TOWARD A THEORY OF MOVEMENT RULES

by

MARK REUBEN BALTIN

B.A., McGill University
(1971)

M.A., University of Pennsylvania
(1975)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(September, 1978)

Signature of Author.....
Department of Linguistics and
Philosophy, July 31, 1978.

Certified by.....
Thesis Supervisor

Accepted by.....
Chairman, Departmental Committee

TOWARD A THEORY OF MOVEMENT RULES

by

MARK REUBEN BALTIN

ABSTRACT

Submitted to the Department of Linguistics and Philosophy on August 8, 1978, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

This thesis attempts to answer two questions. First, given the mechanisms in the Extended Standare Theory of syntax, must we still posit syntactic movement rules? As well as showing that the answer to this question must be in the affirmative in Chapters 3 and 4, an attempt is made to constrain the power of these movement rules. The main focus of the attempt to constrain movement rules is on the position to which moved elements move. An explanation is also given for the absence of rules which extrapose elements to the left in Chapter 2. The main approach to these attempts to constrain the power of movement rules lies in a notation for writing movement rules devised here.

Thesis Supervisor: Noam Chomsky

Title: Institute Professor

This thesis is dedicated to

my parents,

David and Elizabeth Baltin,

for love, encouragement, and support
all these years, and never pushing
law school all <u>that</u> hard.

"In the absence of a thecry of what rules can and cannot state, whatever a rule states, it cannot explain."

David Perlmutter and
Janez Oreznik,
"Language-Particular Rules
and Explanation in Syntax".

ACKNOWLEDGEMENTS

During my first year at M.I.T., in the spring of 1976, it was my good fortune to take a phonology course with Morris Halle. Halle's persistent concern with the claims which different notations and formalisms in phonology make impressed me as being a true insight into the way in which to appropriately constrain the theory of language, and I decided to try to apply this methodology to syntax. Whether or not I've succeeded is something which others must judge, but I would like to register my appreciation to Halle in any event for truly fine teaching, and, in his capacity as founder and first chairman of our department, for creating an atmosphere at M.I.T. which is the most congenial for intellectual productivity that I have ever encountered. M.I.T. is a place where the distinction between faculty member and student tends to blur into the superordinate term colleague. I would like to thank Jay Keyser for helping to continue this ambience, as well as for much help to me personally and extremely stimulating discussions about syntax.

For matters directly related to this thesis, I would like to thank my thesis advisor, Noam Chomsky, who would instantly find the chink in the armor, and would usually, a split second later, find a way of repairing the chink. His encouragement and accessibility, as well as his contributions to linguistics, are truly unsurpassed; the questions which he was the first to pose in print as well as in lectures and

discussions were landmarks that I have tried to keep sight of through the wanderings in this thesis. Ken Hale, who graciously shared his legendary knowledge of the languages of the world, would always come up with the insight from Language X that was right on target. Joan Bresnan, whom I regard as the syntactician's syntactician, was always quite incisive, and could slice through the bull to the central point in a second; the high standards of Bresnan's work, particularly with regard to inherent interest of the positions taken and level of argumentation, should be models to us all.

I would like to thank my friends for making my three years in Cambridge extremely pleasant ones, in particular Wilson Gray, Ann Lekach, Shelly Lieber, Guy Carden, Jill Carrier, Dave Duncan, John McCarthy, Jane Grimshaw, Randy Hendrick, Jean-Yves Pollock, Jacqueline Guéron, Alan Glazer, Norbert Hornstein, Amy Weinberg, and Ivan Sag. In addition I would like to thank Annie Zaenen, Shelley Lieber, Jacqueline Guéron, and Bob Freidin for reading and criticizing particular chapters. Ivan Sag made a suggestion about the presentation of Chapter 4 which proved invaluable.

I would like to acknowledge a couple of debts from a somewhat greater distance in time. Bill Labov, ever since I walked into his office in the fall of 1971 after dropping out of the University of Michigan (where I was a grad student in linguistics for one week) has been a great teacher, mentor, and friend; I would like to thank him for exposing

me to one of the most creative minds in the field, and for much help and friendship over the years.

At a still greater remove, I would like to thank David Lightfoot for introducing me to the fascinating world of syntax while I was an undergraduate at McGill. A measure of Lightfoot's teaching ability can be gleaned from the fact that his course, which met at 9 A.M., was probably the only class in the university which had full attendance regularly at that hour.

Chapter 3 is a much improved version of a paper which was presented at NELS VIII in 1977 under the title PP As a Bounding Node.

Versions of Chapter 2 were presented as invited talks at Yale University, Swarthmore College, and New York University. I would like to thank the audiences at those talks for helpful comments; at Yale, Guy Carden, Donca Steriade, and Barry Schein were among the more active participants.

I'd like to thank Dave Duncan for a splendid job of typing this thesis; whatever you may think of the content, you must admit it looks really good.

TABLE OF CONTENTS

TITLE PAGE..... 1

ABSTRACT..... 2

DEDICATION..... 3

QUOTE..... 4

ACKNOWLEDGEMENTS..... 5

CHAPTER 1: Introduction..... 10

CHAPTER 2: On the Status of Extraposition Rules..... 20

SECTION I: Rule Interaction Arguments..... 25

SECTION II: On Formulating the Construal Schema.. 35

SECTION III: Further Arguments..... 40

SECTION IV: Two Final Modifications..... 61

SECTION V: Concluding Remarks..... 66

FOOTNOTES TO CHAPTER 2..... 70

CHAPTER 3: Bounding and Construal..... 71

SECTION I: Theories of Bounding..... 71

SECTION II: PP As a Bounding Node..... 81

SECTION III: On the Nature of Bounding..... 116

FOOTNOTES TO CHAPTER 3..... 128

CHAPTER 4: On the Notion "Structural Change" in Movement Rules..... 133

SECTION I: Relative Clause Extraposition..... 138

SECTION II: Extraposition..... 140

SECTION III: On Comp-Substitution..... 154

SECTION IV: On Movements to VP Boundaries..... 171

SECTION V: On Position I..... 186

SECTION VI: Landing Sites in Universal Grammar... 187

FOOTNOTES TO CHAPTER 4..... 193

BIBLIOGRAPHY..... 194

BIOGRAPHICAL SKETCH..... 203

CHAPTER 1: INTRODUCTION

Between the mid-1960's and the present time, a sort of balance of power has been fought between the syntactic and semantic components within the framework of generative grammar. One extreme position, which has come to be known as generative semantics, has maintained that there is no principled division whatever between syntactic and semantic rules; a clear exposition of this viewpoint can be found in Póstal (1972), who argues on general grounds of theoretical simplicity that the theory with the least amount of devices is preferable. Therefore, a theory such as generative semantics is considered to be homogeneous, and hence to be valued more highly by the criterion of Occam's Razor, i.e., theoretical parsimony.

The other position, that of Chomsky (1970), Jackendoff (1972), and others, is that the attempt to maintain the unity of syntax and semantics by generative semanticists has led to a conception of the syntactic component which has rendered the position that there is no such distinction an impregnable fortress. That is to say, the notion of possible syntactic rule is such an amorphous notion in such a theory that there would be no possible way to refute such a theory; therefore, Occam's Razor in this case, according to the opponents of generative semantics, has elevated the tenet of the unity of syntax and semantics from an empirical hypothesis to an almost theological act-of-faith. In Chomsky's (1972) words:

> If enrichment of theoretical apparatus and elaboration of conceptual structure will restrict the set of possible grammars and the class of sets of derivations generated by admissible grammars, then it will be a step forward... although one wants the 'simplest' linguistic theory, in some sense that cannot be specified in a useful way, elaborations and complications of linguistic theory are all to the good insofar as they narrow the choice of grammars and the range of admissible languages (i.e., sets of derivations).

Adherents of the second position therefore maintain that there is a principled distinction between syntactic and semantic rules. Much of the work within the second position, which has variously been called the extended standard theory, interpretive semantics, or lexicalism, has been devoted to constraining the syntactic component (Emonds 1970, 1976; Jackendoff 1977; Bresnan 1976). However, while particular semantic rules have been proposed (by, for instance, Jackendoff 1972; Lasnik 1976; Reinhart 1976; Williams 1977), a full picture of the semantic component has not really emerged within the framework of the extended standard theory.

This thesis is couched within the framework of the extended standard theory, which claims that a given dependency cannot automatically be handled within the syntax. The phenomena dealt with here have traditionally been handled by types of syntactic rules which are called <u>movement rules</u>. Examples of such movement rules are the following:

(1) (a) <u>That John is a fool</u> is obvious.

(b) It is obvious <u>that John is a fool</u>.

(2) (a) John likes somebody.

(b) I wonder who John likes.

(3) (a) John murdered Bill.

(b) Bill was murdered by John.

(4) (a) John gave a book to Martha.

(b) John gave Martha a book.

(5) (a) All of the men have arrived.

(b) The men have all arrived.

(6) (a) The man who is from Boston has just arrived.

(b) The man has just arrived who is from Boston.

(7) (a) I like Max.

(b) Max I like.

Most linguists agree that there is a relationship of some sort between the (a) and (b) sentences above, but there are two questions to ask once this point has been established:

(i) What is the basis for the linguist's intuition that there is a relationship?

(ii) What is the nature of the relationship?

I think that (i) can be answered in advance of a specific analysis, but (ii) cannot. Among the facts that might lead linguists to posit a relationship is the fact that the (a) and (b) sentences are synonymous and share co-occurrence restrictions in the sense of Z. Harris (1964).

For example, the verbs in (3b), (6b), and (7b) are all transitive verbs whose direct objects, in these sentences, are not in their normal, immediately post-verbal position; while the direct object is missing, there is a compensation, in that an NP is present initially, and this NP can be construed as a direct object in another version of the sentence, i.e., the versions of the sentences in (3a), (6a), and (7a).

Traditionally, dependencies of this sort have been described by syntactic rules which would move an element from one position in the sentence to another position in the sentence.

Many of the arguments for movement rules, however, are undercut by the question of the division of labor between syntactic rules and semantic rules. Given the existence of semantic rules, one of the initial facts that would lead a linguist to posit a relationship, that of synonymy, is now of questionable relevance, since a semantic rule could account for synonymy in the above cases, just as it must do so for pairs like:

(7) A bachelor entered the room.

(8) An unmarried male adult entered the room.

There does not seem to be an answer to (ii), then, based exclusively on synonymy. There must be other criteria, then, to enable us to decide.

One criterion might be to decide that, in the cases of

(1-6), a movement rule would enable us to simplify sub-categorization restrictions, in the sense of Chomsky (1965). For instance, one would not want to say that an otherwise obligatorily transitive verb such as like does not have to be transitive when an NP is generated in pre-subject position, as in (7).

Kayne (1975) gives an argument of this form for French when he discusses the status of a rule which he calls L-tous. In French, certain quantifiers, among them the universal 'tout', which means 'everything', can occur either in the position of direct object or before the verb. The alternation can be seen in the following pair:

(9) Il a mangé tout.

(10) Il a tout mangé.

As Kayne notes, one must somehow state that every transitive verb in the language can allow the direct object to be missing just in case tout, which could have appeared in the position of direct object, appeared pre-verbally. One way of directly expressing this regularity would be to posit a syntactic rule which moved quantifiers such as tout from the direct object position to pre-verbal position.

However, again there is enough independently motivated machinery within the extended standard theory to allow one to consider the possibility of an alternative statement of the generalization. Jackendoff (1975), Oehrle (1975),

Bresnan (1978) have argued for a device known as a lexical redundancy rule. This device is a type of rule which operates within the lexicon to state regularities about lexical entries. For example, Oehrle (1975) shows convincingly that, instead of a transformation of dative-movement, which would relate the sentence pairs in (4), the appropriate mechanism would be a rule which states that if a verb subcategorizes for prepositional datives, it also subcategorizes for a double object construction, after showing that both constructions must be base-generated in at least some cases. Crucially, this rule does not operate in the syntax directly. Therefore, with just this amount of data about L-tous, suppose we could motivate a pre-verbal expansion of some sort, and then generated the QP there; we might then propose a lexical redundancy rule which would state that a verb which was transitive had another subcategorization in which the QP was pre-verbal.

Actually, this picture is highly over-simplified; Kayne gives more arguments for the syntactic nature of this rule, which, in conjunction with the argument given above, make his case for the syntactic nature of L-tous quite convincing. I return to some of these arguments in Chapter 2. I am simply mentioning the form of some arguments by themselves, to review some of the criteria which have been adduced for movement rules.

Moreover, the device of lexical redundancy rules cannot be used carte blanche. Because they relate subcategori-

zation frames of a verb, the two positions of the element which can occur in either one must be positions generable by the phrase-structure rules, and the relationship must be clause-bound, since the rule is dependent upon subcategorization frames of a particular verb. Therefore, if the dependency is unbounded, the relationship could not be handled by a lexical redundancy rule.

Although an unbounded dependency would rule out the use of a lexical redundancy rule, a bounded dependency may, obviously, be due to any one of a number of factors. For instance, Chomsky (1973) has postulated a series of constraints which jointly ensure that the only position through which a moved element may escape from a clause is through the complementizer position. This unique escape hatch, in conjunction with the Baker-Bresnan comp-substitution universal, which states that only languages with clause-initial comps permit comp-substitution rules, guarantees that rightward movement rules will be bounded (but see Kaufman 1974). The effect of all of this is that a dependency which might be handled by a rightward movement rule will be bounded, and thus might invite a semantic or lexical treatment; however, this boundedness may be due to extraneous factors.

This thesis will try to determine, for certain dependencies, whether they are best handled in the syntax or in the semantics. Among the criteria which would enable us to decide, there are two chief considerations.

One has been briefly alluded to earlier. If an element in position A could also have occurred in position B, and nothing occurs in position A which could not have occurred in some other position, which happens to be empty when position A is filled, we would postulate movement to position A. In this case, the movement would be non-structure-preserving, in the sense of Emonds (1976).

A second criterion would be whether or not the dependency in question obeys conditions (only) which are best viewed as conditions on logical form, as opposed to conditions on syntactic rule applicability. This will be the subject matter of Chapter 3; one of the points made in that chapter will be that for at least one condition, that of subjacency, two sentences can have structures that are identical in the relevant respects at the level of logical form, but could plausibly differ only with respect to a putative syntactic derivation, and subjacency must make the distinction. The conclusion, therefore, is that subjacency is a condition on syntactic rule applicability, and if a dependency obeys subjacency, the dependency must arise as the result of a movement rule.

A third criterion should be mentioned here. If a dependency exhibits split control, i.e., if the element which is putatively moved corresponds to more than one position in the sentence, a movement analysis would be totally impossible. One such case arises in the case of result clauses, discussed in Chapter 3:

(11) Bill is so hungry that he'll eat anything.

Intuitively, there is a dependency between the degree word so and the finite complement which one might wish to capture by an analysis which co-generates them under the same node, and which moves the result clause to clause-final position. However, Liberman (1975) has shown cases where the result clause would have to correspond to as many as three so's in the sentence:

(12) So many kids read so many books so often that it's hard to keep up with them.

Since there would be no single source for the result clause in such a case, the argument goes, a movement analysis would be totally out of the question. This argument is quite valid. Recently, some (e.g., Andrews 1975) have attempted to apply this form of argumentation to extraposed relatives, as in (13):

(13) A man entered the room and a woman left who were similar.

Therefore, extraposed relatives would have to be base-generated, according to this line of argument, in the same way that result clauses would have to be.

However, there is a crucial difference between the result clauses and the extraposed relatives; in the case of the result clauses, split control can be within a simple sentence, whereas one can never get split control within

a simple sentence for extraposed relatives. Thus, although (12) is acceptable, sentences like (14) are not:

*(14) A man saw a woman who were from Boston.

Therefore, in the case of extraposed relatives, the data is rather unclear, so that the split control argument is rather inconclusive.

Once we have agreed to posit a class of syntactic movement rules, we will attempt to impose severe limitations on their form; this attempt forms the heart of Chapters 2 and 4.

A word should be said here about the scope of this work. Little has been said here about the so-called cyclic NP movement rules such as passive, dative, raising, etc. As I remarked earlier, Oehrle (1975) has cast severe doubts on the existence of a transformation known as dative-movement; recently, Bresnan (1978) has argued against all of these transformations. In any event, these would seem to pose no problem for the system developed here, as will be seen in Chapter 4. The main focus of this work is on wh-fronting, topicalization, Q-Float, adverb movements, and various extraposition rules.

CHAPTER 2: ON THE STATUS OF EXTRAPOSITION RULES

In the literature on generative syntax, one frequently finds reference to the term extraposition rule, as distinct from the more general term movement rule. One never finds a definition of the term, but one knows that the sentences in (1-4) exemplify extraposition phenomena, while the sentences in (5) and (6) do not:

(1) (a) A man who was from Philadelphia came in.

(b) A man came in who was from Philadelphia.

(2) (a) A review of Chomsky's latest book has just appeared.

(b) A review has just appeared of Chomsky's latest book.

(3) (a) All of the men have arrived.

(b) The men have all arrived.

(4) (a) The claim that Fred killed his father has never been proved.

(b) The claim has never been proved that Fred killed his father.

(5) (a) John killed who?

(b) Who did John kill?

(6) (a) I really like Max.

(b) Max I really like.

The rule that relates (1a) and (1b) is called extraposition from NP, the rule operative in (2) is

termed extraposition of PP, the rule in (3) Q-Float[1], and the rule in (4) is termed complement extraposition. The rule in (5), however, which is termed wh-movement, and the rule in (6), topicalization, would not be considered extraposition rules. Quite simply, then, extraposition rules can be defined as rules which move elements out from under the domination of an NP. The rules in (5) and (6), however, move the entire NP in the unmarked case. The question I wish to pose in this essay is whether the notion extraposition rule is simply a term that linguists have invented, or whether there is something in the nature of language that gives extraposition rules a separate status. In other words, is the distinction between extraposition rules and the rules operative in (5-6) ontological or terminological?

I will argue that the distinction in ontological. That is, when one takes the tack that extraposition rules have the form they do for a reason, one makes predictions about a class of phenomena which have hitherto, to my knowledge, escaped investigators' notice. Crucially, there are generalizations which one can make about extraposition rules if one gives them a separate status from other types of movement rules in the metatheory.

The generalization I wish to make is the following: All extraposition rules in English, and a few other languages, move elements to the right. That is, while we can write rules using the notations for transformations in works

such as Chomsky (1965), like (7-9):

(7) $_{NP}[_{NP}\ \bar{S}]_{NP}$ -X
1 2 3⟶ 1 3+2

(8) $_{NP}[_{NP}\ PP]_{NP}$ -X
1 2 3⟶ 1 3+2

(9) $_{NP}[_{QP}\ NP]_{NP}$ - Aux - VP
1 2 3 4⟶ 2 3 1+4

we do not find rules like (10-12):

*(10) X- $_{NP}[_{NP}\ \bar{S}]_{NP}$
1 2 3⟶ 3+1 2

*(11) X- $_{NP}[_{NP}\ PP]_{NP}$
1 2 3⟶ 3+1 2

*(12) X- $_{NP}[_{QP}\ NP]_{NP}$
1 2 3⟶ 2+1 3

Thus, we never get sentences like (13-16), which are parallel to (1-4) except that the movement has taken place in the opposite direction:

*(13) Who was from Philadelphia a man came in.

*(14) Of Chomsky's latest book a review has just appeared.

*(15) All definitely the men have arrived.

*(16) That Fred killed his father the claim has never been proved.

