INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. 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 material submitted.

The following explanation of techniques is provided to help you understand markings or notations 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 through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

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

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
HENDRICK, RANDALL JAMES
ON NESTING AND INDEXICAL CONDITIONS IN
LINGUISTIC THEORY.

UNIVERSITY OF CALIFORNIA, LOS ANGELES, PH.D.,
1979

COPR. 1979 HENDRICK, RANDALL JAMES
University Microfilms
International

© Copyright by
Randall James Hendrick
1979

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
UNIVERSITY OF CALIFORNIA
Los Angeles

On Nesting and Indexical Conditions
in Linguistic Theory

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Linguistics

by

Randall James Hendrick

1979
The dissertation of Randall James Hendrick is approved.

Stephen R. Anderson

Carlos Otero

A. Carlos Quicoli

Joseph E. Emonds, Committee Chair

University of California, Los Angeles
1979
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>vi</td>
</tr>
<tr>
<td>Vita and Publications</td>
<td>vii</td>
</tr>
<tr>
<td>Abstract</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1. Prospects and Results</td>
<td>1</td>
</tr>
<tr>
<td>The Context of Inquiry</td>
<td>1</td>
</tr>
<tr>
<td>Overview</td>
<td>11</td>
</tr>
<tr>
<td>2. Subcategorization and the Association</td>
<td>26</td>
</tr>
<tr>
<td>Principles</td>
<td></td>
</tr>
<tr>
<td>Contextual Features for Lexical Insertion</td>
<td>26</td>
</tr>
<tr>
<td>On Restricting the Interpretation of Contextual Features</td>
<td>27</td>
</tr>
<tr>
<td>Certain Problems with the Strict Locality Condition</td>
<td>33</td>
</tr>
<tr>
<td>Selecting COMP and Tense and the Broadly Local Condition</td>
<td>33</td>
</tr>
<tr>
<td>Comparative Clauses</td>
<td>44</td>
</tr>
<tr>
<td>Manner Adverbs</td>
<td>50</td>
</tr>
<tr>
<td>Complements to Adjectives</td>
<td>60</td>
</tr>
<tr>
<td>Presentation of the Association Principles</td>
<td>68</td>
</tr>
<tr>
<td>Narrowing the Broadly Local Interpretation of Contextual Features</td>
<td>69</td>
</tr>
<tr>
<td>Disambiguating Certain Broadly Local Structures</td>
<td>75</td>
</tr>
<tr>
<td>Multiple-Headed Comparatives</td>
<td>80</td>
</tr>
<tr>
<td>Deletion of That in Complements to Adjectives</td>
<td>84</td>
</tr>
<tr>
<td>Footnotes to Chapter 2</td>
<td>93</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
## Chapter 3. Clause Internal Bindings and the Association Principles

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>WH-Movement and Passive</td>
<td>95</td>
</tr>
<tr>
<td>Clitic Pronouns</td>
<td>97</td>
</tr>
<tr>
<td>Multiple Clitic Pronouns in Spanish</td>
<td>97</td>
</tr>
<tr>
<td>Comparison with Previous Analyses</td>
<td>110</td>
</tr>
<tr>
<td>Clitic Pronouns in French</td>
<td>116</td>
</tr>
<tr>
<td>Left and Right Quantifier &quot;Movement&quot;</td>
<td>120</td>
</tr>
<tr>
<td>Movements of Adverbs</td>
<td>127</td>
</tr>
<tr>
<td>Clitic Pronouns and the Strict Frame Requirement</td>
<td>135</td>
</tr>
<tr>
<td>&quot;Optional&quot; Subcategorization</td>
<td>142</td>
</tr>
</tbody>
</table>

## Chapter 4. The Clause Boundedness of Conditions in Linguistic Theory

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limiting the Association Principles to a Single Clause</td>
<td>153</td>
</tr>
<tr>
<td>Other Principles of Universal Grammar Limited to a Single Clause</td>
<td>155</td>
</tr>
<tr>
<td>Case Concord and the Function of the Passive Morphology</td>
<td>155</td>
</tr>
<tr>
<td>Unifying the Nominative Island Condition and the Opacity Condition</td>
<td>166</td>
</tr>
<tr>
<td>The Status of the Sonata-Violin Paradigm</td>
<td>186</td>
</tr>
<tr>
<td>Remarks on Chomsky's Counter-Examples</td>
<td>194</td>
</tr>
<tr>
<td>Footnotes to Chapter 4</td>
<td>198</td>
</tr>
</tbody>
</table>

## Chapter 5. Clitic Pronouns and the French Causative Construction

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Problem</td>
<td>199</td>
</tr>
<tr>
<td>The French Causative as Base Generated</td>
<td>203</td>
</tr>
<tr>
<td>Subcategorization</td>
<td>203</td>
</tr>
<tr>
<td>Reflexives</td>
<td>211</td>
</tr>
<tr>
<td>Subject Oriented Adverbs</td>
<td>221</td>
</tr>
<tr>
<td>The Effects of the Association Principles</td>
<td>225</td>
</tr>
<tr>
<td>Asymmetries of Clitic Movement, the SSC and the Association Principles</td>
<td>226</td>
</tr>
<tr>
<td>Asymmetries of Quantifier Movement, the SSC and the Association Principles</td>
<td>248</td>
</tr>
<tr>
<td>The Marked Nature of the VP Complement</td>
<td>263</td>
</tr>
<tr>
<td>Footnotes to Chapter 5</td>
<td>270</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

This study was done in bleak times. I have no doubt that under different material and psychological conditions the particular analyses presented here could have been improved. All in all, though, I consider the completion of this study itself an accomplishment.


In addition I owe a debt to the following people for material and psychological support: S. Anderson, M. Baltin, J. Barzman, J. Emonds, P. Herreshoff, J. Hendrick, W. Hendrick, C. Miller, C. Piera, R. Schwanner, E. Torrego, and D. Wheeler.

Lastly I want to thank J. Emonds for the intellectual respect that made my years of being a student at UCLA bearable.
VITA

April 29, 1951--Born, Chico, California

1973 --B.A. in Comparative Literature, University of California, Irvine

1974 --M.A. in Comparative Literature, University of Toronto

1974-1976 --Instructor, Moorpark College

1976-1977 --Teaching Assistant, University of California, Los Angeles

1979 --Research Assistant, University of California, Los Angeles

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

On Nesting and Indexical Conditions in Linguistic Theory

by

Randall James Hendrick
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 1979
Professor Joseph E. Emonds, Chair

This study presents a number of proposals for linguistic theory which function to constrain the output of the core syntactic component of a generative grammar. These conditions can be divided into a "nesting" condition on the one hand, and indexical conditions on the other hand.

The "nesting" condition which is proposed is formulated as follows:

The Association Principles
Given the class of rules R, if R_i involves two terms A_i and B_i in a string A_i - X - B_i - Z containing A_j and B_j related by R_j, then

viii
1. if $A_i - X$ contains $A_j$, $X$ contains $B_j$ as well, and

2. at least one phrasal node dominating $B_i$ does not dominate $B_j$ or conversely

where $B_i$ is structurally non-distinct from $B_j$.

A major source of corroborating evidence for the Association Principles comes from contextual features governing lexical insertion. In Chapter 2 it is argued that of any two elements involved in a contextual feature, at least one must be in the domain of the other. Several subsidiary conventions are adopted to make this proposal descriptively adequate. Among these conventions is the Association Principles which serve to structurally disambiguate certain configurations (e.g., multiple headed comparatives) permitted by the formalization of contextual features just mentioned. A second major source of evidence corroborating the Association Principles (detailed in Chapter 3) comes from clause internal anaphoric bindings. Passive, WH-movement, adverb movement and topicalization in English, as well as clitic pronoun movement in Spanish and quantifier movement in French, provide evidence favoring the Association Principles.

Other major indexical conditions are considered in Chapter 4. It is argued that the Opacity Condition, the Nominative Island Condition and the Subjacency Condition can be collapsed as the Index Suppression Condition. This
condition has certain empirical advantages over the original formulation of the Opacity Condition. Another condition, the Case Concord Condition is proposed.

The Case Concord Condition has the effect of explaining why the passive morphology is associated with the movement of an NP into subject position. These conditions are limited to a single clause and it is argued that limiting the Association Principles to a single clause removes various counter-examples to nesting conditions that have been noted in the literature.

The fifth chapter examines the interaction of clitic pronouns and the French causative construction. Treating this construction as a base generated VP complement permits one to avoid complications with respect to conditions in linguistic theory. Asymmetries in the operation of the clitic placement rule are shown to follow from the Association Principles. The claim that the causative construction involves a VP complement is reconciled with the general claim that all infinitives in English involve sentential complements.

The sixth chapter returns to the notion of structural non-distinctness employed in the formulation of the Association Principles. Certain theoretical advantages are obtainable by defining the matrices \([N^+]\) and \([-N^-]\) as non-distinct in Italian. In particular "strong" readjustment rules, that is operations involving more than 

x
Chomsky-adjunction of adjacent constituents, can be avoided. Similarly evidence is presented suggesting that the matrices \([+N]_{-V}\) and \([-N]_{-V}\) are non-distinct in English.
Chapter 1

PROSPECTS AND RESULTS

1. The Context of Inquiry

Linguistics aims ultimately at providing a characterization of natural language. This task reduces to the subsidiary goals of defining Universal Grammar and delineating the parameters in which the variability of particular grammars may fall. There is no particular reason to believe that our scientific ability is commensurate to this ultimate goal, although this recognition says nothing about the legitimacy of the undertaking. As in other sciences, it is useful to identify problems, a solution to which bears on the ultimate objectives. Such a problem is often called a research program (cf., Lakatos 1970).

Prior to the Chomskyan revolution in linguistics, the research program adopted by most linguists was the construction of a set of procedures which would enable a researcher to extrapolate an observationally adequate grammar of a language from a set of primary speech data. The resulting description of the particular language was not realistic in the sense that it had no necessary relation to the mind of speakers. In addition, because the research program did not aim at a realistic
characterization of natural language, it identified the task of characterizing natural language with providing at least one description for each known natural language. No other goal was left once realism was abandoned: the notion of a possible but non-existent human language was empty and trivial since there was no basis on which to say that one grammar was possible while another was not. Similarly, given two descriptions which are observationally adequate there was no principled grounds on which to prefer one over another. Insofar as the research program provided no reason to expect that the notion of a possible natural language was constrained and that the variation between languages had relatively narrow parameters, it became important to describe every natural language.

The empiricism of this program is unmistakable. The guarantee that a particular description constituted knowledge was provided by the procedure for the construction of descriptions. This procedure was intended to proceed inductively from "primary sense data" to larger generalizations without reference to other generalizations not already based on the "primary sense data." The "primary sense data" just referred to was constituted by "observable" behavior, that is, behavior embedded in the acting out of the existing social order. Even the procedures postulated to construct generalizations were necessarily data bound: the only permissible procedures were
the establishment of elements and the illumination of their
distribution relative to one another. In this way theore-
tical constructs were intended to always be verified by
sense data and any theoretical category or principle of
distribution which could not be verified by constructing
procedures to derive them from such "observable" behavior
were dismissed as "mentalist." It is no accident that in
this general research program observational adequacy was
the most ambitious goal that could be envisioned. Explan-
tory adequacy was beyond the grasp of this research program
because in order to achieve explanatory adequacy a pro-
cedure to select one of several descriptions would need to
be devised and yet such a procedure cannot have its moti-
vation in the analysis of primary sense data, in the estab-
lishment of elements and the examination of the distribu-
tion of those elements.

The creation of generative grammar marked a
qualitative shift away from the research program outlined
above. Some of the essential properties of that shift are:
(a) the claim of constructing a realistic theory, (b) an
extension of the domain of data in the purview of the
theory, (c) a reduction in the stringency of the guarantee
of knowledge from necessary verification to potential
falsification, and (d) setting explanatory adequacy as a
goal of linguistic theory. I want to make some brief
remarks on each of these properties.
One of the most basic aspects of the Chomskyan revolution that produced generative grammar was the demand for a realistic characterization of natural language. Linguists interested in the possibility of such a theory have developed two problems that research attempts to resolve. One problem is that of defining a linguistic theory which can be related in a well-defined fashion to an independently motivated performance model. For examples of work carried out in this paradigm see Chomsky and Miller (1963), Bresnan (1978), and Fodor (1978) among others. A somewhat different, though not contradictory, problem is the characterization of how children learn language in the first place (sometimes called the projection problem). This problem is the one approached by Chomsky throughout his work. I return to this topic again below.

In addition, generative grammar threw over the criterion that evidence be behavior observable in the acting out of the existing social order. Many people both inside and outside the field misunderstand this change in linguistics. Generative grammar makes no a priori theoretical distinctions in data, nor a distinction between psychological and linguistic data, nor a distinction between semantic and syntactic data, nor a distinction between sociolinguistic and "intuitive" data, even though these distinctions might provide convenient ways of speaking in an informal way. Generative grammars need not be
constructed exclusively on intuitive judgements of acceptability contrary to a popularly held misconception. The predominance of such intuitive judgements at the present moment in linguistic research is related to the fact that the set of actually occurring sentences that can be realistically gathered is tiny when compared to the infinite set of possible sentences, and the relative ease with which that intuitive judgements can be made in contrast to the task of collecting actually occurring sentences.

However, pre-Chomskyan linguistics did restrict the domain of data a priori. Moreover the American Structuralists restricted that domain in a very conservative fashion, a fashion that limited consideration to functional language exchange within narrow sociohistorical parameters. In this way when generative grammar took intuitive judgements about acceptability to be part of the domain of data that it undertook to explain, it freed linguistics from only examining a narrow class of data.

Another aspect of this shift is the reduction of the guarantee of knowledge from necessary verification to postential falsification, that is, from the identification of knowledge with proven theories to the identification of knowledge with theories not disproven. Prior to the advent of generative grammar it was assumed methodologically that the certainty of the empirical base, existent primary speech data, could be transferred directly to
theoretical propositions if one employed "inductive logic" and ascribed to the general rule that any given fact could disprove a universal theory. From the very beginning generative grammar exhibited its willingness to forego the identification of knowledge with proven theories, theories which could have the certainty of the empirical base, by postulating "core" statements that after having been established as plausible the field would not be willing to let be disproved and by hypothesizing auxiliary statements to protect those core statements. The principle that rules of grammar only involve constituents, established from the very outset of the development of generative grammar in Chomsky (1957), is a case in point. However it is worth noting that the clarity of understanding on this point is low and by no means is there hegemony in the field on this topic. The controversiality in some people's minds of Chomsky's (1977b) statement that only theories (and not data) disprove theories is representative of the persis-
tence of this aspect of the early structuralist empiricism.

Lastly let me make a general remark about explanatory adequacy. This goal has a great deal to do with the demand for a realistic linguistic theory. At the same time the goal of explanatory adequacy, of selecting one descriptively adequate grammar from among numerous observationally adequate accounts, has been a major impetus in the criticism and growth of linguistic theory in generative
grammar. Indeed one of the superficially most striking characteristics of generative grammar, its formal nature, is a result of the aspirations to an explanatory model: it is only in terms of explicit, formal descriptions that competing theories can be compared. One should not underestimate the importance of the place of explanatory adequacy in the development of generative grammar. Science or knowledge is not a static set of statements but a continually changing set of statements with an expanding empirical domain. It is just for the lack of this historically developing character that Lakatos (1970) identifies post-Freudian psychoanalysis and post-Leninist Marxism as non-scientific.

As mentioned above this study is aimed to be a contribution to the solution of the projection problem. Classically research in this area attempts to limit in a significant way the class of grammars that a language learning procedure can formulate. The motivation behind these attempts is the recognition that the learning procedure hypothesizes a close approximation of the adult grammar at an early point and yet the class of grammars permitted by a version of linguistic theory like that presented in Chomsky (1965) is unlimited. It is necessary to narrow this class drastically in order to provide a solution to the projection problem. The A/A condition of Chomsky (1968), Ross's (1967) constraints on variables in
syntax, Emonds' (1970) structure preserving constraint, the
development of the Bar Notation (most explicitly treated in
Jackendoff 1978) and Chomsky's (1973) conditions on trans-
formations and subsequent refinements, are all representa-
tive of this general line of research. These studies
expand the empirical domain of linguistic theory. They
constitute auxiliary statements or hypotheses to the overall
proposals of Chomsky (1965). They function to remove evi-
dence apparently falsifying the core proposals of Chomsky
(1965). It is the proliferation of these auxiliary state-
ments and the expanding empirical base of the theory that
they entail which shows the productiveness of the research
problem and which justifies our treating generative grammar
as knowledge or science.

Of course studies in this area are still rather
basic in that they simply restrict the class of grammars
defined by linguistic theory. Little thought has been
given to the question of constructing an idealized pro-
cedure that will actually postulate a grammar. Wexler and
Culicover (in press) and Culicover and Wexler (1977) are
important attempts to move in this direction. In their
discussion one finds that the general conditions used to
restrict the class of grammars are involved in the learning
procedure in one of two ways: on the one hand there are
conditions which limit the data the learning procedure can
operate on, and on the other hand there are conditions
which limit what kinds of rules the learning procedure is capable of hypothesizing. The Subjacency Principle of Chomsky (1973) is an example of the former, the requirement that terms involved in rules of grammar necessarily be constituents is an example of the latter. This distinct is parallel to the distinction drawn in Chomsky (1973) between conditions on the function of rules on the one hand and conditions on the form of rules on the other hand. The general principles of Universal Grammar proposed in this study are formulated as principles constraining the functioning of rules and might be involved in restricting the functioning of the learning procedure. More direct relations to a learning procedure have not been explored but would no doubt be of theoretical interest.

I believe that there are important convergences between generative grammar as it relates to the projection problem and psychoanalytic theory. Besides the interest in early childhood, both theories intimately deal with and presuppose non-conscious "knowledge," both theories hypothesize that biology plays a significant role in this "knowledge," both theories fought competing theories with empiricist characteristics in order to establish themselves, both make crucial use of binary oppositions, and a range of other properties. This is not the appropriate place to examine in any detail the relation between linguistic and psychoanalytic theory; in Hendrick (in progress) some of
these questions are dealt with. However it is worth noting that it was precisely psychoanalysis' attention to childhood, its biologism, its use of binary oppositions, its recognition of non-conscious knowledge and its general antiempiricist origin that made it controversial and gave momentum to the reaction among the "post-Freudians" (e.g., Horney, Fromm, Reich, Adler) to eliminate precisely these characteristics from the theory as it was developed by Freud (cf., Marcuse 1955, Jacoby 1975). This revision in psychoanalysis pushed from a variety of direction by the post-Freudians was made in the name of therapeutic results in "real life situations." Meanwhile orthodox Freudians in the face of these attacks extended the irrefutable core to such an extent that psychoanalytic theory became inflexible and unwilling to enlarge its empirical domain.

Linguistics should extract a lesson from the history of psychoanalysis in this respect. I am afraid that the performance model problem and questions like "How is language actually used?" "How can language be simulated by existing technology?" "How can we obtain therapeutic results with language-deficient individuals?" will encourage the subordination of theoretical developments to practical or technological results with the same deleterious effects that similar concerns had on psychoanalysis. The performance model problem will give rise to results which though not any more realistic than the results of research on the
projection problem in any well defined way will be more "realistic" in the sense of having technological implications. Such developments and the subsequent confusion of realistic models with models having immediate technological applications would lead to a narrow dogmatism on the part of researchers working on the projection problem in the name of orthodoxy. Insofar as the projection problem is less directly related to practical handling, it permits theoretical developments, the essence of science, until a point where the gains of generative theory are firmly entrenched and widely recognized.

In the next section I provide a fairly full recapitulation of the content of this study. This overview, organized chapter by chapter, is designed to give an overall perspective to the wide variety of phenomenon treated here, especially in light of the rather detailed argumentation that often is involved. Hopefully this perspective will facilitate the reading of this study and help the reader to an early evaluation.

2. Overview

This study is concerned with the existence in Universal Grammar of a condition requiring certain structures to be "nested." For the moment we can say that $B_j \ldots G_{j+n}$ is nested in $A_1 \ldots A_{i+n}$ if every $B_k$ has $A_i$ to its left and $A_{i+n}$ to its right. So for example the pair $B_j, B_k$. 
is nested in \( A_i, A_1 \) in (1a) but \( B_j, B_1 \) in (1b) is not nested in \( A_i, A_k \).

\[
\begin{align*}
(1) \text{a.} & \quad [A_i \ldots [B_j \ldots ]_{B_k} \ldots ]_{A_1} \\
\text{b.} & \quad [A_i \ldots [B_j \ldots ]_{A_k} \ldots ]_{B_1}
\end{align*}
\]

Nesting, in this sense, has been of some importance since the earliest studies in generative grammar. Chomsky's (1956) demonstration of the inadequacy of a finite state grammar for the description of natural language depends crucially on nested dependencies like either...or..., and if...then... Similarly Chomsky (1961) and Chomsky and Miller (1963) focus some attention on the effect that nesting of relatively long elements has on reducing the acceptability of a sentence.

It is Bordelois (1974) that first uses the notion of nesting to restrict the functioning of rules of grammar. Bordelois proposed the restriction in (2).

(2) The No-Crossing Condition

Prevent the application of any rule resulting in the following configuration:

\[
...X...Y...X...Y...
\]

where \( X \) and \( Y \) are constituents under a given sentential node, relating to their respective traces by means of a rule and where obligatory control is operative within the scope of the crossing paths.
For Bordelois "traces" are understood somewhat differently from Chomsky (1976). A "trace" for Bordelois is any anaphoric element while current terminology treats "traces" as an empty node exhibiting the same properties as elements like each other or herself, that is, cases of bound anaphora, or "obligatory control" in Bordelois' formulation.

Bordelois presents several arguments in favor of the No-Crossing Condition. The first and central argument involves extraction from the embedded complement in the Spanish causative construction. Bordelois hypothesizes that the underlying structure for a causative construction is V-S-NP where the NP controls the subject of the embedded sentence. This means that a Spanish sentence like (3a) has the underlying structure in (3b).

(3a). dejé comprar cigarrillos a Juan.
    I let Juan buy cigarettes.

b. dejé [\text{[S-PRO} _i comprar cigarrillos] (a) Juan,].

(4a). dejé a Juan comprar cigarrillos.

b. dejé a Juan comprar cigarrillos [\text{[NP]} e].

Variants of (3a), like (4a), are generated by postulating an intraposition rule moving Juan to the left and resulting in a structure like (4b). These assumptions allow Bordelois to use the No-Crossing Condition to account for the unacceptability of sentences like (5).

(5a). *Qué libro lo dejaste leer.
    Which book did you let him read?
b. Qué libro lo dejaste leer [NP\textsubscript{e}][NP\textsubscript{e}].

(5a) has the structure in (5b) where the lines representing the trace bindings crossover. Consequently (5) is blocked by the No-Crossing Condition.

A second argument involves the reference of the pronoun, la, in (6)-(7).

(6a). Juana invitó al hombre a sequírla.

b. Juana\textsubscript{1} invitó [S, al hombre\textsubscript{2} a sequírla\textsubscript{1}].

(7a). *Juana ordenó al hombre sequírla.

b. Juana\textsubscript{1} ordenó al hombre [S, PRO sequírla\textsubscript{1}]

[NP\textsubscript{e}].

The sentence in (6a) is acceptable because no crossing lines are involved. However in (7) this is not the case. The lines representing the bindings do crossover, and the No-Crossing Condition blocks the co-reference reading indicated in (7a).

In addition, Bordelois presents three other areas in which the No-Crossing Condition is relevant. One involves a rule changing a feature on an embedded verb and which limits the possibilities for extraction. Unfortunately the argument involved is too complex to summarize briefly here. The second area involves raising to subject and the by now familiar contrast between (8) and (9).

(8) what violin\textsubscript{1} would this sonata\textsubscript{2} be easy to play t\textsubscript{2} on t\textsubscript{1}
(9) *what sonata₁ would this violin₂ be easy to play t₁ on t₂

(8) is acceptable but (9) is not. Bordelois attributes this fact to the crossing of the bindings in (9) but not in (8). The third area involves quantifier movement in French. Bordelois argues that the contrasts involved in (10) and (11) can be accounted for by the No-Crossing Condition.

(10)a. Il lui a laissé manger tout.
   b. Il lui a tout laissé manger.
   He let him eat everything.

(11)a. Il l'a laissé manger tout.
   b. *Il l'a tout laissé manger.
   He let him eat everything.

Bordelois accounts for the contrast in (11b) and (10b) by arguing that a crossing structure is involved in (11b). She provides the diagrams in (12) and (13) to illustrate her case.

(12) 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
This argument depends on lui becoming le immediately to the left of manger in order to produce a non-nested configuration.

The No-Crossing Condition has been subject to various attacks. On the one hand some of the empirically based arguments in its favor have been undercut by subsequent analyses which account for the relevant phenomenon by other conditions in Universal Grammar. Quicoli (1976) provides an alternative account of the French quantifier phenomenon which depends crucially on Chomsky's (1973) and (1976) Specified Subject Condition (SSC). Chomsky (1977b) outlines an alternative account for the sonata-violin cases in (8)-(9).

On the one hand, the No-Crossing Condition has been attacked on the basis of its overlap with Chomsky's SSC and evidence indicating that the empirical content of the No-Crossing Condition in excess of the SSC is incorrect. The No-Crossing Condition and the SSC, reproduced as (14), overlap in the sense that both predict that (15) is ungrammatical.
(14) SSC

Given the structure

...X...[α...Y...]...X...

no rule can involve X and Y where α contains a subject distinct from Y and not controlled by X

(15) Jane promised the women PRO to visit each other.

This overlap is somewhat disconcerting insofar as it is suggestive that a generalization is being missed. However when we examine whether the SSC can be dropped from the inventory of Universal Grammar we encounter two problems. In order for a theory T₁ to replace another theory T₂ it is necessary (1) that T₁ predict the same facts at T₂, (2) that T₁ have some empirical content in excess of T₂, and (3) that some of the excess empirical content of T₁ is corroborated (cf., Lakatos 1970). The No-Crossing Condition fails to replace the SSC on either property (1) or (3). The No-Crossing Condition fails to predict the same facts as the SSC because it is incapable of excluding (16), which the SSC does.

(16) The men₁ asked Bill₂ [S₁,PRO₂ to visit each other₁].

Furthermore, Chomsky (1976) demonstrates that some of the excess empirical content of the No-Crossing Condition was not corroborated. So for example, the No-Crossing
Condition would predict that each of the sentences in (17) are ungrammatical.

(17)a. The men₁ asked Bill [S₁ whom₂ [S₁ PRO₁ to see e₂]]

b. Whom₁ did you₂ ask e₁ [S₁ what₃ [S₂ PRO₂ to read e₃]]

Sentences like (16) made it appear unlikely that the No-Crossing Condition could replace the SSC. Moreover sentences like (17) cast doubt on the correctness of the No-Crossing Condition and led Chomsky (1976) to question whether any such condition is tenable. Subsequent research on a nesting requirement, in particular Fodor (1978), has failed to respond to these basic attacks. The overall objective of this study is to frame an adequate nesting requirement, demonstrate that its excess empirical content is in fact substantially corroborated and to explain examples like (17) which apparently falsify nesting conditions.

The second chapter of this study argues that previous accounts have overlooked the fact that the nesting requirement holds over contextual features as well as bound anaphora. The chapter provides justification for believing that subcategorization rules are subject to the nesting requirement. The proposal made in Chomsky (1965) that subcategorization rules can only involve elements
which are sisters can be shown to be too strong. It is incapable of accounting for dependency relations in a non-ad hoc fashion between an element in the specifier system of a head and an element in the complement system of that head (e.g., comparatives and so...that+S constructions). It is also incapable of accounting for dependency relations between a verb and the COMP or AUX of a sentential complement; the dependency relationship between a verb and a manner adverb generated outside V; or the dependency relationship between an adjective and a post-head complement generated outside the AP itself.

In place of Chomsky's restriction, it will be proposed that for two elements A and B involved in a subcategorization rule, either A is in the domain of B or B is in the domain of A, where A is said to be in the domain of B if the first branching node dominating B also dominates A. This proposal will leave some structures ambiguous with respect to the application of subcategorization rules. It will be argued that the nesting condition disambiguates such structures. Evidence in favor of this use of the nesting condition is taken from multiple headed comparatives and comparative adjectives with complements. Some subsidiary conventions are proposed to preserve the notion of "frame" in Chomsky (1965). Such conventions are motivated by a need to retain Chomsky's explanation for the ambiguity of (18).
The third chapter presents a number of phenomena which require a nesting condition in order to approximate any level of explanatory adequacy. These cases include the interaction of WH-movement and passive operating within sentences capable of double passives; the interaction of leftward and rightward quantifier "movements" in French; cases of double adverb movement in English; and instances of double clitic "movement" in the Romance languages. All of these examples involve operations internal to a single $S^2$, and therefore outside the scope of Chomsky's conditions. As such they constitute fairly strong evidence in favor of a nesting condition of some sort.

Certain examples from the clitic systems of French and Spanish are considered which apparently falsify the nesting condition. It is suggested that the nesting condition holds at a level of structure termed surface structure by Chomsky and Lasnik (1977). Subsequent to this level of structure a restricted class of operations may apply to yield non-nested configurations. These operations, it is claimed, are a special subset of what Emonds (1976) terms "local" rules and may be expected to be the locus of much language variation. Finally it is argued that in order to provide a sufficient account of sentences like those in (19) where (19a) is ambiguous but (19b) is not, it is necessary to assume the existence of the nesting
conditions as well as a condition prohibiting optional subcategorization.

(19)a. Mary seems courteous to strangers.

b. To strangers, Mary seems courteous.

It is argued that apparently optional items are in fact obligatory items which may be realized as PRO's with arbitrary reference.

The fourth chapter considers movements and construals over domains apparently larger than a single $S^2$. Although the nesting condition does account for sentences like (20), just as the SSC does, it fails to account for

(20) $\text{John}_1$ promised the men$_2$ PRO$_1$ to see each other$_2$

(21) The men$_1$ asked John$_2$ PRO$_2$ to see each other$_1$

Because the nesting condition will not account for (21), it appears that Chomsky's conditions, or a version of them, are still required. However there is a redundancy between the conditions with respect to (20). It has often been a useful strategy to eliminate such redundancies. Rather than permit this redundancy then, I will propose that the nesting requirement holds over a single $S^2$ domain only. This refinement eliminates the apparent counter-examples like (17). Restricting the nesting condition to a single $S^2$ domain is generalized and it is proposed that all conditions involve a single "clausal" domain.

Chomsky's latest proposals in "On Binding" are

21
considered in this respect. Several revisions in Chomsky's analyses are proposed which have certain empirical and conceptual advantages. It is argued that rather than assigning abstract case in the base (for oblique case) and at surface structure (for nominative and objective case), all case is assigned cyclically in a uniform manner. A case concord condition will be proposed requiring bound elements to bear non-distinct case. These proposals allow an account of passive within the move-NP framework where the function of the passive morphology is to prevent case from being assigned to the object NP thereby both allowing and insuring its movement into subject position. Sentences where move-NP applies but where the passive morphology is absent will be blocked by the case concord condition.

The notion of α being free is also revised so that α is free if and only if no other constituent bearing case appears in the same NP or S' domain with the same referential index. This revision overcomes the difficulty noted by Chomsky of WH-elements with nominative case in COMP. For Chomsky such elements are "free" and necessitate complications in linguistic theory. With this revised notion of "free" however these complications disappear. This revision also gives a unified treatment of the Nominative Island and Opacity Condition. Under Chomsky's assumptions, the Nominative Island Condition is sensitive to case while the Opacity Condition is sensitive to the notion "subject."
Within the revision that will be proposed, these conditions will be unified and only sensitive to case. It will have the advantage of predicting the Subjacency Condition and eliminating certain recalcitrant data in Chomsky's treatment.

The fifth chapter provides an analysis of the French causative construction using the proposals advanced earlier to eliminate the problematic character of this construction with respect to linguistic theory. It will be argued that no sentential complement is involved in these structures. Rather in an analysis similar to Strozer (1977) a VP complement will be proposed for such constructions and reconciled with Chomsky's claim that VP complements do not exist in English. The "Uniqueness Principle" of Chomsky and Lasnik (1977) can be used to motivate a filter blocking VP complements in English since there is no positive evidence falsifying such a filter. In French however this filter encounters falsifying data from clitic and quantifier movement and can not be part of an adult grammar. In this way VP complements are possible in French but not in English. Certain facts about cliticization in the causative construction will be explained in terms of the nesting condition. Because the causative verb is subcategorized for an NP to its right, dative clitics must be nested in this construction. The so-called pro-PP's, y and en, need not be nested in this construction because
they are of a distinct category and do not create any structural ambiguity. Direct object clitics can be non-nested because, it is argued following Emonds (1976), they involve a local rule applying subsequent to surface structure. This analysis has the advantage of avoiding any problems with respect to Chomsky's conditions. In addition the general thrust of the analysis can be extended to account for a problem in Quicoli's (1976) analysis of apparent violations of the SSC posed by clitic movement and quantifier movement in Romance languages outside sentential complements to certain verbs. Quicoli's solution was to suggest that the subject of such sentential complements is deleted prior to these movements so that SSC is not applicable. This proposal, however, is outside the current model where movements apply prior to deletions. If the verbs permitting movements of quantifiers and clitics outside their complements have VP rather than S complements, the apparent counter-examples to the SSC dissolve and we can still maintain that movements apply prior to deletion, as in Chomsky and Lasnik (1977).

The last chapter returns to the nesting condition as a means of disambiguating certain structures. The possibility of a parameterized version of linguistic theory in which structural non-distinctness might be defined somewhat differently in various languages is considered. Up to this point it has been assumed that two categories
must be identical in order to produce structural ambiguity and trigger the nesting condition. The possibility of a language defining two different categories as systemati- cally non-distinct for the purposes of the nesting condi-
tion is considered and shown to have some empirical advantages. Evidently, the crossing of clitic pronouns and verbs in Italian produces unacceptable sentences. Simi-
larly the crossing of prepositions and nouns in English produce unacceptable sentences. These facts can be accounted for by extending the nesting condition so that it applies to them.
Chapter 2

SUBCATEGORIZATION AND THE ASSOCIATION PRINCIPLES

1. **Contextual Features for Lexical Insertion**

   Any grammar of English must be capable of expressing the fact that the verb *put* requires a following noun phrase (NP) as (1) demonstrates, while a verb like *blossom* cannot appear with a following NP as (2) demonstrates.

   (1)a. Mary put the plant on the table.

   b. *Mary put on the table.

   (2)a. The plant blossomed.

   b. *The plant blossomed a flower.

   Of course these are the traditional notions of transitive and intransitive verbs. In Chomsky (1965) it is suggested that such facts can be captured by constructing a lexicon. This lexicon pairs a set of phonological distinctive features with a set of specified syntactic features. The phrase structure rules introduce lexical and grammatical categories as well as syntactic features on them like [± V], [± N], [± COUNT], etc. Groups of the latter constitute "complex symbols." The output of the phrase structure rules is a preterminal string. A terminal string is constructed by "inserting" a lexical item for every lexical
category in the preterminal string. This insertion process obeys the following convention.

(3) If Q is a complex symbol of a preterminal string and (D,C) is a lexical entry, where C is not distinct from Q, then Q can be replaced by D (Chomsky 1965:84).

By allowing features of the form [X _____ Y], that is "contextual features," in the lexicon, Chomsky observes that it is possible to restrict the insertion of lexical items to the appropriate environment. So, for example, a transitive verb like put carries in its lexical entry the contextual feature [+ _____ NP]. This feature guarantees that put will not be placed, in accordance with the convention in (3), into a preterminal string where no object NP is present. Within Chomsky's (1965) model of the base component it is possible for contextual features to refer to category symbols like S, NP, V, etc., or it is possible for them to refer to syntactic features like [± ANIMATE], [± COUNT], etc. Contextual features making reference to category symbols Chomsky terms strict subcategorization rules while contextual features making use of syntactic features are termed "selectional rules."

2. On Restricting the Interpretation of Contextual Features

I want to explore in more detail the formal nature of such contextual features. Formally it is possible to
construct a contextual feature expressing a categorization on a verb in terms of the determiner of the subject NP. In other words it is possible to write an entry like (4).

(4) elapse, +V, +[+DEF N ___ ]

However, there apparently is no such categorization in natural language. It is desirable therefore to limit the formalism of contextual features so as to restrict as narrowly as possible the class of grammars compatible with natural language. Recognizing facts like this, Chomsky (1965) places the following condition on the form of such contextual features.

(5) Each such rule of the form A → Complex Symbol/α ___ β where α Aβ is a σ, where furthermore α is the category symbol that appears on the left in the rule σ → ... A ... that introduces A.

