But there is no detectable difference in meaning between the verbs in (94a) and (94b).

(94) a. Emily bought a rose.

b. Emily bought a rose for Marsha.

To state a verb is an n-place predicate, the n must be specified for some single value, and the existence of (94a) indicates that for buy that value is two. How then, do we account for (94b)? Postal implicitly assumes that the deep structure of (94b) is like (95).

(95)

\[
\text{NP} \rightarrow S \rightarrow VP \rightarrow \text{NP} \rightarrow \text{NP} \rightarrow \text{PP}
\]

Emily \( \text{bought} \) a rose for Marsha

But I think a more correct structure is like that in (96).
The second argument, $NP_2$, is a 'nominally-complex' NP. There is good evidence for this analysis. Transformations operate on single constituents. Now notice that the following examples involving movement rules show that $NP_2$ above is a single constituent.

(97) a. Q: What did Emily buy?
A: A rose for Marsha.

b. What Emily bought was a rose for Marsha.

c. It's a rose for Marsha that Emily bought.

d. The rose for Marsha which Emily bought was an American Beauty.

Compare (97) to parallel examples involving the three-argument predicate give:

(98) a. Q: *What did John give?

b. *What John gave was a cigarette to the cop.

c. *It was a cigarette to the cop that John gave.

d. *The car to his wife which John gave was a Cadillac.

I conclude that the deep structure of (94b) is as in (96). Now why is it that (99b,c) are acceptable, while (99d) is not? I offer the following speculation. Speakers of English misanalyze the parsing of
(99a) to be as in (95) instead of (96). Based on this misanalysis, the speaker incorrectly applies either Passive (to get (99b)), IO Movement (to get (99c)), or IO Movement and then Passive (to get (99d)). In other words, the speaker assumes that there are two arguments in the VP instead of one, and applies rules on that basis. But why is (99d) starred? I think this sentence is grammatical but unacceptable. Bever and Langendoen (1973) and Grosu (1972) argue that perceptual strategies play an important part in marking as unacceptable sentences which are grammatical (in the sense that they can be generated by the grammar). There must be a strategy which assigns the superficial subject of a passive sentence to object position. This strategy will assign the NP a rose in (99b) as the direct object of bought. But when this strategy is applied to (99d) the NP Marsha is immediately marked as the direct object in toto of the verb bought. But this creates a problem, for one is speaking of buying a flower, not a person. Furthermore, the leftover NP a rose could only be assigned as the indirect object of the verb (since the strategy already gives us the subject and direct object). But an inanimate NP cannot be the indirect object of the verb buy.

In summary, I suspect that (99d) is unacceptable because speakers have no strategies which will allow them to effectively recover the underlying structure of the sentence. I have tried to base the remarks just offered on a search for insight into grammatical phenomena rather than an attempt to find a formalism which 'handles the data'. Whatever the exact nature of the solution, it should be clear that a proper understanding will not involve an ad hoc
global constraint.

3.11 **Coordination Reduction.** In a (to my mind) dubious analysis Postal (1972a) claims that (100a) is derived from (100b) by Coordination Reduction.

(100) a. Mary and John Smith (both) have jobs.

b. Mary Smith and John Smith (both) have jobs.

These two sentences differ in that in the a sentence there is a presupposition that the individuals named Smith are related, but there is no such presupposition in the b sentence. Postal claims that Coordination Reduction applies to phrases which are the names of human individuals only when in the semantic representation there is a presupposition that the individuals have the same last name because they are related. Thus, '... the constraint is naturally $[! -RN]$ stated as an ad hoc filter which is not part of Coordination Reduction as such, a filter which throws out all derivations in which there is a semantic representation with the relevant names but without the relevant presupposition and a later tree in which the names have been smashed together by Coordination Reduction$^6(143)$.

I do not believe that a global constraint is necessary here. The problem is that there is little reason to think that (100a) is derived from (100b) by Coordination Reduction. Postal would have to map (101a) into (101b).

(101) a. $\text{NP} \{\text{Mary Smith}\}_{\text{NP}}$ and $\text{NP} \{\text{John Smith}\}_{\text{NP}}$

b. $\text{NP} \{\text{Mary and John}\} \{\text{Smith}\}_{\text{NP}}$
The same process would map (102a) into (102b).

(102) a. John's bicycle and Mary's bicycle are on the porch.
    b. John's and Mary's bicycles are on the porch.\(^{12}\)
    c. *John's and Mary's bicycle are on the porch.

But the underlined NP's in (102a) must become plural when the reduction process applies, as shown by comparing (102b) with (102c).

If this reduction process truly maps (101a) into (101b), the name Smith should be pluralized, giving *Mary and John Smiths. While a proper name can appear in the plural in generic contexts such as the Smiths, it certainly cannot be plural in the context of (100).

Thus, the mapping of (100b) into (100a) is blocked by the impossibility of pluralizing proper names here, and no global constraint in necessary.

If (101b) is not derived from (101a) by Coordination Reduction, where does it come from? Most likely it is an instance of phrasal conjunction, as discussed in Lakoff and Peters (1969). There must be cases where conjoined NP's are generated by the phrase structure rules rather than derived transformationally. John and Mary are alike cannot be derived from *John is alike and Mary is alike.