If the generalization that leftward extraposition rules do not occur is in fact correct, one may draw a number of alternative conclusions concerning the implications for a theory of grammar. One conclusion which a linguist might reach is that the generalization I have made is spurious, and hence that one does not have to, as it were, design a theory around it. Another conclusion might be that one may keep a standard notation for writing transformations like Chomsky (1965) or Peters & Ritchie (1973) and simply add a negative condition to the metatheory stipulating that leftward extraposition rules do not occur. Another conclusion is that the ban on leftward extraposition rules indicates an inadequacy of the standard theory (Chomsky 1965). I think that the choice among these three alternatives can be decided on general grounds. The first alternative, that of claiming an epiphenomenon, is always open, but is in general fraught with methodological peril, since it is a dead end as far as deepening our understanding goes. The second tack can also be rejected on general grounds. One can always allow a stipulation to the effect that a certain class of rules, although formally possible, does not occur in natural language. However, if there are two accounts which would both exclude the class of rules, one simply stating that the class of rules does not occur, and the other stating a more general principle from which the exclusion would follow, I think that the latter course is obviously the preferable one.

It seems to me that there is in fact a more general

explanation for the absence of this class of rules. In the spirit of Chomsky (1977a, b), let us suppose that there exist rather severe restrictions on the form of transformations. Rules in this framework are of the form "Move NP" or "Move WH". Going hand in hand with this proposal, although logically independent of it, is the proposal that all sentence-level interpretation is done at surface structure. If this is the case, moved elements must be put back in the positions from which they came in order to reconstruct the right grammatical relations. Let us hypothesize that this is done by actual rules of construal, and that one such construal rule, which associates heads with modifiers which are discontinuous in surface structure, is the appropriate place to make the generalization about leftward extraposition being a banned phenomenon. I would propose the following universal schema for modifier-head construal:

(17) 

This schema says that at the level at which construal takes place, the NP must always be to the left of the modifier. Now, suppose we use the uninterpreted element convention, which would say roughly that every semantically contentful element must receive an interpretation, or the sentence is filtered out. Therefore, if the head were to the right of the modifier, the modifier could no longer be linked to it, and if there were no other possible head for the modifier, it would not be able to be linked to any-

thing, and therefore could not be interpreted. In the next section, I will show that this state of affairs does occur in English, French, Rumanian, and Persian, and does in fact lead to ungrammaticality, a fact which would be predicted by (17) but which would need global constraints which make reference to semantic representation in other frameworks.

At the end of section IV, we will see that the universality of (17) must be modified somewhat, since Ken Hale (personal communication) has pointed out to me that several Malayo-Polynesian languages, as well as Papago, violate the schema. However, we shall see that these languages are in fact the exceptions which prove the rule.

I. Rule Interaction Arguments

In this section, the extraposition rule with which I shall be primarily concerned is the rule of Q-Float. The reason for this is that the other extraposition rules mentioned above tend to move the modifiers to final position within the clause, and so failure to move the original head past the modifier could be due to the way that the rules are formulated; i.e., the landing site for the moved NP, for instance, in the sense developed in Chapter 4 of this thesis, could be right bracket VP, and the modifier (extraposed relative, complement, etc.) could be right bracket S. In the case of Q-Float, however, the extraposed element (the quantifier) does not move to the end of the clause, but rather to some position before the main verb,

and so a rule which moved NPs to the end of the clause should in principle be able to apply.

A. English

In English, there does exist such a rule of subject-postposing, the rule of Presentational There-Insertion, discussed by Aissen (1975). This rule maps the structures underlying sentences like (18) onto the structures for sentences like (19):

(18) A portrait of Picasso hung on the mantelpiece.

(19) There hung on the mantelpiece a portrait of Picasso.

Now, suppose that Presentational There-Insertion were to relate (20) and (21):

(20) All of the portraits by Picasso hung on the mantelpiece.

(21) There hung on the mantelpiece all of the portraits by Picasso.

The structure for (20) is also the input to Q-Float:

(22) The portraits by Picasso had all hung on the mantelpiece.

In principle, Q-Float and Presentational There-Insertion should be able to apply in the same sentence. Testing the prediction, we find that the sentence is ungrammatical:

*(23) There had all hung on the mantelpiece the portraits by Picasso.

We see, then, that Q-Float and Presentational There-Insertion cannot apply in the same sentence. Of course, one can rule out this interaction by extrinsically ordering the rules in the order in (24):

(24) (a) Presentational There-Insertion
(b) Q-Float

Presentational There-Insertion, since it would move the entire NP to the end of the sentence, would bleed Q-Float if it were to precede it, under this alternative account, since the quantified NP is to the right of the Aux. The device of extrinsic ordering, however, makes the claim that the ungrammaticality of (23) is due to a phenomenon which is totally particular to the interaction of Q-Float and Presentational There-Insertion, and that there is not a more general principle which would rule out this interaction. Construal schema (17), however, would predict the rule interaction which extrinsic ordering (24) merely stipulates, since it could explain both the ungrammaticality of (23) and the absence of leftward extraposition rules. Moreover, we will see that the same device is needed in the grammars of at least three other languages, to which we now turn.

B. French

Kayne & Pollock (to appear) describe in detail the operation of a rule in the grammar of French known as

stylistic inversion. This rule relates (25) and (26):

(25) Je voudrais savoir ce que les hommes ont mangé.

(26) Je voudrais savoir ce qu'ont mangé les hommes.
'I'd like to know what the men have eaten.'

Sentence (26) shows the application of stylistic inversion, a rule which, under certain conditions, moves the subject NP to the end of the clause. The French analogue of English Q-Float, dubbed R-tous by Kayne (1975), can be seen in (28):

(27) Tous les hommes ont mangé le gâteau.
'All the men have eaten the cake.'

(28) Les hommes ont tous mangé le gâteau.
'The men have all eaten the cake.'

Again, we can test the prediction made by the construal schema, this time by applying R-tous and stylistic inversion and seeing the result:

(29) Je voudrais savoir ce que tous les hommes ont mangé.

(30) Je voudrais savoir ce que les hommes ont tous mangé.

*(31) Je voudrais savoir ce qu'ont tous mangé les hommes.

Sentence (31) shows the application of the two rules in question, and the result is ungrammatical. Again, one could rule out the interaction by applying stylistic inversion before R-tous, so that stylistic inversion would bleed R-tous. The comments about extrinsic ordering made in

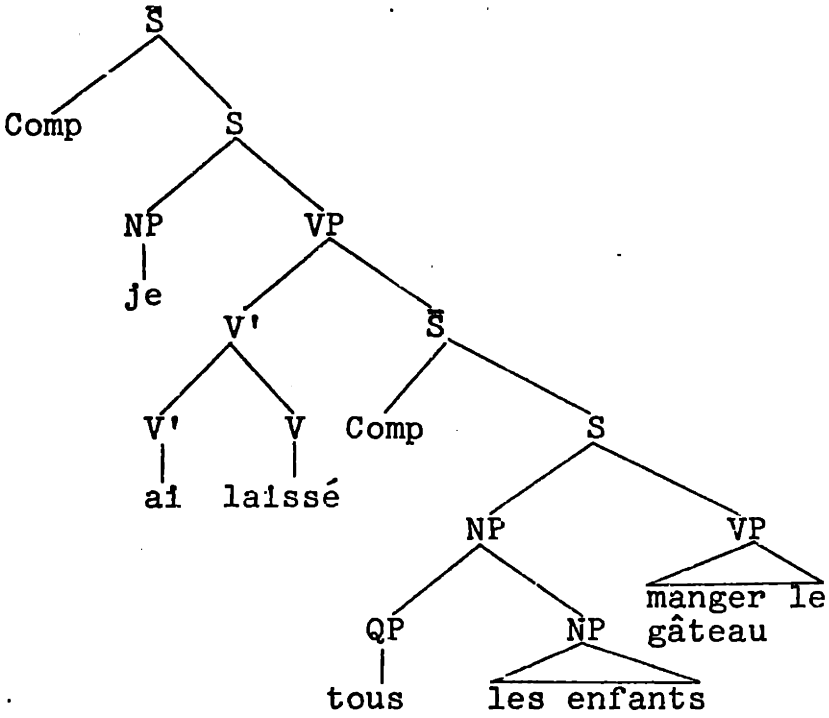
section I.A still apply, causing us to view the device (in this instance at least) with even more suspicion, since we see another ad hoc use of it to circumvent the effect of (17).

In fact, the extrinsic ordering device would have to be used in French for yet another rule, the rule of faire-attraction. This rule derives sentences like (32):

(32) J'ai laissé manger le gâteau à tous les enfants.

from structures like (33):

(33)



Now, we can see that the subject NP ends up in final position in (32). However, when we apply R-tous, we cannot apply faire-attraction:

*(34) J'ai laissé manger tous le gâteau aux enfants.

The ungrammaticality of (34) cannot be traced to a possible stipulation that the position in which the floated quantifier ends up is not a possible position for floated quantifiers, since quantifiers moved by R-tous can end up after the infinitive, as in the following sentence from Quicoli (1976):

(35) Mes amis lui laisseraient manger tous de la salade.
'All of my friends would let him eat salad.'

Of course, there is an alternative here. Kayne (1975) formulates the rule of faire-attraction as follows:

(36) X - faire - NP - V - (NP) - Y
1 2 3 4 5 6 ⟶ 1 2 4 5 3 6

Therefore, one could use this formulation and apply R-tous before faire-attraction, placing the tous after the verb manger. As can be seen from the formulation of faire-attraction in (36), this would have the effect of bleeding faire-attraction.

However, if one looks at the rule in (36), one sees that it is extremely complex, mentioning four constant terms, while only one of those, term (3), is actually affected by the rule. This violates some otherwise reasonable and well-supported restrictions on the form of transformations, such as Chomsky's (1977a) minimal factorization condition, which says that if two constants are mentioned in a structural description, one must be the constituent

affected by the rule; this would say that either term (4) or term (5) would have to be crucially affected by the rule. If we wish to keep minimal factorization, then, we are forced to adopt something like Quicoli's (1976) formulation of faire-attraction, in which the subject is not being postposed, but rather a verb phrase consisting of the verb and direct object preposes to the front of the complement. (Alternatively, if we wished to view faire-attraction as subject postposing, we could say that the subject moves to a VP-final position.)

Moreover, an empirical argument can be levelled against (36). As noted in Kayne (1975), adverbs can appear in the complement of a faire-construction. For example, Kayne cites the following example:

(37) Elles feront toutes les trois soigneusement controler leurs voitures.
'They will all three have their cars checked carefully.'
[Kayne (1975), ch.3, ex.46]

As Kayne notes, floated quantifiers and adverbs occur in the same positions in French, If this is true, it would seem that (36) would have to be revised to allow for adverbs in the complements of faire-type verbs. Once this is done, it would seem impossible to use the formulation of faire-attraction to rule out prior application of R-tous. Therefore, it seems reasonable to eschew the formulation of faire-attraction in (36), and view the ungrammaticality of (34) as another instance of the construal schema at work.

C. Persian

Persian is an SOV language with an unmarked order in which the indirect object follows the direct object. For example, (38):

(38) All the men gave the book to Zhala.

would be translated as:

(39) Hamé mardan kitab-rā be žala dadam.
all the men book-d.o. i.o. gave

Q-Float operates to niche the quantifiers between major constituents. Represented schematically:

(40) S^O^IO^V

All of the following, then, are paraphrases of (39):

(41) mardan hamé kitab-rā be žala dadam.
(42) mardan kitab-rā be žala hamé dadam.
(43) mardan kitab-rā hamé be žala dadam.

Interestingly enough, as Moyne & Carden (1974) show, a rule which they call subject movement has the effect of niching subjects in the same places as floated quantifiers. Thus, (44) and (45) are paraphrases of (39):

(44) kitab-rā hamé mardan be žala dadam.
(45) kitab-rā be žala hamé mardan dadam.

The interaction of Q-Float and subject movement, as

it turns out, is not free. That is, we can niche the subject between the object and indirect object, and float the quantifier between the indirect object and verb:

(46) kitab-rā mardan be žala namé dadam.

but we cannot float the quantifier between the direct and indirect object, and the subject between the indirect object and verb:

*(47) kitab-rā hamé be žala mardan dadam.

Again, these facts are predictable by the construal schema, which says that the head must be to the left of the modifier in order to be linked to it. Again, one could order Q-Float to follow subject movement, but it seems that the weight of the evidence is militating against this, since we now have three (possibly four, depending on whether or not faire-attraction is really a subject-postposing rule which can be collapsed with stylistic inversion in French) orderings in three different languages, all of which have the same effect, so that it would seem that extrinsic ordering would cause a generalization to be masked.

D. Rumanian[2]

In Rumanian, the form of the floated quantifier, for reasons that are unclear to me, takes a sort of linker. Thus, a paraphrase of (48), the variant without floating:

(48) Toti oamenii s-au adunat in piata'.
all the men refl. pron.-have gathered in marketplace

'All the men have gathered in the marketplace.'

would be

(49) (a) Oamenii s-au adunat cu totu in
the men refl.-have gathered [link] all in
piata'.
marketplace

(b) Oamenii s-au adunat in piata' cu totu.

The difference between the (a) and (b) versions in (49) is that the locative follows the floated quantifier in (49a), but precedes it in (49b). The exact formulation of the rule that effects the movement that relates (49a) and (49b), (i.e., which order is basic, whether the rule is simply a local interchange, etc.,) is immaterial here.

Now, interestingly enough, there exists in the grammar of Rumanian a subject-postposing rule, so that a variant of (48) is (50):

(50) S-au adunat in piata toti oamenii.
refl.-have gathered in market-place all men

With the description of Q-Float and subject-postposing as background, we are now in a position to test the applicability of our construal schema. The prediction is that subject postposing and Q-Float should be incompatible.

The data, according to my informant, support this:

(51)*(a) S-au adunat cu toti in piata' oamenii.
refl.-have gathered all in market-place men

*(b) S-au adunat in piata' cu totu oamenii.

Thus, the prediction made by the construal schema is borne out.

E. Summary

We have seen that rule interactions in four different languages which would have the effect of placing a head NP to the right of an element which is discontinuous to and modifies it, are impermissible. Since we have rejected extrinsic ordering solutions as being observationally but not descriptively adequate, we are forced to use something like the construal schema.

In the next section, we will presuppose, therefore, the viability of the schema, and try to formulate a bit more precisely the notion we need of head. After clarifying that notion, we will return to further arguments for this construal schema.

II. On Formulating the Construal Schema

In Section I, I tried to advance a principles explanation for the non-existence of a class of movement rules which made heavy use of the uninterpreted element convention. The proposed construal schema has interesting implications

for the analysis of missing subjects in such sentences as (52):

(52) To all work on adverbs would be difficult.

Under most accounts, the subjects of the infinitives in sentences such as (52) have undergone a rule known as EQUI which deletes them (Rosenbaum 1967; Postal 1970; J.A. Fodor 1975). However, a recent alternative analysis of this phenomenon has been proposed by Joan Bresnan, in class lectures at M.I.T. (Spring 1978). She proposes that there are two ways for subjectless infinitives to arise, one being by the base rule in (53):

(53) $S \longrightarrow (NP)\ \overline{VP}$

These would be the infinitive phrases which occur in NP positions, as in the subject position in (52), and would be the source for infinitival questions and relatives, as in (54) and (55):

(54) What to work on was unclear to them.

(55) They were looking for a topic on which to work.

On the other hand, some infinitives which do not occur in NP positions, and which are dominated by VP, occur as a result of the base rule (56):

(56) $VP \longrightarrow V\ \ \overline{VP}$

This would be the source of the complements of so-called "obligatory EQUI" verbs, such as <u>try</u>.

However, this analysis of missing complement subjects is inconsistent with the construal schema as it stands, since there would be no NP present with which to construe the floated quantifier. Therefore, either the construal schema must be modified, or this analysis of missing complement subjects must be abandoned.

Interestingly, not all types of infinitivals with unexpressed subjects permit Q-Float. Thus, although (52) is possible, (57) and (58) are not:

*(57) What to all work on was unclear to them.

*(58) They were looking for a topic on which to all work.

There seems to be a correlation between the ability of Q-Float to occur in a complement and the ability of that complement to take a lexically present subject. Thus, Q-Float can occur in infinitival complements such as (52), and such complements can take lexically present subjects, as in (59):

(59) For them to all work on adverbs would be difficult.

However, Q-Float, as we see from (57) and (58), cannot occur in infinitival questions and relatives, which also cannot take lexically present subjects:

*(60) I'm looking for a topic on which Bill to work.

*(61) What Bill to work on is unclear.

The account proposed here relies heavily on the analysis of Chomsky & Lasnik (1977) and Chomsky (forthcoming). Chomsky and Lasnik designate positions which can never be filled as "positions of obligatory control", and argue that there are subjects in these positions, but that the NP node is unexpanded (their term for such a noun phrase is PRO). When a lexically specified subject can appear, but doesn't, they argue that it is actually deleted as a reflexive. This second situation corresponds to what has traditionally been called EQUI.

This account of missing infinitival subjects receives support from Q-Float data, whether Q-Float is analyzed as a syntactic movement rule or a rule of construal. If Q-Float is a syntactic movement rule, then the inability of Q-Float to occur in infinitival questions and relatives follows automatically, since an unexpanded noun phrase node would contain no Quantifier phrase node beneath it, and hence no spot from which a floated quantifier could be launched. If sentences like (52) actually had syntactic subjects, on the other hand, this difference would fall right out.

At any rate, postulation of a rule of EQUI would make the required distinction between those infinitival complements which can allow floated quantifiers and those which cannot. Once we allow this rule of EQUI, we can allow it to operate after the construal schema, and hence sentences like (52) cease to be problematic.

Noam Chomsky (personal communication) has suggested an alternative account of the inability of Q-Float to occur in infinitival questions. He has suggested that if one views the wh-word of an infinitival question as a quantifier which obligatorily takes wide scope, and if one views a floated quantifier of a PRO as obligatorily taking wide scope, the two requirements would conflict, ruling out the sentence. This would rule out (57), but would have nothing to say about the ungrammaticality of (58). In (58), the wh-word, being a relative pronoun, does not function as a quantifier, and hence cannot take any scope; there would thus be no conflicting scope requirements, and yet the sentence is still out. Therefore, the analysis in which the inability of Q-Float to occur is traced to the presence of a PRO subject in such constructions is the more general explanation.

The proposed correlation between ability to take a lexically specified subject and ability to take a floated quantifier makes an interesting cross-linguistic prediction. It has long been known that the ability of an infinitive to take a lexically specified subject is largely an idiosyncracy of English. Thus, in French, for example, infinitives must be subjectless, and thus, by Chomsky & Lasnik's account, must take a PRO subject (see Chomsky, forthcoming, for a proposal that would explain this difference between French and English.) Now, the correlation I have suggested would predict, then, that R-tous in an infinitival subject

would be impossible. According to my informants, Marie-Thérèse Vinet and Martine Dorel, the prediction is realized:

*(62) Partir tous serait difficile.
'To all leave would be difficult.'

What I have been arguing up until now is that one can never launch floating quantifiers from PRO, a fact which would follow trivially if one assumed that PRO is simply an unexpanded NP and that Q-Float is a syntactic movement rule. Once we accept this conclusion, we must say that the source for sentences like (63):

(63) They tried to all come home.

would be (64).

(64) All of them tried to come home.

Obviously, these two sentences are not synonymous. In particular, the quantifier in (64) can take a distributive reading that is lacking in (63). Therefore, to account for this lack of distributive reading in (63), we might propose that the complement of <u>try</u> is interpreted at surface structure as a sort of complex predicate <u>all come home</u>. As far as I can see, this move will allow the scope differences to fall out naturally.

III. Further Arguments for the Construal Schema

As we have seen in section II, floated quantifiers are

construed with lexical NPs, i.e., NPs which have actual lexical content, and are not construed with PRO subjects. Now, consider a sentence like (65):

(65) They're likely to all get ticked off.