This restriction requires contextual features to be "strictly local" in the sense of Chomsky (1965): that is, the contextual features and the lexical item in question must be immediately dominated by the same node. In the notation of (5) α and β must bear the structural relationship to A exhibited in (6), that is, they must be "sisters" to A.
(6) \[ \alpha \overset{\sigma}{\longrightarrow} \alpha \overset{A}{\longrightarrow} \beta \]

Given this restriction on the form of contextual features, the contextual feature in (4) can no longer be given expression. This is due to the fact that the base would yield a structure like (7).

(7) \[ S \overset{NP}{\longrightarrow} NP \overset{DET \ N^1}{\longrightarrow} N \overset{\text{an}}{\longrightarrow} Y \overset{\text{hour}}{\longrightarrow} \overset{\text{elapsed}}{\longrightarrow} VP \]

In this structure V and DET are not immediately dominated by the same node. For this reason the restriction in (4) is not strictly local and is precluded by the condition in (5).

The condition in (5) holds over all contextual features, strict subcategorization rules and selection rules as well. Chomsky (1965:112) writes:

(8) Moreover the subcategorization rules that are context-sensitive are in effect strictly local transformational rules. . . . Among the context sensitive subcategorization rules we have distinguished two important subtypes, namely strict subcategorization rules that subcategorize a lexical element in terms of the frames of category symbols in which it appears and selection rules which subcategorize a lexical category in terms of syntactic features that appear in specified positions in the sentence.
This passage makes clear the fact that the condition in (5) constrains both types of contextual features.

Chomsky further restricts the class of selectional rules by requiring that such rules involve heads of phrases. So, for example, he writes:

(9) ... the selectional features must involve the lexical categories that are the heads of grammatically related phrases, in the sense outlined earlier (Chomsky 1965:122).

We will return to the notion grammatically related phrases shortly.

The condition in (5) serves to constrain the class of grammars defined by the general theory of the base component articulated in Chomsky (1965). It is also argued that condition (5), a condition on the form of contextual features, constrains in a significant way the class of possible functional notions definable in linguistic theory.

In the model of the base component presented in Chomsky (1965) functional notions like "subject of," or "object of" (that is notions often referred to as "grammatical relations"), are not primitives and are not introduced by the phrase structure rules. Rather they are interpreted from phrase markers and do not interact with syntactic processes. Of course, the motivation behind this general schema is once again to restrict the class of grammars as tightly as possible.
In Chomsky (1965) functional notions are defined by ordered pairs. For example, the grammar in (10) will generate the tree in (11).

(10) \( A \to Y - B - Z \)

(11) \[
\begin{array}{c}
A \\
Y \\
B \\
Z
\end{array}
\]

In this hypothesized grammar it is possible to define functional notions like \([A,B]\), \([A,Y]\), and \([A,Z]\). In more concrete terms, this conception permits a definition of the notion "subject of a sentence" through the statement in (12).

(12) \([S,NP] = \text{subject of a sentence}\)

(11) identifies as subject of a sentence that \(NP\) immediately dominated by \(S\).

It is also possible in this formalization to construct composite pairs. Given a tree structure like (13) a relation between \(B\) and \(E\) can be expressed as (14a) and a relation between \(B\) and \(F\) can be expressed as (14b).

(13) \[
\begin{array}{c}
G \\
A \\
C \\
B \\
E \\
F \\
C
\end{array}
\]

(14a) \([G, [A,B]], E]\)

b. \([G, [A,B]], [E,F]\]

Expressing functional notions in this way raises a
problem that Chomsky recognizes as a serious one: it is possible to define "pseudo-relations." For example, Chomsky notes that by using composite pairs one can define a relation between the subject of the sentence and the object of the verb, as in (15).

(15) [[S,NP], [VP,NP]]

However the relation between subject and object does not appear in natural language. Facts of this nature present us with the problem of constraining our definition of functional notions in such a way that a pseudo-relation like (15) is precluded by linguistic theory.

Chomsky attempts to resolve this problem by proposing that functional relations may be defined only between a lexical element and some second element satisfying a contextual features. He writes:

(16) We shall say that a selectional rule ... defines a selectional relation between two positions in a sentence. ... Such selectional relations determine grammatical relations, in one of the senses of this traditional term. ... The same notion of grammatical relation could, in fact, have been defined in terms of major categories but the definition in terms of selectional relations seems somewhat more natural. (Chomsky 1965: 113-114)

This passage seems to suggest that relations like (15), while formally expressable, are not possible. This is due to the fact that no selectional restriction can hold between a subject and object in a structure like (17).
(17) S
    NP  VP
      Jane  saw  NP
             Mary

Jane and Mary are not immediately dominated by the same node and any restriction between them would violate the requirement that contextual features be strictly local. Constraining the class of contextual features is thus doubly important in that it also restricts the class of possible grammatical relations.

3. Certain Problems with the Strict Locality Condition

There exists a certain number of problems with the claim that all contextual features are strictly local. I will refer to this claim as the "Strict Locality Condition." This section introduces several of these problems and argues in favor of a slightly looser formulation, the "Broadly Local Condition," which makes crucial use of the notion "domain" rather than sister constituency.

3.1. Selecting COMP and Tense and the Broadly Local Condition

It is well known that certain verbs select the form of complementizer (COMP) and even tense of a sentential complement (cf., Bresnan 1970). The sentences in (18)-(19) are illustrative of these restrictions.
(18)a. Mary wondered which article to write first.
    b. *Which article did Mary wonder to write first.
    c. *Mary wondered to write which article first.

(19)a. Mary knows that only a theory disproves another theory.
    b. *Mary knows that only a theory disprove another theory.
    c. Mary required that he disprove some aspect of structural anthropology.
    d. *Mary required that he disproved some aspect of structural anthropology.

In (18) we find that the verb wonder requires a WH element in the COMP of the embedded sentence. In (19) we find that know and require differ in the kind of tense they permit a sentential complement to exhibit: know selects a finite complement while require selects a sentence with the subjunctive (i.e., no overt) tense. Now if the relevant phrase structure rules introducing COMP and TENSE are (20), as is generally assumed, there is some difficulty expressing these subcategorizations.

(20)a. $\$ \rightarrow \text{COMP} - S \ldots$
    b. $S \rightarrow \text{NP} - \text{TENSE} - \text{VP} \ldots$

The verb wonder and the COMP will not be immediately dominated by the same node, nor will the verbs know and
require on the one hand and the TENSE of their sentential complements on the other hand be strictly local. This fact means that within the present condition on contextual features and given the rewriting rules in (20), the restrictions on COMP and TENSE are not expressable.

Grimshaw (1979) has presented a proposal for complement selection which might be thought to resolve the difficulty just observed, at least for the selection of COMP. Grimshaw observes that exclamatives and questions have similar syntactic forms but are not selected by the same verbs. So for example even though exclamatives and questions can undergo WH-movement, as the sentences in (21) illustrate, they do not necessarily both appear as complements to the same verbs, as demonstrated by the sentences in (22)-(23).

(21)a. How tall John is!
      b. How tall is John?
(22)a. John found out what a fool Bill was.
      b. *John wondered what a fool Bill was.
      c. It's amazing what a fool Bill is.
(23)a. John found out whether Bill had left.
      b. John wondered whether Bill had left.
      c. *It's amazing whether Bill had left.

From these facts Grimshaw concludes that verbs have a semantic frame as well as a syntactic frame. The verb
wonder, for instance, requires the entry in (24) whereas a verb like find out has the entry in (25) and amazing has the entry in (26).

(24) [____ Q]

(25) [____{Q},\{E\}]

(26) [____ E]

In (24)-(26) Q indicates a question while E indicates an exclamative.

Grimshaw goes on to suggest that given (24) it is sufficient to specify the syntactic frame (27) for wonder.

(27) [____ S\text{\textsuperscript{1}}]

Grimshaw's reliance on (27) seems to stem in part from her acceptance of the Strict Locality Condition and from her belief that (24) will provide the question word in COMP to the exclusion of a that-clause. If this is true it would remove the problematic character of at least COMP selection for the Strict Locality Condition. However, I believe that (27) is not sufficient, even if one is willing to accept, as I am, the legitimacy of semantic frames like (24)-(26).

Some lexical statement will be required to mention a WH element in COMP independent of restrictions like (24). To see this point consider echo questions like (28).
(28) John is how tall?

Constructions of this sort appear to naturally class together semantically with other questions employing WH-words and yet they do not satisfy the selection requirement of wonder as (29a) illustrates.

(29)a. *Bill wondered John is how tall?
   b. Bill wondered whether John is how tall?

On the other hand, when a COMP containing a WH-word is present the sentence is acceptable as an echo question, as (29b) shows.

A similar phenomenon can be observed with respect to exclamatives. Not all exclamatives need undergo WH-movement. So for example the sentences in (30) are well-formed exclamatives.

(30)a. John is such a fool!
   b. John was so angry!
   c. You've got so much money!

Some verbs selecting exclamatives can select exclamatives of the form (30). So for example we can observe (31).

(31)a. It's amazing that John is such a fool.
   b. I can't believe that John is so angry.
   c. I can't believe that you've got so much money.

Of course the verbs in (31) can also co-occur with
exclamatives using WH-words. Yet other verbs selecting exclamatives can only co-occur with exclamatives having undergone WH-movement. The sentences in (32)-(33) are indicative of the existence of such verbs.

(32)a. *You can't imagine that John is such a fool.
     b. You can't imagine what a fool John is.

(33)a. *I don't give a damn that you've got so much money.
     b. I don't give a damn how much money you've got.

Verbs like give a damn evidently have a semantic frame like (34) and yet require a syntactic frame like (35).

(34) [_____ E]
(35) [_____ WH S]

Similarly, wonder would need a syntactic frame like (35) in order to account for the facts in (29).

Of course one could begin to subdivide questions or exclamatives into subclasses, E₁ and E₂ for example, where E₁ has a WH-word in COMP and E₂ does not. However, such a proposal would simply be a notational variant of the proposal sketched above unless it was shown to have some excess empirical content.

It thus appears that when a wider range of exclamatives and questions are considered there is still a
necessity of subcategorizing verbs for WH-words in COMP, even after we admit the notion of semantic frames proposed in Grimshaw (1979).

Another source of problems for the Strict Locality Condition is restrictions that a verb places on the subject of a sentence. The verb *like for example requires a [+ANIMATE] subject noun, as the contrast in (36) shows, despite the fact that the verb and the noun itself are not immediately dominated by the same node.

(36)a. Jane likes Mary.
   b. *The rock likes Mary.

We can begin to overcome these difficulties if we introduce the notions of "head of phrase" and "maximal projection of a head of phrase."

Let us define "head of phrase" as in Emonds (1978), that is as (37).

(37)a. Definition:
   In the bar or prime notation let \( B^n = B \)
   with \( n \) bars or primes.

   b. Definition:
   The daughter \( B^j \) of \( B^n \) that has the fewest bars (primes) is the head of \( B^n \).

Using these definitions we can define the maximal projection of \( B \) as in (38).
(38)a. Definition:

$B^n$ is a projection of that $B$ such that $B = B_1, B_2, \ldots, B_k = B^n$ in which each $B_i$ is the head of $B_{i+1}$.

b. Definition:

$B^n$ is the maximal projection ($B^{\text{max}}$) of $B$ if and only if $B^n$ is not the head of any other $B^j$.

In order to see an example of a "maximal projection" consider the phrase the enemy's destruction of the city. This phrase has a structure like (39).

(39)

Given the definitions in (37)-(38) it is the case that only destruction is the head of phrase of $N^2$, and the maximal projection of destruction is $N^2$. The formative enemy cannot be identified as the head of the highest $N^2$ because $N^2$ and enemy are separated by a node of the same category and same number of bars.

40
Utilizing the notion head of phase and maximal projection of a head permits the expression of selection restrictions between the subject and the verb of a sentence. In structure (40) the maximal projection of both the head of the subject noun phrase and the head of the verb phrase are circled.

(40)

\[
\begin{array}{c}
S \\
\text{DET} \\
\text{John's mother likes} \\
\text{Mary}
\end{array}
\]

The maximal projection of these nodes are strictly local in that they are immediately dominated by the same node, S. For this reason we might consider reframing the Strict Locality Condition so that it was defined on the maximal projection of heads as being immediately dominated by the same node.

However, introducing "maximal projection" in this way is not able to resolve the problems that we have noted with the Strict Locality Condition. The selection restrictions between the verb and its object provide a simple example. The maximal projection of the verb, i.e., \( V^2 \), will dominate the maximal projection of the object, \( N^2 \), as one can observe in (41).
In this case it is not possible to say that the maximal projection of the verb and of the object are immediately dominated by the same node. Consequently, there is still a problem with the Strict Locality Condition.

Moreover, the use of "maximal projection" would leave verb-COMP and verb-TENSE dependencies unexpressable within the framework of the Strict Locality Condition. The maximal projection of WH in COMP or subjunctive tense will not be S$^1$. As a result we will still be unable to express such dependencies.

A notion that would allow us to surmount these difficulties is "domain," defined in (42).$^1$

(42) A is in the domain of B if and only if the first branching node dominating B dominates A as well.

This notion is sufficient to account for the verb-object restrictions: in (41) the first branching node, i.e., S, dominating the maximal projection of V, i.e., $V^2$, also
dominates the object. However, (42) alone without the additional theoretical constructs of head of phrase and maximal projection of a head is insufficient to account for the difficulties encountered with the Strict Locality Condition. (42) would allow the statement of verb-object or verb-COMP dependencies but would not permit the formulation of restrictions between a subject N and a verb. In (40) for example, the first branching node dominating the N or the V does not dominate the other element. For these reasons the definitions in (37)-(38) as well as (42) are necessary to properly express contextual features. Recognizing this, we need to reframe (5), the Strict Locality Condition, as (43). I will term (43) the "Broadly Local Condition."

(43) The Broadly Local Condition

Each rule of the form A → Complex Symbol/
α _____ β requires that either \( \{\alpha_{\text{max}}, \beta_{\text{max}}\} \)
is in the domain of A, or A is in the domain of \( \{\alpha_{\text{max}}, \beta_{\text{max}}\} \).

This permits us to account for the three major configurations that we have been examining: verb-COMP or verb-TENSE dependencies; verb-object dependencies; and subject-verb dependencies. The maximal projection of the COMP or TENSE will be in the domain of V. As for verb-object dependencies, in (41) the maximal projection of the object is in
the domain of V. Similarly in the case of subject-verb dependencies, the verb will be in the domain of the maximal projection of the subject, as can be observed in (40).

3.2. **Comparative Clauses**

Within the traditional analysis of comparatives there is no particular problem for the Strict Locality Condition. The relationship between more and than+S\(^1\) is accounted for by positing that both originate in the specifier system of the phrase they modify. So for example, Bowers (1968) and Bresnan (1973) would assign a structure something like (45) to the sentence in (44).

(44) Mary has more marbles than Susan has.

(45)

The motivation behind postulating the sentential complement originates in the specifier system is precisely to make the expression of the subcategorization facts easier. In this sense the analysis in (45) follows the general line of
Smith (1961) where it is argued that restrictive relative clauses also originate within the specifier system of a noun phrase in order to capture some generalizations about the co-occurrence of restrictive relatives and certain determiners.

There are however a number of problems with generating such clauses within a specifier system. First, it is necessary to postulate an obligatory extraposition rule. Such a rule must be obligatory because sentences like (46) are clearly ill-formed.

(46) *Mary has more than Susan has marbles.

Within the general theory of grammar postulated in Chomsky and Lasnik (1977) all transformational rules are optional. Cases where rules apply in an apparently obligatory fashion are actually due to a surface output filter blocking the case where the rule has not applied. These surface output filters are local in the sense that they involve no internal variables. In order to make the comparative extraposition rule obligatory without unduly increasing the class of grammars by permitting obligatory rules, we would need to construct a universal output filter blocking the sentential clause in the specifier system. Emonds (1976) proposes such a filter. It is reproduced as (47).

(47) Surface Recursion Restriction

Given a surface configuration of the form
[\_i \max \ldots A \ldots H_i \ldots], if the base rules permit right sisters \_i \max to \_i, then A \not \equiv XSY, A \not \equiv XPPY where PP dominates a lexical preposition, and A \not \equiv WAZ where W and Z \not \equiv \emptyset. In such cases we say that A does not exhibit "free recursion."

This surface filter will block (46), that is, instances where extrapolation of sentential complements has not operated. Moreover, notice that this filter will also block a sentence like (48) to guarantee the acceptable (49) under the assumption that than+NP constitutes a prepositional phrase (cf., Hankamer 1973 and Hendrick 1978) for argumentation in favor of this assumption.

(48) *Mary has more than Susan marbles.

(49) Mary has more marbles than Susan.

In general it seems that something along the lines of the surface recursion restriction is correct. The difficulty with (47) however is that it could apply to restrict the postulation of phrase structure rules as well if it were not for the apparent necessity of generating comparative clauses within the specifier system of a noun phrase. That is to say, a more restrictive version of grammar could be obtained if free recursion to the left of a head was blocked not only at the surface, but in the phrase structure rules as well, assuming of course that the

46

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
co-occurrence restrictions can be treated in a constrained fashion. The appropriate generalization might be something like (50).

(50) No rule of grammar may define a configuration of the form \([H_{i_{1}}^{\text{max}} \ldots A \ldots H_{i_{2}} \ldots]\) if there is another rule of grammar permitting right sisters \(H_{k}^{\text{max}}\) to \(H_{i}\), where \(A = \text{XSY}, A = \text{XPPY}\) where PP dominates a lexical preposition, or \(A = \text{WAZ}\) where \(W\) and \(Z \neq \emptyset\).

This generalization would require generating sentences like (44) and (49) as base structures.

A second somewhat different argument against generating comparative clauses within the determiner system of a phrase appears in Andrews (1975). Andrews observes that there are multiple-headed comparatives like (51).

(51) As fair a woman and as foul a man as I have ever seen together are coming towards us.

The underlying structure for (51) is something similar to (52).

(52) As much fair a woman and as much foul a man as \([S_{I} \text{have ever seen} [NP_{x} \text{much fair a woman}] \text{and} [NP_{x} \text{much foul a man}] \text{together}]\) are coming toward us.

Andrews considers the possibility that (51) is derived via
right node raising where each of the conjoined NP's has a comparative clause. However he dismisses this possibility because the clause *I have ever seen x much fair a woman together would need to underlie one of the conjuncts despite the fact that it is clearly ungrammatical and together could not co-occur with a singular NP. Andrews (building on Vergnaud 1974:82-83) goes on to argue that it is not possible to counter his objection by claiming that only the surface structure is important and that such deep structure anomalies can be discounted because the comparative clause would need to bind a trace in each conjunct and semantic interpretation would proceed as if the movement had not occurred.

I will assume that if there were some way to express co-occurrence restrictions in a restricted way, then on syntactic grounds it would be preferable to generate the sentential complements of comparatives in place rather than in the specifier system. It is important to note that the same argumentation applies to subcategorization facts pertaining to as, too, so, etc. All these elements govern the choice of post-head complements.

(53)a. Mary is as tall as Jane is.
   b. *Mary is as tall (that) Jane is.
   c. *Mary is tall as Jane is.
   d. *Mary is as tall.
(54)a. Mary is too smart (for anyone) to trick.
   b. *Mary is smart (for anyone) to trick.
   c. *Mary is too smart to me.

(55)a. Mary is so bored that she has finally read Mildred Pierce.
   b. *Mary is so excited about Mildred Pierce.
   c. *Mary is bored that she has finally read Mildred Pierce.

These sentences demonstrate that the specifier of the adjective restricts the choice of complement to the right of the adjective.²

The relation between a quantifier and its modifying phrase therefore also presents some difficulty for the Strict Locality Principle. For example, consider the sentence in (56).

(56) A more elegant hypothesis than the one proposed by Jane would be hard to find.

The structure of the relevant phrase would be (57).

```
(57)
```

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
While (42) is incompatible with the Strict Locality Condition, it is perfectly compatible with the broadly local formulation. That is to say, the quantifier is in the domain of $S^1$ inasmuch as the node immediately dominating $S^1$ dominates the quantifier as well.$^3$

A somewhat different demonstration of the need for the Broadly Local Convention as opposed to the strictly local one can be constructed from facts involving the multiple-headed quantifier phrases. The sentence in (58) is a case in point.

(58) *Tristram Shandy* has so much more complexity than *Pilgrim's Progress* does that it is difficult to compare the two.

The relevant phrase for the purposes of this discussion has the structure indicated in (59). The co-occurrence restrictions between so ... that it is difficult to compare the two and more ... than *Pilgrim's Progress* does cannot be stated in terms of the Strict Locality Condition since in each case the two constituents involved are not immediately dominated by the same node.$^4$ The broadly local formulation however is capable of capturing these facts: both quantifiers are in the domain of the sentential complements in (59).

3.3. **Manner Adverbs**

Another instance providing corroboration for the
so more much complexity than Pilgrim's Progress does that it is difficult to compare the two
broadly local formulation comes from the categorization that manner adverbs place on certain verbs.

Emonds (1976) observes that there exists a small class of verbs which require manner adverbs. For example in (60) the verb word requires a manner adverb as the unacceptability of (60b) indicates.

(60)a. We worded the letter carefully.
    b. *We worded the letter.

Conversely, Chomsky (1965) notes that other verbs are unable to co-occur with manner adverbs. For example in (61) we find that resemble cannot co-occur with such adverbs.

(61)a. Mary resembles Louise.
    b. *Mary resembles Louise \{carefully\}
        \{terribly\}

Within the Strict Locality Convention manner adverbs would, thus, need to be generated as daughters to $V^1$ or the lowest VP, and indeed Chomsky (1965) proposes the following phrase structure rule.

(62) \[ VP \rightarrow V - (NP) - (PP) - (PP) - (\text{Man(adv)}) \]

Chomsky (1965) suggests that analyzing manner adverbs as daughters of $V^1$ provides a structural explanation for why manner adverbs fail to prepose although other adverbs do, as exemplified in (63)-(64) below.\footnote{5}

(63) *Carefully Mary worded her reply.
(64)a. Unfortunately Jane took the wrong off-ramp.
   b. Frequently Jill buys records there.

However there is evidence that manner adverbs are generated outside V\textsuperscript{1}, just like the other adverbs. Lakoff and Ross (1966) propose a transformation like (65) to deal with the so-called do so construction.

(65) \[ X - VP - Y - VP - Z \Rightarrow 1 - 2 - 3 - \text{do so} - 5 \]

where 2 = 4 and 2 begins with non-stative verbs.

Whether or not a transformational rule like (65) is the appropriate formal device to deal with the presence of do so is irrelevant here; what is important is that do so substitutes for a verb phrase of some level. Note that if we apply (65) to the sentence in (66a) we are able to derive (66b).

(66)a. Ideally, students should word their answers carefully, but in general they word their answers carelessly.
   b. Ideally students should word their answers carefully but in general they do so carelessly.

If the adverb was inside the lowest VP, we would not expect to be able to generate the sentence in (66b). Its acceptability argues against generating the manner adverb inside the lowest VP.
Before proceeding to present the advantages of analyzing manner adverbs outside the $V^1$, it is worth noting that in Jackendoff (1978) sentences like (66b) are judged unacceptable, and from this judgement it is concluded that manner adverbs are within $V^1$. This conclusion is of some importance to Jackendoff since he wants to argue that $X^1$ is the only level for strictly subcategorized arguments. The relevant fact to keep in mind when analyzing the dialect in which (66) is acceptable is that the possibility of a grammar in which manner adverbs are generated under $V^1$ is not denied here and Jackendoff's judgements could be accounted for without sacrificing the argument based on the dialect which does generate (66). However, it is also worth noting that Jackendoff's judgement on sentences like (66) can be integrated into a grammar where manner adverbs are generated outside $V^1$. Although do so replaces a constituent, it does not follow that the inability to substitute do so is necessarily a consequence of non-constituency. Numerous alternatives suggest themselves: (a) do so is subject to a restriction prohibiting it from co-occurring with a manner adverb in Jackendoff's dialect, or (b) do so can only substitute for $V^2$ in Jackendoff's dialect. Either of these alternatives explains why sentences like (66b) are unacceptable for Jackendoff without making the assumption that manner adverbs are generated inside $V^1$. Moreover, there are several advantages to
generating these adverbs outside $V^1$.

One advantage is that there is a measure of economy in the statement of subcategorizations that can be gained if manner adverbs are generated outside $V^1$. Recall the convention adopted by Chomsky (1965) whereby subcategorization rules specify only the environments in which an item can appear and a convention redundantly specifies the item negatively for every other environment (cf., Section 4). If manner adverbs are generated as daughters to $V^1$, the great majority of verbs need to be specified as optionally taking a manner adverb. Unless they are so specified, the convention above guarantees that manner adverbs may not appear. If, however, manner adverbs are generated outside $V^1$ then their optional status in the great majority of cases would be captured as a consequence of linguistic theory. The convention mentioned above predicts that the great majority of items outside $V^1$ should be optional, because to subcategorize obligatorily for such an item is more costly. The analysis which assumes manner adverbs to be generated outside $V^1$ then predicts that few verbs obligatorily require the presence or absence of a manner adverb. This prediction corresponds nicely to the observable facts.

A second advantage of assuming that manner adverbs are generated outside $V^1$ relates to the existence of a separate category ADVERB (ADV) and the formulation of a
rule to derive ADV's from A's. There is a theoretical premium to categorizing adverbs as members of A. It is traditionally acknowledged that adverbs and adjectives are related by a productive morphological process of suffixing ly. Generating manner adverbs outside V₁ facilitates the writing of the ly suffixing rule.

Of the positions in which A's and ADV's can appear we find that A's can only appear as left sisters of N or right sisters of V. Adverbs on the other hand appear as left sisters to A and V, and right sisters to V^n and (following Dougherty 1971) N^n where n ≥ 1. For example, we find that the (a) structures below are acceptable, while the (b) structures are not.

(67)a.

```
NP
  /\  N^2
 /   \    N^1
DET A   N
\     /  \      the simple question
       /    \\
      N
```

(b.

```
NP
  /\  N^2
 /   \    N^1
DET A   N
\     /  \      the simply question
       /    \\
      N
```

56
(68)a.  
\[ \text{the women simultaneously} \]

b.  
\[ \text{the women simultaneous} \]

(69)a.  
\[ \text{simply saw Mary} \]

b.  
\[ \text{simple saw Mary} \]

(70)a.  
\[ \{\text{simply, utterly}\} \text{ boring} \]

b.  
\[ \{\text{simple, utter}\} \text{ boring} \]
We can express these distributional facts by assuming that the phrase structure rules generate A's in the positions above and that there is no category ADV. We will further postulate a rule like (71c) which collapses the rules in (71a) and (71b).

\[(\text{71a})\quad A^1 - [+V] \Rightarrow 1 + 1y - 2\]
\[\text{b. } X^{l+1} - A^1 \Rightarrow 1 - 2 + 1y\]
\[\text{c. } (X^{l+1}) - A^1 - ( [+V]) \Rightarrow 1 - 2 + 1y - 3\]

Unlike other attempts to derive adverbs from adjectives in a higher predicate, the analysis here has no difficulty in expressing subcategorization differences between adjectival and adverbial forms.

Jackendoff (1977) objects to the identification of A and ADV on the grounds that such a move leaves unexplained the fact that ADV's have no post-head complements. The fact that the \textit{ly} suffix is incompatible with complements to an adjective can be accounted for in terms of the lexical transformations proposed by Roeper and Siegel (1978). Define as a verb requires a following object, however, when \textit{able} is suffixed to define, no following object is possible. Roeper and Siegel propose to treat this kind of phenomenon with a transformation operating to delete a portion of the subcategorization of define when \textit{able} is present. A similar treatment could be given to adverbs where \textit{ly} causes the deletion of other subcategorized
material. This would certainly be observationally ade-
quate. Below I will suggest a second possibility which makes the complementary distribution between ly and comple-
ments to the adjective fall out from principles of linguis-
tic theory.

This analysis predicts that adjectives not exhibiting ly but to the right of V are sisters of V. For example, we predict that the adjectives in the following sentences, because they do not exhibit ly, are not sisters to V1.

(72)a. Mary ate the meat raw.
    b. Sermons make us nervous.
    c. They beat the students bloody.
    d. Jane hit Bill hard.

Evidence corroborating this prediction is presented in Green (1974). Among other things, Green points out that the do so test discussed above indicates that the adjectives in (72) are within the lowest VP. In the sentences below, the (a) but not the (b) sentences are acceptable.

(73)a. Mary ate the meat raw but we wouldn't have done so.
    b. *Mary ate the meat raw but we would have done so cooked.

(74)a. Sermons make us nervous but I never thought graveyards would do so too.
b. *Sermons make us nervous but I never thought graveyards would do so scared too.

(75)a. They beat the students bloody but John stopped short of doing so.

b. *They beat the students bloody but they stopped short of doing so senseless.

(76)a. Jane hit Bill hard and enjoyed doing so.

b. *Jane hit Bill lightly but would have enjoyed doing so hard.

The (b) sentences indicate that the A must be replaced by do so. Since do so only replaces VP's, it follows that analyzing the A's in these sentences as being sisters to V is consistent with the data in (73)-(76).

I conclude from the foregoing discussion that it is advantageous to analyze manner adverbs as outside V

1 at least for some grammars of English speakers and most probably for all. In such a case the selection of manner adverbs poses problems for the Strict Locality Condition. However, within the Broadly Local Convention no such problem arises inasmuch as the verb will always be in the domain of the manner adverb.

3.4. Complements to Adjectives

Yet another case where the Strict Locality Condition seems overly restrictive and where the Broadly Local Convention is required is the relationship between an
adjective and its complement. Consider the sentences below.

(77)a. It was arrogant for Mary to see the prisoners.
   b. Mary was arrogant with the prisoners.
   c. *Mary was arrogant that the prisoners were tortured.
   d. Mary was unhappy \{ with her job \\
         to see the prisoners \\
         that the prisoners were tortured \}

(78)a. Mary is blind.
   b. *Mary is blind that she was robbed.
   c. John is blonde.
   d. *John is blonde with his friends.

The sentences in (77) seem to indicate that the adjective arrogant can co-occur with a prepositional phrase or a tenseless sentential complement. It may not co-occur with a tensed sentential complement. The adjective unhappy, however, may co-occur with a prepositional phrase and either a tensed or tenseless sentential complement. Other adjectives are more restricted. Blind, for example, may not co-occur with a sentential complement and blonde may not co-occur with a prepositional phrase complement. In this way there are non-trivial co-occurrence restrictions on adjectives that need specifying somewhere in the grammar.

61
The fact that arrogant selects a prepositional phrase would argue, given the Strictly Local Convention, that the PP was dominated by A¹ and the traditional analysis of adjective phrases is that such complements are immediately dominated by A¹. However, in Hendrick (1978a) I provide several arguments to suggest that adjectives have no post-head complements in the base. I summarize briefly some of those arguments.

The arguments are based on facts of distribution as well as movement. First, of all the positions that adjectives (and adverbs) occur, they have complements only in the predicate adjective position. At first glance one may suppose that this fact is a reflex of the surface recursion restriction reproduced in (47) above. However, this restriction permits non-lexical prepositional phrases to appear in pre-head position as in (79).

(79) The man from Chicago's hat

No adjective however shows such prepositional phrase complements prenominally. A second argument is that in the predicate adjective position, we do not observe an adjective with a full set of complements co-occurring with complements to the verb. That is, we do not find structures like (80) where both the adjective and verb have sentential complements. A third distributional fact of some importance is that the head of the adjective can be
(80) 

NP  

S  

VP  

V  

AP  

A  

Spl  

Jane felt happy that she got a car that she found a job
separated from its complement by an adverbial phrase of one sort or another. In this respect the sentences in (81)-(82) are of interest.

(81)a. John felt uncomfortable from the beginning with the direction of the investigation.

b. Jane was suspicious in the end that Jill was guilty.

c. Mary appeared sympathetic at various points with the prisoners.

(82)a. *John described his discomfort from the beginning with the direction of the investigation.

b. *?There were some suspicions in the end that Jill was guilty.

c. *Mary expressed sympathy at various points with the prisoners.

The sentences in (81) show that adjectives permit an adverbial phrase to intervene freely between it and its complement. Nouns however show a much greater variability with which they permit an intervening adverbial phrase, suggesting that the somewhat ill understood rule of extraposition from an NP is distinct from whatever accounts for (81).

In light of these facts, I have proposed that adjectives have no post-head complements in the base. This means that in order to account for the sentences in
(83) we require an adjunction rule creating a constituent of the adjective and its complement.

(83)a. How disappointed that he couldn't go was John.

b. How unhappy with the bank was Susan.

The appropriate rule involved in this process may be (84).

\[(84) \quad U_1 - A' - \{PP\} - U_2 \Rightarrow 1 - 2 + 3 - 4\]
\[1 - 2 - 3 - 4\]

This rule is a local adjunction operation creating structures like (85) optionally.\(^6\)

\[(85)\]

I am assuming (84) works on the principle of "Chomsky"-adjunction.

Such an account of the relation between an adjective and its complements would have the ability to explain the distributional facts noted above. We no longer expect pre-nominal adjectives to have any type of post-head complement for the simple reason that there is no \(S^1\) or PP immediately to the right of A for (84) to operate on. At the same time we can account for the non-existence of a structure like (80) by noting that \(V^1\) can only contain one
sentential complement. In this way the verbal and adjectival complements are mutually exclusive since they occupy the same position in deep structure. The behavior of the adjectival complements with respect to the adverb can easily be explained. The structure could be derived via the PP-"intraposition" rule discussed in Emonds (1976: ch. 5).

Evidence in favor of the rule in (84) can be observed in (86)-(87).

(86)a. Louise {felt} angry at the restaurant about {was }
      the bill.

   b. How angry at the restaurant
      {did Louise feel}
      {was Louise }

   c. *How angry at the restaurant about the bill
      {was Louise }
      {did Louise feel}

(87)a. Mary became angry with Martha that she hadn't paid the bill.

   b. ?How angry with Martha did Mary become that
      she hadn't paid the bill.

   c. *How angry with Martha that she hadn't paid
      the bill did Mary become?

(86) and (87) do not appear to be acceptable. That is to say, for the purposes of fronting rules, the adjective can form a constituent with at most one other phrase. This
follows as a direct consequence of the local character of (84) but is anomalous within the traditional phrase structure account. 7

Moreover, the hypothesis that adjectives have no post-head complements allows us to explain the fact that adverbs with ly exhibit no complements as well, if we make some other minimal assumptions about adjectives. As noted in the previous section, Jackendoff (1977) observes that adverbs cannot select post-head complements even if a morphologically related adjective can select such complements. In this respect he compares phrases like eager to leave with eagerly to leave. Of course if ADV's simply replaced adjectives and there was a phrase structure rule generating PP's and S's as daughters to A^1, we would expect to see phrases like eagerly to leave. Moreover, simply assuming that any complements to adjectives are generated outside the adjective phrase will not in and of itself explain why we do not find sequences like eagerly to leave, where eagerly and to leave and each daughters to, say, V^2. Suppose however that ADV's are created as described in the last section so that eagerly, has the structure in (88).

(88) \[
\begin{array}{c}
A^1 \\
\mid \\
A \\
\mid \\
A \\
\mid \\
eager
\end{array}
\]
Suppose in addition that adjectives are cyclic nodes (following a suggestion in Baltin 1979). If we assume that contextual features obey the Subjacency Condition, it will be impossible for eager to be in a relation with a complement outside A₁ without violating that condition. This approach, if correct, would have an advantage over the possible lexical transformation discussed in the previous section. The lexical transformation would presumably be language-particular whereas the explanation just outlined would make the facts at hand fall out from general principles of linguistic theory.

If complements to adjectives are generated outside the adjective phrase proper, then the co-occurrence restrictions associated with an adjective will be problematic for the Strict Locality Convention. However, the Broadly Local Convention is consistent with these facts. A prepositional or sentential complement generated at some level of the verb phrase will always have the adjective in its domain. So for example, the sentence in (89) will have the structure of either (90a) or (90b). In either case, the PP has the adjective in its domain.⁸

(89) Jane is happy with her job.

4. Presentation of the Association Principles

In this section the ramifications of permitting contextual features that are not strictly local are
discussed. Certainly it cannot be the case that there are no restrictions at all on contextual features.\(^9\) In order to restrict the class of grammars permitted by linguistic theory and thus approximate the goal of explaining how a child learns a language in the first place, we want to restrict contextual features as narrowly as possible, just as we want to restrict other rules of grammar.

4.1. **Narrowing the Broadly Local Interpretation of Contextual Features**

As suggested earlier, one possible way to begin this task of constraining contextual features is to require that one of the two elements involved be in the domain of the other. However, this more lenient restriction is not without some difficulty. Within the framework of Chomsky
(1965) the association between a head and its argument (i.e., the relation between a head and its contextual feature) is non-ambiguous.

Consider the ambiguous sentence in (91), noted in Chomsky (1965).

(91) John decided on the boat.