Suppose the deep structure of the subject NP of (100a) is as in (103), and the deep structure of the subject NP of (100b) is as in (104).

\[
(103) \quad NP \left[ \left[ \text{Mary and John} \right] \right] \quad Smith \quad NP
\]

\[
(104) \quad NP \left[ \left[ \text{Mary Smith} \right] \right] \quad \text{and} \quad NP \left[ \left[ \text{John Smith} \right] \right] \quad NP \quad NP
\]
We now have a natural basis to distinguish the semantics of (100a) and (100b). Since only one name Smith appears in (103), that name must refer to one family, and the individuals bearing the first names mentioned must belong to that family. In (104), two names Smith appear, so each Smith may refer to a different family. Postal has no such natural way to distinguish the meaning. On these grounds the analysis presented here is superior, in addition to not requiring a rule-specific global constraint.

3.12 Tough Movement. Berman (1973) postulates a constraint on Tough Movement. This rule derives (106) from the structure underlying (105).

(105) Albert is tough (for me) to get along with.
(106) To get along with Albert is tough for me.

Now consider the rules shown in (107)-(110) below.

Dative Movement:
(107) a. It is impossible to buy presents for John.
→ b. It is impossible to buy John presents.

About Movement:
(108) a. It is difficult to talk to Mary about such things.
→ b. It is difficult to talk about such things to Mary.

Passive: 13
(109) It is unpleasant to be kicked by John.

Raising:
(110) a. It is difficult to believe [John made such a mistake]
b. It is difficult to believe John to have made such a mistake.

Now notice that if an NP has been moved by any of the above rules it may not be moved by Tough Movement.

Dative Movement:

(111) a. It is impossible to buy John presents.

→ b. *Presents are impossible to buy John.

→ c. *John is impossible to buy presents.

About Movement:

(112) a. It is difficult to talk about such things to Mary.

→ b. *Mary is difficult to talk about such things to.

→ c. *Such things are difficult to talk about to Mary.

Passive:

(113) a. It is easy to be accepted by that group.

b. *That group is easy to be accepted by.

Raising:

(114) a. It is impossible to expect John to know the answer.

b. *John is impossible to expect to know the answer.

Berman proposes that the relevant constraint is that no NP may be moved by Tough Movement if it has earlier been moved by some rule.

There are several difficulties with Berman's analysis. To begin with, the constraint is in many instances more general than one which would only block Tough Movement from applying Notice that Wh Rel Movement cannot apply to a sentence where Dative Movement has applied:

(115) *Sally is a girl I would give my last dime.
Another problem is that it is hard to see what explanation there could be for a constraint on re-movement of an NP. The condition could not be general, for there are many cases where an NP is successively moved by different rules: for example, an NP can be moved by Dative Movement and then by Passive, a raised subject can be passivized and then raised again, etc. I think Berman's data can be explained by other means. First, I shall consider the case of Dative Movement.

Hankamer (1973) provides extensive discussion of the notion of structural recoverability, primarily in regard to deletion rules. I can only present the major relevant points of his work, and urge the reader to go to the original for details and some very interesting discussions. Hankamer first discusses (with regard to deletion rules) certain restrictions which function to insure recoverability of deletion.

A deletion is recoverable if, given only the statement of the rule effecting deletion and the output of a particular application of the rule, the input to the rule can be uniquely determined. In order to meet this condition, a deletion rule would have to be so formulated or so constrained that it could never map two distinct inputs into the same output. Any rule which so neutralized the distinction between two different underlying structures would introduce ambiguity, and a derivation which introduces ambiguity is not recoverable (Hankamer (1973), 39).

One way to prevent ambiguity from arising is by a Structural Recoverability Condition: 'if a deletion rule operating over a variable would introduce structural ambiguity by yielding the same output upon application to two different sources, both applications of the rule are blocked' (41).
Hankamer goes on to show that this condition can be extended to chopping rules. It is just such a condition which will account for Berman's Dative Movement cases. The Structural Recoverability Condition precludes (115), for movement renders the chopping site unrecoverable. The chopping in (115) could have moved an element from either of the chopping sites indicated by dashes in (116).

(116) *Sally is a girl I would give ____ my last dime ____.
The structural ambiguity which would otherwise exist is disallowed, for the chopping site would not be uniquely recoverable without a Structural Recoverability Condition on chopping rules. Such a condition will also block (111b,c). The chopping site in (117) is not uniquely recoverable, for on structural grounds there are two possible chopping sites, as indicated.

(117) Presents are impossible to buy ____ John ____.
One of the readings (that indicated by the leftmost 'gap') would be blocked by selection restrictions, but Hankamer (p. 30) is quite clear that it is structural ambiguity that is blocked, and the matter of selection restrictions is irrelevant here.

Conditions on structural recoverability also seem to be at play in the Raising case. (118) has two possible chopping sites.

(118) *John is impossible ____ to expect ____ to know the answer.
A more remote structure of (118) could be either (119) or (120).

(119) It is impossible [to expect John [to know the answer.]]
(120) It is impossible [for John to expect [to know the answer.]]
One might raise the objection that the chopping site is recoverable
because the restriction against Tough-moving subjects of embedded sentences would prevent (120) from being recovered. But recall the motivation for restrictions on recoverability. To repeat what Hankamer said, where we may substitute 'chopping' for 'deletion':

'A deletion is recoverable if, given only the statement of the rule effecting deletion and the output of a particular application of the rule, the input to the rule can be uniquely determined.'

But the condition against chopping subjects is not part of the statement of the rule. Therefore, given only the statement of the rule, the chopping site is not uniquely recoverable.