Our claim is that the all is construed with they, and has been floated into an adjective phrase. Now, adjective phrases in English can prepose to the front of the clause, and if a preposed adjective phrase contained a quantifier floated off of the subject, the quantifier would end up to the left of the subject. The construal schema would rule out such a situation since, by our account, the quantifier would not be linked to anything. Testing this prediction, we find that it is supported by the data:

*(66) Just how likely to all get ticked off are they?

(67) Just how likely to get ticked off are they?

*(68) So likely to all get ticked off were they that I decided not to tell them.

(69) So likely to get ticked off were they that I decided not to tell them.

Another possible argument for the construal schema can be made, but this argument is based on some assumptions which are at the time of writing somewhat controversial. First I will give the argument in this framework, and then I will point out the areas of controversy and, where possible, some directions to pursue in resolving the issues.

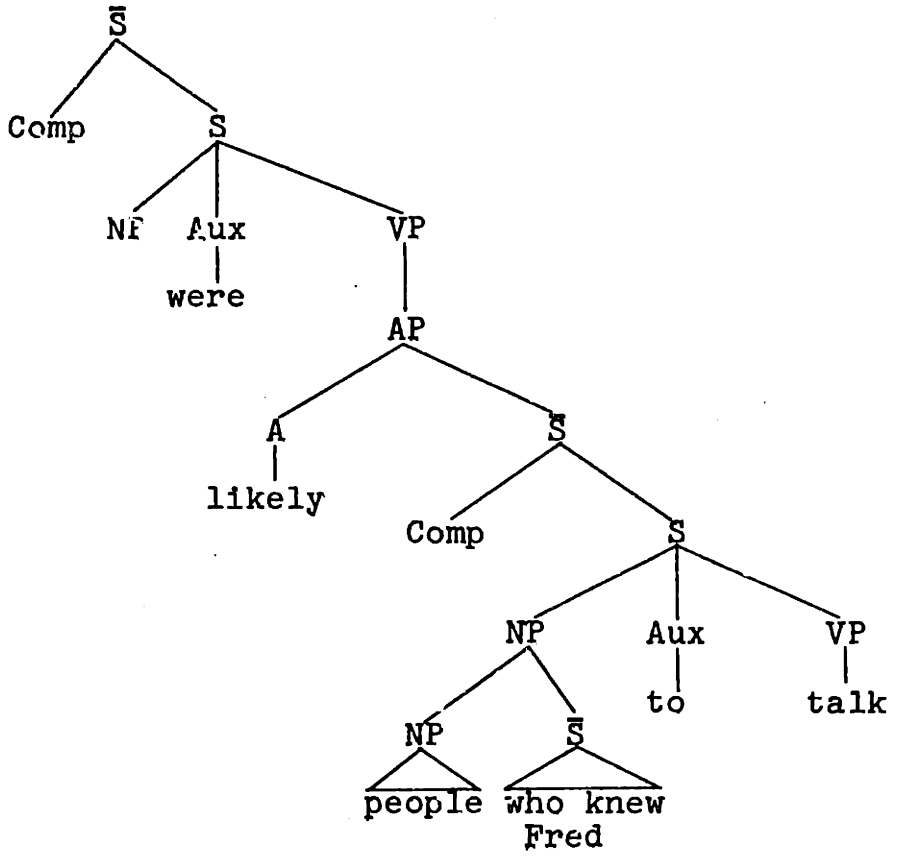
Up to this point, all of the evidence for the construal

schema has come from Q-Float, and it has been assumed that Q-Float is a member of a natural class which includes extraposition from NP, extraposition of PP, and complement extraposition. For the reasons discussed at the beginning of section I, the practical difficulties associated with the other extraposition rules forced a restriction in the focus of our attention. However, there is a possibility that one can make an argument for the construal schema based on extraposed relatives that would be parallel to the argument made about the ungrammaticality of (66-68). In this regard, consider a sentence like (70):

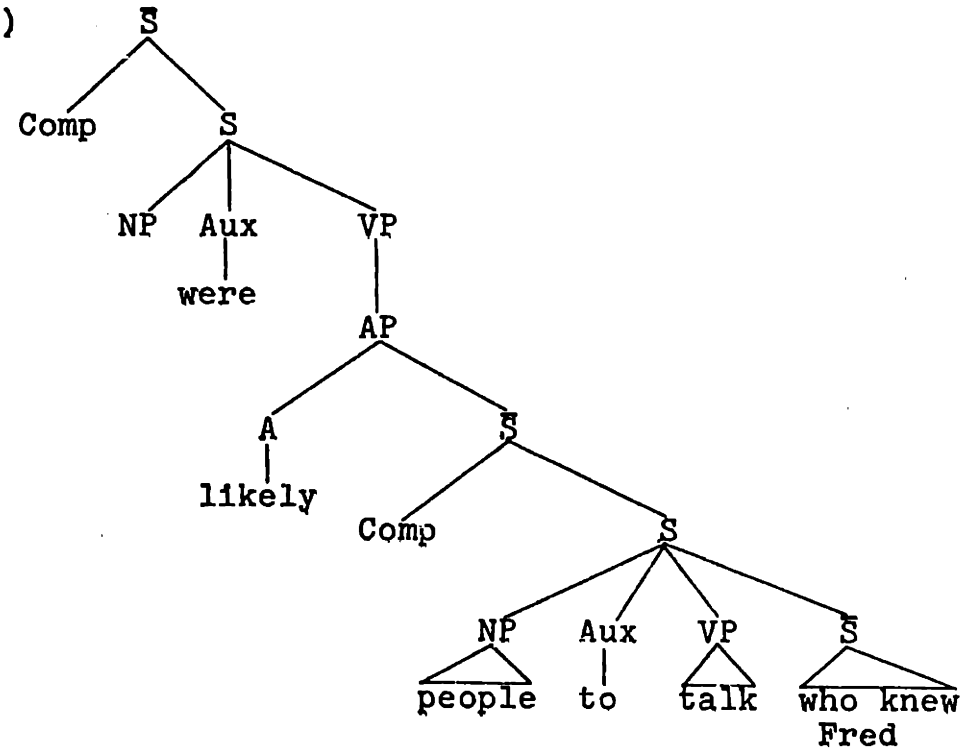
(70) The people were likely to talk who knew Fred.

Now, one can show that likely is a subject-to-subject raising predicate, so that the subject of likely is, assuming a syntactic rule of NP-Preposing, the underlying subject of talk. Given this, if we assume that there is a syntactic rule which extraposes the relative which is cyclic, it should be able to apply on the cycle before NP-Preposing. Therefore, (71) would be the underlying structure of the sentence and (72) would be the derived constituent structure after relative clause extraposition applied:

(71)



(72)



Crucially, these assumptions would lead us to posulate a derived constituent structure as in (72) in which the

extraposed relative clause is a constituent of an S which is in turn a constituent of an AP. Therefore, just as one can front such infinitival complements in (67) and (69), one should, in principle, be able to front the following infinitival complements:

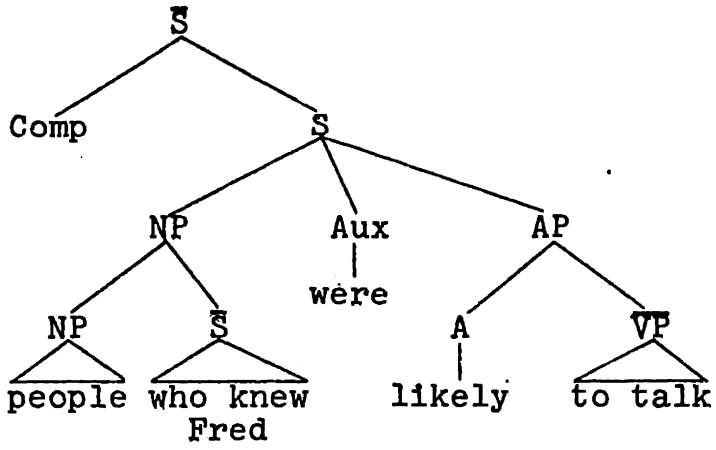
*(73) (a) Just how likely to talk who knew Fred were the people?

(b) So likely to talk who knew Fred were the people that I decided to keep quiet.

With all of the assumptions which I have made until now, this would seem to be a strong argument for the construal schema based on extraposed relatives, since the extraposed relative in (73) and (74) end up to the left of their original heads, and therefore no longer can be linked with them.

However, there are two challengeable assumptions which are at the very foundation of this line of reasoning. One is the assumption that extraposition of relatives is cyclic, and hence would apply before wh-movement. The other controversial point is the existence of a rule of NP-preposing. In particular, Bresnan (1978) has argued against the existence of such a rule. If this is the case, suppose the underlying structure of (70) is (74):

(74)



Notice that the complement of likely is a bare $\overline{VP}$, and is not dominated by S. In Chapter 4, I will argue that relatives which are extraposed from subject position move to the end of S, rather than VP. If this is the case, one could argue that there is no position in the adjective phrase which would be a possible landing site for a relative extraposed from subject position; rather, the extraposed relative in (70) would be dominated directly by S. Therefore, the relevant movement rules operative in (73) and (74) would never front the extraposed relative anyway, since it would be outside the constituent fronted by these rules, namely AP. Therefore, the construal schema would be irrelevant to the ungrammaticality of (73) and (74).

I will now try to deal with these alternatives. The first is based on a proposal by Noam Chomsky made in class lectures at M.I.T. in Spring '78. According to this proposal, the grammar is organized in the following fashion:

(75)

Base Rules

↓ Transformations

Surface Structures

Deletion Rules	Rules of Construal
Filters	Bound Anaphora
Phonology	Opacity
Stylistic Rules	

(See Chomsky & Lasnik (1977) for an exposition of this proposal.) Chomsky has suggested that extraposition rules apply at the level of stylistic rules. One of the motivations for this suggestion is that it would eliminate a stipulation on one of his and Lasnik's filters:

(76) *[$_{\bar{S}}$ that [$_{NP}$ e] ...], unless $\bar{S}$ or its trace (emphasis mine: MRB) is immediately dominated by NP.

[C&L (1977), no.68]

This filter is supposed to account for data supporting a constraint originally proposed by Bresnan (1972) called the "Fixed Subject Constraint". (Bresnan (1977) has a much improved version of this constraint, which she calls the Complementizer Constraint on Variables. Since a discussion of the differences between her constraint and Chomsky & Lasnik's filter would take me too far afield, I simply refer the reader to these papers.) This filter rules out sentences like (77) but allows sentences like (78-80):

*(77) The girl who I pretended that liked pizza visited me.

(78) The girl who I pretended liked pizza visited me.

(79) The girl that liked pizza visited me.

(80) The girl visited me that liked pizza.

The stipulation in the "unless" part of the filter allows (79) and (80) to be excluded from the filter, since the 'That $[_{NP}$ e]' configuration in (79) (a relative clause) is dominated by NP, and the one in (80) is dominated by an $\bar{S}$ which binds a trace dominated by NP. Chomsky has recently proposed to modify this filter by replacing the unless condition in (76) with the following rider:

(81) ...unless the $\bar{S}$ is locally controlled.

Local control would mean that the head of the construction would be adjacent to the configuration ruled out by the filter. If extraposition were a stylistic rule, applying after the filter, then at the stage at which the filter applied, we would have local control in sentences like (80).

Crucially, for our interest, if the extraposition rule follows the filter, and the filter, due to the way the components are set up, follows wh-movement, then extraposition rules must follow wh-movement, and if the rule of extraposition from NP were formulated as in (7), then (73) and (74) could be blocked by the fact that the rule that extraposes relatives is explicitly formulated as a rightward movement rule. This argument for the leftward construal

schema would then not go through.

There are a couple of points to make about the proposal that extraposition rules are stylistic, however. One is that in some languages, these rules must be syntactic. In particular, Taraldsen (to appear), in an extremely interesting paper, has shown that one can perform topicalization and wh-movement out of extraposed relatives in Norwegian. Thus, corresponding to (80), we get sentences like the following:

(82) Per slipper jeg ikke noen inn som liker.
Peter let I not anybody in that likes

*(83) Per slipper jeg ikke noen som liker, inn.

The position of the particle inn in these examples shows whether or not the relative has been extraposed; if the particle intervenes between the head and the relative, the relative has been extraposed, and if the particle follows the head and the relative, as in (83), the relative has not been extraposed. Taraldsen goes on to show that this fact, as well as others, can be explained if one postulates $\bar{S}$, rather than S, as a bounding node in Norwegian, and one postulates extraposition from NP as a syntactic rule in that language, since configurations created as a result of that rule remove subjacency violations.

Chomsky has therefore suggested that extraposition rules may be stylistic in some languages, and syntactic in others. If this dimension is a parameter for language variation, the rider (81) on the Chomsky & Lasnik

'That [$_{NP}$ e]' filter makes an interesting cross-linguistic prediction, given the universality of (81) which would be assumed; the prediction is that in a language in which extraposition from NP is syntactic, the Language X analogue of a sentence like (79) will be grammatical, but the word-for-word translation of a sentence like (80) will not be, since the configuration submitted to the filter's inspection will no longer be locally controlled.

I have no idea whether this prediction is ever realized; I have never heard of such a situation, but that, of course, is irrelevant. I simply point it out as a testable prediction. If such a situation never does occur, we would probably want our theory of grammar to account for such a fact, by modifying the notion of local control.

Therefore, let us put aside this advantage of the proposal that extraposition rules are stylistic, except to note that one can use subjacency to rule out the English equivalent of examples like (82):

*(84) Peter I didn't let anybody in that likes.

Given that S, as opposed to $\bar{S}$, is the bounding node in English, we would be violating subjacency if we attempted to topicalize on the matrix $\bar{S}$ cycle, since we would have to cross two S nodes. I should say here that Taraldsen notes that the Italian and French sentences in which one attempts to extract out of an extraposed relative are still ungrammatical. Therefore, since $\bar{S}$, and only $\bar{S}$, must be the

bounding node in these languages (see the next Chapter of this thesis for elaboration), it is quite possible that extraposition is stylistic in Italian and French.

However, as shown by my proposal to rule out (84), the exact nature of extraposition in English is somewhat unclear. If we consider additional data, I believe that the scales tip in favor of a syntactic rule of extraposition from NP. Thus, looking more closely at the organization of grammar schematized in (75), we see that surface structure is viewed as being operated on by two types of rules which operate independently of each other. (Opacity in (75) corresponds to the propositional island condition and specified subject condition.) If these rules operate independently, however, one would expect stylistic rules to be free to violate the opacity condition. One can readily see that this is not the case. For example, one cannot form (85) from (86):

*(85) The man said that it was the book who was from Boston that was on the table.

(86) The man who was from Boston said that it was the book that was on the table.

Baltin (1977) shows that one must independently be able to extrapose relative clauses to a position before cleft clauses. If we say that extraposition from NP obeys opacity, we have an explanation immediately at hand for the ungrammaticality of (85). However, if extraposition from NP were a stylistic rule for the purposes of

(75), one would have no reason to expect it to obey opacity.

Noam Chomsky (personal communication) has suggested an alternative explanation for the ungrammaticality of (85). He suggests that stylistic rules are all clause-internal; therefore, if extraposition from NP were clause-internal, it could not possibly lower the relative clause into the complement of say. However, we can get sentences like (87) (Chomsky's example):

(87) Who persuaded us who comes from Boston that he would come to an end on Tuesday?

(88) Who persuaded us that he would come to an end on Tuesday who comes from Boston?

In Chapter 4, I will discuss in more detail a rule which I call detachment; this rule moves complements to clause-final position and is, according to my argument there, stylistic. Detachment would derive (87) from (88). Given that detachment is independently motivated, one would have to say that it is also clause-bounded, in order to prevent an alternative derivation of (85).

If we assume that extraposition from NP and detachment are both clause-internal, we can probably still argue that extraposition from NP is a syntactic rule in English, by showing that it applies before wh-movement. For example, consider a sentence like (89):

(89) John pretended that it was the guy that was from Boston that everybody liked.

Corresponding to (89), one can construct sentences in

which the extraposed relative remains in the complement clause:

(90) Who did John pretend that it was that was from Boston that everybody liked?

Since we have agreed that extraposition from NP does not lower the moved element into a lower clause, and that detachment is upward bounded, we must conclude that the relative clause that was from Boston, which was originally part of the focus of the embedded cleft, was extraposed before its head was wh-moved to the front of the matrix S.

Incidentally, it may be helpful at this point to consider the possibility that Q-Float is a stylistic rule. This proposal would run into the problem that, just as extraposed relatives must be separated from their heads in a manner which would suggest that extraposition from NP must be able to precede wh-movement, which is clearly syntactic (as in (90)), Q-Float must also be able to precede wh-movement, as in (91);

(91) The girls who I pretended were all from Boston....

(I chose the verb pretend in order to stave off the possible objection that the matrix verb is a parenthetical, since pretend cannot otherwise be a parenthetical verb.)

Another argument that Q-Float must be syntactic, or at least pre-stylistic, can be made, and this argument has extremely interesting consequences. One can show that

quantification can take place after movement. Among the quantifiers that can float is the quantifier phrase all three. Thus, we have the following paradigm:

(92) All the men were admirers of Garbo.

(93) The men were all admirers of Garbo.

(94) All three of the men were admirers of Garbo.

(95) The men were all three admirers of Garbo.

However, a curious gap in the paradigm exists. Thus, while we can float the quantifier in (96):

(96) All men were admirers of Garbo.

(97) Men were all admirers of Garbo.

we cannot float the quantifier in (98), so that (99) is ungrammatical:

(98) All three men were admirers of Garbo.

*(99) Men were all three admirers of Garbo.

If one believers that Q-Float is a syntactic rule, one will not want to restrict the rule itself so that all three cannot float when the resulting noun phrase is determinerless, but all can. However, a natural explanation suggests itself. If we assume that the noun phrase which takes a certain syntactic shape post-movement receives a semantic interpretation which is the same as other instances of the noun phrase, so that the underlined NP in (100) gets the same interpretation as the corresponding

NP in (93):

(100) The men were admirers of Garbo.

We can suggest the following procedure for semantic interpretation for Q-Floated sentences. Suppose we assume that noun phrases in Q-Floated sentences receive two quantifications. One quantification would occur for the bare noun phrase, and the other would occur with the QP in floated position. (I am assuming, along the lines of Chomsky (1974) and Milsark (1974), that the receives a treatment as a universal quantifier, and so would claim that all noun phrases are quantified.) Thus, suppose we assume a semantics of bare plurals, for instance, along the lines of Carlson (1977), who argues that bare plurals such as "people", "dogs", etc., are treated as natural kinds, so that it is a property of the class that is being referred to. The reason (98) would be out, then, is that the two quantifications would be in conflict. That is, semantically, cardinality (i.e., all three) is inappropriate because it conflicts with the requirement that bare plurals are treated as natural kinds, in Carlson's terms.

A possible objection to this account of why the quantifiers which float are the ones they are, rather than some others, suggests itself. One might note that every, which is, as far as I can tell, synonymous with each in being a distributive universal quantifier, does not float:

(101) The men had each picked up a glass.

*(102) The men had every picked up a glass.

However, it seems to me that every is the exception which proves the rule. While it is true that every cannot float, every one can. Thus, consider the following contrast:

(103) The men were every one admirers of Garbo.
*(104) The men were every admirers of Garbo.

This contrast, as we can see, reduces to the possibilities for this quantifier within the noun phrase itself:

(105) Every one of the men was an admirer of Garbo.
*(106) Every of the men was an admirer of Garbo.

Notice now the possibilities for each in both floated and non-floated position:

(107) Each of the men was an admirer of Garbo.
(108) Each one of the men was an admirer of Garbo.
(109) The men were each admirers of Garbo.
(110) The men were each one admirers of Garbo.