Replacing the Strictly Locality Condition with the Broadly Local Convention permits us to assign two initial structures to (91), both allowing on the boat to satisfy the object restriction on decide. These two initial structures are represented in (92) and (93).

(92)\[
S \\
NP \rightarrow \text{John} \\
VP \rightarrow \text{decided} \\
Vl \rightarrow \text{on} \\
PP \rightarrow \text{the} \\
P \rightarrow \text{boat}
\]

(93)\[
S \\
NP \rightarrow \text{John} \\
VP \rightarrow \text{decided} \\
Vl \rightarrow \text{on} \\
PP \rightarrow \text{the} \\
P \rightarrow \text{boat}
\]
Both (92) and (93) satisfy the domain requirement in (43) because the first branching node above the PP in both cases dominates decided. The problem here is that these claims force us to give up the explanation of the ambiguity of (91) that was proposed in Chomsky (1965).

Chomsky's explanation was that the Strict Locality Condition insured that when decide is subcategorized for an object it would appear as a daughter to $V^1$ while the place-PP would always appear as a sister to $V^1$. In this way there was a structural explanation for the ambiguity in (91). As our analysis stands, we have no explanation for this ambiguity.

In Emonds (1976) it is proposed that:

(94) In a local transformation the following relation must hold between the two constituents $C-C'$ affected by the rule: there exists a node that immediately dominates one of these constituents that also dominates the other.

Emonds' use of "immediate dominance" is very close to the use we made of domain in formulating the Broadly Local Condition. Suppose we were to reframe (94) as (95).

(95) In a local transformation the following relation must hold between two constituents $C-C'$ affected by the rule: the first branching node dominating one of these constituents
also dominates the other (and conversely).

In (95) we will assume that the parentheses have their usual formal interpretation. (95) states that a local rule will be interpreted first as applying to C and C' if they are sisters and only more broadly when either C is in the domain of C' or C' is in the domain of C if the converse relation is not possible.

In the transformational sub-component (95) does not seem to make any empirical differences from (94). However, in the base the modification in (95) is important and provides at least two advantages. First, it permits us to eliminate the reference to domain in our formulation of the Broadly Local Condition by making it follow from an even more general linguistic principle, i.e., (95). As a result of postulating (95) we can simplify (43) to (96).

(96) Given a rule of the form $A \rightarrow \text{Complex Symbol/}$$
\alpha \ldots \beta$ such a rule is local and $\alpha$ and $\beta$
may be interpreted as $\alpha^{\max}$ and $\beta^{\max}$.

Second, we are able to maintain Chomsky's explanation for the ambiguity of (91), as outlined in the following paragraphs.

(95) makes it possible to say that $A \rightarrow CS/ \ldots B$
is more highly valued if $A$ and $B$ are sisters as in the
strictly local convention. However, this condition may be overridden at a cost to the grammar if the
subcategorization entry is made slightly more complex. In the remainder of this study I will follow the convention that a subcategorization restriction of the form (97), when there is no rule of the form (98), is interpreted as (99a) or (b).

(97)  A, + _____ B  
(98)  C → ... A ... B ...
(99)a.  A, + _____ [D B]  
b.  A, + [D _____] B

If, on the other hand, there is such a rule as (98) it will be possible to subcategorize A for B within a different category only by a rule of the form (99). In this case, if there is a contextual feature like (97) and a rewriting rule like (98), (97) can only be applied when A and B are immediately dominated by the same node.

Now return to (91). Under the interpretation where **decide** takes an object we can simply require (100).

(100)  decide, +V, + _____ PP  
(101)  decide, +V, + [_____] PP

(100) can only correspond to the structure in (92). In order for the object to appear as in (93), the more complex restriction in (101) would be required. The structural explanation for the ambiguity in (91) is once again available to us.

73
This convention entails that verbs, in the most highly valued case, are not subcategorized by the subject of an embedded sentence, since an NP position appears in V₁ which is a sister to V. Thus, any apparent dependencies between a verb and a subject of a sentential complement would involve a PRO in the sentential complement bound to an NP inside V₁. However, no similar restriction appears on COMP or TENSE since neither of those elements can appear as a sister to V.

Another problem area somewhat different from (91) is how to block generating indirect objects on verbs like see as in (102).

(102)a. John saw Jane.
   b. *John saw Jane to Bill.

Chomsky (1965) solves this problem by adopting a convention whereby subcategorization rules specify only the environments in which an item _ - a appear and that the item is furthermore negatively specified for every other context. Thus the entry in (103) implies the entry in (104).

(103) see, + ____ NP
(104) see, - ____ PP

Of course we are unable to adopt this convention within the framework of the Broadly Local Convention without some revision. It is untenable to claim that the redundant restriction in (104) applies to all PP's that have see in
their domain. In order to avoid this problem we will adopt the following convention.

(105) An item A is redundantly specified - _____ B
  if + _____ B is not mentioned in the lexical entry for any B which is a sister to A in the phrase marker.

One of the effects of (105) is that while sisters to A need explicitly to be mentioned in a lexical entry in order to occur, it is the unmarked case for elements outside A\textsuperscript{1} to be optional. Thus it takes no special entry in the lexicon to express the fact that see may or may not occur with a place PP as in (106).

(106)a. Jane saw Martha at the movies.
  b. Jane saw Martha.

However, to require a complement that is a sister to V\textsuperscript{1} will be costly since the entry will need to specify the particular item as + _____ D.

4.2 Disambiguating Certain Broadly Local Structures

Let us now pursue the ramifications of those cases where the contextual feature must be taken in the broadly local interpretation. Although we have managed to reduce the ambiguity in decide on the boat, numerous other ambiguities may exist given the Broadly Local Condition. In particular, consider the hypothetical grammar that contains the rules in (107).
(107a) \[ X \rightarrow A - B \]
   \[ B \rightarrow C - D \]
   \[ D \rightarrow E - F \]
   \[ F \rightarrow G - (E) \]

b. \[ A \rightarrow CS/____ E \]

c. \[ C \rightarrow CS/____ E \]

(108)
```
  S
 /\   /
A  B /\   /
 C D /\   /
E  F G  E
```

(109)
```
  S
 /\   /
A  B /\   /
 C D /\   /
E  F G
```

There are ambiguities in both (108) and (109). In (109), for example, it is possible as things stand for both (107b) and (107c) to operate using the same node E as the appropriate contextual feature. Similarly, in (108) it is possible for the rules in (107b-c) to operate as illustrated in (110) or (111). The lines in (110)-(111) indicate which E serves as the context for the expansion of each element. When two elements, A and B, are involved in a member of a class of rules R, we will say that A and B

76
are "R-associated," i.e., "contextually associated," "transformationally associated," "anaphorically associated," etc.

(110)

(111)

Just what exactly the class of rules is in each case is an empirical question; however for the time being I will assume that it contains at least subcategorization rules. For purposes of exposition we will call lines like those in (110)-(111) "association lines" since they serve to show which elements are involved in rules of grammar. In addition, configurations like (110) where the association lines do not intersect we will term "nested structures" while configurations like (111) where the association lines do intersect, we will term "crossing structures."

Rather than permit linguistic theory to treat (108)
as either (110) or (111) for the purposes of (107b-c) we can restrict linguistic theory, and thereby improve it, if we have some principle which will disambiguate the structure in (108) as either (110) or (111). Of course it is important to keep in mind that at each juncture we face decisions which have empirical consequences. It might empirically be the case that (108) is structurally ambiguous as either (110) or (111). If not, it might empirically be the case that (111) is how (108) is disambiguated. I am going to hypothesize that neither of these options is warranted and that in fact linguistic theory disambiguates (108) as (110) and not (111).

In order to capture this hypothesis in a formal way let us postulate the condition in (112) as a principle of Universal Grammar.

(112) The Association Principles

Given the class of rules R, if \( R_i \) involves two terms \( A_i \) and \( B_i \) in a string \( A_i - X - B_i - Z \) containing \( A_j \) and \( B_j \) related by \( R_j \), then

1) if \( A_i - X \) contains \( A_j \), \( X \) contains \( B_j \) as well where: \( B_i \) is structurally non-distinct from \( B_j \).

Notice that I have left "the class of rules R" and the notion "structurally non-distinct" undefined for the moment. The reason for this omission is that I will want to devote

78
more attention to these notions at other points in this study. For the moment, however, I will proceed under the assumption that \( B_i \) and \( B_j \) are "structurally non-distinct" from each other if their heads of phrases are non-distinct. I take "sentence" to have no head of phrase. It will therefore be structurally non-distinct from all other categories. I will further assume that the "class of rules \( R \)" has as a subset the rules of selection and local subcategorization. In subsequent chapters I will treat \( R \) as including not only contextual rules but anaphoric rules and the core syntactic rules as well. The condition in (112) requires nesting and disambiguates structures like (108) as (110) rather than as (111). The condition in (112) also serves to prevent the \( E \) in (109) from satisfying the requirement for both (107b) and (107c): if \( X \) needs to contain \( B_j \), then \( B_i \) and \( B_j \) cannot be the identical node.

Before proceeding we should observe one other fact. Within the model of Chomsky (1965) it is possible to speak of "frames" of strict-subcategorization since for any head of phrase \( C \), all the categories \( C_j \) that are sisters to \( C \) are associated with \( C \). In order to retain the notion of "frame" we will postulate that (112) has the further restriction in (113).

(113) At least one phrasal node dominating \( B_i \) does not dominate \( B_j \) or conversely.
(113) will require that \( B_i \) and \( B_j \) not be immediately dominated by the same phrasal node. That is to say, \( B_i \) and \( B_j \) cannot be sisters.

In summary then I am proposing the condition in (114).

(114) The Association Principles

Given the class of rule \( R \), if \( R_i \) involves two terms \( A_i \) and \( B_i \) in a string \( A_i - X - B_i - Z \) containing \( A_j \) and \( B_j \) related by \( R_j \), then:

1). if \( A_i - X \) contains \( A_j \), \( X \) contains \( B_j \) as well, and

2). at least one phrasal node dominating \( B_i \) does not dominate \( B_j \) or conversely,

where \( B_i \) is structurally non-distinct from \( B_j \).

I will call (114) the Association Principles. Sometimes I will find it convenient to refer to the restriction in (114.1) as the "nesting condition" and the condition in (114.2) as the "strict frame requirement."

4.3. Multiple-Headed Comparatives

In order to provide some evidence corroborating the Association Principles we can return to the multiple-headed comparative constructions. Recall that earlier we
gave some consideration to the multiple-headed comparatives noted in Andrews (1975). An example of such a construction is reproduced in (115).

(115) As many more people than you expected as I predicted came to the party.

(116) *As many more people as I predicted than you expected came to the party.

Sentence (115) may be contrasted to the sentence in (116) in which the order of the clauses has been reversed. In Andrews (1975) this fact is accounted for in a relatively narrow and ad hoc fashion. Andrews proposes what he calls the mirror-image constraint requiring the clauses to appear in the opposite order as their respective determiners.

While this constraint is observationally adequate for the sentences at hand, it is clearly inferior to the more general formulation of the Association Principles if the latter prove to be descriptively adequate.

The structure of the relevant NP's in (115)-(116) are as indicated in (117) and (118) respectively where the association lines have been drawn as well as the factorizations relevant to the formulation of the Association Principles.

(117) as many more people [than you expected] [as

I predicted]
(118) as many more people [as I predicted] [than
you expected]

The Association Principles block (118) and require (117). Only in (117) are the association lines nested.

Another source of evidence in favor of the Association Principles is contained in sentences like (119).

(119)a. This flower is fragrant to smell.
   b. The soup is pleasant to prepare.
   c. Bill is tough to get to know.

Evidently we want the lexicon of English to capture the fact that the underlined adjectives in (119) may occur with an optional sentential complement in order to distinguish them from the underlined adjectives in (120)-(122) which cannot appear with such sentential complements.

(120)a. This flower is wilted.
   b. *This flower is wilted to smell.

(121)a. This soup is cold.
   b. *This soup is cold to eat.

(122)a. Bill is dashing.
   b. *Bill is dashing to get to know.

The lexicon of English will then contain entries something like (123).

(123)a. fragrant, +A, + _____ (S1)
   b. pleasant, +A, + _____ (S1)
c. tough, +A, +_____ (S1)

Certainly (123) is schematic and leaves unspecified crucial statements that depend on how one decides to derive the sentences in (119), that is, either by a deletion as in Lasnik and Fiengo (1974) or via WH-movement as in Chomsky (1977b). What is of interest for our purposes is the behavior of the adjectives and their complements when another head-complement relation is added to their structure. So for example, it is worthwhile noting the sentences in (124)-(126).

(124)a. This flower is more fragrant to smell than that rose is.

b. *This flower is more fragrant than that rose is to smell.

(125)a. This soup would be more pleasant to prepare than the quiche would be

b. *This soup would be more pleasant than the quiche would be to prepare.

(126)a. Mary is no more difficult to get to know than Susan is.

b. *Mary is no more difficult than Susan is to get to know.

In each case above we find the Association Principle at work guaranteeing a nested construction. A factorization of (124) is presented in (127) in order to illustrate this.
(127)a. This flower is \[more \text{ fragrant} \] [PRO to smell] [than that rose is] [PRO to smell] [B₁] [B₂]

b. This flower is \[more \text{ fragrant} \] [than that rose is] [PRO to smell] [B₁] [B₂]

In (127a) nothing violates the Association Principles, that is, the variable X separating A₁ and B₁ contains both A₂ and B₂. This means that the construction in (127a) is nested. The structure in (127b) however, is completely different: notice that the variable X separating A₁ and B₁ contains A₂ but not B₂. The structure therefore is a crossing one and is blocked by the Association Principles.

4.4. Deletion of That in Complements to Adjectives

The examples just considered provide evidence that ambiguous structures are interpreted as nested. It is also important to provide evidence in favor of the claim in
the Association Principles that there are unique frames for contextual features, i.e., restriction (114.2). One source of such evidence is the possibility of deleting that in structures where adjectives have sentential complements.

In Chomsky and Lasnik (1977) it is proposed that there is a rule of free deletion in COMP the output of which is subject to certain negative output filters. One of the filters proposed by Chomsky and Lasnik is (128).

\[ (128) \star [\alpha_{NP} \text{TENSE VP}], \text{unless } \alpha \text{ is adjacent to and in the domain of } [+V], \text{ that or NP} \]

The term "domain" in (128) is understood as in (129).

\[ (129) \text{A is in the domain of B if A is dominated by the first branching node that immediately dominates B.} \]

This filter is intended to account for the contrast in acceptability between the sentences in (130) and (131).

\[ (130)a. \text{The fact that John was here surprised me.} \]
\[ b. \text{It came as a surprise to me that John was here.} \]
\[ c. \text{It is unlikely that John is here.} \]
\[ d. \text{That John is here, I have no reason to think.} \]
\[ e. \text{I pleaded (argued) with Bill that he shouldn't fire her.} \]
f. John believes (that) Mary saw Sam, and Bill, that Sue saw Harry.

(131)a. *The fact John was here surprised me.
b. *It came as a surprise to me John was here.
c. *It is unlikely John is here.
d. *John is here I have no reason to think.
e. *I pleaded (argued) with Bill he shouldn't fire her.
f. *John believes (that) Mary saw Sam, and Bill, Sue saw Harry.

The sentences in (131) correspond to those in (130) except in (131) deletion of that in COMP has operated. The filter in (128) should block deletion of that in each case of (131).

In order to guarantee that the that in (130c) cannot be deleted, Chomsky and Lasnik point out that it would be necessary to recognize a structure like (132) for the sentence.

(132)
In this structure, that John is here is not in the domain of unlikely and hence that cannot be deleted without contravening (128). If we were to recognize either (134) or (135) as the relevant structures, then we would predict that COMP deletion should always be able to apply.

\[(134) \text{it is } \underset{A^1}{\text{unlikely}} \underset{S^1}{\text{that John is here}}\]

\[(135) \text{it } \underset{V^1}{\text{is }} \underset{A^1}{\text{unlikely }} \text{ that John is here}\]

In both cases, that John is here is within the domain of unlikely and since adjectives are [+V], the filter in (129) would not block deletion of that. The unique frame requirement in the Association Principles guarantees that these sentential complements will be outside V^1 since the AP and S^1 can only share the same frame if they are associated with a common head (cf., restriction 113), assuming with Hendrick (1978a) that adjectives have no post-head complements in the base. In this case we have a principled explanation for the facts in (130)-(131).

Before leaving this topic I want to respond to some criticisms of this analysis made by P. Schachter and R. Stockwell (personal communication). Their criticisms amount to the fact that they disagree with the judgment that (131c) is unacceptable. Nevertheless they both agree that in questions the COMP may not delete as (136) illustrates.
(136)a. *How unlikely John is here is it?
   b. How unlikely that John is here is it?

Similarly they agree with the judgments in (137).

(137)a. *How unlikely is it John is here.
   b. How unlikely is it that John is here?

The judgments in (136)-(137) are easily accounted for
within a framework like Hendrick (1978a) where adjectives
have no post-head complements in the base. In order to
account for sentences like (136b), an adjunction rule is
postulated, Chomsky-adjointing the sentential complement to
the adjective. The resulting structure is (138).

(138)

\[ \begin{array}{c}
A_1 \\
A \\
Q \\
\end{array} \]

\[ \begin{array}{c}
S_1 \\
A \\
COMP \\
\end{array} \]

\[ \begin{array}{c}
S \\
NP \\
\end{array} \]

\[ \begin{array}{c}
VP \\
V \\
NP \\
\end{array} \]

how unlikely that John is here

In this structure, that John is here is outside the domain
of unlikely and hence deletion of that is blocked. Simi-
larly, deletion of that is also blocked in (138) since I
assume that adjectives leave traces and that the trace
keeps the sentential complement outside the domain of the
verb be. This fact can be seen in the trees in (139)–
(140).
(139) S1
  COMP
  NP
  V1
  AP
  A1
  How unlikely it is

(140) S1
  COMP
  NP
  V
  V
  A1
  A
  Q
  How unlikely it is

How unlikely it is e that John is here

How unlikely it is that John is here
The structure in (139) corresponds to the instance where adjectives leave traces while the structure in (140) corresponds to the structure where they do not. In (139) the first branching node above _is_, that is the node defining its domain, is \(V^1\), and \(S^1\) is outside that domain. Deletion of COMP is therefore impossible. The structure in (140) however, is different. The first branching node above _is_ is VP and therefore \(S^1\) is in the domain of _is_ in (140). For this reason we will assume that adjectives leave traces, and we have a natural explanation for the judgments in (137).

It appears then that the "dialect" represented by Schachter and Stockwell differs only with respect to (131c), the other judgments being completely predictable within the general framework defended here. I propose to account for the variation in judgments concerning (131c) in one of two ways: (a) there is a difference between dialects in the formulation of the local adjunction rule: for speakers accepting (131c) the adjunction rule is as formulated in (141) while for speakers rejecting (131c) the adjunction rule is as in (142), (b) both dialects have an adjunction rule like (141) but they differ in that speakers not accepting (131c) have the additional output filter in (143).

\[
(141) \quad X - A^1 - \{S^1\}_{PP} - Y \Rightarrow 1 - 2+3 - 4
\]
(142) \[ X - WH+A - \{S^1 \} Y \Rightarrow 1 - 2+3 - 4 \{PP\} \]

(143) \[*[a \ A \ \beta]\] where \(\beta \neq \emptyset\)

(141) but not (142) can operate in (131c) to yield a structure like (144).

(144)

\[
\begin{array}{c}
S \\
NP \\
V \\
S \\
A \\
\text{COMP} \\
NP \\
V \\
NP
\end{array}
\]

In (144), that John is here is within the domain of unlikely since the first branching node dominating unlikely also dominates the sentential complement. For this reason that may be deleted without violating the filter in (128).

When how is present and (141) operates, the structure is as in (138) and deletion of COMP is blocked for the reasons previously discussed. Now in the dialect assuming (142), it is not possible to do an adjunction in (131c), hence it is not possible to derive a structure like (144) where the clause is within the domain of unlikely. (142)
can operate however to yield the correct results with respect to the sentences like (136)-(137) where a WH item is involved.

Alternatively we could assume that both dialects permitted the adjunction represented in (141), the only difference between the dialects being whether (143) applied or not. In the dialect where (143) applied adjunctions like (144) to a bare adjective would be blocked so that on the surface the complement would not be in the domain of the adjective and that would never delete. In the dialect without (143), that is, the dialect of Stockwell and Schachter, nothing blocks (144) and consequently we would generate sentences like (131c). Since the filter in (143) applies only to bare adjectives, in those cases involving a WH+ADJ sequence, cases like (138), the filter would be inapplicable and the two dialects would have the same judgments.

Either of the above explanations for the variability of judgments surrounding that deletion avoids recourse to vague "perceptual strategies" but it also fits in well with Emonds' hypothesis that language variation is restricted to local rules.
FOOTNOTES TO CHAPTER 2

1 This definition is taken from Chomsky and Lasnik (1977) and bears a very close similarity to "c-command," a notion developed in Reinhart (1976).

2 I assume with Chomsky and Lasnik (1977) that than and as are not in COMP. I have not treated selection of COMP or TENSE by the specifier items like too. Careful consideration of such questions may pose problems for the formulation of the Broadly Local Condition and suggest possible refinements.

3 I have adopted Jackendoff's (1978) analysis of the specifier system. The particular phrase structure rules for (57) are not crucial for the conclusions here however. All that is important for the argument to go through is that more is in the adjective phrase.

4 Once again, I have adopted Jackendoff's analysis of the specifier system for constructing the tree. However, similar results follow from Bresnan's (1973) analysis of the specifier system. As an aside, Jackendoff's (1978: ch. 6) claim that different and alike subcategorize quantifiers, is, if true, another case where the Strict Locality Condition is overly restrictive and where the Broadly Local Condition is preferable.

5 I do not have an explanation for why manner adverbs fail to prepose. In Chapter 3, however, certain restrictions on the preposing of manner adverbs are shown to fall out from the Association Principles proposed in this chapter. It may be possible to extend that analysis of adverbs to these cases using an analysis similar to the one sketched in Section 6 of Chapter 3 for a slightly different preposing phenomenon.

6 The formulation of this rule is somewhat different from that of Hendrick (1978). The earlier formulation would not permit a correct account of the that deletion phenomenon discussed at the end of this chapter.

7 Although rule (84) affects only adjacent conditions, it departs from Emonds' (1976) requirement that local rules involve at least one non-phrasal category. This
might indicate that Emonds' restriction is too narrow. I am assuming a version of linguistic theory where local rules are not iterative.

8 In fact we will argue below that only (90b) is a well-formed structure.
Chapter 3

CLAUSE INTERNAL BINDINGS AND THE
ASSOCIATION PRINCIPLES

As mentioned earlier, previous attempts to establish a nesting requirement in linguistic theory have based their conclusions on phenomena which had alternative explanations by other conditions in Universal Grammar. This chapter presents several phenomena which have an elegant explanation if one assumes that a nesting condition like the Association Principles is part of Universal Grammar. These phenomena involve bound anaphora, that is, the association of an antecedent and an empty node (a trace or PRO depending on whether a movement is involved or not, cf., Chomsky 1976). All of the phenomena treated in this chapter are clause internal. As such they do not overlap other conditions on rules and by showing that a nesting condition like the Association Principles expands the empirical domain of linguistic theory, these phenomena strongly suggest that the Association Principles are on the right track.

1. WH-Movement and Passive

Consider the sentences in (1)-(3).
(1) Mary took unfair advantage of Jane.
(2) Unfair advantage was taken \([NP^e]\) of Jane.
(3) Jane was taken unfair advantage of \([NP^e]\).

The passive transformation can apply to (1) to generate either (2) or (3). Similarly it is possible for WH-fronting to apply to (1) in which case we derive either (4) or (5).

(5) \{What kind of\} advantage did Mary take \([NP^e]\)
\{How much \}
of Jane?

(5) Who did Mary take unfair advantage of \([NP^e]\)?

The sentence in (1) provides a rare example where two NP's may be moved within an \(S^1\) by the same rule. When both passive and WH-fronting are applied to (1) it is possible to derive either (6) or (7).

(6) Who was unfair advantage taken \([NP^e]\) of \([NP^e]\)?

(7) *\{What kind of\} advantage was Jane taken \([NP^e]\)
\{How much \}
of \([NP^e]\)?

The sentence in (6) is acceptable while that in (7) is not. This fact does not follow from any condition on rules of grammar that I am aware of. Conditions like the Specified Subject Condition or the Tensed-S Condition are not applicable because they constrain movements outside \(S^1\). Conditions like the Superiority Condition or Wilkins' (1976) Variable Interpretation Condition which do deal with clause
internal movements are equally inapplicable. However, the Association Principles provide a completely natural account of these facts. In order to see just exactly how these predictions are made I provide the structures and factorizations in (8)-(9).

\[(8)\]
\[\begin{array}{c}
S^1 \\
\downarrow \\
\text{COMP} \\
\downarrow \\
S \\
\downarrow \\
\text{NP} \\
\downarrow \\
\text{V} \\
\downarrow \\
\text{NP} \\
\downarrow \\
\text{PP} \\
\downarrow \\
\text{P} \\
\downarrow \\
\text{NP} \\
\end{array}
\]

\[\begin{array}{c}
\text{who} \\
\downarrow \\
A_1 \\
\end{array}
\]

\[\begin{array}{c}
\text{unfair advantage} \\
\downarrow \\
A_2 \\
\end{array}
\]

\[\begin{array}{c}
\text{was taken} \\
\downarrow \\
B_2 \\
\end{array}
\]

\[\begin{array}{c}
\text{e} \\
\downarrow \\
B_1 \\
\end{array}
\]

\[\begin{array}{c}
\text{of e} \\
\end{array}
\]

The structure in (8) corresponds to (6) while the structure in (9) corresponds to (7). Note that in (8) the variable X contains both \(A_2\) and \(B_2\) but in (9) the variable only contains \(A_2\). This means that only (8) is a nested structure and consequently only (8) is well formed.

2. Clitic Pronouns

2.1. Multiple Clitic Pronouns in Spanish

In Perlmutter (1970) a number of interesting facts concerning the order of Spanish clitic pronouns are presented. Perlmutter observes the sentences in (10)-(11).
(10a). Me lo recomendaron.
They recommended him to me."

b. Te lo recomendaron.
They recommended him to you.

c. Se lo recomendaron.
They recommended him to him/her them.

(11)a. *Me le recomendaron.
They recommended me to him.

b. *Te le recomendaron.
They recommended you to him.

Perlmutter is essentially concerned with placing output constraints or filters on clitic sequences. These sentences, however, cannot be so easily accounted for by a filtering device. Perlmutter proposes to account for these sentences by postulating the following.

(12) If the direct object of recomendar is first or second person, use of the clitic form of the indirect object results in an ungrammatical sentence.

This restriction, while apparently observationally adequate, is unsatisfying because of its ad hoc character. What is interesting is that the data in (10)-(11) can be predicted from the Association Principles, conditions of significant generality.

Consider for example (10a). It has a structure like (13).
(13) me le recomendaron \[NPe\] (a) \[NPe\] \\
(14) me le recomendaron \[NPe\] (a) \[NPe\]

On the other hand, (11a) has a structure like (14). Clearly whether one takes clitic placement to be a movement phenomenon resulting in a trace, or a construal phenomenon involving associating a base generated clitic with a PRO as in Strozer (1977), (13) but not (14), is a nested construction. The Association Principles, because they block crossing of association lines, prevent the generation of (11).

Another important fact about Spanish clitic pronouns can be observed in the following sentences, adapted from Strozer (1977).

(15)a. *Me le \{presentaron \} \{recomendaron \} \{describieron \}

They introduced/recommended/described me to him.

b. *Te le \{presentaron \} \{recomendaron \} \{describieron \}

They introduced/recommended/described you to him.

(16)a. *Le me \{presentaron \} \{recomendaron \} \{describieron \}

They introduced/recommended/described me to
him.

b. *Le te \{presentaron \}
   \{recomendaron\}
   \{describieron\}

They introduced/recommended/described you to him.

Of course the sentences in (15) are excluded by the Association Principles since they involve a crossing structure, as (17) illustrates.

(17) \{me\} le presentaron [NP\(e\)] (a) [NP\(e\)]

That is to say, the Association Principles predict that the unmarked case for pre-verbal clitics is indirect object to the left of the direct object.

The Association Principles, because they exclude the sentences in (15), reduce the amount of work borne by the output filters. Now the problem rests only in excluding sentences like (16). In order to account for these sentences, we can postulate the following general filter.

(18) *\[
\text{PRON} \\
\text{-REFLEX} \\
\text{III}
\] - [PRON]

This filter is important not only in that it explains why the sentences in (16) are unacceptable but it also permits us to simplify the rule generating the so-called "spurious se" rule in Spanish.

Perlmutter observes the sentences in (19)-(20).
(19)a. Lo recomendé a ti.
    I recommended it to you.
b. Te lo recomendé.
    I recommended it to you.
(20)a. Lo recomendé a él.
    I recommended it to him.
b. *Le lo recomendé.
    I recommended it to him.
c. Se lo recomendé.
    I recommended it to him.

What Perlmutter notes is that while te and me may precede lo we do not find instances where le precedes lo; rather se does. The same can be observed of the respective plurals of le and lo, i.e., les and los. We do not find any examples of the sequences in (21).

(21) *le lo  *le la  *le las  *le los
    *les lo  *les la  *les las  *les los

Perlmutter uses these facts to motivate the rule in (22).

(22) Spurious se rule (obligatory)

\[
\begin{array}{c}
\text{PRO} \\
\text{III} \\
\text{DATIVE}
\end{array} - \begin{array}{c}
\text{PRO} \\
\text{III} \\
\text{ACC}
\end{array} \rightarrow \text{se} - 2
\]

We are in a position, given our preceding analysis, to simplify (22) considerably. We are able to reduce (22) to (23).
(23) le(s) - PRON $\Rightarrow$ se - 2

1 - 2

The filter in (18) guarantees that (23) is obligatory, since if it does not apply (18) will filter out the structure. In addition we do not need to make use of any feature like DATIVE in (22) since any direct object as in (16) will be blocked by the Association Principles.

There are some apparent counter-examples to the claims just advanced that the "clitic placement" phenomenon is within the domain of the Association Principles. I want to examine these apparent counter-examples and demonstrate that in fact they do not detract from the general explanatory value achieved by letting the Association Principles govern the clitic placement phenomenon.

To begin with, we can observe sentences like (24) where there is an apparent crossing of association lines.

(24)a. Te nos describieron.
    They described you to us.

b. Te nos ensuciaron
    They dirtied you on us.

c. Te me ensuciaron.
    They dirtied you on me.

d. Te me descibiste.
    You described yourself to me.

(25) Te nos describieron $^{[NP_e]} (a) [NP_e]$
The sentences in (24) all seem to have a crossing structure like that illustrated for (24a) in (25). It is an interesting fact about Spanish clitics, however, that a sequence \( \text{nos} - \text{te} \) yields an ungrammatical structure. So for example, Perlmutter notes the sentences in (26)-(27).

(26)a. Me escapé.
    I escaped.

b. Me le escapé.
    I escaped from him.

c. *Me te escapé.
    I escaped from you.

(27)a. Nos escapamos.
    We escaped.

b. Nos le escapamos.
    We escaped from him.

c. *Nos te escapamos.
    We escaped from you.

Perlmutter proposes to account for these facts by postulating a filter blocking \( \text{nos} - \text{te} \) sequences. His proposal, however, will not remove examples like (24) as apparent counter-examples to the claim that the Association Principles hold over clitic placement. In Strozer (1977) the possibility of a local permutation is considered in order to account for the non-appearance of \( \text{nos} - \text{te} \).
sequences. Such a local rule might be formulated as in (28).

(28a) \[
\begin{array}{c}
Pron_I \\
1 \\
\end{array} \rightarrow \begin{array}{c}
Pron_{II} \\
2 - 1 \\
\end{array}
\]

b. \[
\begin{array}{c}
Pron_{+I} \\
1 + \\
\end{array} \rightarrow \begin{array}{c}
Pron_{-Person} \\
2 - 1 \\
\end{array}
\]

(28b) is meant to express the same thing as (28a): it assumes a somewhat different way of distinguishing clitics. It opposes third person on the one hand from first and second person on the other hand. The argumentation in favor of this division is made in Strozer (1977). (28) opens up a possible avenue of explanation. The Association Principles hold at surface structure. Suppose that, in accordance with Chomsky (1973), we permit a class of local (in the sense of Chapter 2) or "housekeeping" rules to apply subsequent to surface structure. Let us furthermore begin to restrict this class of "late" rules so that (a) such housekeeping rules are strictly local and (b) such housekeeping rules involve no phrasal nodes. These housekeeping rules will include deletions and operations like (28). Furthermore we expect such rules to be subject to language variability insofar as they are outside core grammar. In this case, we maintain that at the level of structure where the Association Principles are applied,
i.e., surface structure, (28) has not applied yet and in fact the relevant structures for sentences like (24) are as in (29).

(29)a. nos te describieron \([\text{NP}_e](a)\) \([\text{NP}_e]\)

b. nos te ensuciaron \([\text{NP}_e](a)\) \([\text{NP}_e]\)

c. me te describieron \([\text{NP}_e](a)\) \([\text{NP}_e]\)

d. me te describiste \([\text{NP}_e](a)\) \([\text{NP}_e]\)

Subsequent to the checking of the Association Principles at surface structure (28) is able to apply to generate the sentences like (24).

Sentences like (30) are ambiguous for many speakers.

(30) Te me\{describieron\}
    \{presentaron\}
    \{recomendaron\}

They described/introduced/recommended me to you/you to me.

At the level of surface structure (30) will have either the structure of (31) or (32). The double structure accounts for the ambiguity in these sentences.

(31) Me te describieron \([\text{NP}_e](a)\) \([\text{NP}_e]\)
(32) Te me describieron \([_{NP}e]\) (a) \([_{NP}e]\)

In (31) the rule (28) applies to yield the acceptable (30) which is homophonous to (32).

The claim that (28) exists as a language particular rule operating after surface structure where we only find structures like (29) is given some corroboration by the fact that some speakers of Spanish do not require the \(nos - te\) interchange (as pointed out to be by A. Rivas, personal communication).

The reflexive clitic pronouns also raise some apparent counter-examples to the claim that the Association Principles govern clitic placement. There exist sentences like (33) which appear to be crossing structures, noted in Strozer (1977).

(33)a. El niño se me lavó porque se lo había pedido.
The child washed himself for me because I had asked him to.

b. Mamá pidió que me lavara y me le lavé.
Mom asked me to wash myself and I washed myself for her.

c. Te le lavaste.
You washed yourself for him.

(34) El niño se me lavó \([_{NP}e]\) (a) \([_{NP}e]\) ....
The structure in (34) illustrates the crossing character
of the association lines in (33). The fact that a crossing structure is present in (33) suggests that a local permutation has operated subsequent to the level where the Association Principles are applied, i.e., subsequent to surface structure. This rule might be formulated as (35).

\[(35) \quad \text{[PRON]} - \text{[PRON} + \text{Reflexive]} \Rightarrow 2 - 1 \]

1 \quad - \quad 2

A further simplification of the clitic system of Spanish is now possible. Compare (35) to the rule inverting the first and second person clitics, that is, rule (28). There is considerable overlap here and we can collapse the two rules as (36).

\[(36) \quad \text{[PRON} <\text{+I}> [\text{PRON} \{<-\text{Person}> \{[+\text{reflexive}]\}] \Rightarrow 2 - 1 \]

1 \quad - \quad 2

The postulation of this rule means that we can treat sentences like (33) as having a nested structure, as in (37), at the point where the Association Principles apply.

\[(37) \quad \text{El ni\ñ o me se lav\textcircled{a}} [\text{NP}\text{e}] (a) [\text{NP}\text{e}] \ldots \]

I conclude that the reflexive clitics pose no obstacle to extending the domain of the Association Principles to include clitics, and in fact supports them.

Before proceeding I should note that sequences of

108

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
clitics with the same person specification like se se or me me are impermissible. I assume here that such examples can be dealt with adequately by a condition of disjoint reference for clitics proposed in Strozer (1977). Since such sequences are not crucial to this discussion I will not consider them further. Similarly I have not considered clitics that express the interest of the speaker, that is, those clitics marked [+SPK] in Strozer (1977). An example of such a clitic can be found in (38).

(38) Me le dieron un helado.

They gave him an ice cream on/for me.

Such clitics are, as Strozer suggests, base generated and cannot be bound to any element within sentence grammar. It therefore presents no problem for the Association Principles and does not concern us here.

A brief summary of the mechanisms proposed here to account for clitic placement phenomena in Spanish follows.

(39)a. The Association Principles govern clitics

b. le(s) - PRON \[\rightarrow se - 2\]

c. \[
\left[
\begin{array}{c}
\text{PRON} \\
\langle +I \rangle \\
\end{array}
\right] - \left[
\begin{array}{c}
\text{PRON} \\
\langle [-\text{Person}] \rangle \\
\{ [+\text{reflexive}] \}
\end{array}
\right] \rightarrow 2 - 1
\]

d. Strozer's disjoint reference for clitics.