Even if the approach developed here can't block the Raising cases, there is another way they can be blocked. Tough Movement always moves the rightmost element of a sentence, save that a preposition-al phrase of adverb may follow. But there is never a following S, NP, or VP. In (114a) John is followed by the VP (or maybe NP or S, depending on what one thinks about pruning) to know the answer, and application could be blocked on those grounds.

There seems to be a different principle at work in the Passive case. First, consider the form the derivation of (113) would have to take.

\[(121) \begin{align*}
  a. & \text{[for that group to accept one}_i \text{] is easy for one}_i \\
  \rightarrow & b. \text{[for one}_i \text{ to be accepted by that group] is easy for one}_i \\
  \rightarrow & c. \text{[to be accepted by that group] is easy for one}_i \\
  \rightarrow & d. \text{[to be accepted by that group] is easy} \\
  \text{(Passive)} \\
  \text{(Equi)} \\
  \text{(Indefinite Deletion)}
\]
If we then applied Tough Movement to (121d), we would derive (113b). But consider what the rules of Passive and Tough Movement do. They function to topicalize certain NP's, and the rest of the sentence is part of the comment. It is easy to see that there is a difference in meaning between (122a) and (122b), depending on what is topic.

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

Application of Passive topicalizes one in the embedded sentence. But application of Tough Movement topicalizes that group. If we are to preserve the requirement that transformations preserve meaning, then one must somehow be marked as topic, so Passive can apply to it, and that group must also be marked as topic, so Tough Movement can apply. But a sentence cannot have two topics, which is why (113b) is odd. Another way to look at it is that Passive throws that group into the background and therefore it cannot be moved into the foreground, or topic position, by Tough Movement.

Berman's constraint runs into problems with the About Movement case. To begin with, there is a question about the data, for most speakers that I have questioned find (112c) acceptable. Furthermore, whatever blocks (112b) is more general, for it blocks any leftward movement rule.

(123) a. *It's Mary who I want to talk about such things to.
   b. *Mary is the girl who I want to talk about such things to.

Given the acceptability of (112c), what could block movement of
Mary to the left in (112b) and (123)? I think we don't actually have a condition blocking rule application, but rather a length- and-complexity output condition on stranding the preposition to. The more intervening material there is between Mary and the preposition to, the worse the sentence sounds:

(124) a. Mary is difficult to talk about such things to.
   b. Mary is difficult to talk about these distressing things to.
   c. Mary is difficult to talk about things which affect her family to.
   d. Mary is difficult to talk about those things concerning the office of the Presidency to.

Notice that the same phenomenon appears in sentences which do not involve About Movement at all:

(125) a. Mary is impossible to speak to.
   b. Mary is impossible to give presents to.
   c. Mary is impossible to try to speak to.
   d. Mary is impossible for anyone to begin to try to speak to.

Notice also that if the to is carried along with the moved element, the sentence sounds much better:

(126) a. Who is it difficult to try to talk about such things to?
   b. To whom is it difficult to try to talk about such things?

Berman's article is interesting, for it points out the problem of spurious generalization. To be sure, a crucial part of science is discovering generalizations. But one can be too quick to group a selected set of data together and draw a conclusion. Such a move is
especially suspect when the conclusion offers no hope of providing an explanation, i.e. when it is purely descriptive. In the case at hand, the descriptive device is unnecessary, for subsets of the data can be explained by reference to certain natural principles of grammar. When an explanation is sought the apparent generalization turns out to be spurious. This example points out one of the problems with purely descriptive formalism in a theory.

3.13. **Auxiliary Reduction.** King (1970) notices that the rule of Auxiliary Reduction, which gives contracted forms of auxiliary is, has, would, had, have, will, are, and am, is blocked from applying if an element immediately following the auxiliary is moved or deleted. Thus we have the distinction between the a and b sentences below.

(127) a. Tell Harry that the concert's at two o'clock.
   b. *Tell Harry where the concert's at two o'clock.

(128) a. Bill's rich these days.
   b. *Sam's richer than Bill's these days.

My concern here is with whether a rule-specific constraint is needed. It is of course possible that this constraint is nonglobal. An attempt at a nonglobal formulation is made in Baker and Brame (1972). However, Lakoff (1972) presents what I think are serious problems with their attempt, so the question of reformulation is still open. So far as the question of generality is concerned, we need to know whether we must specifically mention the rule Auxiliary Reduction. Baker (1971) discusses this problem, and indicates
that general restrictions on stress-lowering come into play. Since Auxiliary Reduction is dependent on prior stress-lowering, we need not single out any particular rule for mention.

Baker discusses the rule Auxiliary Shift, which (in his words) 'positions the finite auxiliary at the left hand side of the verb phrase, to the left of a variety of different types of preverbal elements' (167). If the auxiliary is emphasized, it may not be re-positioned to the left, as (129) and (130) show (examples from Baker, 169).

(129) a. We often HAVE heard those allegations.
    b. The money never WOULD have been found by the police if Jones hadn't lost his key.
    c. Murphy never IS angry.

(130) a. *We HAVE often heard those allegations.
    b. *The money WOULD never have been found by the police if Jones hadn't lost his key ring.
    c. *Murphy IS never angry.

Auxiliary Shift is also restricted from applying when a constituent following the auxiliary is moved or deleted, even when the adverb rather than the auxiliary bears emphatic stress.

(131) a. I wonder where Gerard USUALLY is ___ at this time of day.
    b. I wonder where Gerard is USUALLY at this time of day.