Thus, it seems that the possibilities of a quantifier's occurrence in floated position reduces to its possibilities within the noun phrase. As far as I can tell, there is no semantic explanation for why every must be followed by one in the partitive construction, while each simply may. There are two syntactic accounts of this fact. In the first,

we set up a subcategorization restriction in which distributive quantifiers must co-occur with one in the QP when in the partitive construction, and posit a local deletion rule like the following:

(111) one $\longrightarrow$ $\emptyset$ / each___

The second alternative would be to make the subcategorization restriction optional in the case of each. I have no way of deciding between these two alternatives. In any event, we can see that every does not damage the generalization.

Ken Hale (personal communication) has informed me that this proposed device of double quantification receives interesting cross-linguistic support. There are languages like Japanese which differ from English and French in that any quantifier can float, even a numeral. These languages all have the property that they lack articles; therefore, the definite-indefinite distinction, which I am assuming is a corollary of different types of quantifications, is neutralized in these languages, and so noun phrases in such languages would be vague with respect to these types of quantifications. Therefore, since noun phrases without quantifiers would receive no quantification, noun phrases which have launched floated quantifiers would receive only a single quantification in such a language (e.g., Japanese) and thus no conflict would result from conflicting quantification. Thus, no device would exist in such a language

to restrict the operation of Q-Float, and so Q-Float would be unrestricted.

Notice, incidentally, that if this explanation for which quantifiers can float is the correct one, it provides a fairly strong argument for the autonomy of syntax, for consider a theory such as that of generative semantics (see Lakoff 1970, Postal 1970, McCawley 1968), in which semantic representation, which is usually given in a sort of predicate calculus notation, is identified with underlying syntactic structure. In such a theory, the logical scope relations of a sentence such as (98) are determined at underlying structure, and therefore, obviously, before any syntactic movement rules, such as Q-Float, apply. Therefore, the shape of the noun phrase after Q-Float should be totally immaterial for semantic interpretation, since semantic interpretation would have already applied. That is, there would be no reason at all, under that theory, to expect the noun phrase men to behave like an instance of a bare plural in a sentence lacking a floated quantifier. A procedure which performed quantification post-movement, which is crucially relied on here, would be totally unavailable in such a framework. (The implications of this account for trace theory will be dealt with in Chapter 3.)

Crucially, also, if Q-Float applied after filters, on the left-hand side of the grammar in (75), there would be no reason to expect the movement to affect logical form, whereas the account proposed here for the ungrammaticality

of (99) demands that movement be able to feed quantification.

Given these arguments, I think that there is some indication that extraposition rules in English occur in the transformational component.

I will now discuss the question of what node dominates infinitival complements of adjectives such as likely. Suppose it is simply $\overline{V}P$, as in (74). The consequences for the effect of the construal schema to rule out (73a & b) have already been discussed. The claim that such complements are not dominated by S directly predicts that elements which do not occur under $\overline{V}P$, but under S, will not be able to co-occur with these adjectival complements; for example, result clauses (Andrews 1975, Williams 1974, Sag 1976) have been shown to occur under S. For example, in the following sentence, the result clause is most plausibly generated outside of the coordination of two S's:

(112) John is so fat, and Bill is so short, that neither of them will make the basketball team.

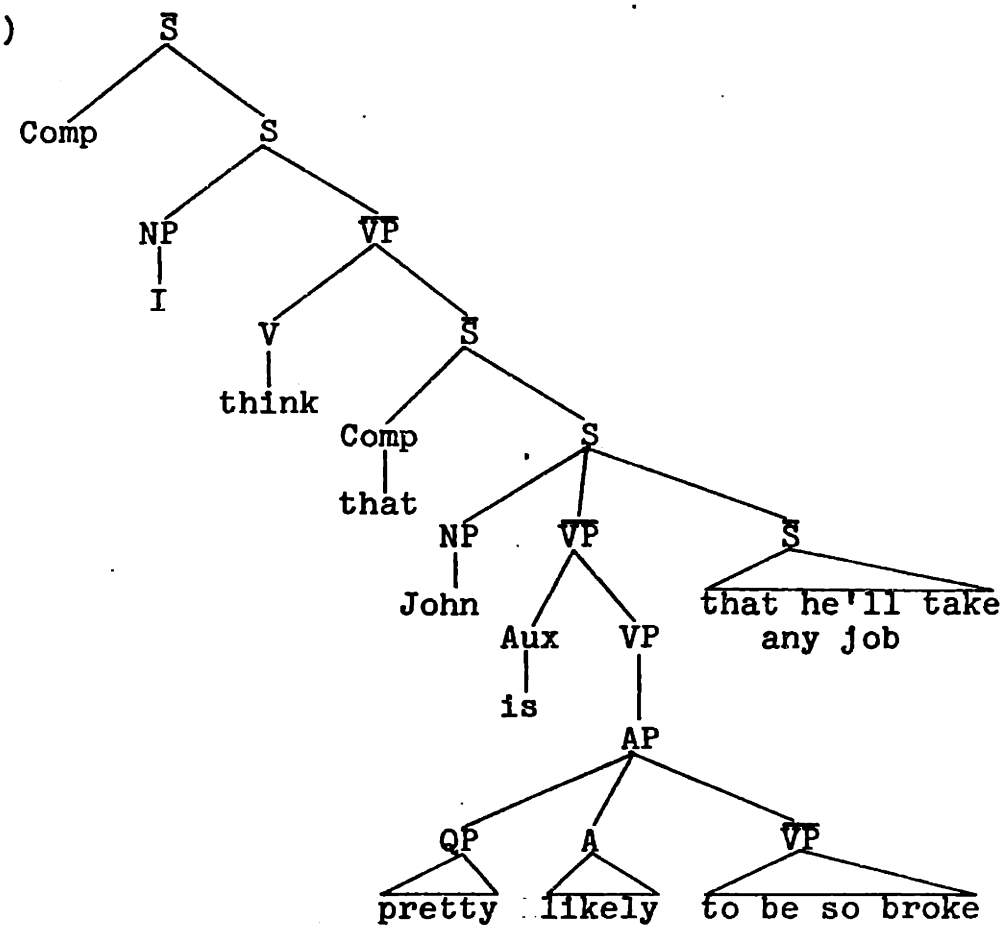
(For further arguments, see Andrews 1975, Williams 1974, Sag 1976.) Given this, the position that the complement of likely is a bare $\overline{V}P$ (for ease of exposition, I shall henceforth refer to this as the $\overline{V}P$ position) would lead one to predict that in a sentence like (113):

(113) I think that John is pretty likely to be so broke that he'll take any job.

the result clause would be attached to the matrix S, and

hence out of the adjective phrase. The structure would then be as follows:

(114)



If the result clause were attached as in (114), and hence out of the adjective phrase, then rules which front adjective phrases should obligatorily strand the result clause. Hence, (115) should be good and (116) ungrammatical. I think that the data go in the opposite direction, however:

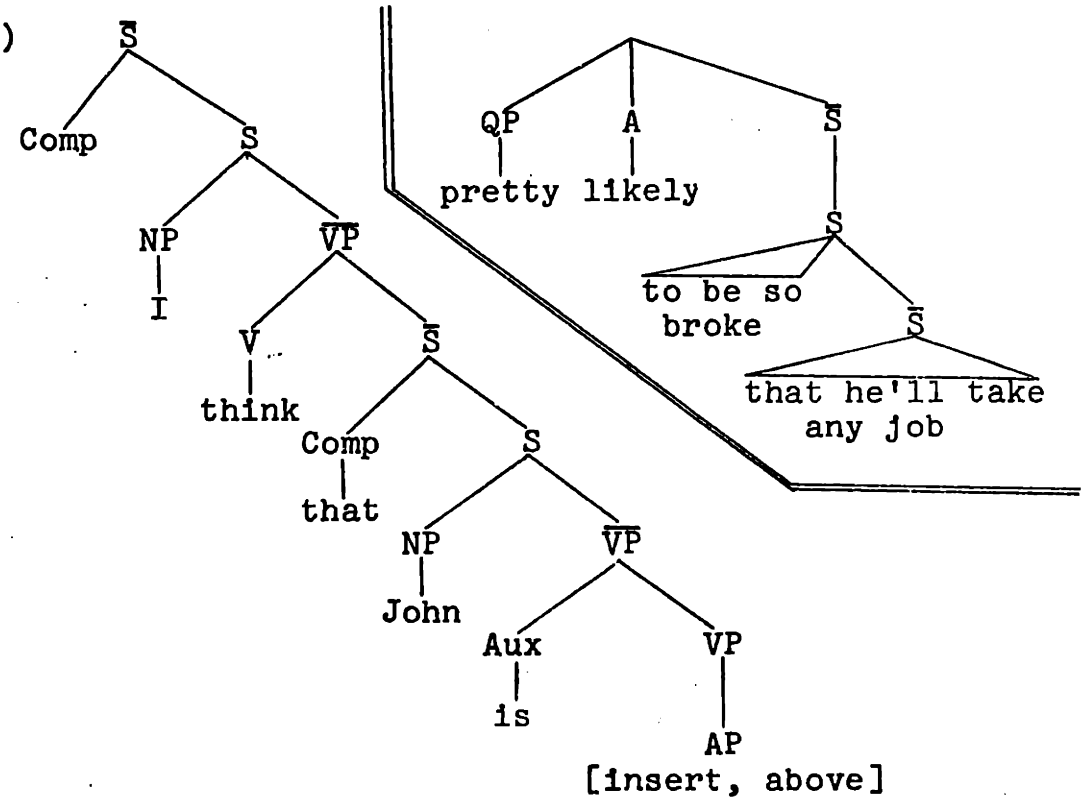
*(115) Just how likely to be so drunk do you think that John is that he'll take any job?

(116) Just how likely to be so drunk that he'll take any job do you think that John is?

If we posit an S above the infinitival complement of likely, we can then provide a home for the result clause within the AP, and hence allow the result clause to be fronted when wh-movement moves the AP to the front of the clause.

Therefore, I will henceforth assume that the infinitival complement of likely is dominated by S. Once we make this assumption, we must then ask again what would rule out (73a & b). As we have seen from (63), the insertion-prohibition of Chomsky (1965), which prohibits morphological material from being inserted into a lower clause, does not seem to be operative. Therefore, even if one were to accept the arguments of Bresnan (1978) that there is no need for a rule of NP-Preposing, so that the underlying structure of (70) would be as in (117):

(117)



the extraposed relative would be able to move into the infinitival complement, and hence could be subsequently fronted as part of the AP to yield the ungrammatical (73a & b). The construal schema, then, would block (73a & b) by placing the modifier (the extraposed relative) to the left of its head, and thus not in a position to be linked with the head.

IV. Two Final Modifications

Two final modifications must be made here. In discussing the ungrammaticality of various examples in the preceding sections, the reason given was that the modifiers ended up to the left of their heads, in violation of the construal schema. However, there is a case in French in which this state of affairs can occur. Kayne (1975) discusses in detail the rule of L-tous, which, in certain situations, moves quantifiers optionally to the left over a verb. The rule relates (118) and (119):

(118) Je les ai tous vu.

(119) Je les ai vu tous.

This rule cannot operate out of full NPs, so that (120) is ungrammatical:

*(120) J'ai tous lu les livres.

(121) J'ai lu tous les livres.

In fact, the only time it can operate is when the ante-

cedent is cliticized or wh-moved, or when the quantifier occupies the entire noun phrase. Thus, the following sentences are grammatical:

(122) Qu'avez-vous tous lu?
'All of what have you read?'

(123) J'ai tout lu.

(124) Qu'avez-vous lu tous?

(125) J'ai lu tout.

Interestingly, cliticization and wh-movement have the effect of moving heads to the left of the quantifier, so that the clitic in (118) and (119), and the wh-element in (122) and (124) precede the quantifiers that modify them, which is exactly what the construal schema predicts. However, there are sentences in French which superficially violate the construal schema, as in (126):

(126) J'ai tous voulu les lire.
'I've wanted to read them all.'

Here, the tous, which Kayne shows can move over an indefinite number of contiguous verbs, precedes the clitic which it must modify.

There are two ways out of this dilemma, one of which is due to Noam Chomsky (personal communication), who has suggested that if one assumes that floated quantifiers leave traces, one can maintain that the quantifier which ends up to the left of the clitic in (126) can bind the position the quantifier originally moved from, and the clitic can be

construed with the trace.

A second alternative would be to apply (17), the construal schema, at the end of each syntactic cycle, in the same manner as Lasnik's (1972) rules for scope of negation. Therefore, (126) will result from L-tous applying on a higher cycle after construal.

At this point, I have no way of deciding between the two alternatives. At the end of the next chapter, I shall show that the proposal that QPs leave traces has consequences for the nature of semantic rules that first seem intolerable, but that may be necessary in any event. Therefore, we can opt for the trace approach, and do interpretation at a unified level.

We began this chapter by asking why a certain class of movement rules didn't exist. One implication of the construal schema is that a notation which allows leftward extraposition rules is too powerful. Given the construal schema, we can dispense with brackets in the formulations of rules (7-9). It also enables us to collapse L-tous and R-tous, since the differences which Kayne (1975) pointed out (i.e., that L-tous does not operate out of full NPs, but R-tous does, etc.) are now predictable. Therefore, Q-Float in French, if a movement rule, if bidirectional, and the existence of this class of rules is predicted by a notation for movement rules given in Chapter 4.

As Ken Hale (personal communication) informs me, the universality of (17) must be modified somewhat. For ex-

ample, in Cebuano (Bell 1976), and Maori, Tongan, and Samoan (Chung 1976), as well as Papago, floated quantifiers end up to the left of their heads. The relevant examples are as follows:

(127) (a) Cebuano [Bell's (39b)]

Gibasa tanan sa mga istudiante ang mga libro ni Rizal.
(tanan is the floated quant.)
'Rizal's books were all being read by the students.'

(b) Tongan [Chung's (175c)]

na'e tafulu'i kotoa ia 'e he kakai vaivai.
past scold all him erg. the people old
'The old people all scolded him.'

(c) Samoan [Chung's (186b)]

pe sa omai 'uma lava i Oceanside tagata?
Q past come- all Emp to " " man?
pl.
'Did the people all come to Oceanside?'

(d) Papago [Hale, personal communication]

nt o hema ha-cepos hegam wipsilo.
AUX Fut. one them-brand(sg.) those calves
'I will brand one of those calves.'

Hale informs me that in Papago, the language with which he is most familiar, the other extraposition rules obey the construal schema (the necessary data is lacking in the other languages) and goes on to suggest that one perhaps should not consider quantification in the case of floated quantifiers to be a case of modification.

These languages, however, have a crucial property which distinguishes them from English, French, Rumanian, and Persian. English, French, Rumanian, and Persian reconstruct grammatical relations on the basis of word order, while the Malayo-Polynesian languages mentioned here and Papago do not; Maori, Tongan, and Samoan have an extremely rich system of case marking, and Papago makes heavy use of verbal agreement. Notice, however, that case-marking and verbal agreement, by their very nature, are perceptually salient markers of grammatical relations; therefore, it is quite possible that floated quantifiers, in such languages, are construed directly with the marker of a particular grammatical relation. Since I do not know the details of these languages, these remarks are speculative in nature, but do receive some support from a recent paper by Shibatani (1977), who treats Q-Float in Japanese. Shibatani shows that noun phrases which are marked by the nominative always launch floating quantifiers, regardless of whether or not they are subjects, whereas noun phrases which act as subjects with respect to other processes, such as subject honorification and triggering reflexives, will not launch floating quantifiers if they are marked by a case other than the nominative. Therefore, in Japanese, it seems that we must restrict Q-Float to noun phrases marked by particular cases. Rather than formulating a rule of Q-Float in Japanese which explicitly mentions case, suppose we simply say that construal in such a language operates with case rather

than word order.

Therefore, we would be led to the following prediction: Languages which violate the construal schema (17) will be those which have perceptually salient, phonologically specified markers for NPs, i.e., case-markers or agreement markers. Languages which lacked these phonologically specified markers, however, would rely on fixed word-order to reconstruct grammatical relations, and so the construal schema would fulfill a useful function in such languages. Therefore, the generalizations which encompass the absence of leftward extraposition rules, and the rule ordering facts of sections I and III, in a class of languages, are actually explained by viewing these generalizations to be corollaries of a construal schema. That is to say, the construal schema can explain why these generalizations hold for the class of languages that they do, rather than some other, arbitrary set. Therefore, the violations of the construal schema in Papago, Maori, Tongan, Samoan, and Cebuano could actually support the notion of a construal schema, contrary to a first glance at the facts.

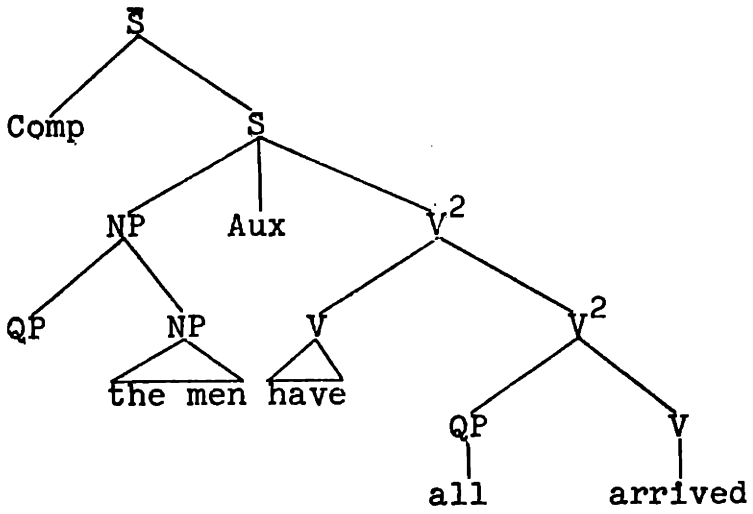
IV. Concluding Remarks: The Schema in the Scheme of Things

One of the premises of trace theory is that the relation between a moved element and its trace, which Chomsky (1977) defines as the category label which is left behind when the contents are moved away, is one of anaphora.

Thus, Fiengo (1974) argued that a trace cannot precede and command its antecedent. Now, the relation between a trace and its antecedent is defined on a notion which Reinhart (1976) calls C-Command (I go into more detail about this notion in Chapter 4). This condition on traces has come to be known as proper binding (see Dresher and Hornstein (to appear) for relevant discussion).

Implicit in discussions of proper binding has been the notion that one does not have to bind an antecedent to its trace by a special rule; the normal conditions on anaphora suffice. While this may work for rules like wh-movement, it definitely does not work for a rule like Q-Float. Suppose that the derived constituent structure for a sentence like (3b) is what I argue that it is in Chapter 4, namely this:

(128)



We can see that the trace of QP in subject position, the empty QP, does not C-Command its antecedent, which is under the lower V^2; in fact, neither is in the domain of the other, in Reinhart's sense. Therefore, in the case of a

rule like Q-Float, it does not seem possible to appeal to general conditions on anaphora to determine the original position for the floated quantifier. To do so, it seems necessary to use a rule. If we accept trace theory, then, suppose we flesh in the construal schema in the following way:

(129) $[_{NP} \ldots[_{\bar{\bar{X}}} e]\ldots] \; Y - \bar{\bar{X}}$

where e is the designated empty symbol.

In other words, this rule would bind modifiers which begin within an NP, $\bar{\bar{X}}$ being the modifier in this case (or whatever the appropriate category symbol would be). Needless to say, this would not be the only way in which elements moved out from under the domination of an NP can be bound to their traces. Thus, in some languages, such as French, complements of nouns can be extracted from within a NP by wh-movement, as in (130):

(130) La problème dont j'ai trouvé la solution...
'The problem for which I have found a solution...'