In the next subsection I intend to compare this analysis of Spanish clitics to its major competitors.

109
2.2 Comparison with Previous Analyses

The analysis just sketched has several advantages which recommend it over the analysis presented in Perlmutter (1971). On the one hand, we are able to dispense with such filters as (12) which are at best ad hoc and at worst formally unstatable. In addition we were able to avoid postulating the positive output filter (40).

(40) \[ SE \sim II - I - III \]

This is an advantage in two ways: first, we avoid permitting positive output filters in the class of grammars, thereby constraining linguistic theory (cf., Emonds 1976), and second, we restrict the class of possible filters to local ones. In addition we found it possible to significantly simplify the spurious se rule.

I want now to compare the analysis I am proposing with the general framework of Strozer (1977). As I have mentioned I have built on Strozer's work considerably, however there are some important differences between her work and my own.

Strozer permits in her version of linguistic theory the possibility of non-local filters. As an example, consider the sentences in (41)-(42).

(41)a. Le hablé a Lola.
     I spoke to Lola.
b. Le gusta a Lola.
   It is pleasing to Lola.

(42)a. Hablé a Lola.
   I spoke to Lola.

b. *Gusta a Lola.

Strozer notes that hablar and gustar apparently select different kinds of indirect objects. She distinguishes these objects with the feature \([\pm \text{INV(olved)}]\) and proposes a number of tests for determining which class an indirect object belongs to. The verbs selecting a \([+\text{INV}]\) indirect object require a clitic pronoun. Gustar is a verb requiring a \([+\text{INV}]\) noun phrase to its right and consequently the clitic must also be present as (42b) illustrates. Hablar however requires a \([-\text{INV}]\) noun phrase and the clitic is only optional, as (42a) shows. Now Strozer proposes to account for these facts through the filter in (43).

(43) Strozer's filter 3
   The following structure, where \(X \not\in \text{PRON}_i\) and \(\text{NP}_i \in +\text{INV}\), is excluded:

\[
X \rightarrow \begin{cases} \text{NP} \\ \text{-PRO} \\ \text{-DIR}_i \end{cases} - Y
\]

We achieve a more constrained and thus preferable version of linguistic theory if, following Chomsky and Lasnik (1977) or Emonds (1976, chp. IV.4), we assume that filters must be local. In this restrictive version of linguistic
theory (43) is unstateable. In this way the approach in the last section is compatible with a more restrictive version of linguistic theory than Strozer's approach and, if observationally adequate, is to be preferred. In this restrictive version of linguistic theory we are able to capture the facts in (41)-(42) by subcategorizing any \( \text{NP} \) to obligatorily co-occur with a preceding clitic pronoun.

We are in a position to reduce the complexity of Strozer's phrase structure rule if we assume the analysis in the preceding section. Strozer's proposed rules are reproduced as (44).

\[
(44) \quad V^2 \rightarrow (\text{NEG}) - (\text{PRON}^{+\text{REFL}}) - \left( (\text{PRON}^{+\text{SPK}}) \text{PRON}^{-\text{DIR}} - \text{PRON}^{-\text{REFL}} \right) - (\text{PRON}^{+\text{DIR}}) - V^1 ...
\]

(45) \( V^1 \rightarrow V - (\text{NP}^{+\text{DIR}}) - (\text{NP}^{-\text{DIR}}) \) ....

The rule in (44) overlaps to a considerable degree the rule in (45) insofar as the order to the direct and indirect objects (at least when they are not reflexive) have a mirror image in the clitic system. This fact is accidental in Strozer's analysis but is predictable given the Association Principles. This means that at the very least we can reduce (44) to (46).
(46) \[ V^2 + (\text{NEG}) - (\text{PRON}) - (\text{(PRON) PRON}) - \]
\[ +\text{REFL} \quad +\text{SPK} \]
\[ (\text{PRON}) - V^1 \ldots \]

Rule (44) expresses the fact that the [+SPK] pronoun only co-occurs with an indirect object. This fact can be accounted for by simply stating a co-occurrence restriction as in (47).

(47) me, + SPK, + _____ PRON
    DATIVE

This restriction in (47) permits us to simplify (44) even further to (48).

(48) \[ V^2 + (\text{NEG}) - (\text{PRON}) - (\text{PRON}) - (\text{PRON}) - \]
\[ +\text{REFL} \]
\[ (\text{PRON}) - V^1 \ldots \]

The rule in (48) is considerably reduced in complexity. The question remains whether we can eliminate the [+REFL] feature on the first PRON. Strozer argues that the left most PRON must be reserved for reflexive clitics because sentences like (49) are not ambiguous.

(49) Lolita se me compró todavía otro vestido con mi dinero.

Lolita bought herself still another dress on me with my money.

If it were possible to also generate the reflexive to the right of the [+SPK] clitic, Strozer argues, it would be

113

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
possible to find two readings for (49): on one reading se would be interpreted as an "inherent" reflexive clitic and me as an indirect object clitic. On the second reading se would correspond to the reflexive indirect object while me was [+SPK]. The first reading is possible as (50) illustrates.

(50) Que te hizo Lolita? A mi se me compró todavía otro vestido con mi dinero (*a sí misma).
What did Lolita do to you? On me, she bought still another dress with my money.

However, the second reading is not possible. So, for example the sentence in (51) where the reflexive is an indirect object is unacceptable despite the fact that in (52) a reflexive indirect object is acceptable without the [+SPK] clitic.

(51) *Lolita se me compró un vestido a sí misma.
Lolita bought a dress for herself on me.

(52) Lolita se compró un vestido a sí misma.
Lolita bought a dress for herself.

Strozer suggests that if the indirect object se is generated to the right of me and then moved to the left, nothing would prevent the [+SPK] clitic from appearing. In this way (49) would be ambiguous. A simple elaboration of (47) however, would prevent insertion of [+SPK] clitics to the
left of reflexive indirect objects. Such an elaboration is presented in (53).

(53)  me, +SPK, PRON, [PRON] [PRON]

                  -REFL

                   DATIVE

Nothing now prevents the movement of the reflexive pronouns suggested in the last section.

Insofar as the co-occurrence restrictions permit the expression of the appropriate restrictions surrounding the [+SPK] pronoun with respect to reflexive pronouns, we can simplify Strozer's phrase structure rule even further. Now it appears that we can simplify the phrase structure rules to (54) since we no longer need to specify [+REFL] on the left most PRON.

(54)  V^2 — (NEG) — (PRON*) — V^1 ....

(55)a.  V^2 — (NEG) — (PRON^1) — V^1 ....

       b.  PRON^1 — { PRON^1 — PRON^1 }

              { (PRON) — (PRON) }

The four undifferentiated pronoun positions may well be able to be reduced even further to a single PRON position as in (55a) if we account for the multiple PRON positions through a coordination rule as in (55b).

In Rivas (1978) the possibility of establishing a nesting condition for clitic pronouns is considered. It is finally rejected however on the basis of sentences like
(56)-(57) from French.

(56) J'ai mis mon chat sur la table.
I put my cat on the table.

(57) Je l'y ai mis.
I put it there.

(56) pronominalizes as (57). The result is a non-nested configuration as is shown in (58).

(58) je le y ai mis \([NP^e] [pp^e]\)
$$A_1 A_2 \quad B_1 \quad B_2$$

However, it is important to note that Kayne (1975) provides arguments that \(y\) is a pro-preposition while \(le\) is a pronoun. In this case \(B_1\) and \(B_2\) are not of the same category, that is, they are not structurally non-distinct, and the Association Principles are not applicable. It thus seems that Rivas' rejection of a nesting condition operating over clitics is somewhat premature.

2.3. Clitic Pronouns in French

The clitic pronoun system in French differs from that in Spanish in relatively small ways. The clitic pronouns of French are listed in (59)-(63).

(59) me, +N, +PRO, +I, +Ref1, - Pl

te, +N, +PRO, +II, +Ref1, -Pl

nous, "N, +PRO, +I, +Ref1, +Pl

vous, +N, +PRO, +II, +Ref1, +Pl

(60) lui, +N, -Pl, +III, +PRO, -Ref1

leur, +N, +PRO, +III, +Pl, -Ref1
(61) se, +N, +PRO, +III, ±Pl, +Refl
(62) le, +N, +PRO, +III, -Pl, -Refl, -FEM
    la, +N, +PRO, +III, -Pl, -Refl, +FEM
    les, +N, +PRO, +III, +Pl, -Refl, ±FEM
(63) y, +P, +PRO
    en, +P, +PRO

(59)-(62) are clitic pronouns while (63) represents the clitic prepositions or PRO-PP's as Kayne terms them. Reflexive pronouns are morphologically distinguished only in the third person: thus (60)-(62) are different from (59) which may be reflexive or not. In the third person there is also a distinction between dative and non-dative clitics, as the difference between (60) and (62) indicates. This distinction also corresponds to the fact that only in (62) is gender distinguished.

Traditional grammatical descriptions of French point out that the following ordering relation holds among the various clitics just outlined.

(64)a. +Person - +III - +III - y - en - V
      -Dative +Dative

b. {me} - {le} - {lui} - y - en - V
   {te}
   {nous}
   {vous}
   {les}
   {leur}

In his study of these facts Emonds (1976) observes that (64a) need not be stated as an output constraint, that it can be reduced significantly. Emonds argues that there

117
exists a base rule like (65) accounting for the ordering between the clitic pronouns and the pro-PP's.

\[(65) \ V' \rightarrow \begin{cases} \text{[PRO]} & \text{[PRO]} \\ +N & +P \end{cases} \ V\]

He proposes to account for the ordering between the various pronoun clitics in the following way. Evidently no two elements of (59)-(61) may co-occur. Moreover in the imperative form, one finds a slightly different ordering, namely (66) where lui and leur class together with the clitics of (59)-(60) despite the fact that in non-imperatives the ordering is slightly different.

\[(66) \ \begin{cases} \text{lui} \\ \text{leur} \end{cases} \ - \text{soi} - \begin{cases} \text{[PRO]} \\ +P \end{cases} \ - \ V \]

These facts lead Emonds to propose that there is only one base \{PRO\} position which can be filled by a structure preserving movement of any of the clitics in (59)-(61). Le, la, les are preposed before this base PRO position, giving the ordering in (66) by a local rule like (67).

\[(67) \ X - V' \rightarrow \begin{cases} \text{[PRO]} \\ +III \\ -REPL \end{cases} \rightarrow \begin{cases} \text{DEF} \\ \alpha PL \end{cases} + 2 - \emptyset - 3 \]

This rule gives the ordering in (66). Emonds further
provides a local rule inverting the clitics in (59) and (61) when the verb is non-imperative. This rule is reproduced as (68).

\[
\begin{align*}
(68) & \quad \{ \text{le} \} - \text{[PRO] - (PRO) + V} \rightarrow \emptyset - 2 + 1 - 3 \\
& \quad \{ \text{la} \} - \text{[REFL]}
\end{align*}
\]

After this rule applies we derive the orderings in (64).

Emonds' analysis of the French clitic system is of particular interest here because it confirms a prediction of the Association Principles. It is the case in French clitic pronouns that there are apparently non-nested or crossing structures. For example although the structure in (69) is nested that in (70) is not.

\[
\begin{align*}
(69) & \quad \text{il me l'a donne} \quad \text{[NP]}(a) \quad \text{[NP]} \\
(70) & \quad \text{je le lui ai donne} \quad \text{[NP]}(a) \quad \text{[NP]}
\end{align*}
\]

The Association Principles predict that the apparent non-nested configurations must be the result of a local rule that applies after surface structure, that is, after the level at which the Association Principles apply. Emonds' analysis is completely consistent with this requirement imposed by the Association Principles. In Emonds' framework (69) and (70) have the structures in (71) and (72) respectively.

\[
\begin{align*}
(71) & \quad \text{il me a donne le (a)} \quad \text{[NP]}
\end{align*}
\]
(72) je lui a donne le (a) \[ \text{NP}_4 \]

It is only after surface structure that a single local reordering operates to derive (70). We will have occasion to return to French clitics at other points in this study.

3. **Left and Right Quantifier "Movement"**

The grammar of French contains two quantifier movement rules, one moving a quantifier to the right and the other moving a quantifier to the left. The sentences in (73) and (74), Kayne (1975) argues, are related by these rules.

(73)a. Elle a voulu les lire tous.
She wanted to read them all.

b. Elle a voulu tous les lire.
She wanted to read them all

c. Elle a tous voulu les lire.
She wanted to read them all.

(74)a. Les soldats ont tous les deux été présentés à Anne par ce garçon.
The soldiers were both introduced to Anne by this boy.

b. Les soldats ont été tous les deux présentés à Anne par ce garçon.
The soldiers were both introduced to Anne by this boy.

120

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(74)c. Les soldats ont été présentés tous les deux à Anne par ce garçon.
The soldiers were both introduced to Anne by this boy.

d. Les soldats ont été présentés à Anne tous les deux par ce garçon.
The soldiers were both introduced to Anne by this boy.

e. Les soldats ont été présentés à Anne par ce garçon tous les deux.
The soldiers were both introduced to Anne by this boy.

There has been a considerable amount of attention paid to these processes in recent studies (cf., Kayne 1975, Pollock 1978, Quicoli 1976, and Klein 1976). In particular it has been an issue of some contention whether the movement of the quantifier to the right is actually a movement or a construal rule. At the same time it has been debated whether the movement of the quantifier to the left involves a variable or whether it is an iterative local rule. For the purposes of the argument in this chapter however, these issues are somewhat tangential. In this discussion I will assume that the rules in question are in fact movements as in (75a-b).

\[(75)a. \quad X - Q - Y [Q_e] - Z \Rightarrow 1 - 4 - 3 - 2 - 5\]
(75)b. \[ X - [Q_e] - Y - Q - Z \Rightarrow 1 - 4 - 3 - 2 - 5 \]

c. Move \( \alpha \), where \( \alpha \) is a constituent.

The rules in (75) are very similar and probably can be ultimately collapsed together under the rubric of a "move constituent" rule like (75c). However, these speculations go beyond the scope of this study. For convenience I will call (75a) Rightward-Tous movement (or R-Tous) and (75b) Leftward-Tous movement (or L-Tous).

The quantifier movements serve to create structures over which the Association Principles can operate to insure that crossing structures are blocked. The examples which are of interest are those cases in which both L-Tous and R-Tous have applied in the same structure. Consider the phrases in (76)-(77).

\[ (76) \text{les livres, qu'elles ont toutes tous lus } \ldots \]
\[ \text{the books which they have all all read } \ldots \]

\[ (77) \ast \text{les livres, qu'elles ont tous toutes lus } \ldots \]
\[ \text{the books, that they have all all read } \ldots \]

These phrases would have underlying structures resembling (78).

\[ (78) \text{les livres, que [toutes elles] ont lus} \]
\[ [tous WH] \]

L-Tous and R-Tous can apply together in (78) to yield either (76) or (77). However, only (76) is a nested construction and (77) will be blocked by the Association
Principles. To see this, examine the structures in (79)-(80).

(79) \(WH \text{ que } [_q_e] \text{ elles ont toutes tous lus } [_q_e]\)

(80) \(WH \text{ que } [_q_e] \text{ elles ont tous toutes lus } [_q_e]\)

Notice that the Association lines in (80) but not in (79) crossover. For this reason (80) is excluded.

This is a very general phenomenon. It can be observed not only in the sentences in (76)-(77) but also in the sentences below.

(81a) Ils ont tous tout compris.

They have all understood everything.

b. \([_q_e] \text{ ils ont tous tout compris } [_q_e]\)

(82a) *Ils ont tout tous compris.

They have all understood everything.

b. \([_q_e] \text{ ils ont tous tout compris } [_q_e]\)

(83a) Ils prétendent avoir tous tout compris.

They are pretending to have all understood everything.

123
b. \([qe] \text{ ils prétendent avoir tous tout compris } [qe]\)

\[(84)a. *\text{Ils prétendent tout avoir tous compris}
\]

They are pretending to have all understood everything

b. \([qe] \text{ ils prétendent PRO tout avoir tous comprise } [qe]\)

The sentences in \((81)-(84)\) are predictable from the Association Principles. The sentences in \((82a)\) and \((84a)\) are both unacceptable due to the fact that the association lines are crossed as \((82b)\) and \((84b)\) respectively demonstrate. On the other hand \((81a)\) and \((83a)\) are acceptable because the association lines do not intersect as \((81b)\) and \((83b)\) indicate.

Additional evidence from the interaction of these quantifier movements suggests that the general approach of applying the Association Principles to the phenomenon is a productive one. On the one hand the local rule that preposes the clitics \textit{le, la, les} is of some interest because it permits non-nested configurations as \((85)\) shows.

\[(85) \text{ Ils les ont tous toutes embrassées.}
\]

They have all kissed them all.

124
(86) \[ _{[Q]}e \] ils \[ _{[Q]}e \] les ont tous toutes embrassées

(86) represents the association lines that intersect.
Within the framework of the Association Principles, we have seen that the _le-la-les_ rule will operate after surface structure, so that at the level where the Association Principles apply to block certain configurations, a non-nested structure like (86) will not be present. The structure in (87) represents what the surface structure of (85) would look like.

(87) \[ _{[Q]}e \] ils ont tous toutes embrassées \[ _{[Q]}e \] les

Notice that in (87) the association lines do not intersect: nothing blocks the structure and subsequently the local rules can apply to derive (85).

Examine the sentences in (88)-(89).

(88) ?Elle les leur ai tous toutes montrées.

(89) *Elle les leur ai toutes tous montrées.

The Association Principles provide an explanation of these facts. These sentences have surface structures like (90)-(91) respectively.

(90) elle leur ai tous toutes montrées \[ _{[Q]}e \] les

125

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The structure in (91) is blocked because the Association lines intersect. (90), on the other hand, is more acceptable because no such crossing in involved.

No doubt some linguists will attempt to explain the phenomenon here by restricting the structural description of L-Tous. The argument might be made that L-Tous has a structural description like (92).

(92) $X - V - Q - Y \rightarrow 1 - 3 - 2 - 1$

In (92) the quantifier can only be moved to the left of a verb. Since the quantifier cannot move over another quantifier, sentences like (77) would never be generated in the first place. The problem with this argument however lies in the fact that it assumes R-Tous necessarily applies prior to L-Tous. Unless this is the case L-Tous could apply to prepose the quantifier and then R-Tous could operate to derive the crossing structure. The possibility of such an ordering solution does not approach the level of explanation permitted by the Association Principles. The ordering solution increases the class of grammars by permitting ordering: it claims that a grammar with the opposite ordering which would then generate crossed structures is just as likely in linguistic theory as not. For
this reason the attempt to solve the problem of (77), etc.,
by complicating the structural description of L-Tous seems
to be a move in the wrong direction. See Chapter 5 for a
fuller discussion of this topic.

4. Movements of Adverbs

Adverbs also provide some evidence in favor of the
Association Principles. This section details some of that
evidence.

Consider the following sentences.

(93)a. Jane had climbed the wall carelessly.
      b. Jane had climbed the wall once too often.
      c. Jane had climbed the wall carelessly once
too often.

These sentences demonstrate the possibility of having two
adverbs in a single sentence, one manner adverb and one
temporal adverb. It is possible to argue that the struc-
ture of (93c) is as in (94). One argument concerns the
subcategorization imposed by the adverbs. Carelessly
imposes a subcategorization on the verb since it co-occurs
with climb as in (93a) but not in (95).

(95) *Jane {had seen} Mary carelessly.
     {saw}

This fact, given the subcategorization restrictions in
Chapter 2, would indicate that carelessly is a constituent
of the VP. Similarly once too often places a
(94)

Jane had climbed the wall carelessly once too often.
subcategorization on the AUX since, although (93b) is acceptable, (96) is not.

(96) *Jane {is } climbing the wall once too often.
    {was}

Assuming that AUX is a sister to VP, the contrast between (93b) and (96) would seem to argue that once too often was also a sister to VP.

A second argument consistent with (94) is that do so indicates a distinction in constituency between the two adverbs. As was discussed earlier, do so replaces a constituent. The following sentences are of interest in this regard.

(97)a. Jane had climbed the wall carelessly once too often but luckily Mary had done so slowly once again.

b. Jane had climbed the wall carelessly once too often but luckily Mary hadn't done so this time.

c. Jane had climbed the wall carelessly once too often and unfortunately Mary had done so too.

The acceptability of sentence (97b) indicates that the two adverbs are separated hierarchically since otherwise do so would obligatorily replace both adverbs.

A third argument involves VP preposing. It is
possible to prepose a VP constituent as (98) illustrates.

(98)a. John said he would get the job and get the job he did.

b. John said he would give the money to Barry and give the money to Barry he did.

Assuming that VP preposing only moves a constituent, consider the contrast in (99).

(99)a. John said that Jane would climb the wall carelessly once too often, and climb it carelessly she did once too often.

b. *Jane said that Mary would climb the wall carelessly once too often and climb the wall carelessly once too often she did.

In (99b) both adverbs are preposed. The unacceptability of (99b) would be predictable given an underlying structure like (94) since both adverbs do not form a constituent.

Let us assume that the structure in (94) is essentially correct. The following sentences are important for the Association Principles.

(100)a. Jane had climbed the wall carelessly.

b. How carelessly had Jane climbed the wall.

(101)a. Jane had climbed the wall carelessly once too often.

b. *How carelessly had Jane climbed the wall once too often.

130
The sentence in (100b) is acceptable whereas the sentence in (101b) is not. This judgement can be accounted for given the Association Principles and the assumption that (94) underlies (101). Sentence (101b) would have a structure like (102).

\[
(102) \quad \text{\underline{How carelessly} \underline{had} \underline{Jane climbed the wall \underline{e}}}
\]

\[
A_1 \quad A_2 \quad B_1
\]

\[
\text{\underline{once too often}}
\]

\[
B_2
\]

This structure is a crossing one. It will be excluded by the Association Principles since X contains A₂ but not B₂.

Adverbs can be moved to various VP internal positions as well as sentence initially. The sentences in (103)-(104) illustrate some of the positions that adverbs can be moved into.

(103)a. Jane had climbed the wall carelessly.

b. Jane had carelessly climbed the wall.

c. Jane carelessly had climbed the wall.

(\# 103a)

d. Carelessly, Jane had climbed the wall.

(104)a. Jane had climbed the wall once too often.

b. Jane had once too often climbed the wall.

c. Jane once too often had climbed the wall.
d. Once too often, Jane had climbed the wall.

There is apparently a considerable amount of freedom in the positioning of adverbs like the ones we are dealing with here. Sentence (103c) differs from the other sentences in (103) in that it is not understood as a manner adverb. Rather it has what Jackendoff (1972) terms a "speaker oriented" reading and can be paraphrased as "it was careless for Jane to have had climbed the wall" as opposed to "it was in a careless fashion that Jane climbed the wall" which paraphrases the other sentences in (103).

The sentences in (103)-(104) only involve a single adverb. Interestingly enough, however, this freedom is significantly restricted when two adverbs are present in the sentence. Consider sentences like (105)-(106).

(105)a. Jane had climbed the wall carelessly once too often.

b. Jane had carelessly climbed the wall once too often.

c. Carelessly, Jane had climbed the wall once too often. (≠ 105a)

(106)a. Jane had climbed the wall carelessly once too often.

b. Jane had once too often climbed the wall carelessly.
c. Once too often Jane had climbed the wall carelessly.

d. Jane once too often had climbed the wall carelessly.

(107)a. Jane had climbed the wall carelessly once too often.

b. Jane had once too often carelessly climbed the wall.

c. Jane once too often had carelessly climbed the wall.

d. Jane had carelessly once too often climbed the wall. (≠ 107a)

Examine (105) first. Although it is possible to understand the preposed adverb in (103d) in its manner adverb sense, this is not possible in (105c). The Association Principles provide an account for this fact. If (105c) was to have a manner adverb reading, it would have a structure like (108).

(108) \[
\begin{array}{c}
\text{carelessly} \\
A_1 \\
\text{e} \\
B_1
\end{array}
\begin{array}{c}
\text{Jane} \\
A_2 \\
\text{once too often} \\
B_2
\end{array}
\text{climbed the wall}
\]

In (108) the manner adverb reading gives rise to a crossing structure as the factorization in (108) illustrates.

Under the speaker oriented sense, however, no such problem
arises since such adverbs presumably originate sentence initially. The sentences in (105) can be contrasted with the sentences in (106). Although the Association Principles serve to block the manner adverb reading in (105c), the Association Principles are never triggered in (106). In each case the respective adverbs are nested. The sentences in (107) show the Association Principles at work again prohibiting the manner adverb reading for *carelessly* in (107d) but not (107c). This is because in (107c) but not (107d) the adverb *carelessly* can be bound to the manner adverb position without violating the Association Principles. The structures in (109)-(110) corresponding to (107c) and (107d) respectively, deominstrate this fact.

\[ (109) \quad \text{Jane had}\begin{array}{l} \underline{\text{once too often}} \quad \underline{\text{carelessly}} \\ \underline{A_1} \quad \underline{A_2} \end{array}\]

\[ \text{climbed the wall}\begin{array}{c} e \\ B_2 \end{array}\begin{array}{c} e \\ B_1 \end{array}\]

\[ (110) \quad \text{Jane had}\begin{array}{l} \underline{\text{carelessly}} \quad \underline{\text{once too often}} \\ \underline{A_1} \quad \underline{A_2} \end{array}\]

\[ \text{climbed the wall}\begin{array}{c} e \\ B_1 \end{array}\begin{array}{c} e \\ B_2 \end{array}\]

In (109) the variable separating \(A_1\) and \(B_1\) contains both \(A_2\) and \(B_2\): this means that the structure is nested and is well-formed. On the other hand, in (110) the variable

134
separating $A_1$ and $B_1$ contains only $A_2$ and not $B_2$. Consequently (110) is not well-formed and is blocked by the Association Principles.

5. **Clitic Pronouns and the Strict Frame Requirement**

Up to this point the argumentation presented for including movements under the class of rules governed by the Association Principles has involved the prohibition against crossing structures. Further evidence can be found in that bindings do not violate the strict frame requirement either.

In his important and valuable work on French, Kayne (1975) observes the following sentences.

(111)a. Jean est antipathique à Marie.
     Jean is antipathetic to Marie.

b. Jean t'est antipathique.
     Jean is antipathetic to you.

c. Tout le monde croyait Jean antipathique à Marie.
     Everyone thought Jean antipathetic to Marie.

d. *Tout le monde te croyait Jean antipathique
     Everyone thought Jean antipathetic to you.

(112)a. Jean est semblable à Marie.
     Jean is similar to Marie.
b. Jean vous est semblable.
   Jean is similar to you.

c. Moi, je crois Jean tout à fait semblable à ses parents.
   I think Jean quite similar to his parents.

d. *Moi, je vous crois Jean tout à fait semblable.
   I think Jean quite similar to you.

(113)a. Jean est inférieur à Marie.
   Jean is inferior to Marie.

b. Jean lui est inférieur.
   Jean is inferior to her.

c. La plupart des gens croyaient Jean inférieur à Pierrette.
   Most people thought Jean inferior to Pierrette.

d. *La plupart des gens lui croyaient Jean inférieur.
   Most people thought Jean inferior to her.

We can ignore the interesting question of whether a movement rule is responsible for placing the clitic in preverbal position or whether it is generated there in the base and subject to a rule of construal. In the sentences in (111)-(112) it is apparently the case that all pronominal clitics may appear preverbally in ETRE ADJ A constructions while such pronominal clitics are blocked in
CROIRE NP ADJ A constructions. In attempting to account for this phenomenon Kayne suggests that CROIRE NP ADJ A constructions actually derive from complex sentences in which be-deletion has eliminated être. In this way a sentence like (113c) has the underlying structure of (114).

(114) Les plupart de gens croyaient [Jean être inférieur à lui]

In this structure there is according to Kayne (1975) a specified subject intervening between the matrix verb and the pronoun in the embedded S. Any attempt to move the pronoun into preverbal position or, if we assume base generation of clitics, any attempt to construe the preverbal clitic pronoun with a pronoun to the right of à will be blocked by the Specified Subject Condition.

There is a difficulty with this analysis. Consider the following sentences,

(115)a. Jean est fidèle à ses principes.
   Jean is faithful to his principles.

b. Jean y est fidèle.
   Jean is faithful to them.

c. On croit Jean fidèle à ses principes.
   People think Jean is faithful to his principles.

d. On y croit Jean fidèle.
   People think Jean faithful to them.
(116)a. Jean est digne de cet honneur.
    Jean is worthy of that honor.
b. Jean en est digne.
    Jean is worthy of it.
c. Tout le monde croit Jean digne de cet honneur.
    Everyone thinks Jean worthy of that honor.
d. Tout le monde en croit Jean digne.
    Everyone thinks Jean worthy of it.

Notice that in these sentences, the PRO-PP's, y and en, prove capable of fronting in the CROIRE NP ADJ A construction. Of course if these sentences have a structure like (114) it is somewhat difficult to explain this asymmetry: the movement of the PRO-PP's should also be blocked by the Specified Subject Condition. Kayne suggests that the prepositional clitics may be moved to the right out of the embedded sentence prior to movement so that they are then immune to the Specified Subject Condition. Presumably such a movement would be a restructuring rule similar to those proposed in Rizzi (1978). However, Kayne recognizes the ad hoc nature of this suggestion and concludes that the phenomenon must be considered unexplained.

Now there is some reason for supposing that, in fact, no embedded sentence is involved in the sentences (111)-(113) and (115)-(116). In discussing infinitival constructions, Kayne observes that French clitics from an
embedded non-tensed sentence never appear on the matrix verb. This is true of pro-prepositions as well as pronoun clitics. In this regard, observe the following sentences.

(117)a. Elle voudrait le manger.
        She would like to eat it.

   b. *Elle le voudrait manger.
        She would like to eat it.

(118)a. Je tiens à vous revoir.
        I am anxious to see you again.

   b. *Je vous tiens a revoir.

(119)a. Ils essaient d'en avoir.
        They are trying to get some.

   b. *Ils en essaient d'avoir.

(120)a. On va tout y mettre.
        Everything is going to be put there.

   b. *On y va tout mettre.

Let us assume that the Specified Subject Condition is at work blocking the movement of clitics out of the embedded sentences. In these instances Kayne's admittedly ad hoc suggestion for the behavior of the prepositional clitics in (115)-(116) looks even less attractive since it evidently fails to operate in other embedded infinitival sentences. The only reason for assuming an embedded sentence in
(115)-(116) is to invoke the Specified Subject Condition. We see, however, that the croire construction differs from other embedded sentences in that prepositional clitics appear on the matrix verb. Rather than postulating two ad hoc movement rules or somehow weakening the Specified Subject Condition let us see if we can develop a natural account of the sentences in (111)-(113) by assuming that the structure involves no embedded sentence.

Suppose that the structure of the sentences in (111)-(113) and (115)-(116) is something similar to (121).

\[(121) \text{ NP } \begin{array}{c} V \end{array} - \text{NP} - \text{AP}\]

Given this assumption, the Association Principles can predict that the sentences in (111)-(113) are ungrammatical while the sentences in (115d)-(116d) are acceptable. Consider (113d) as a paradigmatic case: it has the structure in (122), assuming that clitic pronouns are generated as part of a non-phrasal verbal complex, following Emonds (1976, 1978). This structure is blocked by the Association Principles. In particular the strict frame requirement is violated in (122). Every phrasal node dominating Jean dominates te and conversely: this means that inasmuch as te and Jean are structurally non-distinct, having \([+N]\) as part of their syntactic feature matrix, they can only be associated with the same element. Yet in (122) these two items are associated with different heads, inférieure and...
Tout le monde croyait Jean inférieure (a) e

(122)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
croire respectively. Consequently (122) is excluded as a well-formed structure.

However, the Association Principles block such structures only where there is "structural ambiguity," where the two categories are structurally non-distinct. In (122) te and Jean are both [+N, -V]; that is, they are non-distinct. In sentences like (115) and (116) where a PRO-PP is fronted however, there is no structural ambiguity since y and en are [-N, -V]. For this reason the Association Principles will not operate to exclude the structure. This provides us with a relatively straight-forward and natural explanation for the contrasts in (111)-(113) on the one hand and (115) and (116) on the other hand without recourse to various ad hoc measures.

6. "Optional" Subcategorization

The discussion of the Association Principle has proceeded as if all associations were obligatory. That is to say we have examined situations where there are association rules like (123) but not (124).

\[
\text{(123) A, \underline{____} B}
\]
\[
\text{(124) A, \underline{(B)}}
\]

It is important at this point to consider what effects "optionality" has on the Association Principles and whether they are in need of revision.

There is an interesting construction in both
English and French which has some light to shed on this issue. In the following sentences, it is demonstrated that be and être do not take prepositional complements, whereas seem and sembler do.

(125)a. Mary is courteous to strangers.
   b. *Mary is courteous to strangers to me.

(126)a. Mary seems courteous to strangers.
   b. Mary seems courteous to strangers to me.

(127)a. Marie est fidèle à ses parents.
   b. *Marie m'est fidèle à ses parents.

   Marie is faithful to her parents.

(128)a. Marie semble fidèle à ses parents.

   Marie seems faithful to her parents.
   b. Marie me semble fidèle à ses parents.

   Marie seems faithful to her parents to me.

Notice that in both French and English the prepositional complements to seem or sembler need not appear lexically, that is, they are optional as the contrasts between (126a-b) and (128a-b) illustrate.

Now there is a very interesting phenomenon which is exhibited in clitic placement in French and PP-preposing (or topicalization) in English.

(129)a. Mary is courteous to strangers.
   b. To strangers, Mary is courteous.
(130)a. Mary seems courteous to strangers.
   (ambiguous)
   b. To strangers, Mary seems courteous. (PP
      is a complement to seem)

(131)a. Mary seems courteous to strangers to me.
   b. To me, Mary seems courteous to strangers.
   c. To strangers, Mary seems courteous to me.
   (# 131a)

In (129)-(131) we find that it is possible to prepose a
preposition which is a complement to an adjective just in
those cases where the verb does not optionally select a
prepositional complement. In other words, nothing prevents
preposing to apply to the PP complement of an adjective
when the verb is be, but when the verb is seem and can
select its own PP complement, the preposed PP is inter-
preted as the complement to the verb, not the adjective. A
similar set of judgements are produced when clitic place-
ment is applied to the French examples involving être and
sembler.

(132)a. Jean est fidèle à ses parents.
   Jean is faithful to his parents.
   b. Jean leur est fidèle.
      Jean is faithful to them.

(133)a. Jean semble fidèle à ses parents.
   (ambiguous)
Jean seems faithful to his parents.

b. Jean leur semble fidèle. (_leur_ is the complement to _sembler_)

Jean seems faithful to them.

These contrasts are very general in both languages. The fact that different rules in the grammars of French and English are involved indicates that the phenomenon is not a property of the rules themselves, but rather a property of linguistic theory. In addition, the possibility that the phenomenon is somehow a reflex of the fact that in the examples considered so far both prepositions are _to_ is negated by the examples below where other prepositions are involved and yet the judgements exhibit the same patterning.

(134)a. John seemed sympathetic with the prisoners to me.

b. To me, John seemed sympathetic with the prisoners.

c. *With the prisoners John seemed sympathetic to me.

(135)a. John seemed grateful for the money to me.

b. To me, John seemed grateful for the money.

(c. For the money, John seemed grateful to me.

(136)a. John seemed afraid of his boss to me.

b. To me, John seemed afraid of his boss.

c. *Of his boss, John seemed afraid to me.

145
This rather complex set of data is significant evidence in favor of the Association Principles governing movements. First consider the difference between (129) and (131), avoiding (130) for a moment. Let us assume that be is subcategorized as in (137) while seem is subcategorized as in (138).

(137) be, + _____ AP
(138) seem, + _____ AP (PP)

The statement in (137) is straightforward. The statement in (138) requires that seem be inserted in the environment of an AP and an optional PP. Compare now the surface structures of (129b), (131b), and (131c), represented here as (139)-(141) respectively.

(139) [to strangers] Mary is courteous [e]

(140) [pp to me] Mary seems courteous [pp to strangers] [ppe]

(141) [pp to strangers] Mary seems courteous [pp e]

Sentence (139) presents no problems, and is grammatical, because it involves only a single association line and

146
nothing is crossed. In (140) there are multiple association lines, one between the PP and its trace as well as between the adjective and its propositional complement. Yet in (140) the sentence is grammatical because these association lines are nested and do not intersect. More formally the variable X containing $A_2$ also contains $B_2$. The situation is different in (141) however: in (141) the association lines between the PP and its trace on the one hand, and the verb and its PP complement on the other hand, crossover. That is to say, although X contains $A_2$ is does not contain $B_2$. Because the association lines crossover, the structure is blocked by the Association Principles.