The common factor in both situations is that the auxiliary has nonlow stress, so we may specify that Auxiliary Shift can apply only when the finite auxiliary is unstressed. We can provide a uniform formulation of the restrictions on Auxiliary Shift and Auxiliary Reduction
if we specify that the auxiliary be unstressed, a condition for which Baker provides some independent evidence in both cases. We can then specify that a general condition on stress lowering prevents application of the rules.

Now we must still face the question of how a deletion site prevents stress lowering. Baker offers the tentative proposal that 'the principle effect of deletion sites is to block the application of phonological rules by intervening between two elements mentioned in the structural description' (177). This, of course, does not make the restriction on stress lowering nonglobal, for a deletion-site mark is just a way of encoding a global constraint.

The problem with Auxiliary Reduction is complex, and no doubt much remains to be said on the subject. I know of no acceptable nonglobal alternative, but the constraint is not rule-specific. I have left this example out of the discussion of general roles because I wish to consider only clearly syntactic global constraints. The issue of the interaction of syntactic and phonological phenomena is outside the scope of this inquiry. I mention this example because it has receive such wide publicity.

3.14 **Raising/Passive/Psych Movement.** Postal (1972a) presents a restriction on interchanging certain NP's. (132b) and (133b) have undergone Raising and Psych Movement in the derivation from the _a_ versions.

(132) a. I seem [Jerry like Lucille]

→ b. Jerry seems to me to like Lucille.
(133) a. I strike Jerry like Lucille
   \rightarrow b. Jerry strikes me as liking Lucille.

(134) differs from the two above sentences in that Passive rather

than Psych Movement has applied.

(134) a. The police found out Jerry was living with Margaret
   \rightarrow b. Jerry was found out by the police to be living with
   Margaret.

There is a constraint on coreference in sentences which undergo

either Passive of Psych Movement:

(135) a. *Jerry seemed to me to like me.
   b. It seemed to me that Jerry liked me.

(136) a. *Jerry struck me as liking me.
   b. It struck me that Jerry liked me;

(137) a. *Jerry was claimed by Pete to have attacked him.
   b. It was claimed by Pete that Jerry attacked him.

(138) a. *Jerry was found out by the police to be criticizing
   them.
   b. It was found out by the police that Jerry was
   criticizing them.

The relevant constraint is:

(139) Throw out all derivations which have both:

(i) an underlying structure of the form:

```
S_1
  \rightarrow X, NP_1 V_1 NP_2

S_2
  \rightarrow NP_3 V_2 W NP_4
```
where $NP_1$ and $NP_4$ are stipulated coreferents; and

(ii) a later derived structure of the form:

This constraint blocks derivations in which $NP_3$ above is raised into $S_1$, and then interchanged with $NP_1$ by either Psych Movement or Passive.

3.15 Comparative Simplification. A constraint discovered by Michael Geis (1973) states that the rule of Comparative Simplification can apply (with the lexical item earlier than) only when the verb modified by at a time is identical to the verb originally modified by at which. Geis derives (141) from (140) by a rule which deletes the underlined material.

(140) John left at a time which was earlier than the time at which you left.

(141) John left earlier than you left.

Notice that (142a) is ambiguous, since at which could modify either say or leave.

(142) a. John left at a time which was earlier than the time at which you said that Pete left.

b. John left earlier than you said that Pete left.

But (142b) is unambiguous since the underlined material can be deleted.
by Comparative Simplification only if *at which* (as well as *at a time*) modifies the verb *leave*. Notice that we must also mention the distinction between the lexical items *earlier than* and *before*, for (143), unlike (142b), retains ambiguity after Comparative Simplification.

(143) John left before you said that Bill left.

3.16 Summary. In this chapter I have presented those global constraints which mention the names of specific rules in their formulation. I have shown that the overwhelming majority of these are not rule-specific global constraints. Either they are instances of general global constraints, or they are nonglobal constraints (deep structure constraints, surface structure structure constraints, constraints on the way some particular rule operates). I showed in the last chapter that the general constraints, however, could not be reduced to nonglobal alternatives. There are only two rule-specific constraints that I have not provided an alternative for: Postal's constraint on moving a raised NP under certain conditions (section 3.14) and the constraint proposed by Geis on Comparative Simplification. Hopefully, future research will show that these are either nonglobal or general, or that there is some basis to explain why these exceptions exist. In any event, the fact that so many of the rule-specific constraints have nonglobal alternatives lends credence to the idea that grammars should not contain rule-specific global constraints.
NOTES TO CHAPTER THREE

1Fauconnier shows that the Andrews-Lakoff global constraint is empirically inadequate, for there are in Greek constructions parallel to the 'accusativus cum infinitivo' construction of Latin (cf. Fauconnier (1971), 149-54). There are infinitival complements in which the subject of the infinitive is in the accusative case, and likewise any predicate modifiers of the subject. Yet if the subject of the infinitive is deleted by Equi, the modifiers take on the case of the controller for Equi. The global constraint cannot account for this, for at no point is the controller NP the derived subject of the infinitive. But Fauconnier's solution can easily account for such cases. Andrews tries to patch up the global constraint by having Equi superimpose the lower NP on the controller. Such a move would introduce an entirely new type of rule into transformational grammar and there is no independent motivation for it.

Furthermore Fauconnier (1973, 17) points out that Andrews's proposal fails in configurations like

\[
S_0 \left[ \ldots, NP_x^1, NP_x^2, \ldots, S_1 \left[ NP_x^3, A \right] \right] \\
[\text{case 1}][\text{case 2}][\text{case 3 = acc.}] 
\]

depending on \( NP_x^3 \) would be superimposed on \( NP_x^1 \) before case marking, and A could only take case 1. In fact, A can take case 1, 2, or 3 except when the controller \( NP_x^1 \) is in the nominative. Examples are in Andrews (1971).