(131) Combien avez-vous donné de pommes à Jean?
'How many have you given of apples to Jean?'

In the case of wh-movement, however, wh-movement is not a rule which moves only modifiers of NPs; it moves entire NPs, PPs (in some languages) which do not have to be part of NPs originally, APs (again in some languages). The reason the modifiers in (130) and (131), then, can end up

to the left of their heads is that they get interpreted by the construal procedure for wh-movement; in effect, they get a sort of "free ride", and can escape from the effects of (129) because they get interpreted in another fashion. On the other hand, since the extraposition rules in this chapter only move elements which are, in a sense syntactic dependents, the schema (129) would be the only option available for the construal of such elements.

FOOTNOTES TO CHAPTER 2

1. Many linguists would not call Q-Float ar extraposition rule. I cannot determine whether or not this is simply due to the fact that it does not have the term extraposition in the name for the rule. At any rate, I am including it because I claim that it is a member of a set of rules which form a natural class.

2. All information about Rumanian comes to me from Donca Steriade, to whom I am greatly indebted. She came up with the idea for the argument in this section.

CHAPTER 3: BOUNDING AND CONSTRUAL

In this chapter, I will discuss various theories of bounding which have been proposed in the literature, trying to shed light on empirical predictions which each makes along the way. I will also argue that operations across PP boundaries must be constrained. After establishing this conclusion, I shall argue that one of the implications of this is that bounding is more appropriately defined as a condition on the applicability of syntactic transformations rather than as a condition on construal. In the first section, I will review major theories of bounding.

I. Theories of Bounding

A. Ross (1967)

Ross, in his dissertation, argues that the A-over-A condition proposed by Chomsky (1962, 1964) should actually be replaced by a number of separate constraints, among which are what have since been called the Right Roof Constraint, the Complex Noun Phrase Constraint, the Sentential Subject Constraint, and the Coordinate Structure Constraint.

The Right Roof Constraint states that a rightward movement rule cannot move an element out of an S-node that dominates the element prior to the application of the movement rule. As discussed in Chapter 2, this would account for the inability of a relative clause to extrapose to the end of the matrix clause as in (1):

*(1) Just how eager to call people up are you who live in Reno?

(2) Just how eager to call people up that live in Reno are you?

This crucially assumes that infinitive phrases are dominated by S, rather than by VP. This point is discussed in Chapter 2, so I will not recapitulate the arguments here.

The Complex Noun Phrase Constraint states that no element may be moved out of an $\bar{S}$ immediately dominated by NP when the NP has a lexical head adjoined to the $\bar{S}$. This would account for the inability to question out of a relative clause or noun-complement construction, as in (3) or (4):

*(3) Who did John meet the girl that liked?

*(4) Who did Fred discuss the claim that Fred liked?

The Sentential Subject Constraint forbids one from moving an element out of a sentential subject. From (5), then,

(5) That Max liked Mary was obvious.

one cannot form (6),

*(6) Mary that Max liked was obvious.

by topicalizing the object of "liked".

The Coordinate Structure Constraint simply says that one cannot remove an element from a coordinate structure. Thus, from (7):

(7) John liked Peter and Mary.

one cannot wh-move and element in the position of "Mary":

*(8) Who did John like Peter and __?

B. Chomsky (1973)

Chomsky (1973) presents a unified alternative to several of Ross's constraints, in particular the Complex NP Constraint and the Right Roof Constraint. Chomsky formulates a principle knowns as Subjacency, which states that movement rules may involve only terms in adjacent cycles. Therefore, if NP and S are both cyclic nodes, sentences (3) and (4) will be ruled out since any movement out of a complex NP will cross two cyclic nodes, thus violating Subjacency. Similarly, the Right Roof Constraint would fall out of subjacency since, for instance, a movement rule which moved an element out of an NP past the S dominating the NP would be violating Subjacency. Apparent violations of Subjacency, as in (12):

(12) Who did Fred pretend that Max said that Mary liked?

are explained by postulating successive cyclic movement is that the movement must be from complementizer position to complementizer position, and that elements cannot move out of Comp. In section III of this chapter, I will provide more evidence for this condition, and argue for some theoretical implications that it has.

C. George Horn (1974)

G. Horn, in his doctoral dissertation, argues for a constraint which he calls the NP Constraint. This constraint states that elements may not move out of NPs. It subsumes the Complex NP Constraint and the Sentential Subject Constraint, and another constraint noticed but not formulated by Ross. Horn observes that PPs cannot be questioned when they are part of an NP. Therefore, we have the following patter of grammaticality:

(13) John wrote a book about Nixon.

(14) About whom did John write a book?

(15) John destroyed a book about Nixon.

*(16) About whom did John destroy a book?

Horn's explanation for this pattern is that the sentence (13) has the bracketing in (17), while (15) has the bracketing in (18):

(17) $[_{\bar{S}}$ Comp $[_{S}[_{NP}$ John$][_{VP}[_{V}$ wrote$][_{NP}$ a book$]$ $[_{PP}$ about Nixon$]_{VP}]_{S}]_{\bar{S}}]$

(18) $[_{\bar{S}}$ Comp $[_{S}[_{NP}$ John$][_{VP}[_{V}$ destroyed$][_{NP}[_{NP}$ a book$]$ $[_{PP}$ about Nixon$]_{NP}]_{VP}]_{S}]_{\bar{S}}]$

Horn's arguments for this difference in bracketing are quite convincing, as is his claim that the difference correlates with differences in wh-movement. However, there are a number of reasons to reject his NP Constraint.

For one thing, extraposition rules, as discussed in Chapter 2, are a straightforward counterexample to the NP Constraint. Horn notes this, and argues that one can violate the constraint only in cases where one explicitly mentions the NP brackets in the formulation of the rule. This proviso on the constraint is not unreasonable, since it defines the constraint as a sort of unmarked condition on rule application. That is, one may formulate rules that are custom-made to violate the constraint, by explicitly flagging the rule as a violation. This approach has been taken by other investigators. For instance, Chomsky (1973), after formulating the A-over-A Condition, makes the following remarks:

> Notice that the condition.... does not establish an absolute prohibition against transformations that extract a phrase of type A from a more inclusive phrase of type A. Rather, it states that if a transformational rule is nonspecific with respect to the configuration defined, it will be interpreted in such a way as to satisfy the condition.
>
> [Chomsky (1973), p. 235]

Nevertheless, if one accept the arguments in Chapter 2, one will not wish to formulate these rules with labelled bracketings, since one can predict the fact that these rules are extraposition rules on other, more general grounds. Also, his constraint does not predict the generalization noted by Akmajian, since extraposition of PP, being one of the rules that doesn't fall under the NP Constraint, should be able to move a PP embedded under an infinitely large number of NPs.

To account for the range of facts that Horn (1974) discusses, Chomsky (1977) proposes that with Subjacency as the relevant conception of bounding, S should be a bounding node, as opposed to $\bar{S}$. Given this, an extraction which moved an element out of an NP past the S would violate Subjacency. However, given that extraposition rules, unlike wh-movement, are strictly S-internal, movement out of the NP by an extraposition rule would not violate Subjacency.

Chomsky's account and Horn's account make different cross-linguistic predictions, in languages in which S could not be a bounding node. For example, Rizzi (1978) and Sportiche (1978) have argued that in Italian and French, respectively, $\bar{S}$ rather than S must be the bounding node, so that in these languages it is possible to violate the wh-island condition, as long as the moved element originates within the wh-clause. Therefore, the following sentences are grammatical in these languages:

(19) L'homme à qui je sais quelle lettre écrire.
'The man to whom I know what letter to write.'

(20) Il solo incarico che non sapevi a chi avrebbero affidato è poi finito proprio a te.
'The only charge that you didn't know to whom they should entrust has been entrusted exactly to you.'

[Rizzi's (6a)]

Rizzi argues convincingly that the fact that one can extract out of a wh-headed clause in Italian, as well as a host of other facts, could be explained by postulating

only $\bar{S}$ as the bounding node. Therefore, if an element is extracted out of a wh-headed clause, it would not violate Subjacency if it moved up to the Comp of the next clause, since a rule which cycled on S would simply be involving one cyclic node. Crucially, however, this explanation for the difference between English, on the one hand, and French and Italian, on the other, predicts that violations of the NP Constraint will be able to occur when the Comp of the complement is not filled by a wh-form. Rizzi's examples are the following:

(21) Questo autore, de cui ricordo che mi hai mostrato il primo libro, mi sembra estremamente interessante.

'This author, by whom I remember that you showed me the first book, seems to me very interesting.'

[Rizzi's (25c)]

(22) Francesca, di cui so che recentemente hai conosciuto la sorella, è una mia buona amica.

'Francesca, of whom I know that you have recently met the sister, is a good friend of mine.'

[Rizzi's (26c)]

Similarly, in French, one finds counterexamples to the NP Constraint like the following:

(23) Combien avez-vous mangé __ de pommes?

'How many have you eaten __ of apples?'

(24) Je voudrais savoir combien vous avez mangé __ de pommes.

'I would like to know how many you have eaten __ of apples.'

(The possibilities of extraction of <u>combien</u> were

pointed out and discussed most extensively by Hans Obenauer; see, for instance, Obenauer (1976).)

It seems, also, that one can extract noun-complements fairly easily in French, as in (25):

(25) La problème dont j'ai trouvé la solution...
'The problem to which I have found the solution...'

Also, as Dominique Sportiche has pointed out to me, it is possible to violate the Sentential Subject Constraint, as in (26):

(26) L'homme avec qui parler serait difficile...
'The man to whom to speak would be difficult...'

These facts indicate that the NP Constraint should not be a primitive of the metatheory, but rather that, where applicable, it should fall out of an independently necessary and more general conception of bounding.

D. Bresnan (1977)

Bresnan (1977) argues for a formulation of Subjacency which is refined in Bresnan and Grimshaw (1978). In this conception of Subjacency, Subjacency is not viewed as a condition on analyzability but rather a condition on trace binding, in which traces are copied into the Comp of each subjacent cyclic node, and then ultimately at the initial gap. The procedure of copying the trace itself, then, is the operation that is bounded by Subjacency. The reason for this conception of Subjacency is that Bresnan has argued

that there are rules which delete over a variable, and which are not appropriately reformulated as movement rules, but which are nevertheless subject to the "island constraints", which Subjacency is designed to explain. Therefore, on this account, syntactic rules themselves are unbounded, so long as a path of subjacent Comps intervenes between the controller of the antecedent-trace relation (the moved element itself in the case of a moved element; the antecedent in the case of a rule which deletes an element under identity).

I do not intend to directly address the merits of the case for deletion over a variable in this thesis, since I have nothing of substance to add to the debate (see Bresnan 1975, 1977a, b; Chomsky 1977). However, even if one believes that deletion over a variable exists, one can still reformulate Bresnan's conception of Subjacency into a condition on proper analysis. In this case, one could analyze a string with respect to a transformation over an unbounded length, just so long as every cyclic node between the constants in the rule had an empty Comp.

There would be one problem, however, with viewing Subjacency as a condition on analyzability in a theory with unbounded rules, which would not arise if one viewed Subjacency as a condition on analyzability in the sense that one simply could not operate on non-adjacent cycles (Chomsky's position) or a condition on trace binding (Bresnan's position). Let us call the third position, the one I have

suggested here, as the Unbounded-Analyzability position (UAP). If one accepts the UAP, one must have some way of preventing two unbounded rules from applying in the same domain. For example, consider (27), in which comparative deletion has applied:

(27) John gave more books to Martha than Fred gave to Sally.

One must have some way of blocking wh-movement of the dative in the comparative clause:

*(28) To whom did John give more books to Martha than Fred gave?

One can block this situation in Chomsky's system by postulating comparative deletion as wh-movement (see Chomsky 1977), which is a comp-substitution rule. Under this account, there is only one position in Comp aside from the position under Comp where the invariant marker hangs, for moved elements to move to. Therefore, filling the Comp with the compared constituent would block movement of the questioned dative into the Comp in the comparative clause, and Subjacency would not allow direct extraction out of the comparative clause into the matrix clause.

Bresnan's system could equally well handle this restriction, by postulating a principle against trace-erasure. She could argue that when one copies a trace into a Comp, (i.e. by comparative deletion) one cannot erase that trace by another operation which would copy a

trace into that Comp. Dresher & Hornstein (forthcoming) have independently argued for such a principle. Therefore, the two systems here make the same predictions.

However, given that the UAP says nothing about traces and allows one to formulate unbounded rules, one would have to stipulate, under this conception of bounding, that the Comp becomes inaccessible if it has been analyzed already with respect to some syntactic rule. I take no stand here on the theoretical intolerability of such a stipulation, and am merely pointing it out as a consequence.

E. Summary of Section I

In this section, I have tried to review some theories of bounding and to draw some differences between approaches. In the next section, I shall argue that PP must be defined as a bounding node for the purposes of subjacency, and then I shall try to draw some conclusions about the nature of bounding in general from this.

II. PP as a Bounding Node

First, consider the rule that extraposes relative clauses in English. This rule relates sentences like (29) and (30):

(29) A man arrived who was from Philadelphia.

(30) A man who was from Philadelphia arrived.

This rule interacts crucially with wh-movement, in

that one can wh-move the head of a relative clause if one extraposes the relative clause, as in:

(31) Who do you know who's from Boston?

(The order of application of wh-movement and extraposition is irrelevant here.) Now, a curious restriction exists on this extraposition rule. When the element which would undergo extraposition is contained within a prepositional phrase, as in (32), extraposition yields unacceptable results;

(32) In which magazine which was on the table did you see it?

*(33) In which magazine did you see it which was on the table?

If both PP and NP were cyclic nodes, subjacency would block extraposition out of the PP.

A second argument for PP as a bounding node can be made if we assume, following Emonds (1976), that adverbial subordinate clauses are introduced by the following base rule:

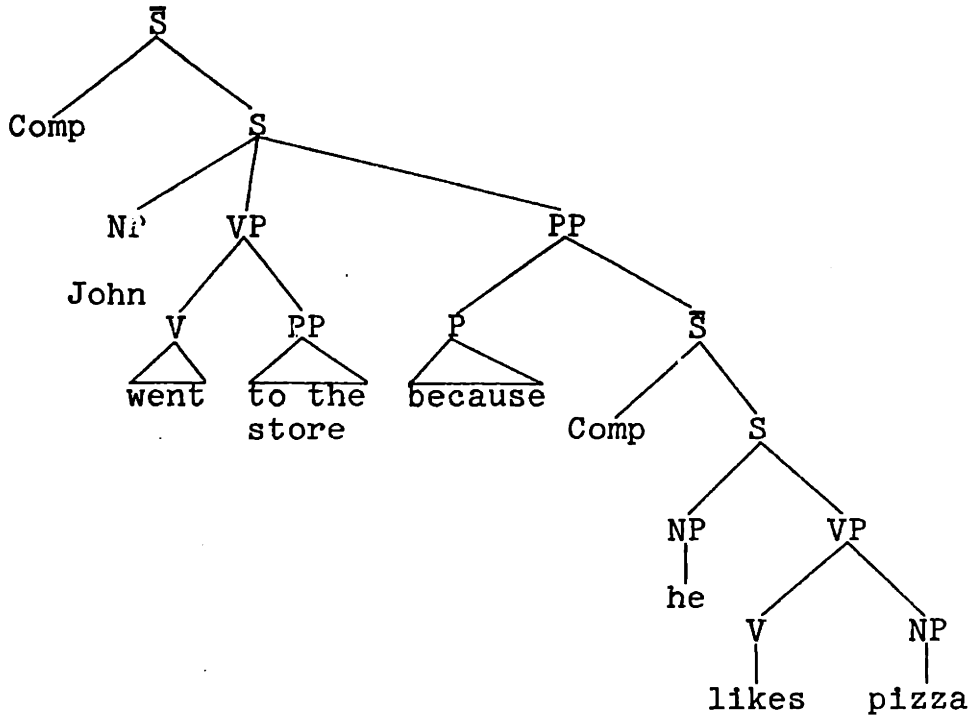
(34) PP $\longrightarrow$ P $\bar{S}$

Therefore, a sentence like (35):

(35) John went to the store because he likes pizza.

would have the following underlying structure:

(36)



Now, as is well-known, adverbial subordinated clauses are extraction islands (Geis, 1969; Bresnan, 1977). Therefore, one cannot wh-move an NP in the position of pizza in (36) to form a question like (37):

*(37) What did John go to the store because he likes?

The PP in (36) is dominated by S, which we have claimed is a bounding node in English. Therefore, if wh-movement were to apply in (36), it would move the wh-element past the PP and the matrix S, violating Subjacency if PP were a bounding node.

In languages like French and Italian, in which S̄, rather than S, has been argued to be the relevant bounding node, the equivalents of (37) would be blocked if PP were a bounding node in these languages and adverbial subordinate clauses were introduced by rule (34). Movement of

the wh-element to initial position ofthe $\bar{S}$ in the PP would be as far as the wh-element could go without violating Subjacency, since movement into the matrix $\bar{S}$, as in (37), would cause the movement to cross the $\bar{S}$ within the PP, and the PP itself, thus operating across two bounding nodes.

Thus, we have two areas of syntax so far which the postulation of PP as a bounding node would explain. There is an apparent problem with this analysis when we consider the internal structure of prepositional phrases. Jackendoff (1973) postulates the following base rule:

(38) PP $\longrightarrow$ P PP

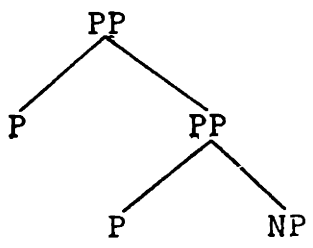
This expansion for PPs, according to Jackendoff, is instantiated by the following examples:

(39) Harpo rode the horse out of the barn.

(40) Sam disappeared down into the barn.

If the hypothesis about PP as a bounding node is correct, then one should not be able to extract the object out of the most deeply embedded prepositional phrase. The reason is that subjacency would block direct extraction in the configuration in (41):

(41)



The results are counter to the prediction, however. Thus, the wh-moved versions of (39) and (40) sound fully acceptable to my ear:

(42) What did Harpo ride the horse out of?

(43) What did Sam disappear down into?

While the base rule which Jackendoff postulated may be a correct one in some cases, however (and I will argue below that it is indeed correct in some cases) I do not think that all sequences of two prepositions followed by a noun phrase are instances of the nested PP structure. In an extremely interesting paper, Hendrick (1976) argues that the following base rules exist:

(44) (a) $P'' \longrightarrow$ [Spec P'] P'

(b) Spec P' $\longrightarrow \left\{\begin{matrix}\text{right} \\ \text{straight}\end{matrix}\right\}$ P'

Among Hendrick's diagnostics is the observation that the intensifier right, which is a specifier of prepositions occurring at the leftmost boundary of prepositional phrases, cannot always occur between the first and second prepositions in the sequence Prep+Prep+Noun Phrase. For example, (45) is ungrammatical:

*(45) Sam disappeared down right into the barn.

The variant (46), however, is grammatical:

(46) Sam disappeared right down into the barn.

One might attempt to counter the force of Hendrick's observation by noting that not all prepositions can co-occur with the specifier right even when the preposition is the first, or only one, in the sequence. For example, of cannot co-occur with right under any circumstances:

(47) He dreamed of Jeannie.

*(48) He dreamed right of Jeannie.

Therefore, one might draw the conclusion that, since there is an uncontrolled variable in the test for the internal structure of these PPs, the distribution of right can shed no light on this question. This conclusion would be drawn too hastily, however, since there are prepositions which can occur as the second member of the two preposition sequence which can co-occur with right independently. Into is one of these prepositions:

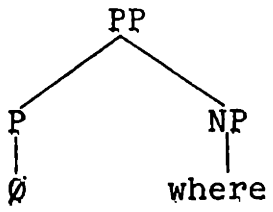
(49) He ran right into a truck.