Consider now the judgements in (130). Here, when the PP complement is to the right of the adjective, the structure is ambiguous and the PP can be interpreted as a complement to the verb or to the adjective. When the PP is preposed, however, it can only be interpreted as being complement to the verb and not the adjective. Suppose that optionality is not expressed in subcategorization: instead the optionality of an element is a consequence of the more general optional expansion of a phrasal node to PRO. In this way **seem** would be obligatorily subcategorized for an adjective phrase and a prepositional phrase, although it would permit the PP to dominate a PRO with arbitrary reference (indicated as $\text{PRO}_1$). In this way (130a) would have surface structure like (142) or (143).

147

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(142) Mary seems courteous [PRO₁][to strangers]

(143) Mary seems courteous [to strangers] [PRO₁]

It follows from this analysis that a string like (130b) could not be like (141), i.e., derived from (143) where the lexical PP is a complement to the verb rather than to the adjective. In order to see this point consider (144)-(145) which corresponds to (142)-(143) respectively, where PP-preposing (or topicalization) has applied.

(144) [to strangers] Mary seems courteous [PRO][e]

(145) [to strangers] Mary seems courteous [e][PRO]

Notice that of (144)-(145), only (144) is well formed given the Association Principles inasmuch as the association lines in (145) crossover. Claiming that "optional" subcategorizations really only involve non-lexical PRO's thus seems (a) to resolve any complication of the Association Principles, (b) constrain the mechanism of subcategorization, and (c) account for the facts in (129)-(131) which would be anomolous otherwise.

The hypothesis that subcategorization permits only obligatory statements and that cases of seeming optionality are actual cases of abstract PRO's with arbitrary reference alternating with lexical material goes a long way in
explaining the French data in (127)-(133). Consider the adjective *fidèle*. This adjective may take as a complement à + NP as evidenced by (127a). This fact can be accounted for by stipulating either (146a) or (146b).

(146)a. fidèle, +A, + _____ {NP {PRO₁}}

b. fidèle, +A, + _____ à + {NP {PRO₁}}

The decision of which statement is correct depends on whether one analyzes the dative à as present in underlying structure or inserted at some later point: if we assume (146a) the grammar of French will require an à insertion rule (as is argued in Kayne 1975) in order to account for (127); if on the other hand we assume (146b) in our description of French we will require an à deletion rule to account for (127b). Whichever assumption is made, the analysis forwarded here is unaffected. Let us assume (146b). Now in (127) no problems arise because *être* is not subcategorized like (146). In (147) however, we find that in addition to (146), *sembler* is also subcategorized to select a complement in addition to the AP. In particular, it has the restriction in (147).

(147)a. sembler, +V, + _____ AP {NP {PRO₁}}

b. sembler, +V, + _____ AP à + {NP {PRO₁}}
Once again whether we adopt (150a) or (150b) depends in part on how we analyze the appearance of à.

Under these assumptions the sentence in (133a) is ambiguous and has the structure of either (148) or (149).

\[
\text{(148)} \quad \text{Jean semble fidèle à ses parents (à) } [\text{NP}_e] \\
\text{(149) } \quad \text{Jean semble fidèle (à) } [\text{NP}_e] \text{ à ses parents.}
\]

The association lines indicate which prepositional phrase is the complement of which verb or adjective. In each case we note that the structures are nested and do not contravene the Association Principles. However, unlike (133a), (133b) is unambiguous: the clitic pronoun must necessarily be interpreted as the complement of the verb and not the adjective. This constraint on the interpretation of the clitic follows directly from the Association Principles.

Consider the structures in (150)-(151).

\[
\text{(150)} \quad \text{Jean leur semble fidèle (à) } [\text{NP}_e] \text{ (à) } [\text{NP}_e] \\
\text{(151) } \quad \text{Jean leur semble fidèle (à) } [\text{NP}_e] \text{ (à) } [\text{NP}_e]
\]

It is clear that the bindings in (150) which represent the interpretation in which the clitic is the complement of the adjective is a crossing structure and is blocked. Only the structure in (151) is consistent with the Association Principles.
Principles and it is precisely this structure which corresponds to the interpretation in which the clitic pronoun represents the complement to the verb. In this way there is a generalization to be captured between the interpretation of the clitic when only one complement is present, as in (133b), and the interpretation of the clitic when both complements are present as in (128b).

The sentence in (128b), like (133b), can only be interpreted in a way that takes the clitic to be the complement of the verb and not the adjective. This follows from the Association Principles: in this respect, the relevant structures parallel (150)-(151).

(152) Jean me semble fidèle (à) [NP] e à ses parents.

(153) Jean me semble fidèle à ses parents (à)[NP] e

The structure in (152) where the clitic represents the complement of the adjective crosses association lines and is, therefore blocked by the Association Principles. On the other hand (153) is consistent with the Association Principles because the structure is nested.

It is important to recognize that first the interpretation of the clitic pronoun in (128) follows from the Association Principle. This means that we have added support for extending the Association Principles to the
class of rules involved in clitic placement facts. It is also important to observe that postulating the abstract PRO's with free indexes for the sentences in (133) allows us to explain why those sentences show the same pattern of association as do the sentences in (128). These explanations contrast sharply with the present state of linguistic theory where such facts seem totally anomalous.
Chapter 4

THE CLAUSE BOUNDEDNESS OF CONDITIONS

IN LINGUISTIC THEORY

In this chapter, bindings over domains larger than a single clause (i.e., $S^1$) are considered. The potential competition between the Association Principles and Chomsky's Specified Subject Condition is dealt with by restricting the application of the Association Principles to a single $S^1$ domain. It is further suggested that all the conditions in Universal Grammar are restricted to a single $S^1$ domain. Chomsky's latest proposals in "On Binding" are considered in this respect. Several revisions are proposed in the "On Binding" framework with the consequence that the other conditions on grammar are also restricted to a single $S^1$ domain. Finally it is shown how restricting the Association Principles to a single $S^1$ domain has the additional advantage of removing the major counter-examples to any condition requiring nested configurations.

1. Limiting the Association Principles to a Single Clause

One of the major factors impeding progress into the understanding of a condition requiring nested configurations is the fact that when one examines sentences
involving domains larger than a single $S^1$, the condition requiring nesting makes some of the same predictions as Chomsky's Specified Subject Condition. The sentences in (1) are important in this respect.

(1)a. John$_1$ asked the men$_2$ [$S$,PRO$_2$ to visit each other$_2$]

b. *John$_1$ promised the men$_2$ [$S$,PRO$_1$ to visit each other$_2$]

c. The men$_1$ asked John$_2$ [PRO$_2$ to visit each other$_1$]

The sentences in (1a) and (1b) are accounted for by either the SSC or the Association Principles. The sentence in (1c) however is only blocked by some formulation of the SSC and not the Association Principles inasmuch as the association lines for (1c) would be nested.

It appears to be the case then that there is a considerable amount of overlap between the Association Principles and the SSC but that this overlap is not complete. The fact that the Association Principles do not block (1c) is not evidence that they are flawed, nor is the fact that the Association Principles make the same predictions as the SSC with respect to (1a-b) evidence that the SSC is flawed. It has sometimes proven a useful research strategy however to reduce the redundancy of various parts of the grammar. This general strategy has prompted various
linguists to reduce this redundancy by eliminating either the SSC or the Association Principles. The possibility of altering the conditions themselves in order to reduce this redundancy has not been attempted up to this point.

In Chomsky's formulation of the SSC, it is always mentioned that the two terms related by the rule of grammar must be separated by a cyclic domain α. Inasmuch as most of the evidence produced to motivate the Association Principles involves terms within the same minimal cyclic domain, it seems possible to reduce the overlap between the conditions by restricting the Association Principles to a single minimal cyclic domain. This revision eliminates the redundancy that we observed between the Association Principles and the SSC. However we might legitimately ask why the Association Principles are restricted to a single minimal cyclic domain whereas the SSC, apparently is not. The next section shows that the SSC can be replaced by conditions consistent with the generalization in (2) which holds for the Association Principles.

(2) C and C' satisfy a condition in linguistic theory if and only if every cyclic node containing C contains C', and conversely.¹

2. Other Principles of Universal Grammar Limited to a Single Clause

2.1. Case Concord and the Function of the Passive Morphology

In this section, Chomsky's latest proposals

155

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
altering somewhat his formulation of the SSC and the Tensed-S Condition are considered. Several revisions are proposed which enable a unified formulation of these conditions and which hold only over a minimal cyclic domain. The fact that the SSC apparently only involves terms in distinct cyclic domains will be a consequence of the indexing convention adopted from Chomsky (1978). In this way I will be arguing that (2) is a generalization about all conditions in Universal Grammar and not simply a subset of the conditions.

Chomsky (1978) is a major attempt at bringing out larger generalizations within the ambitious framework of core grammar set out in Chomsky (1976, 1977b) and Chomsky and Lasnik (1977). In particular, it aims, by introducing certain abstract case features, to reduce the overlap between the Specified Subject Condition and the Tensed-S Condition on the one hand, and to simplify the "NP to VP" filter of Chomsky and Lasnik (1977), considering infinitival relatives in a detailed fashion, on the other hand.

Chomsky's analysis of case marking derives some independent motivation from its ability to deal with certain anomalous facts surrounding the passive construction. The particular direction of Chomsky's analysis necessitates recognizing certain unnatural restrictions on the application of case assignment and is unable to capture certain important generalizations. I want to consider Chomsky's
analysis here and suggest some revisions overcoming these difficulties which I will characterize more fully shortly.

A problem often noted by linguists is the apparent asymmetrical behavior of the move-NP rule creating the passives in (3)-(6).

(3)a. John saw Mary.
   b. Mary was seen \( \text{[NP}_e \) (by John).

(4)a. Washington has slept in this bed.
   b. This bed has been slept in \( \text{[NP}_e \) (by Washington).

(5)a. John gave a book to Mary.
   b. A book was given \( \text{[NP}_e \) to Mary.
   c. *Mary was given a book to \( \text{[NP}_e \).

(6)a. Mary sat \{near \ \{under \ this car for three hours.
   b. *This car has been sat \{near \ \{under \ \text{[NP}_e \) for three hours.

It has often been suggested that the rule responsible for these structures is something like (7). Emonds (1976), for example, presents such an analysis.

\[
\begin{array}{cccccccc}
X - & \text{NP} - & V - & (P) - & \text{NP} - & Y \Rightarrow & 1 - 5 - & \text{be+en+3} - 4 - \emptyset - 6 \\
1 - 2 - 3 - 4 - 5 - 6
\end{array}
\]

Under this hypothesis (3)-(5) are the unmarked cases whereas (6) is marked and must somehow be blocked. However

157
within the framework of a core grammar like the one envisaged in Chomsky (1978), the limitations of the structural description present in (7) must follow from principles of linguistic theory or from surface output filters.

In Chomsky (1978) sentences like (3)-(6) are discussed within a core grammar framework. Chomsky's argument is that recognizing the presence of abstract case permits the general move-NP rule to account for the phenomenon at hand without sacrificing the more general claims to explanatory adequacy implicit in the core-grammar framework.

Chomsky (1978) proposes that abstract case be assigned as in (8).

(8)a. NP in the domain of TENSE = nominative
    b. NP in the domain of V = objective
    c. NP in the domain of P = oblique

Recall that we generally say that a node A is in the domain of a node B if the first branching node C dominating B dominates A as well. In addition to postulating the rules in (6) Chomsky also assumes that oblique case is assigned at deep structure while nominative and objective case are assigned at the end of every S domain. A filter excludes NP's not having been assigned a case. Given these assumptions, the following account of the data in (3)-(6) can be provided. In both (3b) and (5b), every NP will have a single case marking and no difficulties arise. In (6b) and (5c) however the NP's in the subject position will be
assigned oblique case in the base and then nominative case at the end of the S cycle. Because the NP's have double case marking the structures are excluded. We will return shortly to the issue of double case assignment and case discord in a moment. For Chomsky (3) and (5)-(6) are the unmarked cases whereas the example in (4) is more marked. In this instance Chomsky proposes that slept in will be lexically marked not to assign oblique case to the object of the preposition. Consequently the passive movement will be possible in (4) just as it is in (3), but not in (5) or (6).

Several important issues arise when one looks closely at Chomsky's analysis. On the one hand, assignment of case seems somewhat unnatural and we would prefer a theory in which case assignment is done in a uniform manner. In addition, the mechanism blocking double case assignment is left unspecified. Finally it is unclear what blocks movements like (9) in core grammar.

(9) John$_2$ knows e$_2$

(9) differs from (3) only in the absence of the passive morphology. In other words, we want to know what principle of Universal Grammar makes the passive morphology obligatory when NP-movement operates.

Suppose that we were to assume that case was assigned by (8) at the end of every cyclic domain, oblique as well as nominative and objective case. Take S$^1$, S, NP,
and PP as cyclic categories. We have already noted the "common-sense" requirement that no category may be assigned two distinct cases. In order to formalize this intuition, I propose that Universal Grammar contains a condition like (10).

(10) The Case Concord Condition
    Given $C^k_i$ and $C^l_i$ where $i$ is a referential index the case of $C^l_i$ must be non-distinct from the case of $C^k_i$ if $C^l_i$ and $C^k_i$ do not contain distinct lexical material.

This restriction permits us to block examples like (11)-(12) while admitting examples like (13).

(11) * [this car] was sat [near] [e]  
     +nom        +obl  
     [2]          [2]

(12) * [John] knows [PRO]  
     +nom  
     [2]  
     [+obj 2]

(13) [the women] know [each other]  
     +nom  
     [2]  
     [+obj 2]

The examples in (11) and (12) are excluded because the two elements bound together have distinct case markings and are non-distinct lexically. The example in (13) differs from that in (12) because in (12) $C_i$ and $C_j$ do not contain distinct lexical material while in (13) they do. This means that condition (10) is inapplicable to (13), although it is still applicable to (12).

Instances of WH-movement pose no particular
problems for (10). Consider the simple sentence in (14).

\[(14) \text{who} \quad \text{did} \quad \text{Mary} \quad \text{visit} \quad \text{e}\]

Case is assigned on the $S$ cycle, then WH-movement applies on the $S^1$ cycle, leaving a trace which, I assume, is marked with the same case as the moved constituent. The binding represented in (14) is permitted by (10) since the cases will not be distinct.

The Case Concord Condition in (10) is not quite right yet however. In particular we want to know what prevents bindings like (15).

\[(15)a. \text{Jane}_2 \text{ wants PRO}_2 \text{ to know PRO}_2\]
\[b. \text{Jane}_2 \text{ expects [Bill] to know [e]NP}_2\]

In (15a) the possibility of generating two PRO's, each co-indexed, is represented. In (15b) the possibility of moving a case marked NP to the subject of the infinitive where it is assigned case by the verb, as suggested in Chomsky (1978).

In order to avoid these problems let us hypothesize that the cases we have been dealing with are in fact decomposable into the following feature matrices.

\[(16)a. \text{nomina}tive = [\text{+case} \quad [\text{+objective} \quad [\text{+oblique} ] ] ]\]
b. objective = [+case
+objective
-oblige]
c. oblique = [+case
-objective
+oblige

NP's will be [-case
-oblige
-objective] at the output of the rewriting rules. The use of the features outlined in (16) are motivated by the fact that they allow us to extend our analysis to possessive (or genitive) NP's. Of course such NP's require case or they will be filtered out by the no-case filter discussed in Chomsky (1978). However, no proposals have been advanced for the treatment of case assignment to such NP's. We will assume that the case assignment rules are as in (17).

(17)a. NP = [+case] in the domain of a grammatical formative (e.g., TENSE or POSSESSIVE) or a lexical formative (e.g., V or P)
b. NP = [+objective] in the domain of V
-oblige
c. NP = [-objective
+oblige
in the domain of P
d. NP = [-objective
-oblige
in the domain of N'
e. NP = [+objective
+oblige
in the domain of VP

In this framework possessive induces case to the subject of an NP just as tense does to the subject of a tensed clause. I will assume that the so-called raising verbs merely assign the feature [+case] to the subject of the infinitive

162

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
resulting in a matrix like \( \begin{bmatrix} +\text{case} \\ +\text{obj} \\ +\text{obl} \end{bmatrix} \). Furthermore, I assume
\( \begin{bmatrix} +\text{case} \\ +\text{obj} \\ +\text{obl} \end{bmatrix} \)
that the rules in (17) apply at the end of every cycle.

The rules that we have postulated will now give the matrices in (18) for the sentences in (15).

(18)a. Jane wants \( \begin{bmatrix} \text{PRO} \\ +\text{case} \\ +\text{obj} \\ +\text{obl} \end{bmatrix} \) to know \( \begin{bmatrix} \text{PRO} \\ +\text{case} \\ +\text{obj} \\ -\text{obl} \end{bmatrix} \)

b. Jane expects \( \begin{bmatrix} \text{Bill} \\ +\text{case} \\ +\text{obj} \\ +\text{obl} \end{bmatrix} \) to know \( \begin{bmatrix} e \\ +\text{case} \\ +\text{obj} \\ -\text{obl} \end{bmatrix} \)

Because an NP with the feature [+case] can bind PRO's with the feature [-case], we need to revise (10) slightly to (19).

(19) Case Concord

Given \( \begin{bmatrix} C^k_i \\ +\text{case} \\ +\text{obj} \\ +\text{obl} \end{bmatrix} \) and \( \begin{bmatrix} C^l_i \\ +\text{case} \\ +\text{obj} \\ +\text{obl} \end{bmatrix} \), where \( i \) is a referential index, \( a = \gamma \), and \( \beta = \phi \), if \( C^k_i \) and \( C^l_i \) do not contain distinct lexical material.

The Case Concord Condition will now block the sentences in (15). This is due to the fact that when we examine their case features, as in (18), we find a conflict in the value for the feature \( \pm \text{oblique} \) in violation of the requirement in (19). Moreover, the Case Concord Condition is

163
consistent with the requirement (2) which limits conditions
to a single minimal cyclic domain. There is no need to
permit the Case Concord Condition to operate over a larger
domain.

Return now to the passive facts. I am claiming
that (5c) as well as (8) are unacceptable because distinct
cases end up being assigned to the subject and object NP's,
thus violating (19). The issue at hand is why passives do
not yield structures like (20), that is, structures also
violating (19).

\[
\text{(20)} \quad \text{Mary} \quad \text{was seen} \quad \text{e}
\]
\[
\begin{array}{c}
\text{+case} \\
\text{+obj} \\
\text{+obl} \end{array}
\begin{array}{c}
\text{e} \\
\text{+case} \\
\text{+obj} \\
\text{+obl} \end{array}
\]

The condition in (19) seems to require that objective case
not be assigned to the NP to the right of the verb. This
would most naturally be the result if the NP to the right
of the verb is not in the domain of \text{see} in (20), that is,
if the NP is not dominated by the first branching node
above \text{see}. If the NP is not assigned case, the NP movement
will be able to apply without violating (10). The passive
verbal morphology is important in this respect. Compare
(20) to (21) where no such morphology appears.

\[
\text{(21)} \quad \star \text{Mary} \quad \text{saw} \quad \text{e}
\]
\[
\begin{array}{c}
\text{+case} \\
\text{+obj} \\
\text{+obl} \end{array}
\begin{array}{c}
\text{e} \\
\text{+obj} \\
\text{+obl} \end{array}
\]

164
(21) will be excluded by (19): the object is in the domain of see and consequently undergoes case assignment. When NP movement applies, the grammar yields bound NP's with conflicting case assignment and the sentence is blocked. I am claiming that the passive morphology is required in order to prevent the object from being assigned objective case. This will be achieved if the structure of the passive is something like (22).

(22) \[ S_{NP} [V_{y\text{be}} \text{en} [V\text{see}]] \text{Mary} \]

(23)

\[
\begin{array}{c}
S \\
NP \\
V \\
AP \\
\text{Mary} \\
\text{was} \\
\text{see} \\
\text{e}
\end{array}
\]

In (22) Mary is outside the domain of the V's and hence is not assigned case. One possible proposal to make the internal structure of (22) more concrete is (23), which is a variant of the structure proposed in Chomsky (1977a).

In summary, I am arguing that passive verbs are verbs which prohibit their objects from receiving case marking by the rules in (17). This lack of case marking has two functions: (a) it makes NP-preposing obligatory, since otherwise the object will not have case and will be

165
excluded by Chomsky's no-case filter; and (b) it makes co-
indexing between subject and object possible, since if the
object has case a violation of the Case Concord Condition
results.

2.2. Unifying the Nominative Island Condition and the
Opacity Condition

Another function that abstract case performs in
Chomsky (1978) is to permit a reduction of the overlap
between the Specified Subject and Tensed-S Conditions. In
a sentence like (24), both the Specified Subject and
Tensed-S Conditions block the relevant binding.

(24) The men2 said that Mary saw each other2.

In Chomsky (1978) the range of the Tensed-S Condition is
restricted: since TENSE conditions the assignment of the
nominative case, Chomsky modifies the Tensed-S Condition
to (25) and accounts for any violation of the Tensed-S
Condition not involving a nominative (e.g., 24) by the
Opacity Condition in (26).

(25) A nominative anaphor in S cannot be free in
S1 containing S.

(26) If α is an anaphor in the domain of the sub-
ject of β, β minimal, then α cannot be free
in β where β = NP or S1.

Chomsky says that "an anaphor α is bound in β if there is:
a category c-commanding it and co-indexed with it in β."
otherwise α is free in β."

The Nominative Island Condition (25) will block sentences like (27) from being well-formed.

(27) The men said \( [S, \text{that} [\text{each other} \_2^{+\text{nom}}] \_2^{+\text{nom}}] \) was hungry

Here each other is free in \( S^1 \) since there is no category c-commanding it in β which also bears its index. Moreover it carries nominative case. Consequently it falls within the domain of (25) and is excluded.

There is a problem associated with (25) however. Chomsky recognizes that (25) would have a more natural formulation as (28).

(28) A nominative anaphor α cannot be free in \( S^1 \)

This formulation cannot be adopted however because of the problem presented by sentences like (29), under the assumption that Chomsky ultimately accepts in which the case of the moved constituent is carried along.

(29) I wonder \( [S, [\text{who} \_2^{+\text{nom}}] [S, \text{e} \_2^{+\text{nom}}] \text{ was Mary} \)

Notice that (28) but not (25) will block (29). Who in (29) is free in \( S^1 \); if it carries case, as we assume it does, then it violates (28). I want to propose a revision in (25)-(26) which avoids these problems.

In Chomsky (1978) the following proposal is

167
forwarded to make explicit the indexing of categories. At
deep structure every lexical category is assigned a refer-
ential index, i, i greater than 1, from left to right.
The integer 1 is reserved for arbitrary reference. In
addition, these categories are assigned anaphoric indices,
represented as a set of integers \{i, j, \ldots\}, indicating
which indices the category cannot be interpreted as co-
referential with. That is to say, the anaphoric indices
\{i, j, \ldots\} note that the category in question is disjoint
in reference from categories with the referential index,
\{i, j, \ldots\}. For example, the sentence in (30) illustrates
how these indices are used.

(30) The men$_2$ said that Bill$_3^{\{2\}}$ liked Mary$_4^{\{2,3\}}$

Binding of two categories involves suppressing integers in
the anaphoric index. This can be seen in (31) where the
binding of the pronoun with the man is possible because 2
is suppressed in the anaphoric index of the pronoun.

(31) The man$_2$ said that Bill$_3$ liked him$_4^{\{3\}}$

In this framework anaphoric elements, like each other, are
assigned only referential indices by construal rules like
Chomsky's (1978) rule of COINDEX. This means that a sen-
tence like (32a) has the indexing in (32b).

(32)a. John promised Mary to show up at nine
b. John$_2$ promised Mary$_3$$_{[2]}$ [PRO$_2$ to show up at nine]

The referential index of an anaphor and the anaphoric index of a pronoun are termed "designated indices" by Chomsky (1978). The designated index falls in the domain of the binding conditions and captures properties that anaphors and pronouns have in common. The principal property that pronouns and lexical NP's have in common is the manner in which they are indexed.

With this much background, I would like to propose replacing (25)-(26) with (33).

(33) Index Suppression Condition

Erase a designated index, i, of C where i > 1 and C is [+case], if and only if i appears nowhere else in the same minimal α as a referential index.

Let α in (33) range over NP and S$_1$. We will say that C$_i$ and C$_j$ are in the same minimal α if every α that contains C$_i$ also contains C$_j$ and conversely. Like (25)-(26), (33) erases the designated indices of elements blocking disjoint reference (and thus permitting co-reference in the case of pronouns) or eliminating indices of anaphors thus producing open sentences which I assume will be blocked at the level of Logical Form for independent reasons. (33) differs from (25)-(26) principally in altering the notion "free." In
(33), an index is "free" either if there is no co-indexed category which c-commands it or which it c-commands in the same minimal α. Let us examine how (33) operates.

Consider pronouns first. The sentences in (34b) and (35a) are underlying structures prior to the operation of (33).

(34a). Mary$_2$ told Susan$_3$$_2$ [that Jane$_4$$_2$ liked her$_5$$_{2,3,4}$]

b. Mary$_2$ told Susan$_3$$_2$ that Jane$_4$$_{2,3}$ liked her$_5$$_{4}$

(35a). Mary$_2$ told Susan$_3$$_2$ [that she$_4$$_{2,3}$ liked Jane$_5$$_{2,3,4}$]

b. Mary$_2$ told Susan$_3$$_2$ that she$_4$ liked Jane$_5$$_{2,3,4}$

The (b) sentences show the effects of (33). In (34b) we are able to erase 2 and understand her as co-referential with Mary because 2 appears nowhere else in the relevant domain as a referential index. The same explanation is true in (35b). It thus seems that (33) is capable of accounting for the facts surrounding disjoint reference.

Examine now the facts that are accounted for by the Nominative Island Condition. The sentences in (36)-(37)
are blocked by the Nominative Island Condition.

(36)a. John observed the dog was hungry.

b. *The dog₂ was observed [₁[Npₑ]₂ was hungry]

(37) *The men₂ observed [₁ that [each other]₂ was hungry]

In each example the anaphor indexed as 2 occurs in a minimal $S₁$ domain which fails to exhibit another instance of 2 as a referential index: consequently (33) erases the index of the anaphor, resulting in an open sentence.

Compare the sentence in (37) to the sentence in (38).

(38) The women₂ want $[₁PRO₂$ to visit Mary₃{2}]

Condition (33) will not operate in (38) to erase the index of the PRO even though 2 appears nowhere else in the relevant $S₁$ as a referential index. This is because the PRO in (38) does not carry case. In this respect (33) captures a generalization between (38) and (39).

(39) *The men₂ want Bill₃{2} $[₁PRO₃$ to visit

\[
\{\text{each other}\} \quad \{\text{PRO}\} \quad ₂
\]

In Chomsky's account (39) is excluded because the rightmost anaphor is free in the domain of the subject while (36)-(37) are excluded due to some fact concerning their case marking. That is, the reason for their
ungrammaticality is related to completely different syntactic facts. Given (33) however, (39) is no different from (36)-(37): that is, the index of the anaphor is erased because it appears nowhere else in $S^1$ and its associated constituent bears case. The acceptability of (38) is correlated directly to the fact that the lack of TENSE in the embedded S prevents case from being assigned to PRO, leaving it outside the province of (33).

Consider now the sentences in (40) which have been somewhat central to the development of the Specified Subject and Opacity Conditions since Chomsky (1973).

(40)a. They expect each other to win.

b. *They were quite happy for each other to win.

c. They were quite happy for pictures of each other to be on sale.

Normally, tenseless sentences assign no case to their subject NP's and consequently only a PRO can appear as an infinitival subject without violating the condition that all lexical NP's must exhibit case. The sentence in (40a) must have the structure in (41) then.

(41) They expect each other$_2^{\ S_1[NP]\ e}$ to win$_2^{\ NP}$

In this structure each other will be assigned objective case and the index of the empty element will not be erased.
because it does not bear case. The sentence in (40b) is somewhat different. The structure of this sentence is as in (42).

\[(42) \text{*} \text{They}_2 \text{ were quite happy } [s_1 \text{for } [s_\text{each other}_2 \\
\text{to win}]]\]

Following Chomsky (1978) I will assume that for in (42) assigns case to each other. This means that the unacceptability of (42) cannot be attributed to a lack of case marking on the embedded subject and in fact sentences with other NP's as the subject, like (43), are acceptable.

\[(43) \text{They were quite happy for Bill to win.}\]

The explanation for the unacceptability of (40b) derives from the condition in (33). That is, the index of each other in (42) will be eliminated by (33) since no other element with the same referential index appears in the relevant S¹ domain. Such an operation results in an open, and consequently ill-formed, sentence.

The example in (40c) is slightly more complex. In this sentence, each other is assigned oblique case. In sentences like (44), condition (33) will operate to erase the index of each other, causing the sentence to be open and therefore excluded.

\[(44a) \text{*They}_2 \text{ were quite happy for } [\alpha \text{John's}_3 \text{pictures}_4 \text{ of each other}_2 ] \text{ to be on sale}\]
b. *They\textsubscript{2} were quite happy for \([_\alpha\text{the pictures of each other}_{2}]\) to be on sale

Although Chomsky (1973) distinguishes some degree of difference in the acceptability of (44a-b), he notes that both are worse than sentences like (40c). Rather than attempt to explain the difference between (44a-b), I will concentrate on their less acceptable status with respect to (40c). As I mentioned above, (33) will operate on the sentences in (44) to block both sentences. The question quite naturally arises as to how \underline{each other} in (40c) prevents its index from being eliminated by (33), making it different from the sentences in (44). The conditions proposed so far would seemingly require a PRO to be present in the NP. I suggest then that (40c) has the structure of (45).

(45) They\textsubscript{2} were quite happy for \([_\alpha\text{PRO}_{2}

pictures\textsubscript{3}\textsubscript{2}] of each other\textsubscript{2}] to be on sale

The condition (33) will be unable to erase the index on either the PRO or \underline{each other} because in either case there is another NP in the relevant domain with the same referential index.

Several types of phenomena provide support for the general direction of research just outlined. Besides permitting us to collapse the Opacity and Nominative Island

174
Conditions, the general framework defended here has the advantage of eliminating the awkwardness of the Nominative Island Condition that we noted above. Recall that although it was desirable to formulate the Nominative Island Condition as (46) sentences like (47) prevented such a move.

(46) A nominative anaphor cannot be free in $S^1$

(47) \[\text{who} \quad \text{do you think } [S' [^e_2] [S [^e_2 \text{ saw Bill}]]^+\text{nom}^+\text{nom}^+\text{nom}]^+\text{nom}]^+\text{nom}]

If movement leaves case behind on the trace, as Chomsky assumes it does, then in $S^1$ the trace of WH leaves a free anaphor in the COMP of the embedded $S^1$. This fact prevents simplifying the Nominative Island Condition as (46). However, the framework assuming (33) avoids this problem entirely: the index of the trace in COMP will not be eliminated since another NP with the same referential index is present in the relevant domain.

We need to now consider in more detail the process of WH-movement. R. Friedin (personal communication) has pointed out to me that Chomsky's (1976) discussion of the so-called cross-over facts seems to suggest that the trace of a WH element has a different status from other traces. It has often been noted that (48) differs from (49) in that in (49) who and he cannot be co-referential although in (48), he and him can be co-referential.
(48) The woman he loved betrayed him.

(49) Who did the woman he love betray?

This fact seems to be related to the property that WH elements are not referential, as Chomsky (1976) notes. We can begin to account for this general phenomenon if, following a suggestion by Frieden, we assume that the initial trace of WH has a referential and anaphoric index not subject to the condition in (33). This is equivalent to treating the initial trace of a WH as a name. Subsequent movements of the WH element could also leave traces with the same indexing as names, or they could simply bear the referential index of the WH. What is essential however is that the WH-element itself after it is moved would only have a referential index and would be subject to the binding conditions, in particular (33). This means that (48)-(49) have the structures in (50)-(51) respectively, prior to the operation of (33).

\[
\begin{align*}
(50) \quad [_{NP}\text{the woman}_{2} \ [_{he_{3}}^{2}] \text{loved}] \text{ betrayed} \\
& \text{him}_{4}^{2,3} \\
(51) \quad \text{who}_{4} \text{ did } [_{NP}\text{the woman}_{2} \ [_{he_{3}}^{2}] \text{loved}] \text{ betray} \\
& \text{[ e ]} \\
& +_{WH}^{4}^{3,2}
\end{align*}
\]

In (50) condition (33) will erase the anaphoric index 2 on him permitting the co-reference reading. In (51), however,
because the trace is marked $+WH$, the condition does not apply to erase either anaphoric index. Condition (33) is applicable only to the referential index on who but does not erase this integer because it appears elsewhere in the relevant domain.$^3$

"Bridge" constructions, that is, constructions permitting successive cyclic WH-movement, are important at this juncture. Within Chomsky's analysis where $S$ and $S^1$ are cyclic nodes (cf., Chomsky 1977b), $S^1$ must be lexically suppressed as a cyclic node in order to permit the movement characteristic of the bridge constructions. For Chomsky then, languages like Russian not permitting successive WH-movement are the unmarked case.$^4$

Now consider (52).

(52)a. Who did John say he saw?

b. Who$_4$ did John$_2$ say $[S,[e]_4[S_3^{1}_{2}]$ saw

\[
\begin{array}{c}
e \\
+WH \end{array}_4^{1}_{2,3}
\]

If nothing more is said about this construction, the index of the fronted WH-element will be eliminated by condition (33) since no other instance of 4 appears in the top-most $S^1$ as a referential index. Of course, this is exactly what we want to happen in the non-bridge constructions. The question at hand then is how to prevent (33) from blocking (52). The answer to this query seems to be that either
S or $S^1$ can correspond to $a$ for the purposes of (33). This fact will be lexically marked just as Chomsky lexically suppressed $S^1$ for the purposes of the Subjacency Condition. Allowing $a$ to be $S$ in (52) will not erase the anaphoric indices on the trace of the WH-element because of its name-like property. Within the domain of the top-most WH-element we now have the intermediate trace which blocks the operation of (33).

Another simplification of Universal Grammar that provides some support for this analysis is that we are able to drop the Subjacency Condition from the inventory of conditions in Universal Grammar.

We find that condition (33) is capable of accounting for the various phenomena that were previously accounted for by the Subjacency Condition. Recall that in Chomsky (1973) and Chomsky (1977b) sentences like those in (53) are used to motivate the Subjacency Condition.

(53)a. *John said [$S$ which city [$S$ [NP the man from $[NP_e]$] saw Mary]]

b. *[S, who $S$ did John say [$S$, which book $[S$ Mary gave $[NP_e]$ to $[NP_e]$]]]]

These sentences would have the indexing of (54a-b) respectively, prior to the operation of (33).
(54)a. John$_2$ said [$_S$,which city$_4_1$$_S$$[^Np the man$_3_2$] from [\[+comp\]$_4_{3,2}$] saw Mary$_5_{2,3,4}$]]

b. Who$_5$ did John$_2$ say [$_S$,which book$_4_1$] [$_S$Mary$_3_{2}$ gave [\[+comp\]$_4_{2,3}$] to [\[+comp\]$_5_{2,3,4}$]]

In (54a) the WH-element labeled 4 will have its index erased because 4 is not a referential index in the same minimal domain. Similarly in (54b) the WH-item indexed 5 will have its index erased because 5 appears nowhere else in the relevant minimal domain.

It may be argued that the Subjacency Condition is still required in the framework being outlined because otherwise an NP could use the possessive position as an "escape hatch" analogous to COMP. That is to say, it might be thought that it is possible to generate a well formed string like (55a) by doing the movements indicated in (55b).

(55)a. who$_4$ did John$_2$ think [$_S_1$[e]$_4_1$$_S$$[^Np[e]$_4_1$ books about [e]$_4_{2,3}$] would be interesting]]
b. who did John think $S_1$ COMP$_S$ $S_{NP}[e]$ books about $[e]$ would be interesting

However, the Case Concord Condition blocks such movements. The binding of the empty nodes inside the NP have the form in (56) with the appropriate case features indicated.

\[
(56) \quad [NP \quad \begin{array}{ll} & e \\ +\text{case} & \text{books about} \\ -\text{obj} & \end{array} \quad \begin{array}{ll} & e \\ +\text{case} & \text{about} \\ -\text{obj} & \end{array} \quad _4^4\{2,3\}]
\]

In (56) the value for the feature oblique is not the same in both nodes. As a result, such a binding violates the case concord condition, (19). It thus seems to be the case that the basic facts motivating the Subjacency Condition can be handled by (33).

There are two areas of empirical phenomena which can be captured in the framework being proposed here but which escape the "On Binding" framework.

First consider sentences like (57).

(57)a. The men said that pictures of them would be on sale.

b. The men said that pictures of each other would be on sale.