2See Langacker (1969) and Ross (1967). Roughly, the constraint specifies that an anaphor may not both precede and command its antecedent.

3In a footnote Postal claims that the relevant sentences go through derivations of the following sort:

(i) Harry needed green bananas, and so George bought numerous bananas which were such that they were green.

\( \rightarrow \) (ii) Harry needed green bananas and so George bought numerous such bananas.

\( \rightarrow \) (iii) *Harry needed green bananas and so George bought numerous (such) (ones).

Postal concludes that this is evidence against blocking underlying structures in this case. But this is a non sequitur. All that is required is some statement of the incompatibility of numerous with a following NP containing the item such (or the semantic material in its lexical decomposition).
This rule is separate from the rule (usually called Quantifier Floating) which positions all, both, each, etc. into the post-auxiliary position of the verb phrase.

It is interesting that a movement rule either moves something over only one clause boundary (for example Raising) or else it is unbounded. There are no rules which move something over only two boundaries, three boundaries, odd-numbered boundaries, etc.

The unbounded movement rules operate in a single swoop, and are not successive cyclic. See Postal (1972c 471-2; 1972d).

Fauconnier (1971, Chapter V) claims that QF in French can operate to the left. But the only sentences he gives in evidence of this are those such as (76a), for which he claims the derivation:

(i) a. J'ai mangé chacun (de) NP.
   → b. J'ai chacun mangé NP. (by QF)
   → c. Je les ai chacun mangés (by Clitic Movement)

But we could just as easily allow Clitic Movement to carry a prenominal modifier along with the pronoun, giving (ii):

(ii) Je chacun les ai mangé.

Then QF applies, giving (76a). We thus maintain that QF only that operates to the right.

Grinder (1971) replaces deletion of indefinites by the mechanism of optional lexicalization. This has no bearing on the present issue.

Notice that this strategy supports the misanalysis of (99a), where a rose is assumed to be in toto the direct object argument of the verb.

It is a fact that there are sentences like (i), seemingly parallel to (99a).

(i) John bought a flower for the altar.

But the NP in the prepositional phrase cannot be an indirect object, as shown by the fact that IO Movement cannot apply.

(ii) *John bought the altar a flower.

Such facts indicate that grammatical relations cannot be stated only on structural configurations, which is the import of recent work by Postal and Perlmutter.
It was noticed in the literature by Fillmore and Postal that there are dialects (or perhaps idiolects) in which (99d) is acceptable. Neither Fillmore nor Postal can give a basis for a natural explanation of the dialect differences. Fillmore must postulate alternative rule orderings, a solution not only ad hoc but theoretically shaky in view of the move to eliminate extrinsic ordering. Postal must simply state that his constraint exists in some dialects but not in others. His constraint becomes even more ad hoc in this event. The analysis sketched above seems to me to provide a fruitful avenue for the study of dialect differences, since one expects dialects to differ on the basis of performance and perceptual strategies rather than on the basis of the rules and constraints of the grammar. In the example discussed here (99d) would be acceptable in some idiolects because some speakers would have perceptual strategies which would allow them to delay blockage of structures until a deeper level had been reached by application of other strategies. I suspect that some speakers process sentences at a 'deeper' level than others. The general issue has not been explored in any detail, but I see no reason to think that all speakers have the same uniform set of perceptual strategies.

I am concerned with the reading of (102b) which is synonymous with (102a). There is another irrelevant reading of (102b) in which John and Mary each have more than one bicycle.

In (109) Passive has applied to the lowest clause. The relevant intermediate stage before Passive applies is something like:

(i) It is unpleasant for one [John kick one]
CHAPTER FOUR

TRANSFORMATIONS AND FILTERS

1 General Remarks.

I have shown that while there are strong arguments for the existence of general global constraints, the arguments for rule-specific global constraints are weak. In the great majority of cases, the data motivating rule-specific global constraints can be explained by reference to principles which are nonglobal in nature. If we can eliminate rule-specific constraints from our theory, it becomes much more likely that global constraints will turn out to be universal. It is possible, of course, to speculate that a specific rule might have a global constraint on it in any language which exhibits that rule. But the chance that there could be such a situation is vanishingly small. This is because so much of the argumentation for constraints depends on the details of the way rules are formulated, and the details of rule operation differ greatly in different languages. For example, Raising in Japanese has a very different formal character from Raising in English, due to the fact that Japanese is verb-final, while English is not. Since universal grammar can give only the outlines of rules and not their details, the different ways in which rules operate in particular languages makes it unlikely that there could be a rule-specific
universal global constraint.