However, when into is the second member of the sequence, it still cannot be preceded by right when the first member of the sequence is down or another member of a certain class of prepositions, as is evidenced by the ungrammaticality of (45). There are cases, however, where right can intervene between the first and second preposition, as in (50):

(50) A rabbit appeared from right behind the rock.

One must account for the difference between cases where right can and cannot intervene. If we were to claim that both Jackendoff's structure and Hendrick's structure exist, we might be able to correlate the position of specifiers like right with other facts. For instance, Grimshaw (1977) and Bresnan and Grimshaw (1978) have argued that certain wh-morphemes, such as where, have the following structure:

(51)



This analysis of where captures naturally the fact that where exhibits both PP-like and NP-like characteristics. It acts as a PP in being able to stand for a PP when the verb strictly subcategorizes for a PP. The verb put is such a verb:

*(52) John put it.

(53) John put it on the table.

(54) Where did John put it?

Also, where, like PPs, can trigger inversion:

(55) Where do you believe were found the remains of Judge Crater?

(56) In Boston were found the remains of Judge Crater.

On the other hand, where acts as an NP in being able

to occur as the head of a relative clause:

(57) Where has he ever lived that wasn't a total dump?

As Grimshaw (1977) shows, PPs cannot occur as the heads of relative clauses.

For our purposes, then, in attempting to answer questions about the internal structure of prepositional phrases, Grimshaw's analysis makes an interesting predictions. It predicts that where will be able to function as the entire complement of prepositions which have Jackendoff's structure, but will not be able to replace the sequence after the first preposition in Hendrick's structure. Thus, given that we are claiming that from in (50) is a preposition which subcategorizes for a PP complement, we should be able to substitute where for the sequence that follows it within the PP and wh-move the where. Sentence (58) exhibits all these characteristics:

(58) Where did a rabbit appear from?

On the other hand, since we are claiming that down in (40) is in the specifier slot, we would not expect to be able to replace the sequence which follows it by where and wh-move the where, since we cannot strand specifiers of prepositions. For example, right cannot be stranded in sentence (61):

(59) He ran right up the stairs.

(60) Right up which stairs did he run?

*(61) Up which stairs did he run right?

Therefore, the ungrammaticality of (61) would, under this account, be a consequence of the formulation of wh-movement, which moves the maximal projection of a category:

*(62) Into which hole did Sam disappear down?

There is, however, one variable which has not been controlled for. Not all prepositions co-occur with all specifiers, so that exactly, for example, can co-occur with at, but not with:

(63) He arrived exactly at six.

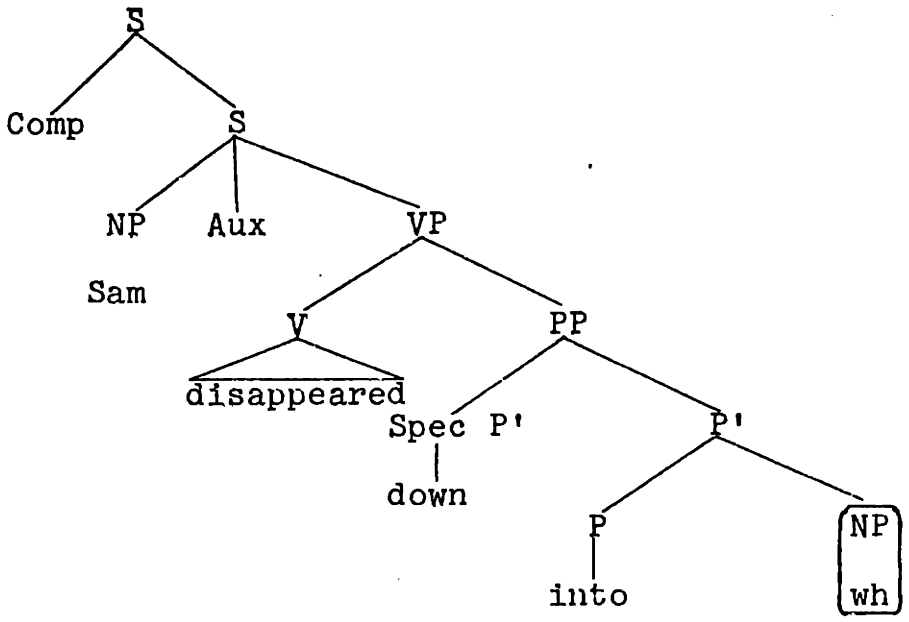
*(64) He arrived exactly with Marcia.

Now, one might object that down cannot co-occur with where in any event, so that even a multiple wh-question in which the where is left behind would be ruled out. However, this is false, since (65) sounds acceptable:

(65) Who disappeared down where?

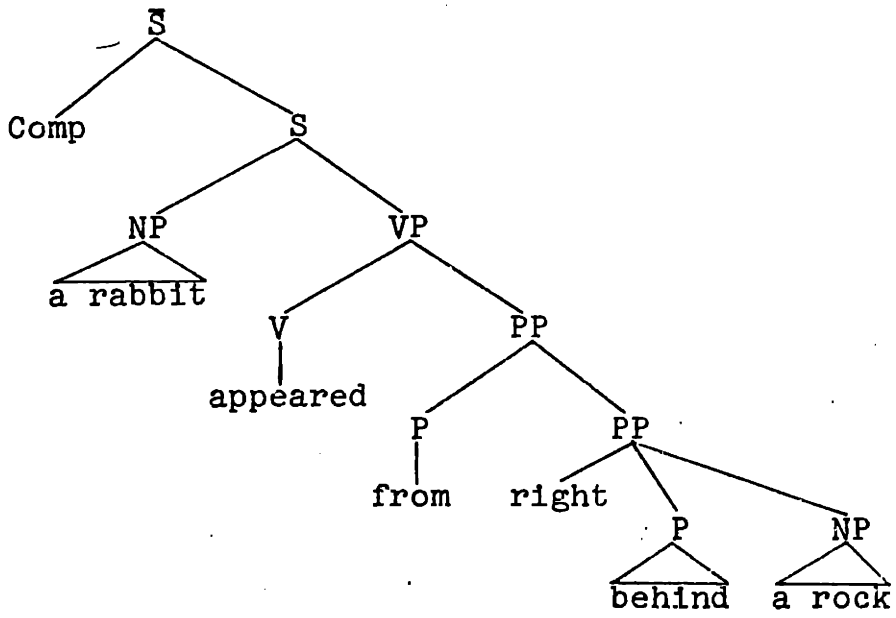
Getting back to the main topic, I would propose that Hendrick's structure is the appropriate one for cases in which the object of the second preposition can undergo wh-movement. For example, the structure of (42) prior to wh-movement would by (66):

(66)



Being neutral for the moment about whether or not S or $\bar{S}$ is the bounding node in English, extraction of the circled NP is only crossing over one PP node, and so wh-movement is possible. Thus, extraction of the prepositional object in Hendrick's structure is consistent with PP as a bounding node. On the other hand, let us consider the pre-movement structure for the sentence corresponding to (50):

(67)



Given that wh-movement would move the wh over the two PPs, if PP were a bounding node, one would predict that wh-extraction from this structure would lead to ungrammaticality, and the prediction is confirmed:

*(68) What did a rabbit appear from bihind?

(69) From behind what did a rabbit appear?

Other instances of the structure in (42) can be found. For instance, consider the following example:

(70) He dragged the thing to behind the house.

The intensifier right can occur between the first and second prepositions in (70), which is one of our main diagnostics for the nested structure:

(72) He dragged the thing to right behind the house.

Therefore, wh-extraction of the object of the second preposition in (72) should be impossible. The results support this prediction:

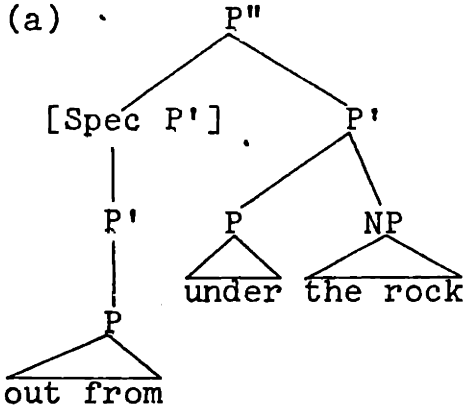
*(73) What did he drag the thing to behind?

A curious problem emerges on further consideration, however, of the ungrammaticality of (68). Joan Bresnan and C.L. Baker (personal communications) have independently pointed out to me that while (68) is deviant for them, extraction is possible when from is preceded by out:

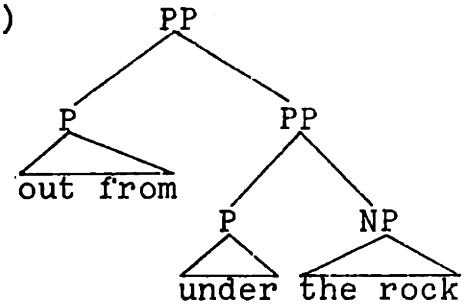
(74) What did a rabbit appear out from under?

There seems to be dialectal variation here, because other speakers whom I have asked do not find (74) acceptable, although some do. For those who do, however, it may be plausible to say that out from is a sort of compound preposition which has two subcategorizations; it can occur in either Hendrick's structure or in Jackendoff's structure. The two structures would be the following:

(75) (a)



(b)



As support for this structural ambiguity, note that right can occur between out from and under, but when it does, wh-extraction of the object of under is no longer possible:

(76) A rabbit appeared right out from under the rock.

(77) A rabbit appeared out from right under the rock.

(78) What did a rabbit appear right out from under?

*(79) What did a rabbit appear out from right under?

There is some evidence for the claim that out from is a compound when one considers the fact that right cannot intervene between out and from, although it can, in other cases, modify from;

*(80) A rabbit appeared out right from under the rock.

(81) He came right from work.

Therefore, I claim that out from is really one word and is a preposition which has two subcategorizations for these speakers; when it can precede right, it must have Jackendoff's structure, the one shown in (74b), and when the object of the preposition which follows it can be extracted by wh-movement, it has the structure shown in (74a). Therefore, when right follows out from, the object of the following preposition cannot undergo wh-movement, since the structure of the PP, by our criteria, must be as in (74b) in such a case, and Subjacency would block extraction. Therefore, (79) is ungrammatical.

Interestingly, nothing said so far would prevent wh-movement of the embedded prepositional phrase, so that (82) and (83) should be grammatical:

*(82) Behind what did a rabbit appear from?

*(83) Behind which house did he drag it to?

I think that the explanation for the inability of PPs immediately dominated by other PPs to undergo wh-movement by themselves may shed some light on the proper formulation of the A-over-A principle, first discussed in Chomsky (1962). Sag (1976) suggests the following formulation of A-over-A:

(84) Immediate Domination Principle (IDP)

Given 2 proper analyses, PA1 and PA2, of a sentence S with respect to a transformation T

(84) (cont.)

which differ minimally with respect to the value of some predicate P in the structural description of T, if

(1) /P/ PA1 (the value of P under PA1) is some terminal substring t_1,

and (2) /P/ PA2 is some terminal substring t_2,

and (3) t_1 is analyzable as A1 ($A1 \varepsilon V_{NT}$),

and (4) t_2 is analyzable as A2 ($A2 \varepsilon V_{NT}$),

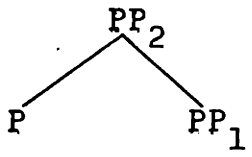
and (5) A1 immediately dominates A2,

then PA2 is an inadmissable proper analysis of S with respect to T.

[Sag's 3.3.15]

The applicability of Sag's principle to (82) should be clear, but I will elaborate somewhat. In the structure (41),

(41)



given that a rule which operated on PPs would have two proper analyses, it would have to choose PP_2 as the relevant target predicate.

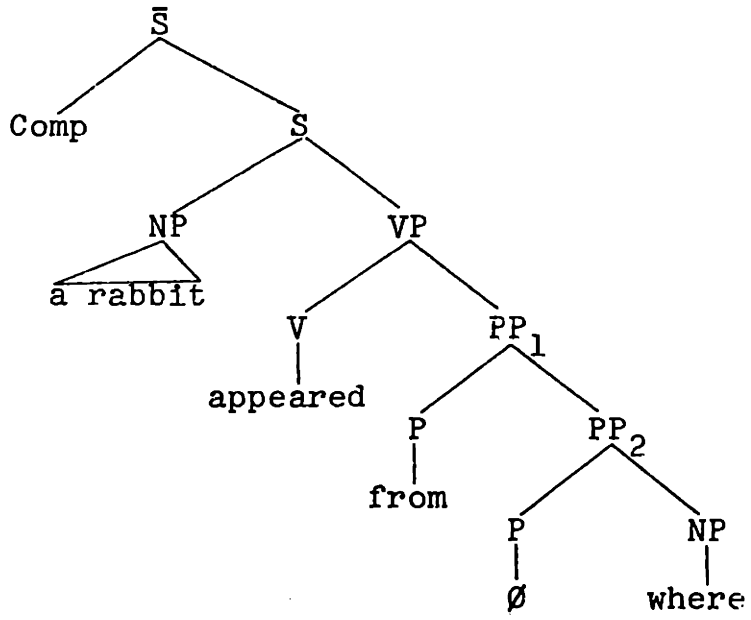
It would be rather instructive, at this point, to compare the unacceptability of (82) with the acceptability of (58), repeated here for convenience:

(58) Where did a rabbit appear from?

Superficially, it would appear that the acceptability of (58) would be a counterexample to Sag's Immediate Domination Principle, if we assume that the fronted constituent

in (58) is a PP. However, there is another possible derivation of (59), in which the fronted constituent is an NP. Thus, consider that the input structure for wh-movement in (58) is (85):

(85)



The claim here is that the NP which is the object of the null preposition is the constituent which undergoes wh-movement. Therefore, the Immediate Domination Principle is not really violated. However, Subjacency would be violated here, if we define Subjacency by the simple counting of nodes, assuming that PP is a bounding node. It appears, however, that simple arithmetic procedures are not enough for the proper formulation of Subjacency in any event. Chomsky (1973) defines Subjacency as follows:

(86) (a) Category "L-contains" category B if and only if A properly contains B and for all C≠A, if A contains C and C contains B, then A=...C..., where '...' contains a lexical item.

(86) (b) B is "subjacent" to A if and only if A is superior to B and there is at most one cyclic category C such that C L-contains MMC(B) [minimal major category; parentheses mine: MRB], and C does not contain A.

[Chomsky's 108]

By this definition of L-contains, the bottom NP where would be subjacent to PP_1, since PP_2 would not L-contain the NP. Therefore, extraction of where in (58) would not violate either Subjacency or the Immediate Domination Principle.

This account of the possibilities of extraction of where by wh-movement has interesting implications for the analysis of topicalization. Chomsky (1977) analyzes a variety of processes which superficially operate unboundedly as covert wh-movement. One of these processes is topicalization, whose unbounded character is exemplified in (87):

(87) Max I promised to Mary that I would visit.

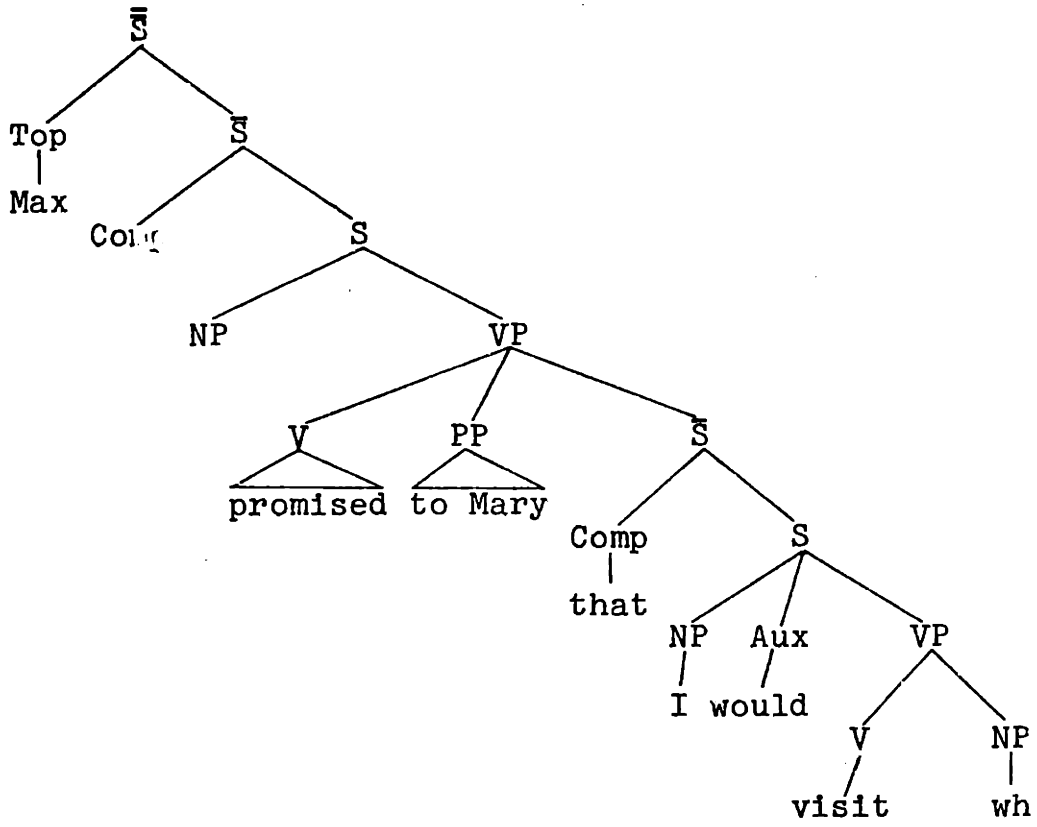
Chomsky postulates the following base rules:

(88) (a) $\bar{\bar{S}} \longrightarrow$ Top $\bar{S}$

(b) $\bar{S} \longrightarrow$ Comp $\begin{Bmatrix} \bar{\bar{S}} \\ S \end{Bmatrix}$

The derivation of (76), then, would be as follows: the underlying structure would be (89);

(89)



Now, the wh would move from Comp to Comp by normal wh-movement to the Comp adjacent to the head, where it would delete. While most discussions of topicalization have centered on topicalization of NPs, it is clear that PPs can undergo this process as well. Consider the following sentence:

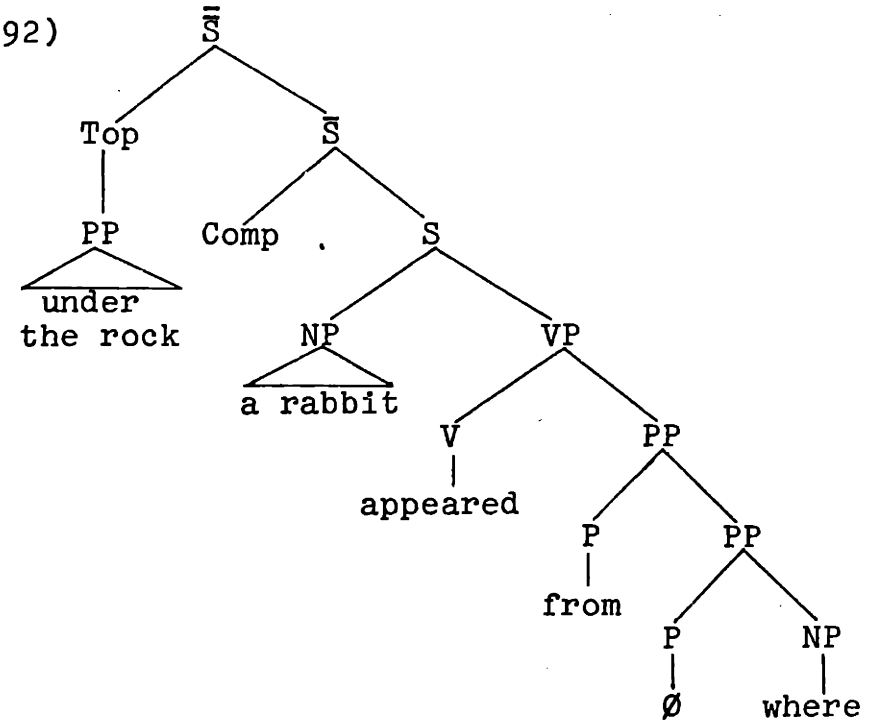
(90) On the table I pretended that I put my wallet.