Both sentences are acceptable. Sentence (57a) is of some interest in that it is possible to interpret them as referring to the men. This phenomenon was first observed,
I believe, by R. Friedin. Within Chomsky's "On Binding" framework these sentences are problematic: them in (57a) should be disjoint in reference from the men for the same reason that each other can be bound to the men in (57b); that is to say neither the Nominative Island nor theOpacity Conditions are relevant to eliminate indices on these two anaphors. However, the framework developed in this chapter finds no difficulty when it confronts sentences like (57). Sentence (57a) has the possible structures in (58) while sentence (57b) has the possible structures in (59).

(58)a. The men$_2$ said that [NP$_3^{2}$ pictures of them$_4^{2,3}$] would be on sale.

b. The men$_2$ said that [NP$_2^{2}$ PRO$_2$ pictures$_3^{2}$ of them$_4^{2,3}$] would be on sale.

(59)a. The men$_2$ said that [NP$_3^{2}$ pictures of each other$_2$] would be on sale.

b. The men$_2$ said that [NP$_2^{2}$ PRO$_2$ pictures$_3^{2}$ of each other$_2$] would be on sale.

At this point the Index Suppression Condition can operate on (58a) and (59a) to erase 2 inasmuch as it appears nowhere else in the NP. This operation will make (59a)
unacceptable but will give the potential co-reference reading with the men in (58a). As a result, for (57a) there is at least one derivation where the men and them can be understood as co-referential. This result is not possible given the Opacity Condition.

The second advantage also involves the presence of the abstract PRO in the specifier system of an NP. J. Emonds has drawn my attention to the contrast between the sentences in (60) and (61).

(60)a. New York {ruined destroyed} itself financially.

b. New York {destroyed ruined} it financially.

(61)a. *New York {revealed made it clear to me} that a story about itself would be totally unacceptable.

b. New York {revealed made it clear to me} that a story about it would be totally unacceptable.

The sentences in (60) conform to the general pattern we have observed: within the same clause two NP's are disjoint in reference. The indexed structures for (60a-b) would be (62a) and (b) respectively.

(62)a. New York\textsubscript{2} ruined itself\textsubscript{2} financially.

182
b. New York{2} ruined it_{3} financially.

Nothing will eliminate the index on the reflexive pronoun in (62a) and the sentence is interpreted accordingly. In (62b) the Index Suppression Condition will not erase 2 in the anaphoric index of it, since 2 appears elsewhere in the S, and as a result it must be disjoint in reference from New York.

Now consider (61). The "On Binding" framework analysis (61a) is no different from sentences like (63).

(63) The men {revealed \hspace{1cm} } that a story
{made it clear to me}

about themselves would be totally unacceptable.

That is to say, neither the Opacity Condition nor the Nominative Island Condition is capable of eliminating the index on the reflexive pronoun in order to block the sentence. Consequently both (61a) and (63) are generated in the "On Binding" framework, although (61a) is unacceptable.

In the framework being developed in this chapter there is a straightforward explanation for the contrast between (61a) and (63). Given the Index Suppression Condition, the only way to maintain a designated index inside an NP is if there is a PRO inside the NP as well, as we have already observed. This means that (61a) has the structure of either (64a) or (b).
(64a). New York\textsubscript{2} revealed that \(\text{NP}\text{stories}_{3[2]}\) about itself\textsubscript{2} would be totally unacceptable.

b. New York\textsubscript{2} revealed that \(\text{NP}\text{PRO}_{2}\text{stories}_{3[2]}\) about itself\textsubscript{2} would be totally unacceptable.

In (64a) the index of itself will be eliminated by the Index Suppression Condition, resulting in an open, and therefore ungrammatical, sentence. Only (64b) will permit the index on itself to remain. Yet the sentence in (64b) is nevertheless unacceptable. The contrast between the sentences in (65) and (66) demonstrate that an NP in the specifier system of story must be animate. This is true for NP's in general, a fact noted in Sapir (1924).

(65a). \{*New York's story \} would be totally unacceptable.

b. It would take years to write

\{*New York's story \}.

\{ the story of New York \}

c. \{*The Bay Bridge's story \} would be totally unacceptable.

d. It would take years to write

\{*the Bay Bridge's story \}.

\{ the story of the Bay Bridge \}
(66)a. \{Che Guevarra's story \} would be 
    \{The story of Che Guevarra\} 
    totally unacceptable.

b. It would take years to write 
    \{Che Guevarra's story \}.
    \{the story of Che Guevarra\}

c. \{Jeff Bridges' story \} would be 
    \{The story of Jeff Bridges\}
    totally unacceptable.

d. It would take years to write 
    \{Jeff Bridges' story \}.
    \{the story of Jeff Bridges\}

This means that the sentence in (64b) is unacceptable for 
the same reason that the sentence in (67) is unacceptable.

(67) *New York \{revealed \} that its 
     \{made it clear to me\} 
     story would be totally unacceptable.

Sentence (67) is ungrammatical under the reading where \textit{its} 
and \textit{New York} are co-referential because a non-animate noun 
appears in the specifier system of \textit{story}, just as in (65).

In this section I have proposed several revisions 
of Chomsky's (1978) proposals. I have simplified the 
mechanism assigning case and provided an explanation for 
the appearance of passive morphology, based on a principle 
of case concord. In addition, I have proposed replacing 
the Nominative Island Condition and the Opacity Condition 
with the Index Suppression Condition. This has the
advantage of bringing out some empirical generalizations lost in the "On Binding" framework. It has also been seen that the Index Suppression Condition probably permits one to eliminate the Subjacency Condition from the inventory of conditions in Universal Grammar.

3. The Status of the Sonata-Violin Paradigm

I have said nothing in this chapter yet about the "traditional" sentences like (68), generally thought to motivate some condition blocking crossing structures.

(68)a. It would be easy to play the sonata on this violin.

b. The sonata would be easy to play on this violin.

c. This violin would be easy to play the sonata on.

d. What sonata would it be easy to play on this violin?

e. What violin would it be easy to play this sonata on?

(69)a. *What sonata₂ would this violin₃ be easy to play [e]₂ on [e]₃ .

b. What violin₂ would this sonata₃ be easy to play [e]₃ on [e]₂ .

186

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(68b–c) illustrate object to subject raising applying to either the object of the verb or of the preposition. The sentences in (68d–e) illustrate how WH-movement can apply to either object as well. The sentences in (69) present the results of applying both movements in the same sentence. At first glance one might assume that these sentences demonstrated the validity of the Association Principles since the non-nested (69a) is unacceptable while its nested counterpart, (69b), is acceptable. Such a conclusion might be overly simplistic. In particular we need to study in more detail the relationship between these sentences and the other general conditions on binding to assure ourselves that these facts do not follow from other conditions. In other words, we want to make sure that these sentences actually provide some evidence in support of the Association Principles.

Chomsky (1977b) analyzes sentences like (68b–c) as being a special subcase of WH-movement. These sentences have structures like (70)–(71) with the appropriate movements indicated by arrows.

\[(70) \text{This sonata is easy } [\overset{\text{COMP [PRO to play WH on}}{\text{S}^1} \text{the violin]}.\]

\[(71) \text{This violin is easy } [\overset{\text{COMP [PRO to play the sonata on WH]}}{\text{S}^1}].\]
When moved to COMP the WH-phrase will be deleted by the free-deletion in COMP rule discussed in Chomsky and Lasnik (1977) and Chomsky (1978). This permits the generation of sentences like (68b-c). Notice that given either the Specified Subject Condition or the Opacity Condition, these sentences must involve WH-movement and not NP-movement as had been traditionally assumed prior to Chomsky (1977b). In order to see the necessity of this fact consider the structures in (72)-(73).

\[
(72) \quad e \text{ are easy } \quad S_1 [\text{PRO to buy presents for children}]
\]

\[
(73) \quad e \text{ are easy } \quad S_1 [\text{PRO to buy presents for children}]
\]

The structures in (72)-(73) correspond to the structures in (70)-(71) in a WH-movement analysis. The movements indicated by the arrows in (72)-(73) would be prevented by the Specified Subject or Opacity Conditions, since the movement in each case crosses a specified subject, i.e., the PRO. Similarly the structures will be blocked by the Opacity Condition inasmuch as the trace of the moved NP will be free and in the domain of a subject.

The necessity of assuming that WH-movement is involved in sentences like (68b-c) makes it difficult to account for sentences like those in (69) where another

188
instance of WH- movement is clearly involved. These sentences would have to have structures like (74).

(74) \[ S,COMP [S,sonatas are easy [S,COMP [S,PRO to play WH on what violin] \]

In (74) because the COMP of the embedded S is filled and COMP's may not be doubly filled, the NP, what violin, must move directly into the COMP of the matrix S. Such a movement violates the subjacency condition of course, since it involves moving across two S cycles, as well as violating the Specified Subject or Opacity Conditions. This then leaves the acceptability of the sentences in (69) anomalous.

Chomsky attempts to account for the acceptability of the sentences in (69b) by arguing that structurally similar sentences are unacceptable and the phenomenon is limited, being best described by assuming that the prepositional phrase is within the matrix rather than the embedded S. For Chomsky then the normal or unmarked case would be if both (69a) and (b) were unacceptable. Chomsky contrasts the judgements in sentences like (69) with those in (75) which appear to be analogous to (69).

(75)a. You found a sonata to play on the violin.

b. *What violin did you find a sonata to play on.
(76a). You found a violin to play the sonata on.

b. *What sonata did you find a violin to play on?

(77) \[ S^1 \text{COMP} [S_{you \text{ found} \[ NP_{a \text{ sonata} \[ S^1 \text{COMP}} \]

\[ S_{PRO \text{ to play WH on what \text{ violin}]]}]]

(78) \[ S^1 \text{COMP} [S_{you \text{ found} \[ NP_{a \text{ violin} \[ S^1 \text{COMP}} \]

\[ S_{PRO \text{ to play what sonata on WH}]]}]]

Both (75b) and (76b) are unacceptable. This fact can be explained by appeal to the Subjacency Condition, inasmuch as a WH phrase moves across an S and NP domain. Alternatively the Specified Subject Condition or Opacity Condition could explain the unacceptability of these sentences. Either explanation is sufficient.

Chomsky goes on to suggest that the acceptability of (69b) is due to the fact that it has a structure like (79) where the PP, on what violin, is outside the embedded S\(^1\). In this way, when the phrase moves, it does not move over a specified subject in the same cyclic domain and consequently the grammar permits the generation of the sentence in (69b).

(79) The sonata is easy \[ S_{WH \[ S_{PRO \text{ to play e } \]}} \]

on what violin

The structure in (79) must be derived by a restructuring
rule.

It is worth asking at this point what would block
the movements in (80) which would permit (69a) to be
generated.

(80)  \[ \text{S}^{1}_{\text{COMP}} [S[e] \text{ is easy} [\text{S}^{1}_{\text{COMP}} [S_{\text{PRO to play}} \\
\text{which \underline{sonata]] on this \underline{violin}}]}

In this structure it is possible to perform the movements
represented by the arrows: that is, NP movement could
replace the subject with the readjusted PP while successive
cyclic WH-movement could front the object of the embedded
clause. In this way nothing would prevent the generation
of the (a) sentence in (69).

Chomsky's proposals for case assignment are rele-
vant at this point. Recall that NP's in the domain of P
are marked oblique in the base. What this entails is that
the movements represented in (80) would yield an ill-formed
structure since this violin would be assigned two cases:
oblique case in the base and nominative case on the sur-
face. This response, however, crucially relies on the
assumption that the entire PP and not just the NP is
readjusted outside the embedded S: otherwise the NP will
not have oblique case and nothing will block NP movement
into the subject position. This means that while we are
able to describe how (69a) is unacceptable, the theory of
grammar assumed by Chomsky (1978) permits in principle a
possible variant of English in which (69a) is grammatical, that is, a grammar where only the NP is restructured, not the entire PP, and NP-movement is subsequently possible.

Just as (69b) challenges the descriptive adequacy of Chomsky's conditions, they also serve to challenge the adequacy of the Case Concord Principle and the Index Suppression Condition. In particular, we cannot assume that the sentence in (81a) derives via an NP-movement since such a proposal violates both conditions.

(81)a. This sonata is easy to play on the violin.

b. \[\text{This sonata}^{+\text{nominative}_2} \text{is easy}^{S,\text{PRO to play}}_{S}\]

\[e^{+\text{obj}_2} \text{on the violin}\]

In (81b) we find the case concord restriction violated inasmuch as the bound elements have distinct case, and the Index Suppression Condition will eliminate the designated index of the trace resulting in an ungrammatical sentence. This means that, just as in Chomsky (1977b), WH-movement must be involved in the generation of these sentences, deriving them ultimately from a structure like (82).

(82) \[\text{This sonata}^{+\text{nominative}} \text{is easy}^{S,\text{WH}^{S,\text{PRO to play}}}_{S}\]

\[e^{+\text{obj}} \text{on the violin}\]

In (82) the WH-element can be understood as co-referential.
to *sonata* and, after free deletion in COMP applies, we derive the sentence in (81a).

Of course the preceding analysis complicates the analysis of sentences like (69b) in precisely the way that Chomsky's analysis was complicated. That is, the WH-element cannot move through the embedded COMP because it is already filled. Additionally it is not possible for the phrase *what violin* to move directly into the matrix COMP without causing the IndexSuppression Condition to erase the index of the WH-element in the topmost COMP. Instead, we are forced to claim, like Chomsky, that *on what violin*, or simply *what violin*, occurs in the matrix sentence.

The advantage that the Association Principles provide in either of these analyses is that they limit the class of grammars permitted by linguistic theory. That is, the Association Principles prevent a variant grammar of English in which sentences like (69a) are grammatical. In order to see this point, assume that just the NP *the violin* is restructured into the matrix sentence. The Association Principles will prevent moving that NP into subject position if successive WH-movement from the embedded sentence also operates. Examine (83).

(83) Which *sonata* is this *violin* easy [\[\begin{array}{c}
A_1 \\
X
\end{array}\] $\text{WH}$ $\begin{array}{c}
A_2 \\
S_{B_1}
\end{array}$

$[S_{\text{PRO to play e on }]} e_{B_2}$

193
In (83) X contains $A_2$ but not $B_2$. This means that (83) is a crossing structure and as such violates the Association Principles. The Association Principles then predict that no variant grammar in which (69a) is grammatical will be part of the class of grammars determined by linguistic theory.

Sentences like (69) provide evidence in favor of the Association Principles but the character of this evidence is much more indirect and limited than has often been assumed (e.g., Bordelois 1974 or Fodor 1978).

4. Remarks on Chomsky's Counter-Examples

The proposals advanced in this chapter now permit us to deal with the examples observed in Chomsky (1976) which apparently falsify any condition requiring nesting. Chomsky notes the sentences in (84).

(84)a. What books$_1$ have those men$_2$ written $t_1$
    about each other$_2$.

b. I told them$_2$[what books$_1$[PRO$_2$ to read $t_1$]].

c. I$_1$ asked them[what books$_2$[PRO$_1$ to read $t_2$]].

d. To whom$_1$ did John$_2$ seem $t_1$[t$_2$ to be referring].

e. Whom$_1$ did you$_2$ ask $t_1$[what$_3$[PRO$_2$ to read $t_3$]]
f. dnes me₁ ji₂ Jana ukazala t₁ t₂.

today to-him her Jana showed.
Jana showed her to him today.

I would like to begin by saying that I will have nothing
to say about the Czech example in (84f). In and of itself
such a non-nested configuration means nothing as my earlier
discussion of clitic pronouns in Romance suggested. That
is, crossing structures like (84f) are permitted but not
at surface structure and only as the product of a
restricted class of local rules. In the absence of a
detailed treatment of Czech clitics, (84f) is somewhat
irrelevant.

The sentences in (84b-e) are not counter-examples
to the Association Principles. In each of the cases in
(84b-e) the crossing elements are not in the same minimal
cyclic domain; that is to say a cyclic boundary separates
them. Consequently the Association Principles are not
triggered. The only sentence that involves a minimal
cyclic domain is (84a).

In considering (84a) it is important to bear in
mind that the crossing of a construal and a movement rule
is involved. It is certainly possible to conceive of a
version of the Association Principles where construal rules
are not ambiguous with respect to movement rules. In this
way the Association Principles would restrict a movement
crossing another movement, or a construal crossing another construal, but not a movement crossing a construal. In discussing (84) we are really considering alternative formulations of the Association Principles and not the existence of such a principle in linguistic theory.

It is not inconceivable that about each other in (84a) may be a complement to book and be postposed by a stylistic rule subsequent to surface structure (cf., Rochemont 1978). Such a case seems to be (85).

(85)a. Which books about nuclear physics did Jane write?

b. Which books did Jane write about nuclear physics?

In order to avoid such problems and keep the relevant constructions clearly applicable to the claims being made here, let us consider (86) in place of (84a).

(86)a. Which students₁ did the teachers₂ recommend t₁ to each other₂.

b. Which courses₁ did Jane and Mary₂ recommend t₁ to each other₂.

These sentences seem to be acceptable despite their crossing structure. The relevant factor seems to be that, as noted by H. Lasnik (and reported in Chomsky 1976), each other requires a referential antecedent not provided by the
WH-phrase. There thus does not seem to be the necessary structural ambiguity in (86) to trigger the Association Principles. When a referential antecedent is provided for the trace, the sentences become unacceptable, as the topicalized sentences below demonstrate.

(87)a. *These writers\textsubscript{1}, the teachers\textsubscript{2} recommend \( t_1 \)
    to each other\textsubscript{2}.

b. *Those two books\textsubscript{1}, the teachers\textsubscript{2} recommended
    \( t_1 \) to each other\textsubscript{2}.

It thus seems that the apparent counter-examples to the nesting requirement are not relevant to the specific formulation of this requirement in the Association Principles.
FOOTNOTES TO CHAPTER 4

1 I take $S^1$ and NP to be the cyclic nodes for the purposes of (2).

2 Nothing depends on the particular features chosen here. Use of the feature [i-case] has the advantage of making the no-case filter of Chomsky (1978) rather elegant.

3 One way that this general proposal about treating WH-elements as "names" is to assume that referential indices are assigned at the base, but that anaphoric indices are assigned at the surface. WH-elements, including traces with the feature WH, would be ambiguous in English as either a bound anaphor or as a name. This would permit one to assign none, some or all of the WH-elements and their traces anaphoric indices, providing them with their names like property. This proposal raises no particular problems that I can see at the moment and will insure at least one derivation where the "unbounded" movement of the WH-element can be accomplished.

4 I am assuming that if $S^1$ is suppressed, $S$ becomes the cyclic node. One way to suppose that this assumption is realized in a grammar is the following: lexical items that are bridge elements have the entry in (i) associated with them,

(i) $\alpha = S$

The lexical restriction (i) is interpreted as replacing (ii) which is the general, unmarked instance.

(ii) $\alpha = S^1$
Chapter 5

CLITIC PRONOUNS AND THE FRENCH CAUSATIVE CONSTRUCTION

In this chapter I direct attention to a complex phenomenon, that is, clitic placement (or interpretation) and the French causative construction. My intention here is to provide a detailed analysis of what has remained a problematic area of research in French syntax and demonstrate that the Association Principles, interacting with the other aspects of the grammar of French, make a more explanatory account of these facts possible. As such, this analysis adds corroborating evidence to the postulation of the Association Principles and provides an example of how the Association Principles might be manifest in other languages.¹

1. The Problem

The construction that is under analysis in this chapter is represented in (1).

(1)a. Jean fait parler Paul à Marie.
    Jean is making Paul talk to Marie.

b. Elle fait entendre cette symphonie à Jean.
    She is making Jean listen to this symphony.
There are a number of characteristics of this construction that we expect any analysis to capture. When the embedded verb is intransitive, as in (1a), the logical agent of that verb appears as the direct object. On the other hand, when the embedded verb is transitive, as in (1b), the logical agent appears as the indirect object. The verb *faire* differs from other so-called causative verbs in that the structural configuration exhibited in (1) is obligatory and not optional. In this respect compare (1) to (2).

(2)a. *Jean fait Paul parler à Marie.*

Jean is making Paul talk to Marie.

b. Jean fait parler Paul à Marie.

Jean is making Paul talk to Marie.

(3)a. Jean laisse Paul parler à Marie.

Jean is letting Paul talk to Marie.

b. Jean laisse parler Paul à Marie.

Jean is letting Paul talk to Marie.

The difference between (2) and (3) has been glossed over in all accounts of this construction. We will require a descriptively adequate analysis to give a principled account of this fact.

Another characteristic of this construction is the behavior of clitics. The sentences in (4) contrasted with those in (5) are designed to show that only reflexive clitics may intervene between *faire* and the embedded verb.
(4)a. *Elle fera lui manger ce gâteau.
    She will make Jean eat it.

b. *Elle fera le lui manger.
    She will make him eat this cake.

(5)a. Jean fait se parler Marie.
    Jean made Marie talk to herself.

b. Jean fait se connaître Marie.
    Jean made Marie know herself.

The sentences in (6) and (7) can be contrasted in that while the "logical subject" of the embedded verb can appear before the matrix verb, the pronoun clitics (me, te, lui, nous, vous, and leur) that are objects cannot.

(6)a. Elle lui fera manger ce gâteau.
    She will make him eat this cake.

b. Je te ferai connaître cette fille.
    I will make you know this girl.

c. On leur a fait boire du vin.
    People made them drink wine.

(7)a. *Je lui fera écrire mon ami.
    I will make my friend write to him.

b. *La peur de la police te fera téléphoner
    Jean.
    Fear of the police will make Jean telephone you.
c. *Cette éclairage vous fait ressembler cette statue.
This lightning makes this statue resemble you.
d. *Les menaces leur ont fait répondre le criminel.
Threats made the criminal answer them.

In this respect, the pronoun clitics differ from the third-person direct object clitics (le, la, les) and the prepositional clitics (y and en) which can quite productively appear before faire as (8) and (9) respectively show.

(8)a. *Marie se fera embrasser à Paul.
Marie will make Paul kiss her.

b. *Marie te fera embrasser à Paul.
Marie will make Paul kiss you.

c. *Paul voulait me faire embrasser à Marie.
Paul wanted to make Marie kiss me.

d. Je le ferai manger à Jean.
I will make Jean eat it.

(9)a. Cela y fera aller Jean.
That will make Jean go there.

b. Elle en fera sortir Jean.
She will make Jean leave there.

Once again these observations must be given a principled account in any account of the construction under analysis.
2. The French Causative As Base Generated

There exists at least two well-articulated analyses of the French faire construction within the extended standard theory assumed here. These works are Kayne (1975) and Quicoli (1976). Both of these works propose movement analyses which derive sentences like (2b) from structures like (2a) containing an embedded sentence. Before considering the details of the movement analyses, let us first examine their motivation. Kayne (1975) forwards two major arguments assumed also by Quicoli (1976), against analyzing the faire construction as base generated. On the one hand, Kayne argues that a base analysis would necessitate postulating redundant subcategorization restrictions on verbs. On the other hand, he also argues that there is a rule of se-placement which must crucially precede the rule, whatever its form, which creates sentences like (2b) from underlying complex forms. A third argument advanced by Kayne is the necessity of being able to identify the underlying subject of the embedded verb in order to facilitate the interpretation of certain adverbs.

2.1. Subcategorization

First consider Kayne's argument concerning the avoidance of redundant subcategorization specifications.

(10)a. Jean part.

Jean is leaving.
b. *Jean part Marie.

   Jean is leaving Marie.

(11) Marie fait partir Jean.

   Marie is making Jean leave.

Kayne argues that to provide a base analysis of (11) it
will be necessary to recognize two subcategorization
restrictions for partir: on the one hand, partir will not
be subcategorized for a following noun phrase in order to
account for the contrast between (10a) and (b). However,
in the environment of faire, partir will take a following
NP. Kayne assumes a redundancy rule would be constructed
to relate these entries. Kayne objects to such a treatment
on two grounds: first, it is merely a notational variant
of the transformational account, and second that redundancy
rules have a less well understood formal character in
linguistic theory. Both of these observations seem reason-
able to me.

Kayne's objections however can be avoided by
employing the Association Principles. Examine (12).

(12) Elle fait entendre cette symphonie à Jean

       She is making Jean listen to this symphony.

We will not need to postulate separate subcategorizations
and redundancy rules. Instead, let us assume the phrase
structure rules in (13) and the subcategorization restric-
tions in (14)-(15).
(13a) \( V^2 \rightarrow V^1 \ldots \text{PP} \ldots \)

b. \( V^1 \rightarrow V' - \{\text{NP} - \text{PP} - \text{PP} - S^1\} \)
   \[ V^2 \]

c. \( V' \rightarrow \begin{cases} V' \\ \text{[PRO] - [PRO]} \\ \text{+_N} \\ \text{+_P} \end{cases} \rightarrow V \)

(14) faire, +V, _____ V NP

(15) entendre, +V, _____ NP

The phrase structure rules in (13) differ somewhat from those of English in that there is a VP complement possible in French and perhaps in the Romance languages more generally (cf., Strozer 1977 and Chapter 6) but not in English. In Section 6.3 I motivate this distinction in some detail.

Given the phrase structure rules in (13) and the subcategorization restrictions in (14) and (15), we will produce a structure like (16). I assume that \([p^e]\) will, at some point subsequent to surface structure, be replaced by the designated element \(\hat{a}\).

The Association Principles predict the correct relation between verbs and their associated NP's to the right. Notice that (17) is not a well-formed structure given the Association Principles. Because \(X\) contains \(A_2\) but not \(B_2\) this is a crossing structure and is blocked by the Association Principles. In (16), however, the association lines do not intersect and the structure is well formed. In addition neither (18) nor (19) could surface as
(16)

```
S -> v2
   /   \
NP  V1
   /   \
V'  V

v2
---

v1
---

V'
---

NP
---

Det N
---

P NP

elle faire entendre cette symphonie e Jean
```
(18)

```
S
  NP
  V
  V'
  NP
  S
  V
  V'
  NP
  Det
  N

elle faire Jean entendre cette symphonie
```

(19)

```
S
  NP
  V
  V'
  NP
  S
  V
  V'
  NP
  Det
  N

elle fait Jean PRO entendre cette symphonie
```
well-formed structures given the phrase structure rules and subcategorization restrictions above. Structure (18) fails to exhibit the dependency relation expressed in (14) since the "agent" NP precedes V. The same observation holds for structure (19).

The Association Principles in conjunction with the rules in (13)-(15) also guarantee the correct output for intransitive verbs as well. Consider (20).

(20) Jean fait parler Marie à Paul.

Jean is making Marie talk to Paul.

We permit (20) to have the structure in (21). Marie may appear in V₁ in the "direct object" position because the Association Principles are not violated: parler is associated with a PP, not an NP, and so there is no structural ambiguity for the Association Principles to resolve.

The reader may have noted that the Association Principles also permit (22) in addition to (20).

(22) Jean fait parler à Paul à Marie.

In fact, such a construction is permitted by some speakers.² For speakers who reject constructions like (22), we may postulate an output filter, (23), suggested in Ruwet (1972) and Bordelois (1974), prohibiting double à-phrases.

(23) *à NP à NP, where both NP's are [+animate]
I conclude from the foregoing discussion that a base analysis of the faire construction can avoid postulating redundant subcategorizations and is, in principle, possible.

2.2. Reflexives

The second argument that Kayne advances in support of a transformational derivation of the faire construction involves demonstrating that another transformation must operate before the transformation which creates the faire construction. Kayne argues at length in favor of a major transformation, SE-placement (SE-P1), as preceding the putative transformation involved in deriving the faire construction. 3

As noted earlier, only reflexive clitics can appear before the embedded verb in the faire construction. For example, (24) is acceptable but (25) is not.

(24) Jean fait se (parler  } Marie.
       {connaître}

       Jean is making Marie {talk to} herself.
       {know }

(25) *Jean fait (le connaître} Marie.
       {lui parler }

       Jean is making Marie {know him } .
       {talk to him}

In Kayne's framework this follows if SE-P1, unlike clitic-placement, is cyclic. The underlying structure of (24) would be (26) for Kayne.

211
The arrow in (26) shows the movement of the PRO on the lower cycle. Since there exists a different rule moving the clitics in the sentences in (25) and that rule is post-cyclic, the sentences in (25) will not be generated.

Kayne points out that ordering SE-P1 before the passive transformation enables him to explain that the passives in (27) are unacceptable.

(27)a. *Tu te seras décrit par ta femme.
You will be described to yourself by your wife.

b. *Jean se sera décrit par sa femme.
Jean will be described to himself by his wife.

(28)a. Elle te sera décrite par ta femme.
She will be described to you by your wife.

b. Ils vous seront présents par Paul.
They will be introduced to you by Paul.
If passive follows SE-P1, the underlying structure of the sentences in (27) will be such that the subject will not be co-referential with the pronoun, and consequently his SE-P1 will not apply. On the other hand, since clitic placement is post-cyclic, nothing will prevent the generation of the sentences in (28).

Let us consider the hypothesis that the reflexive clitics are base generated. There is a class of verbs in French which obligatorily co-occur with a reflexive clitic, commonly called "inherently reflexive verbs." Some examples of these verbs are presented in (29).

(29)a. Marie s'est évanouié.
Marie fainted.

b. Ils s'en sont allés.
They went away.

c. Jean s'imagine cela.
Jean imagines that.

There has been some debate over the question of whether such inherently reflexives are base generated or whether they are derived transformationally. Kayne (1969) and Emonds (1976) argue the first position while Kayne (1975) argues the second one. This issue bears crucially on the base generation of other reflexive clitics as well.

The arguments presented in Kayne (1975) against assuming that these reflexives are base generated relate
to explaining certain gaps in the lexicon of French in a
principled way. In particular, he notes that while there
are structures like $SE \ V \ NP$ and $SE \ V \ Â \ NP$, there are no
structures of the form $*SE \ V \ NP \ Â \ NP$. In addition Kayne
observes that while there are structures like $SE \ V \ (Â) \ NP$
there are no structures $SE \ V \ ADJ$, that is, inherently
reflexive adjectives. Kayne regards this as an accidental
fact, given the base generation of inherent reflexives,
but notes that it would be a natural outgrowth of a move-
ment analysis since adjectives would not have the "accusa-
tive" or "dative" position in which $SE$ could originate.

There are straightforward explanations for both of
these facts within at least one base generation hypothesis.
We will assume Helke's (1971) analysis of reflexivization
where self (or même in French) is base generated as a
pronoun and a subsequent transformation copies the appro-
priate features from the subject NP onto this pronoun to
ultimately give himself, lui-même, etc. Now there is some
reason for supposing further that the reflexive même is
subcategorized for a preceding reflexive clitic rather than
assuming that such reflexive clitics originate from a
movement transformation. In this regard, the sentences in
(30) but not those in (31) are acceptable.

(30)a. Quand on se parle à soi-même ....

When one speaks to himself ....

214
b. Jean s'écrit à lui-même de très longues lettres.
Jean writes very long letters to himself.
c. ?Jean s'imagine lui-même.
Jean imagines himself.
d. ?Il se tapait sur lui-même.
He was abusing himself.
e. ??Il se tirait sur lui-même.
He was shooting at himself.
f. ??Il se courait après lui-même.
He was running after himself.

b. *On lui courait après lui
c. *On lui écrit à lui.

The sentences in (30) indicate that même and a reflexive clitic can co-occur in the same sentence while the sentences in (31) show that the same is not true for non-reflexive pronouns.

In addition to the sentences in (30), we also find that the sentences in (32) are acceptable.

(32)a. Quand on se parle ....
b. Jean s'écrit de très longues lettres.
c. ?Jean s'imagine.
d. Il se tapait sur.
e. Il se tirait sur.
f. Il se courait après.
These sentences indicate that même in the sentences in (30) is optional. We can account for the behavior of même and the reflexive clitics in (30) and (32) by postulating the following dependency relation along with the following transformational rule.

\[(33) \text{même, } +\text{PRO }, +\text{se } +\text{REFL} \]

\[(34) \text{même } \Rightarrow \emptyset \]

Statement (33) requires that même co-occur with se. Rule (34) will optionally delete même. When (34) is not applied we derive the sentences in (30) and when it is applied we derive the sentences in (32).

This analysis is not yet adequate inasmuch as it is incapable of generating sentences like (35).

\[(35) \text{Il pensait à lui-même.} \]

Sentences of this nature do not invalidate the subcategorization analysis utilizing (33). Instead they simply indicate that the deletion rule in (34) is overly restrictive. It appears that (34) should be revised to (36) which deletes any reflexive.

\[(36) \begin{array}{c}
+\text{N} \\
+\text{REFL}
\end{array} \Rightarrow \emptyset \]

Of course (36) will not operate to delete both reflexive markers, SE and même, because to do so would violate the condition on recoverability of deletions.
The essential difference between inherent reflexives and "standard" reflexives is that while the latter permit either SE or même or both to be present, the former only permit SE. Suppose that the inherently reflexive verbs in (29) have the lexical restrictions in (37)-(39).

(37) evanouir, +V, [+SE _____ ]
    [ - _____ même ]

(38) en aller, +V, [+SE _____ ]
    [ - _____ même ]

(39) imaginer, +V, [+SE _____ ]
    [ - _____ même ]

The restrictions above guarantee that only the reflexive pronoun clitic will appear with these verbs. Because both même and the reflexive pronoun clitic cannot both co-occur with these verbs, the deletion rule in (36) will not apply.

We now have an explanation for the non-appearance of SE V ... MÊME .... In addition, unlike Kayne's analysis, we also have an explanation for the relative scarcity of inherently reflexive verbs. In order to produce an inherently reflexive verb, the verb must be marked to necessarily take the reflexive clitic but not même. This specification can be contrasted to the unmarked case of other verbs where there is no special statement needed in order to co-occur with a reflexive pronoun. In this way, inherently reflexive verbs are more costly and we expect
that they should be scarcer. This corresponds nicely to the empirical phenomena. No such principled explanation is possible within Kayne's framework.

In order to prevent generating the sentences in (40) with his movement rule, Kayne places the constraint in (41) on French.

(40a) *Marie m'a évanoui.
    b. *Ils nous en ont allés.
    c. *Jean leur imagine cela.
    d. *Ils en ont allé (à) leurs amis.
    e. *Jean imagine cela à tout le monde.
    f. *Marie evanouit elle.
    g. *Ils en iront moi.

(41) Inherently reflexive verbs may not co-occur with an accusative NP, or in the case of imaginer with a dative à-NP.

Without (41) Kayne's grammar will generate all of the sentences in (40) where non-reflexive clitics or full NP's occur rather than a reflexive clitic. Convention (41) is intended to avoid this. However (41) is a language specific condition concerning co-occurrence restriction of lexical items. As such it belongs most naturally where other restrictions of the same type are expressed: in the lexicon. It is only Kayne's decision to treat SE-Pl as a movement rule that prevents him from being able to express
these restrictions in their proper place within the grammar. Moreover, once we attempt to express a restriction like (41) within the lexicon, we are led back to postulating a deletion rule like (36).^5

Kayne further observes that there are no inherently reflexive adjectives in French. First, it is important to realize that Kayne cannot provide an explanation for the fact that reflexive clitics, in general, do not appear with être. It is clear from our earlier examination of clitics that NP's within the complement system of adjectives may cliticize to a matrix verb. There is, as far as I can see, no principled reason that there could not exist a structure like (42) within Kayne's framework.

(42) \(NP_1 \text{ÊTRE ADJ à PRO}_1\)

This structure could then undergo SE-Pl just as it could undergo clitic placement and produce reflexive adjectives. In this way, it may be more profitable to attribute the lack of inherently reflexive adjectives to the non-appearance of reflexive clitics with être. This generalization is expressed by the surface filter in (43).

(43) \(^{[\text{PRO} - \text{Être}]^{[+\text{REFL}]}}\)

This account gains added support from an independent quarter. Recall that we observed that, as (30) demonstrated, reflexive clitics could not appear before

219
passives. As we noted at that time, Kayne attributed this fact to an ordering restriction (i.e., SE-Pl before passive). However, within the more highly constrained metatheory proposed in Chomsky and Lasnik (1977) where ordering restrictions are not available, a filter like (43) would need to be proposed independently. I think that this is evidence for the correctness of the general approach here. Kayne's alternative suffers from increasing the class of grammars by permitting rule ordering.

There is only one exception to the generalization in (43) and that is reflexivization in the passé composé which exhibits reflexive clitics adjacent to être, as exemplified in (44).

(44)a. Elle se sont donné des coups de pied.

b. Jean se'est photographié.

When not reflexive, these same verbs exhibit avoir and not être in the passé composé. For example, the sentences in (45) are acceptable.

(45)a. Elles ont donné des coup de pied à Jean.

b. Jean à photographié sa soeur.

We can, then, propose that all verbs have avoir in the passé composé and that a later morphological rule changes avoir to être after reflexive clitics and in the presence of V+é. Napoli (1973) and Rizzi (1978) argue for a similar rule in Italian. Since this morphological rule will apply
subsequent to the filter in (43), the output will be grammatical.