It is my opinion that the most profitable line of research in syntax is that presented by Bach (1971). There he seeks strong substantive universals in syntax by viewing universal grammar as containing a list of 'major' rules in their outline form. Each language draws from this list of transformations in forming its grammar. For examples of this brand of research, see Perlmutter and Orelšnik (1973), Hankamer (1971, 1973), and Ross (1970). A grammar must contain more than transformations, for many derivations that violate no rules are in fact ill-formed. Postal (1972a) adds a filter component to grammars:

By filter, I refer to a type of rule which is defined on such sequences of trees relating semantic to surface structure -RN and which has the function of marking as ungrammatical or ill-formed sequences which may be perfectly well-formed as far as the base rules and transformations are concerned. (138)

There has been much interesting work on the topic of global constraints. However, I think that a false analogy has been made between restricting the class of constraints and restricting the class of transformational rules. For example, Postal says:

...it can be seen that the problem of restricting the class of Global Derivational Constraints is not different in kind from the one for transformational rules. In each case, the initial general statement of the rule definitions permits an excessively wide class of rules, the overwhelming mass of which are found in no language. Hence empirical work must move in the direction of restricting both types of rule. (1972b, 36)

The restriction of transformations has been typically carried out by adding constraints on the class of sentences generable by the grammar, such as the Complex Noun Phrase Constraint, the Sentential Subject Constraint, Postal's Anaphoric Island Constraints, Chomsky's
By analogy with such an approach, the approach to global grammar has been to add constraints which limit those derivations which can be blocked by global constraints. A case in point is Dinnsen (1972). He argues that phonological theory should be constrained by the Null Segment Hypothesis: the only case in which derivational history may be referred to is that in which a rule applies to a string which has resulted from the earlier application of a deletion rule. That is, a rule may apply exclusively to a string which is derived from a deletion rule as opposed to an identical string which has not undergone the deletion rule. Dinnsen attempts to extend this hypothesis to syntax as well. Of course, some global constraints involve deletion rules (for example Controller Cross-Over) but some do not (for example the Quantifier Constraint). Therefore we cannot place an equivalent to the Null Segment Hypothesis on global constraints in syntax.

Beyond questions of empirical adequacy, might not the general approach be misguided? If we have constraints on constraints, do we have constraints on constraints on constraints, etc.? This matter is more serious than it perhaps appears at first. The motivation for adding constraints on derivations is that the generative power of transformational grammar has been claimed to be too strong in that many languages are generable which could not be natural languages. Constraints are then added which limit the class of languages that are generable. But it has been claimed by many people (e.g. Baker and Braine (1972)) that this approach only strengthens genera-
tive capacity in that the class of possible constraints is now unconstrained. People than seek constraints on constraints, such as the Null Segment Hypothesis. But we still have not weakened the generative capacity of grammars, for the class of constraints on constraints is now unconstrained. What would prevent one, for example, from formulating a restriction such as (1)?

(1) The Odd Hypothesis

Global derivational constraints which involve deletion rules are applicable only if the deletion rule which applies in the derivation to be constrained has an odd number of items in its structural description.

What I am trying to get at is that the formalistic approach to constraints on derivations does not offer hope of solving its own internal problem: that of excessive strength. A basic problem with the debate over the strength of grammars has been succinctly phrased by Jorge Hankamer (1973, 37-8):

It is somewhat misleading, in fact, to claim that a theory allowing transformations (of the unrestricted sort we now have) and not allowing some other device, is more restrictive than one which allows both devices. Given the power of transformations to derive virtually any conceivable class of sentences (including, unfortunately, a lot of 'sentences' which could never appear in any natural language), it is not at all clear how the addition of a new device could make the theory 'less restrictive'.

Perhaps we can get around the problem of strength by looking at the nature of global constraints. Restrictions on what rules can exist is a part of the type of universal grammar outlined by Bach. Such restrictions would seem to follow from the function of rules, i.e. they form questions, relative clauses, topicalizations, imperatives, etc. Likewise, it seems to me that insight into
global constraints comes from focusing on their function. Lakoff (1970) claimed that global constraints 'state well-formedness conditions on configurations of corresponding nodes in nonadjacent trees in a derivation'. I would like to combine Lakoff's idea with Postal's proposal concerning a filter component of grammars. An interesting approach to global constraints is to view them as filters which throw out the syntactic structures (i.e. derivations) of certain sentences as ill-formed on the basis of information in nonadjacent components of that syntactic structure. One of the basic claims of transformational grammar is that the structural description of a sentence consists of a set of phrase markers rather than a single phrase marker. Nonglobal filters are those which need refer to only a single phrase marker in the syntactic structure of a sentence (e.g. surface structure constraints).

The approach of transformational grammar has been to pare away from the class of grammars those things which cannot be grammars. But perhaps we can learn about grammars by looking at what they are, as well as what they aren't. Too often, linguists present only solutions to a set of data, without seeking an explanation.
2 Description and Explanation

I think that a major problem with discussions of the strength of grammars is that they are conducted within a framework that does not give adequate attention to concepts of explanation. If theories are to be evaluated by simplicity measures, those theories must be descriptive. Stampe (1968) provides a criticism of the limitation of linguistics to description. Chomsky claims that a theory achieves the level of explanatory adequacy when it selects from a set of descriptions the correct description (by means of an evaluation measure). Universal grammar is then viewed as a set of constraints on possible grammars. Stampe notes that this is a rather limited view of the goals of linguistics, for 'it fails to encompass what most of us would consider the main goal of linguistics, which is to furnish us an explanation rather than just a description of grammatical universals, and therefore it limits the role of universal grammar to an essentially descriptive function.' Stampe notes that explanation concerned with the question why languages have the form that they do is a different matter from Chomsky's concern with explanatory adequacy. For we must also decide between alternative formulations of universal grammar, where reference to perceptual or logical explanations is a crucial basis for favoring certain formulations over others. A description which does not offer hope of an explanation is cast into doubt.

One type of description which offers little hope of an explanation is one where we fail to provide motivated restrictions concerning
where that type of descriptive account may be used. For the use of a certain type of description in cases of one type does not entitle us to extend the use of that description to cases of another type.