Since it obeys all of the diagnostics for wh-movement, in being superficially unbounded and being able to leave a gap in a tensed clause, etc., it would be plausible to analyze topicalization of PPs as being covert wh-movement. Now, our previous discussion of where becomes relevant. Consider the sentence (91):

*(91) Under the rock a rabbit appeared from.

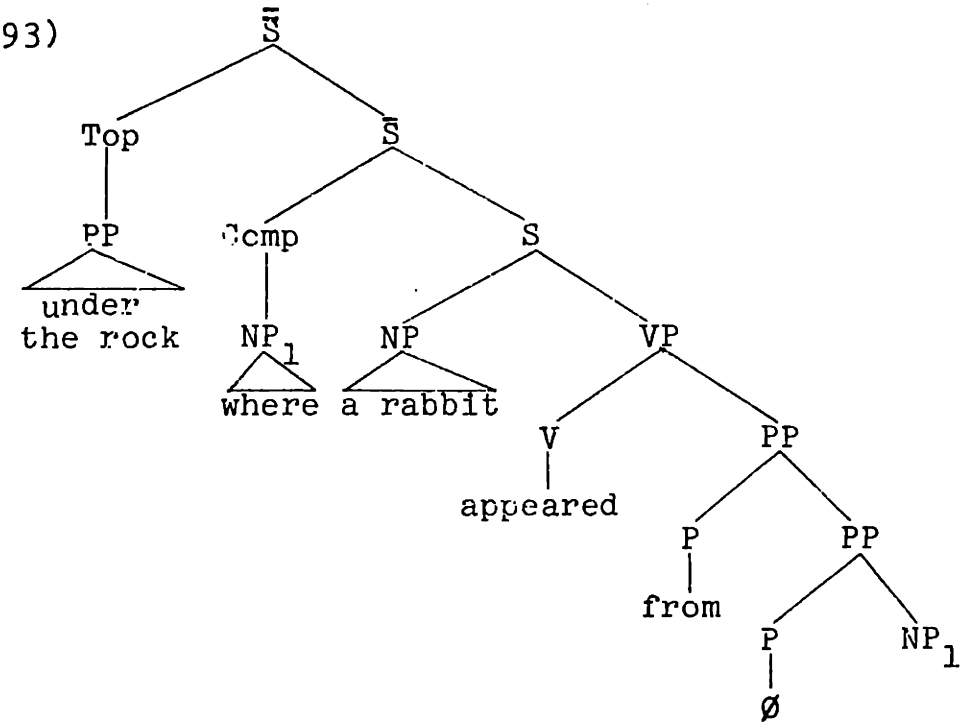
Sentence (91) is ungrammatical, but if topicalization (of PPs) were wh-movement, there would be nothing to prevent the following derivation: The underlying structure would be (92):

(92)



and, by wh-movement, <u>where</u> would move into the Comp adjacent to the Topic:[1]

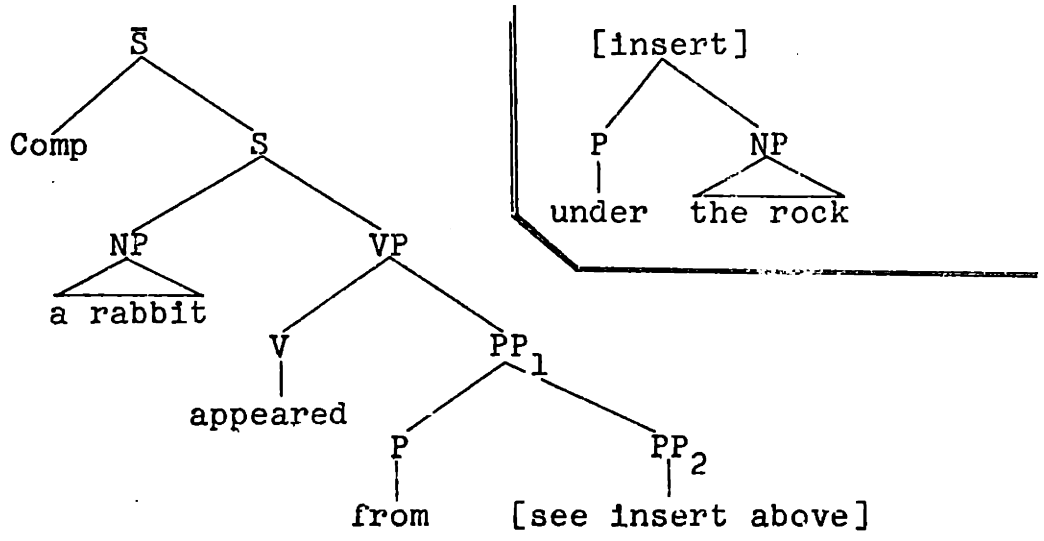
(93)



and then, delete in that position.

Given that we wish to avoid this derivation, I would like to propose an alternative analysis of topicalized PPs. Suppose we were to say that topicalization is movement of the topicalized constituent itself, along the lines suggested in Ross (1967). In this case, the underlying structure of (91) would be (94):

(94)



If topicalization were formulated as follows:

(95) Move $\begin{bmatrix} \bar{\bar{X}} \\ -V \end{bmatrix}$ to Left Bracket S

(More on this way of formulating movement rules will be said in Chapter 4, where I will try to motivate what I call a "Landing Site theory".) Then the Immediate Domination Principle would never allow one to move PP_2. In other words, direct movement of the Topic would allow one to call the Immediate Domination Principle into play. In Chapter 4, I will return to this analysis of topicalization.

This concludes the arguments from English for PP as a bounding node. There are some interesting facts about French which the postulation of PP as a bounding node would help explain. First of all, as Richard Kayne (personal communication) has pointed out, Obenauer (1976) has noted that it is possible to extract a QP by wh-movement from an NP, as in the following case:

(96) Combien avez-vous vu de femmes?
'How many have you seen of women?'

Crucially, it is impossible to extract combien when it originates within the object of a preposition, as in the following sentence:

*(97) Combien avez-vous dansé avec __ de femmes?[2]
'How many have you danced with __ of women?'

Another argument for PP as a bounding node can be

made from the fact that although one can get floated quantifiers associated with bare NP wh-forms, as in (98),

(98) Les livres que j'ai tous lu...
'The books that I have all read...'

one cannot get floated quantifiers with moved PPs, as in (99):

*(99) Les femmes avec qui j'ai toutes dansé...
'The women with whom I have all danced...'

Incidentally, this fact also supports the idea that Q-Float obeys Subjacency.

Another argument from French that PP is a bounding node comes from the fact that although one can cliticize the partitive en, as in (100),

(100) J'en ai vu beaucoup.
'I've seen a lot of them.'

one cannot perform the cliticization when the partitive is contained within the object of a preposition (see Kayne 1975, who draws different conclusions from this fact):

*(101) J'en ai dansé avec beaucoup.
'I've danced with a lot of them.'

Again, PP as a bounding node for French would explain this cluster of properties, since the partitive would be crossing over both a PP and an NP.

The final argument for the bounding status of PP concerns the ability to wh-move NP complements, as in (102):

(102) La solution de la problème a été trouvé.
'The solution to the problem has been found.'

(103) La problème dont la solution a été trouvé.
'The problem to which the solution has been found.'

As can be guessed by now, the ability to wh-move an NP complement disappears when the NP complement is embedded within the object of a preposition. Thus, although (104) is grammatical in French,

(104) Je parlerai avec la sœur de l'homme.
'I will speak with the sister of the man.'

(105) is not:

*(105) L'homme dont je parlerai avec la sœur.
'The man of whom I'll speak with the sister.'

Thus, there are a number of unrelated facts in French which can successfully be explained by the postulation of PP as a bounding node.

In Van Riemsdijk (1977), a number of extremely interesting arguments for the bounding status of PP are presented, with data from Dutch and English. Van Riemsdijk's proposals differ from the one presented here in a number of ways, however. For one thing, Van Riemsdijk claims that the stranded preposition in pseudo-passives, as in

(106) This house has been lived in by many people.

is only apparently stranded, but that the sequence

verb+preposition, when contiguous and comprising a semantic unit, undergoes "reanalysis" such that the unit has a bracketing in which they are both dominated by V. When the verb+adjacent preposition do not form a semantic unit, this bracketing is not permitted, and we do not get pseudo-passives. Therefore, we observe the following minimal contrast (from Bresnan 1977):

(107) This solution was arrived at yesterday.

*(108) This house was arrived at yesterday.

Although this process is not directly germane to my main concerns, while I agree that the verb+preposition sequence is dominated by a single lexical node, I would like to suggest that the way this single domination can come about is due to the operation of a word-formation rule. For example, there seems to be a correlation between pseudo-passives and the ability of these verb+preposition sequences to occur pre-nominally:

(109) She was talked about.

(110) She's the most talked-about person I know.

(111) This house was lived in for many years.

(112) This house has a lived-in look.

(113) This solution was arrived at independently.

(114) The independently arrived-at solution was the subject of much controversy.

Notice that there would be no way of maintaining that the underlined sequences in (110), (112), and (114) have

gotten an extra bracketing after lexical insertion, and thus become a unit, since the base rules would certainly not generate sequences like [$_{NP}$ Det-V-Prep-N $_{NP}$]. However, as Bresnan (1973) has shown, there are base rules like (115):

(115) NP⟶ Det Adj N

(See Bresnan 1973 for arguments against deriving pre-nominal adjectives from underlying relative clauses.) Given this, it would seem that the sequences in (110-114, even) become units due to the operation of a morphological compounding rule operating in the lexicon. If this correlation between pseudo-passives and the ability to occur pre-nominally is general, then pseudo-passives would fail to offer any support for reanalysis. This is not to deny the existence of reanalysis (indeed, I shall consider below a slightly different conception of reanalysis recently put forth by Amy Weinberg), but rather to show that pseudo-passives do not demonstrate it.

However, this view of pseudo-passives does not alter Van Riemsdijk's main point, with which I am in agreement. We both agree that normally one cannot strand the object of a preposition, even by a rule such as NP-preposing, which is certainly clause-internal, so that Subjacency cannot be the cause of the prohibition.

There are two alternatives at this point, if one assumes that there is a rule of NP-preposing. One is to allow context predicates in the statement of this rule (see

Bresnan 1976 for a precise definition of this term). The rule of NP-preposing would, then, be formulated thusly:

(116)[3] np - V - NP - Y
1 2 3 4 ⟶ 3 2 Ø 4

However, Van Riemsdijk eschews the use of context predicates, reasonably (in my opinion) wanting to explore the consequences of a framework in which individual rules are permitted a minimum of descriptive power. This is a valid methodology, as long as one actually succeeds in explaining away restrictions on individual rules by giving independently motivated explanations. Therefore, Van Riemsdijk invokes an extremely interesting constraint on movement rules, originally proposed by Fiengo (1974) but modified by Williams (unpublished) and Zwarts (1975), which he dubs the Head Constraint. This constraint, taken from Van Riemsdijk (1977), is formulated as follows:

(117) The Head Constraint:

No rule may involve X_i/X_j and Y_i/Y_j in the structure

$$\ldots X_i \ldots [_{H^n} \ldots [_{H'} \ldots Y_i \ldots H \ldots Y_j \ldots {}_{H'}] \ldots {}_{H^n}] \ldots X_j \ldots$$

where H is the phonologically specified head and H^n is the maximal projection of H.

[Van Riemsdijk 1977, p. 115, no. (224)]

In effect, what this constraint says is that direct extraction of an element out of a phrase is impossible if

that element originates in the lowest branching level of that phrase. There are some observations which I feel are worth making about this constraint. First of all, notice that it overlaps with Subjacency in a number of crucial cases, even to the extent of stipulating an escape hatch (i.e. any position above the H' level but still within H^n). Therefore, a theory which postulated two such similar bounding constraints which overlap in so many cases would be quite inelegant.

Also, the strong claim that it makes, a claim which is stronger than Subjacency, is essentially that all nodes are bounding nodes. Therefore, it would be interesting to look at other phrasal nodes to see if they obey the Head Constraint. First of all, consider extraction of any direct object, either by passive or by wh-movement:

(118) John was given a book.

(119) What would you like?

If we assume a VP node, the assumption that VP is the maximal projection for verbs would force us to conclude that (118-9) would be ungrammatical. Of course, Van Riemsdijk adopts Jackendoff's (1974, 1977) proposal that actually, what has been labelled S is really the maximal projection of V. However, Hornstein (1976) has levelled strong objections to Jackendoff's proposals, to the effect that actually S functions quite differently from the other phrasal categories. However, if there were other strong arguments

for the Head Constraint, one might draw the conclusion that the Head Constraint and Jackendoff's proposal actually reinforce each other. Therefore, I will hold this objection to the Head Constraint in abeyance temporarily, while I consider adjectival phrases, which must be crucially distinguished from verb phrases in any event. For example, let us assume that the rule often called subject-to-subject raising (Postal 1974) is actually a variant of the rule of NP-preposing, and that there is a rule, AP$\longrightarrow$ A $\bar{S}$. This rule is justified by the fact that adjectives plus infinitives act as constituents in wh-fronting, as in:

(120) Just how eager to please was Fred?

Since there does not seem to be any independent motivation for any structure intervening between the adjective and th- infinitive, the base rule AP$\longrightarrow$ A $\bar{S}$ seems like a reasonable bet. If one believes in NP-preposing and the Head Constraint and the proposed base rule, one would be hard-pressed to explain why it is possible for the subjects of infinitival complements of adjectival phrases to undergo NP-preposing, as in :

(121) He was likely to win.

The adjective <u>likely</u> can be shown to be a raising predicate by the fact that it allows There-Insertion in its complement, as in (122):

(122) There was likely to be a riot.

vs. eager, which does not:

*(123) There was eager to be a riot.

Of course, the straw holder of the three assumptions mentioned in the above paragraph has a possible move. That person could claim that the infinitival complement of likely was extraposed out of the adjective phrase before its subjectunderwent NP-preposing. As evidence for the proposed derivation, one might point to the fact that the infinitive can be left behind when the adjective phrase undergoes wh-movement:

(124) Just how likely was he to win?

However, this would predict that the complement of likely would have to be left behind if its subject underwent NP-preposing. This is clearly false:

(125) Just how likely to win was he?

Another question related to adjectival phrases vis-a-vis the Head Constraint suggests itself. Up till now, nothing has been said about why preposition stranding is permitted at all with wh-movement of complements of adjectives, much less with any PPs at all which are not contiguous to a verb. For example, sentences (126) and (127) are fine:

(126) Who are you happy with?

(127) Who did you talk to Fred about?

Reanalysis in Van Riemsdijk's sense would not be a possible option in (126) since reanalysis is restricted to verbs, and would be inapplicable in (127) since the preposition is not contiguous to the verb. Since reanalysis would be inapplicable in this case, the Head Constraint would predict, contrary to fact, that sentences (126) and (127) would be ungrammatical, since we would be extracting the object of a preposition out of the lowest branching level to a position past the maximal projection of the phrase (i.e., PP).

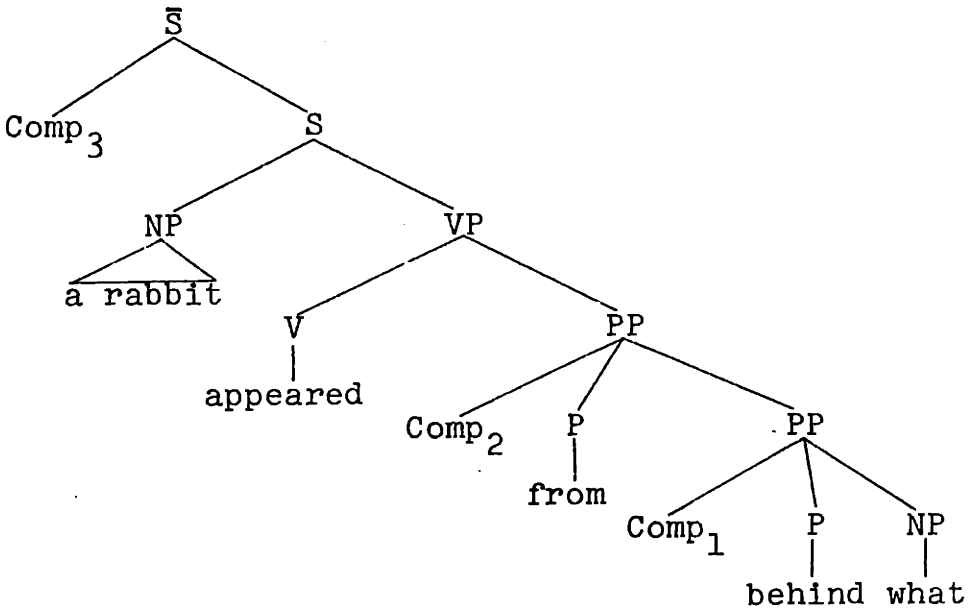
Van Riemsdijk's solution to this problem is extremely ingenious. He proposes that PPs in English have Comps. Noting (probably correctly) that preposition stranding is a marked phenomenon, and that the only real stranding which occurs in English happens as a result of a Comp-substitution rule (wh-movement), the proposal that PPs have Comps as escape hatches, the same as $\bar{S}$, would (so the reasoning goes) account for this.

However, interesting though this proposal may be, I fear that it would overgenerate wildly. Many of the arguments which I have presented for the bounding status of PP would simply not go through if PPs had escape hatches for wh-movement, and the phenomena for which the postulation of Comp-less PP as a bounding node was made would be totally mysterious. For example, sentence (68), repeated here for convenience,

*(68) What did a rabbit appear from behind?

would be generated. On Van Riemsdijk's account, there would be nothing to prevent it from having the following underlying structure:

(128)

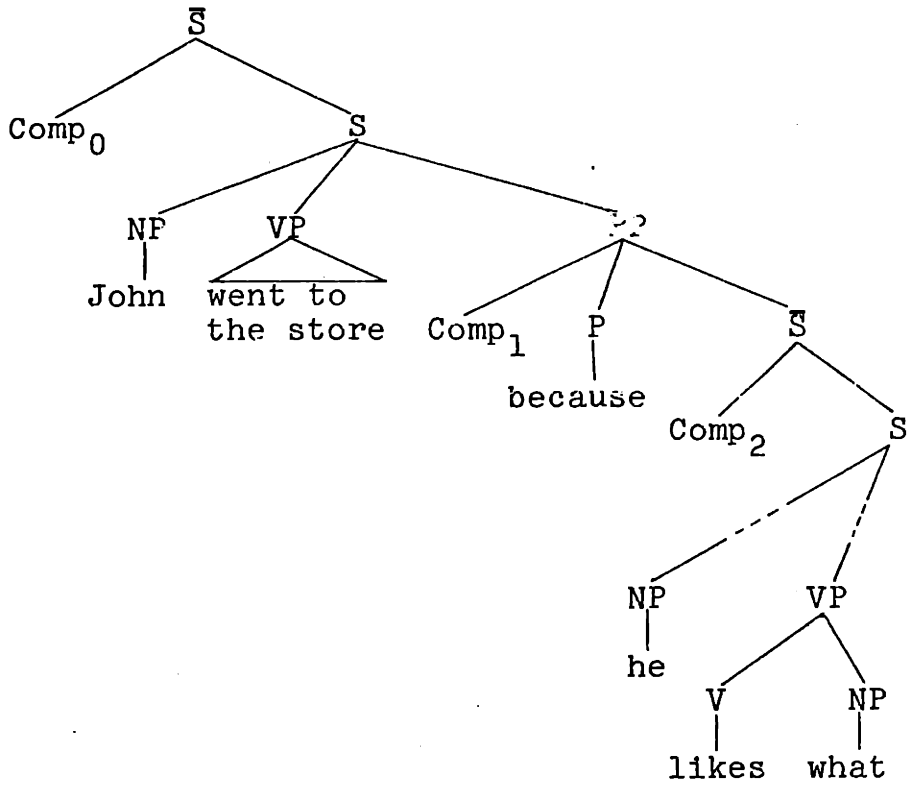


Assuming successive cyclic movement, there would be nothing to prevent the object of <u>behind</u> from moving into $Comp_1$, then $Comp_2$, and finally into $Comp_3$. Similarly, since adverbial subordinate clauses are analyzed as PPs, being generated by rule (34), one would have to ask why sentence (37), repeated here for convenience,

*(37) What did John go to the store because he likes?