I conclude that at least reflexive clitics originate preverbally in the base and that subcategorization restrictions like (33), (37)-(39), coupled with the deletion rule in (36) and the filter in (43), account in a principled way for the otherwise anomalous behavior of inherent reflexives as opposed to normal, or non-inherent, reflexives.

We are now in a position to give a base analysis of sentences like (24). These sentences will have underlying structures as in (46). In both cases, même will be bound to Marie.6 Rule (36) will subsequently remove the même and we produce the sentences in (24).

2.3. Subject Oriented Adverbs

I would now like to examine Kayne's third argument in favor of a transformational analysis of the faire construction. Essentially this argument concerns the ability of the transformational analysis to identify an underlying subject of the infinitive. Kayne argues that this ability is important in order to give the correct interpretation to certain adverbs that can only refer to subjects. In support of this point Kayne looks to the following contrasts.

221
(46)a. 

```plaintext
S
  NP
  V¹
    V'
      V¹
        V¹
          NP
            PRO
                V
            P
                NP

Jean faire se parler Marie à même
```

b. 

```plaintext
S
  NP
  V¹
    V'
      V
        V²
          V¹
            NP
                PRO
                    V
                P
                NP

Jean faire se connaître Marie e même
```

222
(47) Paul s'est hissé d'une seule main sur le cheval.

Paul lifted himself with one hand onto the horse.

(48) Elle a poussé Paul d'une seule main dans l'eau.

She pushed Paul with one hand into the water.

In (47) and (48) the adverb can only refer to the subjects of the sentences. In (49) the adverb can refer to Paul.

(49) La peur a fait se hisser Paul d'une seule main sur le cheval.

Fear made Paul lift himself with one hand onto the horse.

Kayne proposes to explain this fact by constraining these adverbs to being interpreted with subjects and by appealing to the underlying status of Paul as a subject of the infinitive in (49) in order to permit the correct interpretation.

The issue I want to take with Kayne's analysis is that it is not clearly established that it is the subject relation which is crucial to the interpretation of these adverbs. Agency rather than subjectness is the proper relation to be used in mapping the interpretation of the adverb. Support of this counterhypothesis is found if we consider adjectives with agent complements. Consider (50)
and (51).

(50) Le juge pardonnera aux criminels.
The judge will pardon the criminals.

(51) Les criminels seront pardonnés par le juge.
The criminals will be pardoned by the judge.

Kayne (1975:245) argues persuasively that (51) is not derived from (50) by the passive transformation. Instead, he argues, this must be a base construction where pardonné is an adjective with an agent complement. In this kind of case there is no possibility of appealing to the underlying status of juge as a subject to facilitate the interpretation of the an adverb. Yet when an adverb is present, as in (52), it can be interpreted as modifying juge.

(52) Les criminels ont été pardonnés avec enthousiasme par le juge.
The criminals were pardoned by the judge with enthusiasm.

In this case it appears that the interpretation of the adverb must be sensitive to agency. Given this fact, there is no reason to suppose that agency could not permit the interpretation of the adverb in (49).

Note incidently that the same observation has been made with respect to English. Siegel (1975) notes that the (b) sentences below cannot be derived by the passive transformation because there is no corresponding active, as the
(a) sentences demonstrate.

(53)a. *The Eskimos uninhabited Antarctica.

b. Antarctica is uninhabited by the Eskimos.


b. The disturbances were unreported by the press.

In these sentences then the deep structure subject is the same as the surface structure subject. Yet when an adverb is present in the (b) sentences, it is interpreted with the agent complement of the adjective. Note the following:

(55) The natural resources of the area were intentionally untapped by the inhabitants.

(56) The disturbances were unreported by the press for fear of reprisals.

It seems that the initial motivation for preferring a transformational over a base analysis of the faire construction is weak at best. I want now to consider the patterning of the non-reflexive clitics and show that the base analysis is in fact better able to provide an account for that phenomena.

3. The Effects of the Association Principles

The analyses of clitic movement in the faire-infinitive construction presented by Kayne (1975) and Quicoli (1976) attempt to use the SSC to explain the
interesting asymmetries in clitic movement presented in (6)-(9). The attempt to use the SSC in this way was of course the correct move to make in the sense that it is certainly preferable to attribute asymmetries in movements to general properties of linguistic theory rather than complicate particular rules, and the SSC is one of the few conditions proposed up to now that could potentially be applicable. However, there was a cost implicit in the use of the SSC to explain the asymmetries of the clitic movements: Kayne (1975) and Quicoli (1976) are forced to make other theoretically undesirable moves in order to permit the examples in (8d)-(9) which are apparent violations of the SSC. The following two subsections identify the costs of assuming that the SSC is involved in such cases. These costs are then compared with the elegant explanation provided by the Association Principles which, if we assume an underlying VP complement in French, allows us to avoid the complications posed in Kayne (1975) and Quicoli (1976). The third subsection then addresses the issue of assuming a VP complement in French.

3.1. Asymmetries of Clitic Movement, the SSC and the Association Principles

Return to the asymmetries of pre-verbal clitics that we noted at the beginning of this chapter. Recall that the logical agent of the embedded verb is free to cliticize onto faire as (6), reproduced below, illustrates.
(6)a. Elle lui fera manger ce gâteau.
She will make him eat this cake.
b. Je te ferai connaître cette fille.
I will make you know this girl.
c. On leur a fait boire du vin.
People made them drink wine.

However indirect objects of the embedded verb are incapable of appearing to the left of faire as (7), reproduced below, demonstrates.

(7)a. *Je lui ferai écrire mon ami.
I will make my friend write to him.
b. *Je craindrai de la police te fera téléphoner
Jean.
Fear of the police will make Jean telephone you.
c. *Cet éclairage vous fait ressembler cette statue.
This lightning makes this statue resemble you.
d. *Les menaces leur ont fait répondre le criminel.
Threats made the criminal answer them.

Non-third person direct objects of the embedded verb similarly fail to appear to the left of faire. However, the third person direct objects (i.e., le, la, les) are not
restricted in this way giving rise to a curious asymmetry observable in (8).

(8)a. *Marie se fera embrasser à Paul.
Marie will make Paul kiss her.
b. *Marie te fera embrasser à Paul.
Marie will make Paul kiss you.
c. *Paul voulait me faire embrasser à Marie.
Paul wanted to make Marie kiss me.
d. Je le ferai manger à Jean.
I will make Jean eat it.

Finally, the pro-prepositions, y and en, are also free to appear to the left of faire, as exemplified in (9).

(9)a. Cela y fera aller Jean.
That will make Jean go there.
b. Elle en fera sortir Jean.
She will make Jean leave there.

Essentially both Kayne (1975) and Quicoli (1976) want to employ the Specified Subject Condition to block the sentences in (7) while permitting those in (6). In order to make this general proposal work, however, certain complications are required in order to account for the data in (8)-(9). I will argue that these complications outweigh the advantages obtained with respect to (6)-(7) when the Association Principles (in conjunction with the base analysis) provide a unified treatment of (6)-(9).
Kayne (1975:288) summarizes his analysis of the facts at hand when he remarks:

Let us tentatively assume that FI (the transformation creating the faire construction—R.H.) moves the embedded V; thus in the case at hand, it will have the following effect: . . . faire [sNP V PP] + . . . faire V [ NP PP]. . . . In the case where the embedded S contains a non-prepositional object, it will be moved along with the verb: . . . faire [ S_{NP_a} V S_{NP_b} (PP)] . . . faire V S_{NP_b} [ S_{NP_a} (PP)] (and à will subsequently be inserted before NP_a).

In Kayne's analysis then, the embedded verb along with its direct object, if any, moves outside of the embedded S. Kayne argues that this explains why the sentences in (6) are acceptable. In each case, it is the subject of the embedded verb that cliticizes to the matrix verb and nothing runs afoul of the Specified Subject Condition. This contrasts with the sentences in (7). These sentences are unacceptable, in Kayne's and Quicoli's analyses, because in each case the "dative" clitic is moving across a specified subject. These facts are illustrated in (57) where the underlined NP is a specified subject.

(57)a. Elle fera manger ce gâteau à lui

b. *Je ferai écrire mon ami à lui

These sentences illustrate how the contrast between (6) and (7) are to be handled in Kayne's analysis. (57a) is acceptable because the clitic does not cross a specified subject while (57b) is excluded because the movement of the
clitic does cross a specified subject. However, there are important observational and theoretical problems in Kayne's proposal when we attempt to extend the analysis to the sentences in (8)-(9).

Observationally, Kayne's analysis would predict that every sentence in (8) is grammatical. His analysis is apparently falsified by the asymmetry in the acceptability of the third person clitics and the unacceptability of the non-third person clitics. Moreover, the movement of the object NP outside the embedded sentence is a clear violation of the Specified Subject Condition, the same condition used to explain the contrast between (6) and (7). I can see no auxiliary hypotheses that will remove this difficulty.

In addition, the movements of the pro-prepositions also apparently falsify Kayne's analysis since the pro-prepositions move over a specified subject as illustrated below.

(58) cela fera aller Jean y

In (58) the specified subject is underlined. One might attempt to remove the sentences in (9) as problems with respect to the Specified Subject Condition by restricting that condition so that it does not apply to prepositions. Such an analysis however would simply move the problem elsewhere in the grammar by forcing us to give up the
the explanation of the unacceptability of sentences like (59) which were noted in Chapter 3.

(59)a. On va tout y mettre.
   Everything is going to be put there.

   b. *On y va tout mettre,
      Everything is going to be put there.

(60)a. On va \[ \alpha \text{PRO} \downarrow \text{mettre tout y} \]

   b. On \[ \alpha \text{PRO} \downarrow \text{mettre tout y} \]

(59b) is ungrammatical because the movement of y crosses a specified subject, as (60b) shows. (59a), on the other hand, is grammatical because the movement of y is within the domain of the specified subject, as represented in (60a).

Theoretically, Kayne's analysis has a number of difficulties as well. It implies that the verb-moving transformation violates other conditions of linguistic theory. On the one hand, this movement is non-structure preserving. While there may be no particular difficulty with formulating this rule as a local movement in Emonds' (1976) sense, the Freezing Principle of Culicover and Wexler (1977) would make everything inside the embedded sentence unanalyzable for movement. Culicover and Wexler postulate the following principles.

231
(61a. Definition: If the immediate structure of a node in a derived phrase marker is non-base then that node is frozen.

b. The Freezing Principles: If a node X of a phrase marker is frozen, then no node which X dominates may be analyzed by a transformation.

Because the verb movement rule yields a non-base structure, the embedded sentence is frozen. Consequently sentences like (6), (8d), and (9) which involve movements outside frozen structures would be blocked. In order to avoid this then the Freezing Principle which is an otherwise well-motivated condition would have to be eliminated from Universal Grammar resulting in a less restrictive version of linguistic theory.

A second theoretical difficulty with this analysis is that the verb-movement transformation, FI, violates the Specified Subject Condition insofar as it moves a verb outside the domain of a specified subject. Moreover, it is not possible to attempt to eliminate this problem by excepting verbs from the Specified Subject Condition since object-NP's are also moved with the verb outside the domain of the specified subject.

Quicoli's (1976) analysis attempts to overcome the exceptionality of the FI transformation with respect to the Specified Subject Condition. Quicoli proposes that FI
moves the embedded . . . V(NP) . . . sequence to the front of the embedded S so that there is no reordering out of an \( S \), and consequently, no violation of the Specified Subject Condition is involved.

This revision of Kayne's proposal also has several disadvantages. First, there are some problems related to null anaphora which Quicoli's analysis seems unable to deal easily with. In a range of cases, infinitives can be replaced by null anaphora as the examples in (62), taken from Emonds (1978), demonstrate.

(62)a. Marie a vouler visiter le musée, mais moi, je n'ai pas voulu.
   Marie wanted to visit the museum but I didn't want to.

b. Beaucoup de monde osent entrer sans payer, et nous devrions oser aussi.
   A lot of people are trying to enter without paying and we ought to try too.

c. Pierre doit renverser ces tables, mais il ne peut pas.
   Pierre ought to turn these tables over, but he can't.

However, the same null anaphora is blocked in the faire construction.
(63)a. *Jean ne fera pas manger ces gateaux a 
Marie mais moi, je ferai.
Jean will not make Marie eat these cakes, 
but I will.
b. *Jean ne fera pas renverser ces tables a 
Pierre mais moi, je ferai.
Jean will not make Pierre turn these 
tables over, but I will.

In this aspect of its behavior, the infinitive of the 
faire construction acts more like the past participle which 
also fails to be replaced by null anaphora. For example, 
the following sentences are unacceptable.

(64)a. *Marie a visité le musée, mais moi, je n'ai 
pas.
Marie visited the museum but I didn't.
b. *Beaucoup de monde sont entrés sans payer, 
et nous aurions du être aussi.
A lot of people enter without paying and 
we should too.
c. *Pierre a renversé des tables, mais Paul 
n'a pas.
Pierre turned the tables over but Paul 
didn't.

The superficial similarity between the past participle's 
behavior and the behavior of the infinitive in the faire
construction might suggest that the structure of the 
faire + infinitive is a complex verbal structure in some 
way like the past participle. However, this complex 
verbal structure, unlike the past participle, cannot be 
base generated since reflexive clitics do not appear on 
past participles but do appear on the infinitive of the 
faire construction as (65)-(66) indicate.

(65)a. *Jean est se parlé.
    b. *Jean est se connu.
(66)a. Jean fait se parler Marie.
    b. Jean fait se connaître Marie.

Kayne's original proposal explained these facts. Unlike 
other infinitives, the infinitive in the faire construction 
was raised out of the embedded S. The difference between 
(62) and (63) in his analysis could be attributed to the 
fact in (62) an S is given an interpretation by the prin-
ciples of null anaphora while in (63) more than an S is 
null, and no interpretation can be assigned to the raised 
infinite. In Quicoli's analysis, since the embedded 
verb remains within the S, we expect null anaphora to be 
able to operate.

More importantly however is the fact that Quicoli's 
analysis, like Kayne's, is unable to provide a principled 
account for why in (8) only (8d) is acceptable: Quicoli's 
analysis, like Kayne's, proposes the non-third person
object with the verb. From that position it should be free to cliticize onto the matrix verb. However, only the third person clitics (i.e., _le_, _la_, _les_) do.

In addition, it seems difficult to account in a natural way for why prepositional clitics of the infinitive can appear to the left of _faire_ while dative clitics of the infinitive cannot. Of course, it is possible to sort the two movement phenomena into different rules and insure that they apply at different points in the derivation in order to make the relevant distinction between prepositional and dative clitics. However, at least two problems are associated with this kind of solution. First such a solution would constitute an important retreat from the goal of a general move a rule which would have the effect of moving prepositional or dative clitics indiscriminately. Quicoli adopts this alternative and finds it necessary to place the pro-preposition movement rule (i.e., En-Placement) in the cycle and Clitic Placement (which applies to the other non-reflexive clitic forms, in the post-cycle). Such a move certainly increases the class of grammars and is to be avoided if possible.

In addition, Quicoli's analysis makes crucial use of a left to right asymmetry in the version of the Specified Subject Condition proposed in Chomsky (1973). That version blocks movements from the right of the subject but not from the left. Consequently, in Quicoli's analysis, $V^1$
is moved to the left of the subject where the escape of the clitic pronouns is not prevented. However, this left-right asymmetry was eliminated from the Specified Subject Condition as it was further defined in Chomsky (1976) or from its extension as the Opacity Condition in Chomsky (1978). Recall that the Specified Subject Condition was formulated as (67) in Chomsky (1976).

(67) The Specified Subject Condition

Given the structure

\[ \ldots X \ldots [\alpha \ldots Y \ldots] \ldots X \ldots \]

no rule can involve \( X \) and \( Y \) where \( \alpha \) contains a subject distinct from \( Y \) and not controlled by \( X \).

Under this more restrictive version of the Specified Subject Condition, whether one analyzes the movement of the prepositional clitic as a multiple movement as illustrated in (68) or as a single movement, as illustrated in (69), there will still be a violation of the Specified Subject Condition.

(68) \( \text{faire} \ldots [\alpha \ldots V \ldots NP \ldots PP \ldots] \ldots \)

(69) \( \text{faire} \ldots [\alpha \ldots V \ldots NP \ldots PP \ldots] \ldots \)

In each case there is a movement outside of \( \alpha \), that is \( S \), and \( \alpha \) still contains a subject. Consequently, the structure should be blocked. The same objection can be made
with respect to the movement of the third person object clitic in (8d). That clitic too moves outside the domain of a specified subject.

Consider now how these facts can be handled within the base analysis we are proposing. The problem with Kayne's verb-raising analysis disappears in this framework since no specified subject intervenes between faire and the embedded V. We can postulate an adjunction rule like (70) to accomplish this.

\[
(70) \quad X - \left[ V^+ \right] - V - Y \Rightarrow 1 - 2 + 3 - 4
\]

where \( +F \) is a feature or set of features distinguishing the so-called causative verbs.

Rule (70) accounts for the null anaphora facts described above. \textit{Faire} does not permit the infinitive to be replaced by null anaphora because less than a constituent will be null. We return to adjunctions like (70) again in Chapter 6.

The clitic phenomenon can be handled quite elegantly within this framework as well. There is a contrast in acceptability between (6) and (7), repeated below as (71)-(72), because of the Association Principles.

(71) \quad Elle lui fera manger ce gâteau.
     \quad \quad \quad She will make him eat this cake.

(72) \quad *Je lui ferai écrire mon ami.
     \quad \quad \quad I will make my friend write to him.

238

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Taking (71) and (72) as paradigmatic, we will assign them the structures in (73) and (74) respectively. These structures are factored in accordance with the Association Principles. Structures (73) and (74) are blocked since they involve crossing structures. Only (73) is well-formed given the Association Principles. This provides a simple account for why the clitics before *faire* (with the exception of *le, la, les* which we will deal with shortly) are understood as complements to *faire* and not the embedded verb.

Of course the explanation just outlined requires crucially that there be trace immediately to the left of the infinitive. Two alternatives present themselves. On the one hand we can assume that all clitics are base generated pre-verbally, in a way similar to either Strozer (1977) or Rivas (1977). In this case, the clitics will be generated to the left of the infinitive and moved to the left of the finite verb (or "climb") by the general "move α" rule. Alternatively, we can assume that there is in fact a true movement rule, at least for some clitics. In this case, the clitics would need to move from verb to verb in order to insure the presence of the trace to the left of the infinitive. Either of these alternatives also requires that the non-reflexive clitics before the infinitive be excluded by a filter.

The treatment of the *faire* construction being
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
proposed also provides a simple explanation for the acceptability of the sentences in (9), which involve the prepositional clitics. They are able to precede faire because no structural ambiguity is involved: that is, faire is associated with an NP not a PP.

Consider now the sentences in (8) which involve the direct objects to the embedded verb. Recall that sentences like (75) were unacceptable while sentences like (76) were acceptable.

(75) *Marie te fera embrasser à Paul.
(76) Je le ferai manger à Jean.

It is a simple matter to show that (75) is blocked by the Association Principles. Structure (76) provides an analysis of (75) as a paradigmatic example. Note that once again \( A_1 \) and \( B_1 \) are separated by \( X \) which contains \( A_2 \) but not \( B_2 \). Consequently (77) is a crossing structure and is blocked.

The fact that the Association Principles will block (72) raises the question as to why (8d) should be acceptable. The answer to his query is in Emonds' (1976) analysis of French clitics in which all clitic movements are structure preserving (or with minor modifications of Emonds' proposals, base generated) except for \( le, la, les \) which are moved by a local rule. The local status of the \( le, la, les \) rule provides an explanation for why (8d) but
not (8a-c) is acceptable since at the level of surface
structure defined above, a class of local movements and
deletions have not operated (cf., Chapter 3).

As an aside, a related but otherwise anomalous
fact about French falls out of this treatment. There
exists a second construction in which faire partakes where
the agent NP appears in a par-(by)-phrase. In this con-
struction dative and direct objects on the embedded verb
may appear before faire. For example, of the sentences in
(78), (a-b) exhibit dative clitics preceding faire, while
(c-d) exhibit direct objects before faire.

(78)a. Il te fera prêter assistance par son fils.
    He will make his son give you help.

   b. Nous te ferons acheter ce jouet par Jean.
       We will make Jean buy you that toy.

c. Pierre te fera embrasser par Marie.
    Pierre will make Marie kiss you.

d. Pierre s'est fait embrasser par Marie.
    Pierre will make Marie kiss him.

These sentences are unexplained under Kayne's analysis
where the par phrase is derived by postposing the subject
in the embedded S. This NP movement should leave a trace
in the subject position which would block the movement of
the dative clitics in (78a-b) by virtue of the Specified
Subject Condition. The sentences in (78c-d), while
generable, simply complicate the problems that Kayne's analysis has with the data in (8) since the same unacceptable clitic sequences in (8) are acceptable in (78).

Quicoli attempts to overcome this difficulty by suggesting that an indirect object might sometimes be inside $V^1$ and other times not. When such an indirect object is inside $V^1$ it will be preposed and subject to escape from the embedded clause. When it is not inside $V^1$ it will be unable to move outside the embedded clause. This proposal reduces a great deal the explanatory force that Quicoli claims for his overall analysis. The difference between (79a-b) is only that in (b) the indirect object has not been inside $V^1$, that is, has not fronted, and thus is not in a position from which it can escape.

(79)a. Je lui laisserai écrire une lettre par Jean.
I'll let Jean write a letter to him.

b. *Je lui laisserai écrire Jean.
I'll let Jean write to her.

The fact that intransitive, but not transitive, écrire has its direct object outside $V^1$ is a lexical fact and Quicoli's analysis makes (80a) just as possible as (80b).


245
Moreover, the fact that all the direct object clitics can cliticize to faire when a par-phrase is present, a fact unexplained in Quicoli's analysis, suggests that appealing to a hierarchical distinction to explain why indirect object clitics can cliticize to faire is missing an important generalization. Clearly no hierarchical distinction can be made with respect to the direct object clitics and yet they behave just as the indirect object clitics do.

The problems that we have been noting do not arise within the framework of the Association Principles. Let us assume that faire is subcategorized as in (81).

\[
(81) \text{faire, } +V, +_____ \{ \begin{array}{c}
\text{NP} \\
\text{par} \quad \text{NP} \\
\text{PP} 
\end{array} \}\]

This assumption provides a simple explanation for the acceptability of the sentences in (78). Just in these cases, faire is associated with a PP, that is par+NP. This means that the Association Principles will not block the structure since there is no "structural ambiguity." That is to say, the other intervening elements are associated with NP's, not PP's. In this respect, the sentences in (78) are acceptable for the same reason that the sentences in (9) involving pro-prepositional clitics are acceptable. Moreover, this analysis has the advantage of providing a unified treatment of both the direct and indirect object clitics, a generalization missing in Quicoli's analysis.
The last issue we discuss in relation to the faire construction is the difference between it and the laisser construction observed in (2)-(3). Recall that the verb faire required a structure like ... faire V NP à NP ... while laisser could appear optionally in that structure or with a simple infinitival S. Within the transformational proposal of Kayne and Quicoli this fact is glossed over and it is difficult to imagine how, within their framework, this fact could be captured without appealing to rule features and thus significantly increasing the class of grammars. Within the base analysis being proposed here this fact can be captured simply, without increasing the class of grammars, by specifying laisser as in (82).

(82) laisser, V, + \[ \{ V - NP \} \]
    \[ S \]

The specification that laisser takes an S distinguishes it from faire and explains why sentences like (3a), reproduced below, are acceptable.

(3)a. Jean laisse Paul parler à Marie.
    b. Jean laisse parler Paul à Marie.

Jean is letting Paul talk to Marie.

I conclude from the foregoing discussion that the Association Principles make available a number of significant generalizations about clitic placement and the faire construction in French that are uncapturable within other
3.2. Asymmetries of Quantifier Movement, the SSC and the Association Principles

The analysis of French causatives sketched in this chapter derives some additional corroboration from its ability to eliminate certain problems associated with quantifier movement in French.

Quicoli (1976) constitutes an important attempt to show the explanatory value of Chomsky's conditions in relation to phenomena in French. Quicoli begins by noting that a quantifier can be separated from the head it modifies. The quantifier may appear to the right of its head, as the sentences in (83) show, or it may appear to the left of the head, as the sentences in (84) show.

(83)a. Tous les garçons sont partis à la guerre.
   All the boys have left for the war.
   b. Les garçons sont tous partis à la guerre.
      The boys have all left for the war.
   c. Les garçons sont partis tous à la guerre.
      The boys have left all for the war.

(84)a. Jean a manger tout.
      Jean has eaten everything.
   b. Jean a tout manger.
      Jean has eaten everything.

Quicoli assumes, following Kayne (1975), that these
sentences are related by a rightward and leftward quantifier movement rule, R-Tous and L-Tous respectively. The issue of whether these facts are best handled by a movement or construal rule is not relevant here. I will assume for the moment that sentences like (83)-(84) will be related by the general "move α" rule.

Quicoli notes that the quantifier movements possible in (83)-(84) are subject to Chomsky's conditions. So for example, L-Tous may not operate to move a quantifier outside a Tensed Sentence, nor may R-Tous move a quantifier inside a Tensed Sentence. These facts are represented in (85) and (86) respectively.

(85)a. *J'ai tout voulu que Marie mange.
   b. J'ai voulu que Marie mange tout.
      I wished that Mary eat everything.

(86)a. *Mes amis ont voulu que ce garçon mange tous de la salade.
   b. Tous mes amis ont voulu que ce garçon manger de la salade.
      All of my friends wanted this boy to eat the salad.

The unacceptability of (85)-(86) is accounted for by either the Tensed-S or Specified Subject Conditions.

Quicoli contrasts the sentences in (87) with the sentences in (88).

249
(87)a. Elle a voulu tous les lire.
   b. Elle a tous voulu les lire.
   She had wanted to read them all.

(88)a. Elle a semble les avoir tous lus.
   b. *Elle a tous semble les avoir lus.
   She seemed to have read them all.

Quicoli proposes to explain this contrast by capitalizing on the general assumption that (87) involves a PRO in the subject position of the embedded complement while (88) involves a trace. The structure for (87) and (88) would be (89) and (90) respectively.

(89) elle a voulu [S_{PRO tous les lire}].

(90) elle a semblé [S_{NP{e} tous les avoir lus tous}]

Of course tous may not move out of either S in (89) or (90) as long as either subject NP is present. Quicoli proposes to surmount this difficulty by ordering Equi-NP deletion (EQUI), which would apply in (89) but not (90), prior to the quantifier movement. This will mean that at the point where the quantifier movement takes place (87) and (88) actually have the structures in (91) and (92) respectively.

(91) elle a voulu [S_{tous les lire}]

(92) elle a semble [S_{NP{e} tous les avoir lu tous}]

In these structures, the Specified Subject Condition (SSC)
will not block the quantifier from moving in (91) to produce (87) but it will block the quantifier movement from applying in (92) so that we avoid generating (88).

Pollock (1978) provides an elaboration of Quicoli's analysis on this point. Pollock observes that there is a division in verbs selecting infinitival complements (the so-called EQUI-verbs) as to whether quantifiers can be moved outside the infinitival complements. One class, like (93) below, patterns like vouloir in (87) and permits such movements. A second class of verbs does not allow any such movement, as the examples in (94) illustrate.

(93)a. Jean va tous pouvoir les lire.
    Jean will be able to read them all.

     b. Jean va tout devoir lire.
        Jean will have to read everything.

     c. Jean a tout failli lire.
        Jean has almost read everything.

     d. Jean a tous osé les lire.
        Jean has dared to read them all.

(94)a. *Jean a tous certifié les avoir lus.
      Jean has certified to have read them all.

     b. *Jean a tous proclamé les avoir lus.
        Jean has proclaimed to have read them all.

     c. *Jean a tous juré les avoir lus.
        Jean has sworn to have read them all.
d. *Jean a tous couru les mettre sur l'étagère.

Jean has run to put them all on the shelf.

The difference between these two classes of verbs suggests that for the class in (94) PRO is still present when the quantifier movement rule applies so that the sentences are blocked by the SSC. In (93) however, the PRO has already been deleted by EQUI when the quantifier movement rule operates so that the SSC is not applicable. Pollock concludes that there are two EQUI rules, $\text{EQUI}_1$ and $\text{EQUI}_2$, which operate in French: $\text{EQUI}_1$ applies to the verbs in (93) and is ordered before quantifier movement, while $\text{EQUI}_2$ applies to the verbs in (94) and is ordered after quantifier movement.

This general line of analysis has two important difficulties associated with it that we would like to avoid if possible. On the one hand the need to recognize two EQUI rules is problematic. The fact that there are two rules performing the same function suggests a loss of generalization in the analysis. Moreover the ability of linguistic theory to define $\text{EQUI}_1$ and $\text{EQUI}_2$ in such a way that they apply only in the context of their appropriate verbs is far from non-controversial. A more important problem associated with this line of explanation however comes from the ordering of EQUI prior to the quantifier movement. Within the restricted version of linguistic theory outlined in Chomsky and Lasnik (1977) deletion rules
apply subsequent to movement transformations and are completely independent of the construal rules. This means that within that model, whether the quantifier phenomenon is a true movement or a construal rule, there is no way to order EQUI prior to quantifier movement. In this way Quicoli's and Pollock's analyses of the facts at hand require a relatively less constrained version of linguistic theory than the one proposed in Chomsky and Lasnik (1977).

I want to propose that the verbs in (93) which, like vouloir, permit an apparent violation of the SSC actually involve a structure like the faire + infinitive construction. That is, a sentence like (89) has a structure like (95).

(95)

In (95) no specified subject is present to block the movement of the quantifier. These verbs differ from the faire + infinitive construction in two major respects. They require a PRO rather than a lexical agent phrase, and they do not undergo the V-V adjunction rule. Now the sentences like those in (94) which do not permit the quantifier to
move outside the embedded complement only involve structures like (96).

\[
\begin{align*}
&\text{S} \\
&\text{NP} \to \text{VP} \\
&\text{V'} \to \text{V} \\
&\text{P} \to \text{S} \\
&\text{NP} \to \text{VP} \\
&\text{V'} \to \text{V} \\
&\text{PRO} \to \text{V} \\
\end{align*}
\]

Jean a certifié PRO tous les avoir lu

In this structure a specified subject is present. Consequently the movement is blocked. Let me anticipate an objection. It may be argued that analyzing some infinitival complements as VP complements is incapable of accounting for the existence of passive morphology in such cases. This objection assumes that the passive morphology is added transformationally, rather than base generated. Within a general core grammar framework this possibility is precluded. Instead, it seems that the passive morphology should be base generated, and when such morphology is present, NP-movement is obligatory in a way similar to that sketched in the analysis of passives presented in Chapter 4.

The analysis just outlined has the advantage of avoiding the postulation of two EQUI rules and, in addition, preserves the ordering of movement and construal.
rules with respect to deletions. There is, however, independent empirical evidence in favor of this analysis as well. First, it is important to establish than an independent negative may not appear on the embedded verb in the faire + infinitive construction. This fact, noted in Gross (1968), is illustrated by the unacceptability of (97).

(97) *elle a fait ne pas manger ce gâteau à Jean

Assuming that negation cannot appear on an embedded VP but that it can appear on an embedded S, the sentences below are of interest.

(98)a. J'ai voulu ne pas tout lire.
   I have wanted to not read everything.

b. *J'ai tout voulu ne pas lire.
   I have wanted to not read everything.

In order to reconcile (98) with the generalization that negation cannot appear on an embedded VP complement, we must postulate that vouloir can also select a sentential complement as well as a VP complement. The presence of the negative then requires that the underlying structure of (98) be (99). Given the structure in (99), the Specified Subject Condition will block the movement of tout and the grammar will not generate (98). Of course the sentence in (98b) is acceptable without the negative morphology because then it is possible to analyze the sentence as having an
underlying VP, not an underlying S, complement.

The argument just presented depends crucially on analyzing L-Tous as a movement over a variable. Kayne (1975) and Pollock (1978) propose that L-Tous is not a movement over a variable but can instead be analyzed as (100).

(100) \[ X-V-Q-Y \rightarrow 1 - 3 - 2 - 4 \]
\[ 1-2-3-4 \]

Before arguing about the specific reasons forwarded to justify (100), I want to point out that for those who accept the general hypothesis of the "move α" rule, this debate is irrelevant since the grammar will move quantifiers to the left over variables in any event. The question, under those assumptions, is how to block the overgeneration involved which, I believe, the VP analysis in conjunction with the Association Principles does. The arguments in favor of (100) necessitate assuming that core grammar will not move quantifiers to the left over variables in spite of (100). I think this assumption would lessen the general restrictiveness of the core grammar hypothesis. Nevertheless I will consider the individual arguments taken to motivate (100).

Four main arguments are forwarded to justify the formulation in (100). First is the fact that quantifiers cannot move over the negative in sentences like (98). Inasmuch as I have just detailed an alternative explanation
explanation for this fact, I will not consider this an argument in favor of (100). The second argument involves the interaction of R-Tous and L-Tous. So, for example, the sentences in (101) are acceptable but the sentences in (102) where the direct object tout moves over a Z-V sequence is unacceptable.

(101)a. Marie leur a tous tout écrit.
    Marie has written everything to them all.
    b. Marie leur a tous tout dit.
    Marie has said everything to them all.

(102)a. *Marie leur a tout tous écrit.
    b. *Marie leur a tout tous dit.

However, neither of these sentences constitutes evidence in favor of (100) if one assumes a version of linguistic theory incorporating the Association Principles. The sentences in (102) would be excluded by the Association Principles as is argued in Chapter 2.

The third argument favoring (100) concerns the contrast between (103a) and (b).

(103)a. J'ai voulu absolument tout lire.
    I have absolutely wanted to read everything.
    b. *J'ai tout voulu absolument lire.

If L-Tous is formulated as in (100) then the unacceptability of the sentence in (103b) is accounted for inasmuch as
the structural description will not permit the movement over the adverb absolument. The sentences in (104) noted in Kayne (1975) require a complication in this general line of explanation.

(104)a. Il faudrait tous les deux mieux les connaître.

One would have to know them both better.

b. Elle voudrait tous mieux les connaître.

She would like to know them all better.

The sentences in (104) exhibit a structure in which the quantifier has moved over an adverb in contrast to the sentences in (103). The sentences in (104) are part of a larger paradigm in which adverbs like mal, mieux, toujours, beaucoup, and déjà all seem to be capable of separating a quantifier like tous from a verb. The sentences in (105) are representative of this characteristic.

(105)a. Il les a tous toujours appréciés.

b. Il les a toujours tous appréciés.

He has already appreciated them all.

(106)a. Je les ai déjà tous mis à la poubelle.

b. Je les ai tous déjà mis à la poubelle.

I have already put them all in the trash.

(107)a. Il a tous beaucoup appréciés.

b. *Il a beaucoup tous appréciés.

He appreciated them ali alot.

259
(108)a. Elle a tout très mal compris.
   b. *Elle a très mal tout compris.

She understood everything poorly.

In order to account for sentences like those in (105)-(108) Kayne suggests a rule ordering solution. He proposes that the adverbs appear to the right of tous in underlying structure and they are moved to the left of the verb after the application of L-Tous. So, for example, (107a) has the derivation in (109).

(109)a. Il a appréciés tous beaucoup.
   b. Il a tous appréciés beaucoup.
   L-Tous applies
   c. Il a tous beaucoup appréciés
   Adverb movement applies

This solution requires rule ordering in order to account for why the sentences in (107) and (108) are not acceptable. Moreover, it requires some additional movement applicable to toujours but not beaucoup in order to generate (105) and (106).

This solution, while descriptively adequate, is not the solution with the greatest amount of explanatory adequacy. Within the framework of core grammar, rule ordering is avoided so as to restrict the class of grammars compatible with linguistic theory. The facts in (105)-(107) can be explained by the general "move a" rule constrained by
output filters. The fact that a class of adverbs including *beaucoup* may appear only to the right of *tous* would be described by an output filter like (110), in the absence of a more principled explanation.

(110) *Beaucoup tous*

We will therefore prefer the analysis consistent with core grammar over the analysis employing such strong theoretical devices as rule ordering and contextual predicates.

The fourth argument claimed to support (100) concerns the contrast between (111) and (112).

(111)a. Pierre a mis les livres tous sur la même étagère.

Pierre has put all the books on the same shelf.

b. *Pierre a tous mis les livres sur la même étagère.

(112) Pierre a tous pu les lire.

Pierre has managed to read them all.

The argument seems to be that R-Tous applies in these cases to give structures like (113) and (114).

(113) Pierre a mis les livres tous sur la même étagère.

(114) Pierre a pu lire les tous.

At this point (100) cannot apply in either structure
because Q is not contiguous to V. After clitic placement, (114) will become (115), however, and it would then be possible to move the quantifier by (100).