The problem is a general one in science. For example, we may notice in a particular instance that an apple falls to earth and account for this by the force of gravitation. Now if we go elsewhere and observe that a brick falls to earth, we may account for this as being due to the same phenomenon, namely, gravitational attraction between two bodies. The descriptive leap may be made where all factors are equal save for particular instantiation of the elements involved. But suppose we then observe that the ends of two magnets attract each other. We are not entitled to assert that the force of gravitation is responsible for this attraction. To be sure, there is a comparatively infinitesimally small component of gravitation involved, but this does not account for the majority of the attractive force. What we have failed to take into account is the qualitative and quantitative differences between the cases: quantitative in that gravitational and magnetic force differ in strength by orders of magnitude, qualitative in that gravitational force holds anywhere at anytime, whereas magnetic force holds only between the north and south poles of magnetized bodies. Indeed, in the absence of qualitative and quantitative specification, the concept 'gravity' is not an explanation, but only a descriptive device asserting that in certain instances certain bodies attract each other. The situation is parallel in linguistics. We observe in certain instances that
global constraints are to be used. Yet in the absence of further specification, this is only an assertion that some derivations may not be blocked by information in adjacent phrase markers. If we cannot specify conditions where we may and may not have global constraints, the concept is not part of an explicit account of what natural language is. The distinction between rule-specific and general global constraints is the type of qualitative differentiation which will allow for a more adequate explanation of what languages are.

The extension of a type of device from one domain to another appears in a line of argumentation which is implicit in discussions of global constraints, although not explicitly formulated. This line of argumentation is:

(a) Premise 1: a certain descriptive device D is the only apparent way to account for example a.

(b) Premise 2: example b, consisting of a certain array of starred and nonstarred sentences, can be handily handled by descriptive device D.

(c) Conclusion: example b is to be handled by descriptive device D.

But the conclusion does not logically follow from the conjunction of the premises, for we need the additional premise that there is no other way to account for example b than by descriptive device D. It is often claimed that because of the conclusion, we have additional evidence in favor of the descriptive device D introduced in Premise 1. But such reasoning is circular, for the conclusion
is (incorrectly) presumed to follow from the conjunction of
premises 1 and 2. How could a conclusion derived in part from
premise 1 be used in favor of something which must be demonstrated
in adducing premise 1?

In fact, suppose that there is another way to handle example b.
Then the conclusion does not even follow from the premises, and so
could not be used to make a construct mentioned in the first premise
more plausible. Now suppose that there is an alternative account
A for example b which embodies an explanation. Then A must be
preferred to D, for given two accounts of the same domain, the one
which embodies an explanation is preferable, for it tells us more
about the nature of the universe. The alternatives thus cannot
be evaluated solely on grounds of simplicity and elegance. In
particular, there may be cases where a descriptive device D can
provide a uniform description of a set of data, while there is not
a uniform explanation. For example, the data cited by Berman in
support of her constraint on Tough Movement (section 3.12 above)
fall into several different classes, each of which must have a
separate explanation.

3 Rules and Filters

In the last section I noted that unless we can restrict the use
of a certain type of description in some plausible way we cannot
know where we are using that description appropriately. I have
argued that global constraints may not be rule-specific. I think
that this restriction is not something that is arbitrarily imposed
on the class of global constraints, but rather that it follows from an adequate characterization of what global constraints are. If this is true, we avoid the problem that I discussed in section 1, namely that the imposition of an arbitrary restriction does not solve the problem of excessive strength of theories. Based on the general constraints presented in Chapter Two, we may define the notion 'global constraint, as follows.

(2) Global constraints are filters which throw out the syntactic structures (i.e. derivations) of certain sentences as ill-formed on the basis of information in nonadjacent components of that syntactic structure.

In other words, a global constraint is not a condition on the application of any rule. We let transformational rules apply according to their structural descriptions and then check the syntactic structure (derivation) of the sentence generated to see whether it fits the filter. If it does, it is thrown out. In this way, we maintain a distinction between transformational rules and constraints.

Such a distinction seems to be called for. For one thing, consider the very different ontogeny of each type of device. A transformation must be learned by a child on the basis of hearing pairs of synonymous sentences. For example, (3a,b) are synonymous because of their transformational relationship.

(3) a. We believe that John is honest.

b. We believe John to be honest.

It is interesting that no transformations have ever been proposed which do not relate pairs of synonymous sentences. The initial evidence for a transformation must come from the sentences that a
child hears. But the child never hears the sentences that are cited as evidence for a constraint, so constraints must have a different ontogeny. In the present state of psycholinguistic research, we cannot say exactly how constraints are learned. But notice that once a few rules are learned, they can be applied in ways that result in violation of a constraint. How many of the surface structures produced in this way may be unrecoverable by the hearer. (For example, remember Hankamer's Hankamer's No-Ambiguity Constraint, discussed in 3.12.) Since the speaker is himself a hearer, it is quite likely that he will not produce sentences where the deep structure is unrecoverable. There must be some sort of fit between production and perception in order to make communication easier.

The different ontogenies correlate with the fact that while transformational rules account for grammatical sentences, filters account for ungrammatical sentences. This difference was pointed out by Chomsky (1965):

... it is clear that we can characterize unacceptable sentences only in terms of some 'global' property of derivations and the structures they define.... a property that is attributable, not to a particular rule, but rather to the way in which the rules interrelate. (12)

Chomsky's quote expresses what I have argued for here: that filters refer to the syntactic structures of sentences and not to any particular rule.

As further evidence for the distinction between transformations and filters, notice the different effects they have. A transformation consists of a structural description and a structural change.
If some tree is properly analyzable with respect to some transformation, that tree changes form according to the structural change specified. But while constraints have a structural description, they do not have a structural change: their output is not a phrase marker but the creation of a * for a sentence.

There seems to be distinction between a transformational component of grammar and a filter component. If global constraints have the character presented in (2), it follows that they are not rule-specific. A global constraint keeps track of configurations of corresponding nodes at nonadjacent points in a derivation. This is to say that when a derivation is blocked, it is on the basis of some property of one phrase marker in conjunction with some property of a nonadjacent phrase marker. But what could this property be? Just some aspect of the tree structure, which is a hierarchical labelling of constituents. We need to have a case where an earlier and a later tree differ in regard to their structure in order to have a global constraint. In such an event, what difference could it make as to what rule produced the later structure, so long as the relevant configuration occurs? To be sure, there may be cases where only one rule produces an effect of a specific type. But in that event, it would be gratuitous to claim that a constraint is rule-specific.

If there is a distinction between a rule component and a filter component, we can maintain a more restricted definition of what a transformation is. Suppose that global constraints could be rule-specific. This is another way of saying that in some cases they are
not part of the filter component but are part of the transformational component. As an example which may make this more clear, consider the case of Each Shift (section 3.9). Postal proposed that Each Shift may apply only when each-Source precedes each-Target. But we also need to specify that each-Source and each-Target were clause mates earlier in the derivation. We actually have a restriction on the application of the rule Each Shift. That is, we must say that Each Shift can apply only when the relevant conditions obtain.

Remember that a transformation consists of a structural description and a structural change. If we could have global conditions like that presented by Postal, it would mean that no single tree would be properly analyzable with respect to the rule Each Shift. The structural description of the rule would have to mention two phrase markers rather than one. We can't a priori say at what point in the derivation the relevant phrase markers appear (especially in the absence of extrinsic ordering). Thus we would have to scan the entire derivation to see whether there were two trees which together met the structural description of the rule. This means that every time we have to apply a rule that has a global condition we have to scan the entire derivation for rule application. This is a considerable extension of the notion transformation. If global conditions may not be part of a specific rule we can maintain a more restricted definition of what a transformation is. Without this restriction, the transformations would fall into two different classes: those which are global and those which are not. However, each transformation would have to arbitrarily placed into one class or another.
There is no functional or formal distinction which could predict when a rule would have a global condition on application.

Furthermore, if transformations could have global conditions written into their structural descriptions, we would generate the following prediction. Cases have been proposed where a rule may not apply to a certain type of derived structure. For example, Berman claims that if an NP undergoes Raising, it may not later undergo Tough Movement. If we could have such conditions we should also be able to have cases where a rule can apply only to a derived structure. A hypothetical example would be a restriction that an NP could undergo Psych Movement only if it had been raised from a lower clause, but not if it originated as a direct object in deep structure. No such cases have ever been proposed. In all cases where a node is moved into some position, it acts later like a node which originated in that position. Without a restriction against rule-specific global constraints, we make a prediction about language which appears to be false.

4 Summary.

I have tried to find an adequate characterization of what global constraints are. In the process, I have argued that a number of cases for which global constraints have been proposed should reanalyzed. I think this points out the necessity of being careful in the use of formal descriptive devices.

I have presented some evidence in favor of a certain restriction on the use of global conditions in syntax: that they are part of
the filter component of grammar and are therefore not rule-specific. Such a move takes us into the realm of metatheory. As defined by Sanders (1973) metatheory places limits on the class of statements that can be axioms of any particular grammar or theory of grammar. Sanders argues for universally-determined rule application as part of the metatheory: Principles governing the use of the rules of particular languages are universal -- they have a natural basis in the function and goals of human communication.

Sanders, like Postal, draws a distinction between rules and constraints, claiming that constraints (principles governing the use of rules) are universal. Thus, knowledge about a particular language only consists of knowledge about particular rules in that language. Sanders speculates that all global constraints may be universal. While this has not been shown, it is an idea worth consideration. The results of the present essay, if correct, support the idea that global conditions are general in nature. If global conditions are general rather than particular, it becomes more likely that they will turn out to be universal. I hope that the results of this essay can serve as a steppingstone to the discovery of those things which are common to all languages.
NOTES TO CHAPTER FOUR

1See Postal (forthcoming) for details about the operation of Raising in English and Japanese.

2It appears that the Sentential Subject Constraint (discussed in Chapter One) is not part of the Japanese language, and so not universal. Given further research, it may be possible to show that the lack of this constraint follows from some other fact about the Japanese language.
BIBLIOGRAPHY


Gordon, David and George Lakoff (1971) "Conversational Postulates," Papers from the Seventh Regional Meeting, Chicago Linguistic Society, 63-84, Department of Linguistics, University of Chicago.


Hankamer, Jorge (1973) "'Unacceptable Ambiguity,'" *Linguistic Inquiry* 4, 17-68.


Karttunen, Lauri (1969) "Pronouns and Variable," in Papers from the Fifth Regional Meeting, Chicago Linguistic Society, Department of Linguistics, University of Chicago.


Langendoen, D. Terence and Thomas C. Bever (1973) "Can a Not Unhappy Person be Called a Not Sad One?" in Anderson and Kiparsky (1973).


Pullum, G. K. (ms.) "Restating Double-Ing," unpublished manuscript, University of York.