(115) Pierre a pu les lire [NP$^e$ tous

It will never be possible to apply L-Tous in (113) however.

The explanation just outlined places the locus of differentiation between (111) and (112) in the structural description of the quantifier movement rule. I want to sketch an alternative explanation which is compatible with the core grammar framework. Consider the sentences in (116)-(117), observed in Kayne (1975:42).

(116)a. Elle ne connaît que nous autres.
   She knows only us others.
   b. Il ne voit que vous deux.
He only sees you two.

(117)a. *Elle nous connaît autres.
   b. *Il vous voit deux.

These sentences seem to indicate that, as Kayne notes, clitic placement cannot move anything but a bare pronoun. The quantifiers in (111)-(112) form a constituent with the pronoun and therefore clitic placement is not applicable. We might ask then how clitic placement operates in (112). A possible answer is to say that there is a rule readjusting tous, but not deux or autres, outside the NP node dominating the pronoun. Consequently, clitic placement
which applies only to bare pronouns will operate in (112) but not in (114).

It now seems possible to explain the contrast between (111) and (112) by employing the fact that PRO-tous does not form a constituent whereas les livres-tous does. Assume that the readjustment rule moving tous outside the NP only applies when a pronoun is to its left. In this case the quantifier will still be dominated by an NP and Bresnan's (1976) relativized A over A condition could prevent the extraction of the quantifier.

The arguments in favor of a rule like (100) seem less than compelling. I conclude that a general quantifier movement rule which is a sub-case of the "move a" rule operates to generate the relevant sentences we have treated in this section. Given this hypothesis, it also seems preferable to analyze certain EQUI-verbs as having an underlying VP complement in order to reconcile the behavior of moved quantifiers with the conditions on rules of grammar that are being defended in this work.

3.3. The Marked Nature of the VP Complement

In this sub-section I want to suggest how it is that French has a VP complement especially in light of the result of research done in the framework of Chomsky's conditions showing that English has only a sentential complement, a fact that I do not dispute. This asymmetry between the grammars of French and English is curious. Similarly,
my hypothesis that adjectives have no post-head complements in the base has seemed curious to some linguists because it introduces a degree of asymmetry in the base component of the language.

Arguments based on symmetry, while appealing, are only of limited value. On the basis of symmetry one might argue that English should have VP complements since all other categories can apparently be complements. The assumption that English only has sentential complements is asymmetrical in this sense. However, it is well-motivated because it predicts that the SSC will be applicable in all infinitival complements in the language, a prediction that is apparently borne out. In the same way, one might argue on the basis of symmetry that since nouns and verbs can have adjectives in their complement systems, so should adjectives and prepositions despite the fact that we do not find such adjectival complements.

In Chomsky and Lasnik (1977) one of the major filters proposed is something like (118).

\[(118) \ast [\text{NP, NP} \ TENSE \ VP]\]

\[(119) \ [s, \text{NP} \ TENSE \ VP]\]

Chomsky and Lasnik relate the existence of this filter to the fact that there is a structure in the language like (119). In informal terms it appears that if one thinks of the grammar as pairing a surface structure, $s_i$, to a base
structure, $b_i$, then the unmarked case is for only one $b_i$ to correspond to any given $s_i$. In other words, the paring of $s_i$ and $b_i$ is unique. Such a general principle has been called the Uniqueness Principle. K. Wexler has proposed (personal communication) a way of instantiating the Uniqueness Principle in grammars. An informal statement of Wexler's suggestion is that the learning procedure is constructed in such a way that everytime a rule is postulated, a set of filters $F_i$, $F_j$ ... $F_n$ is also postulated to guarantee a unique $b_i$ for any $s_i$. Wexler further suggests that any $F_i$ in this set could be suppressed given a piece of positive evidence demonstrating that $F_i$ is inapplicable in the language. In this framework then postulating a rule like (120) implies creating a filter like (121).

(120) $S \rightarrow NP\ TENSE\ VP$

(121) $[^{C\ NP\ TENSE\ VP}]$ where $C$ is not an $S$

Of course, (121) has (120) as a special sub-case.

This general line of explanation would provide a rationale for why adjectives have no post-head complements. Let us assume that $ly$-adverbs are underlyingly adjectives. With this assumption we find that adjectives appear to the left of adjective phrases, prepositional phrases, noun phrases verb phrases and sentences, as the phrases and sentences in (122) illustrate.

265
(122)a. A $A^1$
John was completely tired.

b. A $V^1$
John finally was tired.

c. A $N^1$
A small boy.

d. A $P^1$
Jane ran completely down the hill.

e. A $S$
Finally John was tired.

Postulating the phrase structure rules like (123) to account for sequences like (122) will entail then postulating filters something like those in (124) as well.

(123)a. $A^2 \rightarrow \ldots A^2 \ldots A^1 \ldots$

b. $V^2 \rightarrow \ldots A^2 \ldots V^1 \ldots$

c. $N^2 \rightarrow \ldots A^2 \ldots N^1 \ldots$

d. $P^2 \rightarrow \ldots A^2 \ldots P^1 \ldots$

e. $S^2 \rightarrow \ldots A^2 \ldots S^1 \ldots$

(124)a. $*[_{C^n} A^n A^n]$ where A is not the head of $C^n$

b. $*[_{C^n} A^n V^n]$ where $C \neq V$

c. $*[_{C^n} A^n N^n]$ where $C \neq N$

d. $*[_{C^n} A^n P^n]$ where $C \neq P$

e. $*[_{C^n} A^n S^n]$ where $C \neq S$
\[ f. \ast \left[ \frac{A^n}{C^n} \ D^n \right] \text{ where } D \text{ is not the head of } C^n \]

The filters in (124a-e) can be collapsed as (124f), where head of phrase is understood as defined in Emonds (1978). The filters in (124) will prevent any structure generated by rules like (125) thus having the effect of prohibiting post-head complements to adjectives.

(125)a. \( A \uparrow \to \ldots A \ldots P \ldots \)
b. \( A \uparrow \to \ldots A \ldots S \ldots \)

Since there is no positive evidence falsifying the filters in (124), they are maintained and form part of the adult grammar.

This general line of inquiry in relation to asymmetries in base structures is explored in some detail in Hendrick (in preparation), especially in relation to PP's. The consequences for VP complements are of immediate relevance. The postulation of a rule like (126), which we have been assuming following Emonds (1978), implies the postulation of a filter like (127).

(126) \( V' \rightarrow \begin{cases} V' \\ \{ \text{PRON}^* \} \end{cases} V \)

(127) \( \ast \left[ \frac{V'}{C^n} \ D^n \right] \text{ where } D \text{ is not the head of } C^n \)

This filter will not block structures generated by (126) but it will block structures like (128).
In English the filter in (127) is not falsified by positive evidence and is part of an adult grammar. In French, however, it will be falsified by the clitic movement phenomenon and quantifier movement phenomenon described above. When the learning procedure encounters sentences like (129) or (130), it will be forced to suppress the filter in (127).

(129)a. Cela y fera aller Jean.
   That will make Jean go there.
   b. Elle en fera sortir Jean.
   She will make Jean leave there.
   c. Je le ferai manger à Jean.
   I will make Jean eat it.

(130)a. Elle à tous voulu les lire.
   She had wanted to read them all.
   b. Jean va tous pouvoir les lire.
   Jean will be able to read them all.

Given the restrictive version of linguistic theory adopted here, it is impossible to construct an analysis of these sentences consistent with the SSC. Suppressing the filter in (127) and assuming that a VP complement is involved in these sentences is consistent with this restrictive linguistic theory and avoids the numerous problems associated
with the attempts to maintain that a sentential complement is involved in (129)-(130). At the same time, this analysis does not force one to admit the existence of VP complements in English since they can only be postulated in very specific contexts.
FOOTNOTES TO CHAPTER 5

1. This chapter is, in part, a revision of Henrick (1976) which was written in 1976 when I first began to formulate the Association Principles. The data presented in this chapter is adapted from Kayne (1975).

2. A slight variation on the general line of analysis presented here would be to assume that sentences like (1a) were derived from (22) by an intraposition rule. Such an intraposition rule seems necessary at least to account for sentences like (i).

(i) Cela fait préférer à Jean la syntaxe à la phonologie.

3. Capitalized SE is a cover term for all reflexive clitics throughout this chapter.

4. These generalizations are true for inherent reflexives, not the productive, "standard" reflexives.

5. The analysis being proposed here for the inherent reflexives would need to stipulate lexically that the inherently reflexive verbs could not appear with certain objects, that is to say, there would have to be a lexical reflex of (41) in the grammar I am proposing. Such restrictions would generalize with the restrictions in (37)-(39) so that the [- _____ même] restrictions would not have to be stipulated separately however. This addition leaves my explanation of the scarcity of the inherently reflexive verbs in tact: such verbs still require a relatively complex lexical entry when compared to other, standard verbs.

6. The mechanism which insures that only Marie can be a controller for the reflexive is not worked out here. Various avenues of inquiry present themselves: one might propose a structure building rule making Marie the logical subject at the level of Logical Form for example, or alternatively one might adopt an analysis of opaque predicates like that presented in Williams (1979). I am not prepared to pursue these alternatives at the present moment however.
Chapter 6

POSSIBILITIES FOR A PARAMETERIZED THEORY

It has been claimed throughout this work that any nesting requirement in general, and the Association Principles in particular, serve to disambiguate structures in the same way that other principles of Universal Grammar like the A/A condition do. This is why the associated B's in the formulation of the Association Principles are required to be "non-distinct." The notion that two elements are structurally non-distinct is not a trivial one. Until now the discussion of the Association Principles has avoided directing attention to this issued by limiting consideration to cases where the two elements are of the same category. In this chapter more attention is paid to this issue and the possibility that the definition of non-distinctness might be somewhat more complex is considered. It is suggested that some aspects of language variation could potentially be accounted for by assuming a parameterized version of linguistic theory in which language particular grammars have a degree of freedom in defining certain primitives like cyclic categories. Similarly languages might define the notion of structural non-distinctness somewhat differently as well, giving rise to a
variety of empirical consequences which constitute a portion of the superficial diversity between languages.

1. V and N As Structurally Non-Distinct in Italian

Rizzi (1978) observes that, in Italian, infinitival complements can be clefted. The sentences in (1) are examples of such cleft sentences.

(1)a. E' proprio a reportargli i soldi che sto andando, stai tranquillo!
   It is just to bring him back his money that
   I'm going, don't worry!

b. Quanto a questa storia, è discuterne con Mario che dovresti.
   As for this story, it is to discuss about
   it with Mario that you should have.

c. La tua disavventura, è proprio a raccontarla a Francesco che ho cominciato.
   Your mishap, it is just to tell it to Francesco that I have begun.

In addition Italian has a "clitic placement" rule, which Rizzi formulates as (2). This rule relates sentences like those in (3).

(2) Clitic Placement

| vbl | V | vbl | PRO | vbl |
| 1   | 2 | 3   | 4   | 5   |
| 1   | 4+2 | 3   | ø   | 5   |

272

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(3)a. Gianni presenterà Maria a Francesco.
   b. Gianni presenterà Maria a lui.
   c. Gianni gli presenterà Maria.

In the environment of certain matrix verbs, clitic placement can operate to move a clitic pronoun from an embedded sentence to the left of the matrix verb. The sentences in (4) are paradigmatic of this process.

(4)a. Gianni deve presentare la a Francesco.
   b. Gianni deve presentarla a Francesco.
   c. Gianni la deve presentare a Francesco.

The underlying (4a) can be transformed to either (4b) or (4c). For the moment I will ignore the fact that the clitic appears to the right rather than to the left of the infinitive in (4b). I will return to the phenomenon in (4c) shortly.

The process of clefting infinitives and the process of moving a clitic pronoun outside an embedded infinitive can not operate in the same sentence. That is, the sentences in (5), which parallel the sentences in (1) with the addition that the clitic placement rule has applied, are unacceptable.

(5)a. *E' proprio a riportare i soldi che gli sto andando ....
   b. *Quanto a questa storia, è discutere con Mario che ne dovresti.
c. *La tua disavventura, è proprio a raccontare a Francesco che la ho cominciata.

Taking (5a) as paradigmatic, it would have a derivation like (6).

(6a). essere [e] che sto andando a riportare i soldi gli
b. essere [e] che gli sto andando a riportare i soldi [e]—clitic placement has applied.
c. E' proprio a riportare i soldi che gli sto andando ... --cleft-sentence formation.

The question at hand is how to block the generation of sentences like (5a).

Rizzi proposes an analysis of this general phenomenon which depends on a "restructuring rule." He proposes that in general infinitives should not permit clitic placement to move a pronoun to the left of a matrix verb because of the Specified Subject Condition. In this way the sentence in (7) is unacceptable for the same reason that the sentence in (8) is unacceptable.

(7a). Piero deciderà di parlarti di parapsicologia.
Piero will decide to speak to you about parapsychology.

b. *Piero ti deciderà di parlare di parapsicologia.
(8a). Credo che Gianni la presenterà a Francesco.
I believe that Gianni will introduce her to Francesco.

b. *La credo che Gianni presenterà a Francesco.

Sentences like (9), however, which do permit a pronoun to move outside an infinitive are problematic for this account of (7), Rizzi assumes that these infinitives have a Specified Subject which should block the movement of the pronoun, everything else being equal.

(9a). Piero verrà a parlarti di parapsicologia.
Piero will come to speak to you about parapsychology.

b. Piero ti verrà a parlare di parapsicologia.

Rizzi suggests that the two verbs in (9) are reanalyzed as a single verbal complex as in (10). Assuming (10), when clitic placement applies it will not be sensitive to the Specified Subject Condition and will move the pronoun to the left of verrà. At the same time it will be unable to move the clitic pronoun to the left of parlare. This account opens up an explanation for why, in sentences like (5), it is not possible to cleft the infinitive and apply clitic placement to move a clitic to the left of a higher verb. Take (1a) repeated as (11) as an example.
(10) 

NP  
|   vverrà  
|   PRO  
|   parlare  
|   ti  
|   di  
|   parapsicologia  

Piero
(11)a. e' proprio a riportargli i soldi che sto andando, stai tranquillo.

b. *e' proprio a riportare i soldi che gli sto andando ....

The clitic pronoun can only appear to the left of the verb *sto if the restructuring operation has applied. Otherwise the Specified Subject will be present and exclude the sentence. The sentence in (11b) must have the structure in (12), that is, a structure where restructuring has applied. It is not possible to extract and cleft the infinitive in (12) due to the fact that the verb *riportare forms a constituent with *sto andando.

Clearly the restructuring solution is an extremely powerful one. Rizzi himself observes this (cf., Rizzi 1978: fn. 35). He notes that restructuring is first proposed in Chomsky (1974) to explain how verbs like take advantage of NP are able to admit double passives. For Chomsky, such a restructuring process was not productive, that is, it was limited to an extremely small class of verbs. In addition, since such an operation did not require any other operation to apply before it, it could be restricted to the level of lexical insertion. Rizzi's restructuring rule however requires productive application and furthermore must apply in the transformational cycle. That is, it cannot be limited to the level of lexical insertion. Couple these powerful characteristics with the
fact that no formalism is produced with which to describe the rule (i.e., the loss of the Specified Subject), and we are left with a less than ideal solution.

It is not necessary to allow linguistic theory access to such a powerful form of restructuring. Indeed it seems possible to account for the phenomenon under discussion without permitting any form of restructuring besides the familiar case of Chomsky-adjunction.

Suppose that V and N (or more precisely, any element having the features $[-N]^+\text{V}$ and $[+N]_-\text{V}$ as part of their syntactic feature matrix, which would include pronouns as well as lexical nouns) were to be treated as structurally non-distinct in Italian. In this event, the sentences like (11) would be explained by the Association Principles. Sentence (11b) is ill-formed because it involves a crossing structure as (13) illustrates.

(13) e' riportare i soldi PRO gli sto andando

$$\begin{bmatrix}
\text{VP} \\
\text{V}^e \\
\text{NP}^e \\
\text{NP}^e
\end{bmatrix}$$

In this way we are able to account for the sentences at hand with only a minor amount of modification in general conditions on rules of grammar. Of course, this analysis assumes in accordance with the analysis in Chapter 4 that the existence of a VP complement permits the extraction of clitic pronouns, a movement blocked in sentential

279
complements by the Specified Subject Condition.

The analysis being proposed here analyzes the so-called restructuring verbs and the causative verbs (and any other verb permitting clitic movement or quantifier movement to operate in apparent violation of the Specified Subject Condition) as both having VP complements. Rizzi (1978) develops an argument however against generating the restructuring verbs in the same manner that the causatives are generated. Since this clearly bears on the adequacy of the analysis that I am proposing, it is worthwhile considering Rizzi's argument in some detail.

The argument is based on the morphological alternation between avere and essere. It is a fact about Italian that most verbs require avere as an auxiliary while other verbs require essere. So for example we can observe sentences like (14).

(14)a. Piero \{ha\} voluto questo libro. \\
\{*è\}

Piero has wanted this book.

b. Piero \{*ha\} venuto con noi. \\
\{è\}

Piero "is" come with us.

An interesting phenomenon arises when we consider embedded infinitives. The selection of the auxiliary can be determined either by the matrix verb or by the infinitive. This means that if the infinitive is a member of the class
of verbs selecting essere rather than avere while the matrix verb selects avere, we expect that either avere or essere may appear. The examples in (15) exhibit this phenomenon and contrast with the sentences in (14).

(15)a. Piero \{ha\} voluto venire con noi. \{e\}

Piero has/"is" wanted to come with us.

b. Piero \{ha\} voluto mangiare con noi. \{*e\}

Piero has/"is" wanted to eat with us.

(15a) but not (15b) can appear with essere as the auxiliary because only in (15a) does a verb appear which belongs to the class of verbs selecting essere.

At first glance it may seem that essere can appear as an auxiliary whenever the verbal complex contains a verb of the class selecting essere. This general characterization turns out to be too broad, however, since, as Rizzi notes, only the right-most member of the verbal complex can trigger the presence of essere as the sentences in (16) demonstrate.

(16)a. Maria li [\{avrebbe\} potuti] stare [+avere] [+essere]

per andare a prendere lei stessa. [+essere] [+avere]

Maria them would have been able to be on the point of going to get ______ herself.
(16)b. Maria ci sarebbe [\(v\) dovuta cominciare ad] [+avere][+avere]
   \[\text{andare}\]
   [+essere]
   Maria there would "be" had to begin to go ________.

Rizzi proposes to account for these facts by the rule in (17).

(17) \(\text{avere} \rightarrow \text{essere} / [v_vbl \ldots vbl V_k]\) where \(V_k\)
   is a verb basically requiring \text{essere}

Rule (17) applies after restructuring has created a complex \(V\), and is only sensitive to the rightmost \(V\) of that restructured complex.

Building on this analysis of the \text{avere} - \text{essere} alternation, Rizzi argues that the causative construction could not have the same structure as the so-called restructuring infinitives since an embedded verb in the causative construction never permits \text{essere} to appear as the auxiliary of \text{fare}. The sentences in (18) are important to substantiate that the \text{avere} - \text{essere} alternation does not apply to verbs similar to the causative \text{fare}.

(18)a. Mario \{ha\} \{fatto\} venire il medico.  \{*è\} \{lasciato\}
   Mario has \{had\} the doctor come. \{let\}
(18)b. Mario lo \{ha\} \{fatto \} \{venire. \}
\{*è\} \{lasciato\}
Mario has \{had\} him come.
\{let\}

Before addressing Rizzi's argument directly, let us return to the formulation of the \textit{avere} \textit{essere} rule. Rizzi himself is somewhat dissatisfied with his proposal, noting that it would be desirable not to have a specific rule for the \textit{avere} \textit{essere} alternation. Rizzi observes that the alternation would best be accounted for by a general principle most probably making crucial use of the notion "head of phrase" as developed in Emonds (1976). Let us assume that only a head of phrase determines the choice of auxiliary, as Rizzi suggests, and let us furthermore assume that sentences like those in (15) involve a VP complement. In addition we will assume that the Italian verbal complex can be analyzed as the French verbal complex is in Emonds (1978). This means that if we postulate a simple adjunction rule like (19), we will transform by Chomsky-adjunction the underlying structure of (15a), which is (20), into (21). In (21) the formerly embedded verb is now the formal head of phrase.

(19) \text{V'} - \text{V} + 1 + 2

When (19) applies, as in (21), the auxiliary will be selected by the "new" head of phrase. The mechanics of this selection will require that the auxiliary be inserted

283
into the tree after the adjunction rule applies. Such a possibility is available in the system of lexical insertion at surface structure developed in Otero (1976) and den Besten (1976). A somewhat different possibility is a stratified process of lexical insertion where grammatical formatives but not lexical formatives are inserted at surface structure. A version of this second alternative is termed a "second lexical hook-up" and outlined in Stockwell, Schachter, and Partee (1973).
This general approach has no difficulty explaining why the rightmost V selects the auxiliary since when it applies left to right it will make the rightmost V of any depth of embedding the head of phrase. The auxiliary will then be selected by this new head of phrase.

Two issues remain. On the one hand, how do we account for the fact that this process does not apply in the causative construction to generate the unacceptable sentences in (18); and second, how do we guarantee that when clitic placement applies from an embedded infinitive, (19) is obligatory. Let us examine the issue of blocking the sentences in (18) first. Recall that in the discussion of the French causatives presented in the fifth chapter, we needed to postulate a V-V adjunction rule. Let us formulate this rule as (22).

\[
(22) \quad \begin{array}{c}
\begin{array}{c}
V \\
+\text{F}
\end{array}
\end{array}
\rightarrow V + 1 + 2 \quad \text{where} \quad +\text{F} \quad \text{is a feature or set of features distinguishing the verbs \{fare, lasciare, \ldots\}}
\]

In (22) we will take \ [+\text{F}] \ to be a feature or set of features uniquely distinguishing the class of causative verbs. This rule will transform (23) which is the underlying form of (18a) into (24). In (24) \underline{venire} \ would not be capable of triggering the change of auxiliaries. We are able to collapse the rules in (19) and (22) as (25). With (25) there is no possibility of generating the unacceptable sentences in (18).

285
Consider now how to guarantee that when clitic placement applies either (25) or (22), is obligatory. Let us assume that clitics are generated as sisters to the auxiliary by a phrase structure rule something like (26), a variant of the phrase structure rule proposed for French by Emonds (1976, 1978).

(26) \[ V' \rightarrow \{ V' \} - V \]

Given the hypothesis that verbs and nouns are ambiguous in
Italian, the Association Principles will require that the clitics and auxiliary are selected by the same verb. This has the effect of guaranteeing that when clitic placement operates on the complement of an embedded infinitive, the adjunction in (25), or (22), must operate so that the auxiliary is in an association relation with the same head.

It thus appears that assuming the presence of a VP complement in Romance in addition to the Association Principles permits an elegant explanation of the so-called restructuring phenomenon. This explanation has the advantage of avoiding the use of powerful restructuring rules. Apparent counter-examples to this analysis based on the essere~avere alternation dissolve when a more principled account of the phenomenon is considered.

2. P and N As Structurally Non-Distinct in English

In this section English cleft sentences are examined. It is argued that treating N and P as structurally non-distinct in English permits us to capture some interesting generalizations about English syntax.

In Chomsky (1977) it is noted that cleft sentences like (27) and (28) bear a strong resemblance to topicalized sentences like (29).

(27) It was Mary that called Claire.

(28) It could have been Claire that Mary called.

(29) Claire Mary called.
For example neither adjectives nor sentential complements are capable of being clefted, as the sentences in (30) illustrate, nor topicalized, as the sentences in (31) demonstrate.

(30)a. *It was happy that she felt.
    b. *It was that Claire was late that she mentioned.

(31)a. *Happy, she felt.
    b. *That Claire was late, she mentioned.

In addition, Chomsky observes that both of these constructions share the major characteristics of WH-movement. That is they exhibit the characteristics in (32).

(32)a. they leave a gap.
    b. where there is a bridge, there is an apparent violation of the Subjacency, the Tensed Sentence and the Specified Subject Conditions.
    c. they observe the Complex NP constraint.
    d. they observe the WH-island constraints.

As a consequence Chomsky hypothesizes that WH-movement is crucially involved in both cleft sentences and topicalized sentences.

Chomsky proposes that the following phrase structure rule be added to our hypothesis of the base:

\[ S^2 \rightarrow \text{TOP} \rightarrow S^1 \]
Assuming this base rule, a topicalized sentence like (29) has a base structure like (33).

(33)

```
  S^2
   TOP
     S^1
       COMP
         NP
           Claire
         S
           NP
             that
           V
             VP
               NP
                 Mary
                 called
                 WH
```

WH-movement will apply to (33), followed by the elimination of *that* and *WH* by the free deletion in COMP rule discussed in Chomsky and Lasnik (1977), to yield (29). Chomsky extends this analysis to cleft sentences so that a sentence like (28) has an underlying structure resembling (34). Once again WH is moved into COMP and subsequently *WH + that* is deleted by the free deletion in COMP rule. I will assume that this analysis of cleft sentences is basically correct.

It is possible to perform various operations in a cleft sentence. Among other things, it is possible to question the topic of the sentences as (35)-(36) indicate.

(35)a. It was Mary that called Claire.

   b. Who was it that called Claire.

(36)a. It could have been Claire that Mary called.

   b. Who could it have been that Mary called.

Presumably these sentences would be derived by successive
(34) 

S → NP AUX VP 

It could have been 

S² → TOP COMP S¹ 

Claire that Mary called WH 

S → NP VP 

V 

NP
WH-movement as the arrows in (37) show. It is commonplace to observe that in English WH-movement can apply either to an entire prepositional phrase, or it can apply to an NP inside a prepositional phrase, "stranding" the preposition in its base position. These possibilities are exemplified in (38b) and (c) respectively.

(38a) Mary talked to Claire.
   b. To whom did Mary talk?
   c. Who did Mary talk to?

If the same operations are applied in cleft sentences, somewhat different results are obtained. Examine the sentences in (39) and (40).

(39a) It was to Claire that Mary talked.
   b. It was Claire that Mary talked to.
(40a) Who was it that Mary talked to?
   b. *Who was it to that Mary talked?
   c. To whom was it that Mary talked?

Sentences like these are first observed, to the best of my knowledge, in Pinkham and Hankamer (1975). The issue at hand is why it is possible to move by WH-movement an entire PP, as in (40c), or to strand a preposition in the lower sentence as in (40a), but not to strand the preposition as in (40b).

If (40b) is generated by successive WH-movement as illustrated by (41), (40b) would involve the movements
indicated by arrow in (41). We need to know then why it is that we are not able to strand a preposition in an intermediate movement site.

This is a very general problem, involved not only in cleft sentences but also in another construction derived from WH-movement. Simple cases of WH-movement for example show the same behavior, as demonstrated in (42) and noted by Postal (1972).

(42)a. John thought that Mary talked to Claire.
   b. Who did John think that Mary talked to?
   c. To whom did John think that Mary talked?
   d. *Who did John think to that Mary talked?

Suppose that we were to assume that P and N are structurally non-distinct in English, though not in French as the analysis of the French clitic system made clear. Under this assumption we need first to assure ourselves that simple cases of movement of a prepositional phrase will not involve a violation of the Association Principles. Suppose the sentence in (38b) has a surface structure like (43a) in addition to or rather than (43b).

(43)a. \[ p_{pp} [p_{to} [n_{NP} whom]] \] Mary talk \[ p_{pp} [p_{e} [n_{NP} e]] \]
   
   b. \[ p_{pp} to whom \] Mary talk \[ p_{pp} e \]

(44) \[ p_{to} [n_{NP} whom] \] Mary talk \[ p_{e} [n_{NP} e] \]

If linguistic theory were to associate to with its empty
node and who with its empty node, as is illustrated in (44), a non-nested configuration would result. This problem would be avoided however if we could insure (a) that associations were defined on phrasal categories and not simply on terminal elements, and (b) that in instances like (43) they must be defined on the phrasal category. We are able to express the requirement in (a) by stipulating (45).

(45) A node A can be in association only if no B dominated by A is in association outside of A.

In addition we can guarantee the (b) requirement if we stipulate (46).

(46) \( A^{n-y} \) defines an association if and only if \( A^n \) cannot.

These conditions have the effect of disallowing (44) as a factorization for the purposes of the Association Principles. Instead only (43b) is a possible factorization, and (43b) does not violate the Association Principles.

Return now to (40c). This sentence has a surface structure something like (47).

(47) who it could have been to [e] Mary talked

\[
\begin{array}{c}
A_1 & A_2 \\
\text{pp} & [pe][Npe] \\
B_1 & B_2
\end{array}
\]

In (43) we were able to define an association only on the
phrasal node PP. In (47) however it is not possible to do so since the PP in COMP does not satisfy the requirement in (46). That is to say, the \[ NP_e \] in COMP defines an association with a third node outside the PP, that is the WH-element, in the matrix COMP. As a consequence, we of necessity have the possible factorization indicated in (47) which will be excluded if N and P are taken to be structurally non-distinct.

While this general approach to the interaction of cleft formation and preposition stranding seems productive, it does raise a minor problem in other areas of the grammar which needs to be diffused. In particular, our conditions now predict that sentences like (48) are ungrammatical because they have a factorization like (49) which violates the Association Principles under the assumption that P and N are structurally non-distinct.

(48)a. Mary told Jane about the film.

b. Who did Mary tell about the film?

(49) Who did Mary tell [e] about the film?

\[
A_1 \quad A_2 \quad B_1 \quad B_2
\]

Confronted with the problem, we have the option of either giving up the explanation of the phenomenon regarding preposition stranding, or of revising the Association Principle in such a way so that (49) is not blocked. Clearly the second alternative is preferable.

The crucial difference between (49) and the
preposition stranding phenomenon reproduced as (50) is that in (49) there is a grammatical relation, an association, between the trace of who and the verb tell, that is, between $A_2$ and $B_1$.

$$\begin{array}{c}
A_1 & A_2 \\
\text{who} & \text{tell}
\end{array}$$

(50) Who is it to $[NP^e]$ that Mary spoke $[p,e][NP^e]$ $A_1$, $A_2$, $B_1$, $B_2$.

In (50) however $A_2$ and $B_1$ are not involved in an association. We can admit (49) while still prohibiting (50) if we assumed that the Association Principles included the addendum (51).

(51) Where $A_2$ and $B_1$ are not R-associated.

The Association Principles can be formulated as (52) which incorporates (51).

(52) The Association Principles

Given the class of rules $R$, if $R_i$ involves two terms $A_i$ and $B_i$ in a string $A_i - X - B_i - Z$ containing $A_j$ and $B_j$ related by $R_j$, then

(a) if $A_i - X$ contains $A_j$, $X$ contains $B_j$,

and

(b) at least one phrasal node dominating $B_i$ does not dominate $B_j$ or conversely where $B_i$ is structurally non-distinct from $B_j$ and $A_j$ and $B_i$ are not involved in some $R_n$.

This addendum is necessary for independent reasons.
Consider once again clitic pronouns in Spanish. The sentences in (53) are of some interest.

(53)a. Me recomendaron a él
They recommended me to him
b. Te recomendaron a él
They recommended you to him

The sentences in (53) would have a structure like (54).

(54) me recomendaron [_{\text{NP}}e] a él
A_1 A_2 B_1 B_2

The question at hand is why the association of the verb and indirect object on the one side, and the clitic me and its association to the PRO or trace on the other side can cross over without resulting in an ungrammatical sentence. The revision of the Association Principles to (52) will prevent the Association Principles from blocking sentences like (54).

Koster (1978) presents an interesting alternative to the explanation just developed for why prepositions cannot be stranded in an intermediate movement site. Koster proposes a nesting requirement, (55).

(55) The Nesting Hypothesis
A marked empty node, $^m e$, is unacceptable when it is within a domain of type $V^i$, $S^j$ (main projection) (where $V^i$, $S^j$ are minimal).

297

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(56)a. Definition: The phrases A and B form a nested construction if A falls totally within B, with some non-null element to its left within B and some non-null element to its right within B.

b. Definition: A marked empty node is an e that is internal to a phrase X" (≠ V^i, S^j).

Koster's Nesting Hypothesis will block (57) but not (58) or (59).

(57) *Who did Jane say to [NP_e] that she forgot her phone number.

(58) Who did Jane talk to [NP_e].

(59) To whom did Jane say [PP_e] that she forgot her phone number.

In (58) e is not nested in the sense of (56a) since to the right of e there is no element. In (57), however, e is nested in this sense and moreover e is marked in the sense of (56b) because it is internal to P". In (59) on the other hand e is not a marked empty node because it is not internal to an X" not equal to V^i or S^j. Similarly, Koster's Nesting Hypothesis will block (60).

(60) *Who did Mary know to e Jane talked e.

In (60) the trace of who is a marked empty node and it is nested, in the sense of (56a), in S^1. Consequently (60) is blocked by the Nesting Hypothesis.

298
I have already provided an analysis in terms of the Association Principles for sentences like (60). Let me briefly indicate how (57)-(59) would be handled by the Association Principles. (58) presents no difficulty because there is in effect only a single association for our purposes. The difference between (57) and (59) can be seen in their respective factorizations, (61) and (62).

(61) Who did Jane say to $[\text{NP}_e]$ that she forgot her

\[
A_1 \quad A_2 \quad B_1 \quad B_2
\]

phone number

(62) To whom did Jane say $[\text{NP}_e]$ that she had

\[
A_1 \quad A_2 \quad B_1 \quad B_2
\]

forgot her phone number

Both (61) and (62) do not satisfy the requirement that $X$ contains $B_2$. Moreover $B_1$ and $B_2$ are structurally non-distinct in the sense that we originally defined (cf., p.79). However, only (61) is a violation of the Association Principles: (62) is not a violation because there is a subcategorization rule associating $A_2$ and $B_1$. No such rule links the V and NP in (61). As a result (61) but not (62) is prohibited by the Association Principles.

Koster argues that his Nesting Hypothesis has the advantage of predicting why stranding of a preposition is possible in Dutch only if that preposition is postpositional. Examine the sentences in (63) and (64).
(63a) Met wie heeft hij [NP e] gepraat?
   With who has he [NP e] talked?
   With whom has he talked?

b. *Wie heeft hij met [NP e] gepraat?
   Who has he with [NP e] talked?
   Who has he talked with?

(64a) Waar mee heeft hij [NP e] gewerkt?
   What with has he [NP e] worked?
   What has he worked with?

b. Waar heeft hij [NP e] mee gewerkt?
   What has he [NP e] with worked?
   What has he worked with?

(63b) has the structure in (65) while (64b) has the structure in (66).

(65)

(66)

300

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Koster argues that his Nesting Hypothesis will block (65) but not (66) because although both involve a marked e only in (65) is that [e] nested in $V^i$.

While the Association Principles do not deal with these facts, an alternative using the binding conditions developed in Chapter 4 might well be capable of accounting for the phenomenon at hand if we treat PP as a domain for the Index Erasure Condition and assume that preposition stranding involves an escape hatch as is suggested in van Riemsdijk (1978). At this point, however, I am not prepared to flesh out this suggestion.

Difficulties for Koster's proposals are posed by Sub-deletion (as noted by Koster himself) and Adjective Preposing. The sentence in (67) is a case of Sub-deletion.

(67) Jane has as many books as I have [ₚₑ] records.

Here [e] is marked since it is internal to $N^2$ and it is nested inside $V^i$. Consequently the Nesting Hypothesis as presently formulated would exclude this sentence. I can see no easy way to refine Koster's Nesting Hypothesis to surmount this problem although one may be possible. Similarly Adjective Preposing as discussed in Bresnan (1974) gives sentences like (68b).

(68)a. I've never seen a really tall tree.

b. I've never seen so tall a [ₚₑ] tree.
If (68) is to be treated by movement, the trace adjacent to tree will be both marked and nested. This would mean that Koster's Nesting Hypothesis would then exclude (68b) as well.

It will be up to future research to determine how the overlap between Koster's Nesting Hypothesis and the Association Principles is to be resolved: whether one or the other should be dropped from Universal Grammar, or whether their domains should be more narrowly restricted, for example by defining N and P as distinct for the Association Principles in English.
FOOTNOTES TO CHAPTER 6

1S. Anderson has suggested that an alternative account of the Spanish facts in (53)-(54) would be to assume a more highly articulated definition of the notion "class of rules R" in the definition of the Association Principles. Under such a proposal it would be possible to prevent the Association Principles from blocking the crossing of movements and construals, or movements and subcategorizations. This would provide an explanation for the facts at hand. It would, in addition, make a number of empirical predictions. For example, all the crossing noted in Chapter 3 would necessarily involve rules of the same type, and some of the analyses would be untenable. In particular, the explanation in Section 6 of Chapter 3 would no longer be possible. Anderson's suggestion is certainly within the spirit of the Association Principles however and might very well turn out to be correct.
REFERENCES


Hendrick, R. (In progress). Food taboos, sex taboos, and the development of primitive social organization.