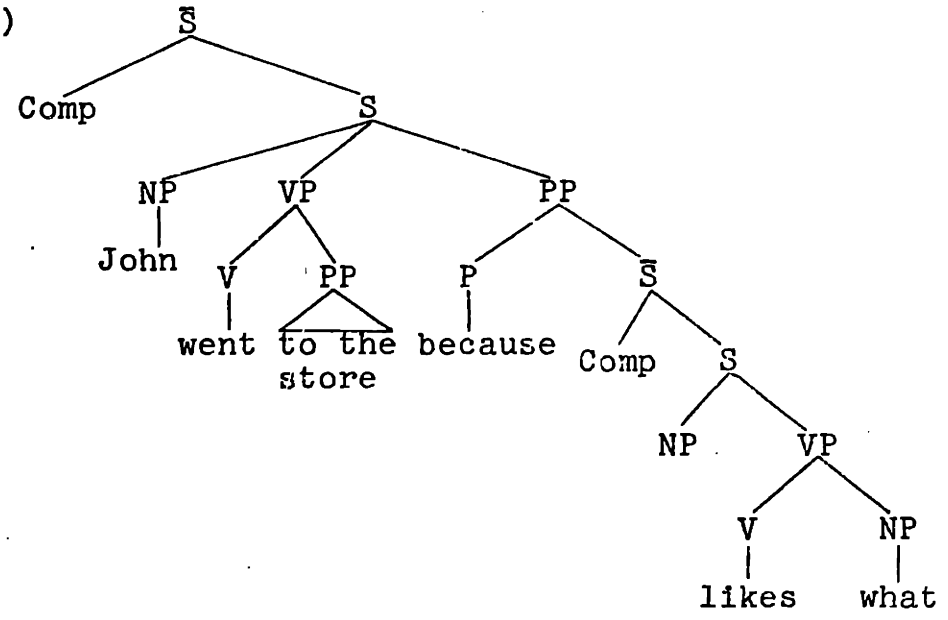
could not be generated if PPs had Comps. The derivation could proceed as follows; the underlying structure would be (129):

(129)



With an underlying structure like (129), the object of <u>like</u> could move, theoretically, from $Comp_2$ to $Comp_1$, and then finally to $Comp_0$. However, if we assume that the underlying structure of (37) is really (130):

(130)



we can account quite neatly for the ungrammaticality of (37), since Comp-less PP and S are both bounding nodes.

A third argument against PPs with Comps comes from a reconsideration of (82), again repeated for convenience:

*(82) Behind what did a rabbit appear from?

As noted, I proposed to explain the ungrammaticality of (82) by appealing to Sag's Immediated Domination Principle. However, if one moves PP_1 into $Comp_2$, PP_1 is no longer immediately dominated by PP_2, and so the Immediate Domination Principle should no longer apply. Therefore, if one assumes this version of A-over-A, the sentence should be grammatical.

A fourth problem with the theory that PPs have Comps was noted by Weinberg (1978), who notes that only prepositions which are dominated by VP can strand, while prepositions which are immediately dominated by S cannot. Therefore, we get the following contrast:

(131) Which solution did you arrive at?

*(132) What time did you arrive at?

(133) I arrived yesterday at six o'clock.

As Weinberg observes, this correlation is totally mysterious within Van Riemsdijk's system. Surely, one will not wish to claim that PPs dominated by VP have Comps, while PPs dominated by S do not. A further consequence of Weinberg's correlation is that it might explain the fact,

noted by Gueron (1978), that one cannot strand the preposition of an extraposed prepositional phrase. Thus (135), formed from (134), is ungrammatical:

(134) A review has just appeared of Smith's latest book.

*(135) Which book has a review just appeared of?

If we assume, as Gueron convincingly argues, that PPs, when extraposed, are dominated by S, we can explain Gueron's observation in a fashion parallel to the explanation for the ungrammaticality of (132).

The thrust of my arguments against the proposal that PPs have Comps is that such a proposal is too weak, in that it says nothing about ungrammaticality which would have to be prevented by independent means. However, the proposal mentioned above, that S and (Comp-less) PP are bounding nodes is, as it stands, too strong, in that it predicts that stranding of prepositions will be impossible in English, although stranding clearly is possible in some cases. Throughout the discussion in the section, I have skirted the issue, by pretending that $\bar{S}$ is the bounding node. However, as we now see, that assumption is untenable (recall also the discussion of the NP Constraint in Section I of this chapter).

Therefore, if we wish to keep both S and PP as bounding nodes for English, we must find some way of making the PP node count as a bounding node in some instances, but not in others. Weinberg (1978) proposes a solution which we shall

accept here. She proposes that reanalysis of prepositions takes place within the verb phrase, but that contiguity of the verb and the preposition is not a necessary condition. In other words, hierarchical structure rather than linearity is the deciding factor for reanalysis. Therefore, given that a preposition in the S is not in the VP, it cannot strand. To rule out stranding in languages like French, in which $\bar{S}$ is the bounding node, she proposes the following filter:

(136) * $[_{OBL}$ e] where e is the empty symbol left by trace.

A word is necessary about the term OBL (short for oblique). Chomsky (1978) proposes that all NPs are universally assigned case, so that subjects of tensed clauses are assigned nominative case, objects are assigned objective case, and objects of prepositions are assigned oblique case. Therefore, (136) states an output condition which bans oblique NPs which are empty. Given that reanalysis would remove the preposition from consideration before case marking applied, the NP which would be the object of a preposition in the VP would not be marked as oblique (although it could be, since the reanalysis would be optional).

While Weinberg does not give an exact formulation of her reanalysis rule, which turns a verb and preposition dominated by VP into a unit, there are some considerations which might be taken into account in stating it more pre-

cisely. For instance, one would not want to reanalyze the top preposition in a nested PP structure in the VP, since one would then be able to front the entire sister PP which would underlyingly follow the top preposition, and (82) would be able to be generated that way. Similarly, if both P_1 and P_2 were reanalyzed, (68) would be able to be generated, since the PP bounding nodes would be gotten out of the way. We would want reanalysis to apply in such a way that sentence (58) is generable, however. Therefore, suppose we flesh in the reanalysis rule in the following way:

(137) In the context $[_{VP}$ V...$[_{PP}$ P NP], reanalyze P provided that its phrasal node is immediately dominated by VP.

The reanalysis rule makes some testable predictions. Thus, although we cannot extrapose relative clauses from inside of fronted prepositional phrases, as witnessed by the ungrammaticality of (32), we should be able to extrapose a relative clause from inside of a prepositional phrase dominated by VP. The sentences in question which exhibit the relevant property would be the following:

(138) I saw it in a magazine yesterday which was lying on the table.

(139) I saw it in a magazine which was lying on the table yesterday.

(140) I arrived at a solution yesterday which I found totally satisfying.

(141) I arrived at a house yesterday which was in a shambles.

These sentences seem acceptable to me, which would be explained by the assumption that reanalysis has occurred.

To sum up this section, I have argued that (Comp-less) PP and S are both bounding nodes, and have accepted Weinberg's proposal that reanalysis occurs when a preposition is stranded within the verb phrase. Since we have established the bounding status of PPs, we will, in the next section, use this conclusion to draw some inferences about the nature of bounding in general.

III. On the Nature of Bounding

With the postulation of PP as a bounding node, we are in a position to head toward an answer to the question which has been the main concern of this chapter; namely, is bounding a condition on logical form, or a condition on rule applicability? First of all, we must consider the fact that semantic rules do not in general obey Subjacency. As Chomsky (1973) has pointed out, if we assume that the rule responsible for the distribution of reciprocals, as in (142):

(142) John and Mary like each other.

is a semantic one, we note that it violates Subjacency. For example, sentence (143) is perfectly grammatical:

(143) John and Mary could hardly believe the claims $[_{PP}$ about $[_{NP}$ each other]].

Therefore, if we considered Subjacency a constraint on construal, and construal part of logical form, we would be claiming that a constraint exists on some semantic rules and not others. However, there are some semantic rules for which a movement analysis has in the past been postulated, and which would therefore seem to be subject to Subjacency, for which Subjacency is actually inapplicable. For example, consider the dependency relationship which exists between the degree word so and consecutive clauses. This dependency relationship is exhibited in the following paradigm:

(144) Mary is so hungry that she'd eat anything.

(145) John is too angry to do anything.

*(146) Mary is so hungry to eat anything.

*(147) Mary is too angry that she would do anything.

*(148) Mary is hungry that she would eat anything.

*(149) John is angry to do anything.

It is obvious that there are co-occurrence relations which exist between the degree word and its complement, such that the degree word so selects for finite complements, and too selects for infinitivals. Also, these complements cannot occur at all except in the presence of these degree words. This dependency has led previous investigators (Selkirk 1970, Bresnan 1973, Baltin 1975) to postulate an underlying structure in which the degree word and its complement form a (QP) constituent, and a movement rule which moves the complement of the degree word to the end of the sen-

tence. However, Liberman (1975) has introduced some considerations which demonstrate that this alternative is incorrect, by showing split control of these degree word complements which occur sentence-finally. For example, while we get sentences like (150),

(150) So many people read so many books so often that it's hard to keep up with them.

we can get at most one result clause at the end of the sentence, while the movement rule analysis would predict the possibility of one result clause per so at the end of the sentence. However, the dependency relationship exemplified by sentences (144-149) must be accounted for in some fashion. Therefore, the natural alternative is to posit a semantic rule of construal which would amalgamate the degree word and its complement at the level of logical form. Now, if we believe that Subjacency is defined on rules of construal, we would predict that we could never construe a result clause in sentence final position with a degree word so if the degree word were a constituent of a noun phrase embedded in a fronted prepositional phrase, since PP is a bounding domain. The facts run counter to this prediction. Thus, Liberman cites the following sentence as grammatical, and I concur:

(151) With so many legions did Caesar advance that resistance was impossible.

It would seem, therefore, that we would want our con-

ception of bounding to distinguish between base-generated construal and cases of discontinuous dependency which plausibly arise as the result of movement. The natural way to make the distinction would be to say that Subjacency is a condition on analyzability, in that it rules out logically possible proper analyses of a string before a syntactic rule applies, rather than a constraint on construal.

Of course, one could also propose that Subjacency is a constraint on trace-binding, or co-indexing an empty node whose contents have been evacuated by a movement or deletion rule (Bresnan & Grimshaw 1978). In this regard, the proposal about semantically conditioning Q-Float which was made in Chapter 2 is of some interest.

In Chapter 2, a proposal was made as to how to explain why the quantifiers which floated were the ones that they were, rather than some arbitrary set. The proposal was that noun phrases which have undergone Q-Float receive two quantifications, one quantification in which the noun phrase which has never had the quantifier present to begin with, and another quantification when the floated quantifier is encountered. If the two quantifications are compatible, the sentence is acceptable; if they clash, the sentence is filtered out on the semantic level. This procedure would explain why (152) is accepbable, but (153) is not:

(152) The men were all eager to work.

*(153) The men were many eager to work.

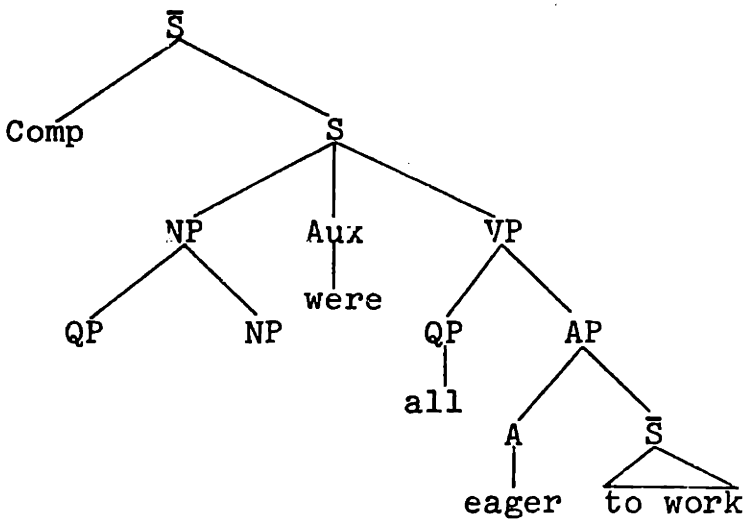
If we assume, following Chomsky (1974) and Milsark (1974), that the definite article receives a semantic treatment in which it is a universal quantifier over a set in the universe of discourse, then the floated quantifier, encountered later in the sentence, will quantify the noun phrase again. The logical form, then, will be a conjunction of the two quantifications. In the case of (152), the two quantifications, both being universal, will be considered an acceptable match, while in the case of (153), one quantification will be a universal over the set in the universe of discourse, while the floated quantifier is not a universal. With this explanation for which quantifiers can float and which quantifiers cannot, we do not have to stipulate in the statement of the movement rule which floats quantifiers, which ones will float and which ones will not. One can simply formulate a rule of Q-Float as in (154):

(154) Move QP to Left Bracket VP

(More will be said about this notation for movement rules in the next chapter.)

However, the procedure of double quantification proposed here has interesting implications for the question of whether non-NP nodes leave traces. Suppose we assume that they do, so that the floated quantifier will leave an empty QP node in the NP when it vacates that noun phrase. Thus, the derived structure of (152) would be (155):

(155)



If (155) were the correct derived structure, the semantic rule which treats the as a universal quantifier would have to ignore the empty QP node in subject position; in other words, the rule would be a string semantic rule. If one were to consider the consequence that traces would be invisible to semantic rules an intolerable consequence, one might conclude that Q-Float does not leave a trace. However, as we saw from the ungrammaticality of (99), repeated here for convenience, Q-Float does obey Subjacency:

*(99) Les femmes avec qui j'ai toutes parlé...
'The women with whom I have all spoken...'

since both PP and NP are bounding nodes. Therefore, if one were to deny that Q-Float left a trace, one would be forced to conclude that Subjacency was not a constraint on trace-binding, since Q-Float does obey Subjacency.

Unfortunately, given our at-present fragmentary knowledge of the semantic component, it would be somewhat rash to conclude that the notion that traces are invisible to

semantic rules is an intolerable consequence. Therefore, I do not feel that any firm conclusions can be drawn from the observation that, in this case, traces would have to be invisible to semantic rules.

There is, however, a somewhat more solid argument that Subjacency must be defined as a condition on analyzability, rather than on trace-binding. As was noted in Chapter 2, floated quantifiers can be construed with heads which have undergone wh-movement, as in (156):

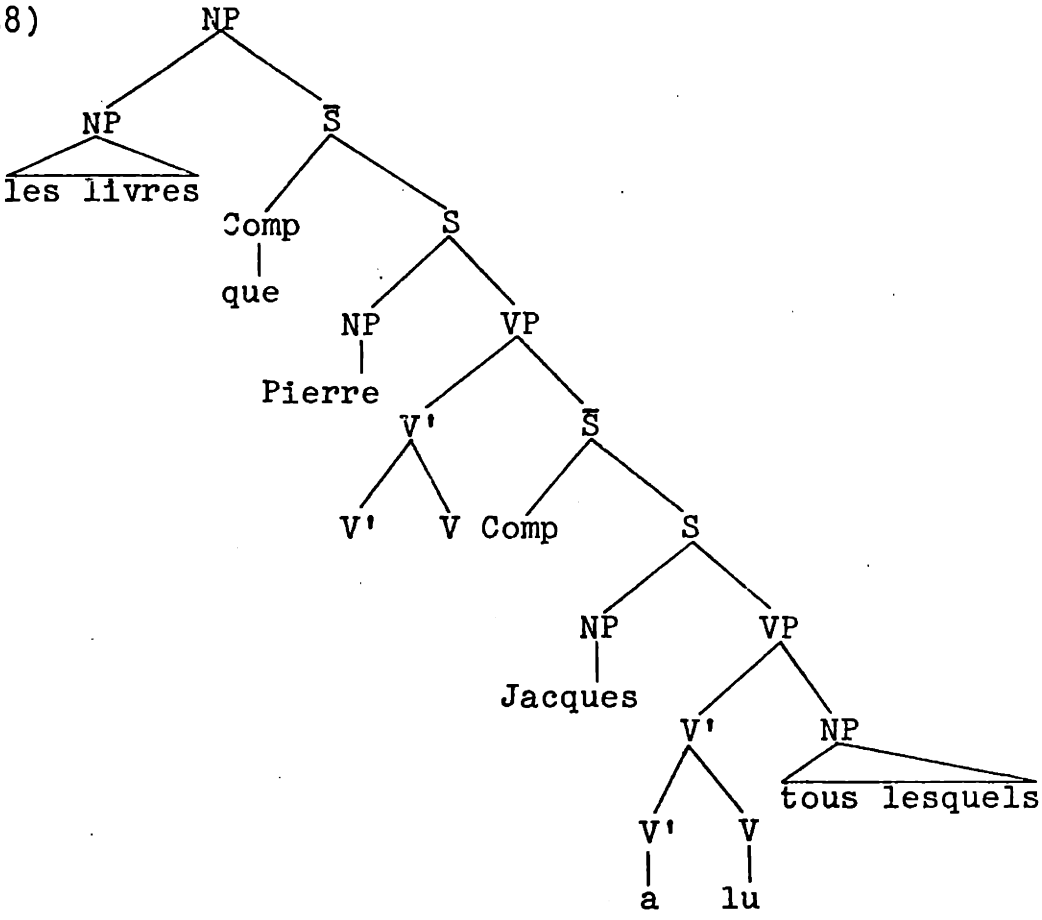
(156) Les livres que j'ai tous lu sont bons.
' the books that I have all read are good.'

Actually, there are two possible accounts of the derivation of (156). On one account, Q-Float could apply before wh-movement, and Q-Float could then apply in Comp, moving the quantifier rightward. When we consider further data, we see that we must choose the first account, in which Q-Float applies before wh-movement. Consider the following ungrammatical string:

*(157) Les livres que Pierre a tous cru que Jacques a lu sont bons.
'The books that Pierre all believed that Jacques read are good.'

In a recent paper, Quicoli (1976) discusses sentences like (157) and argues that they are ungrammatical because Q-Float must obey the Tensed Sentence and Specified Subject Conditions of Chomsky (1973). Therefore, since the source of (157) would be (158),

(158)

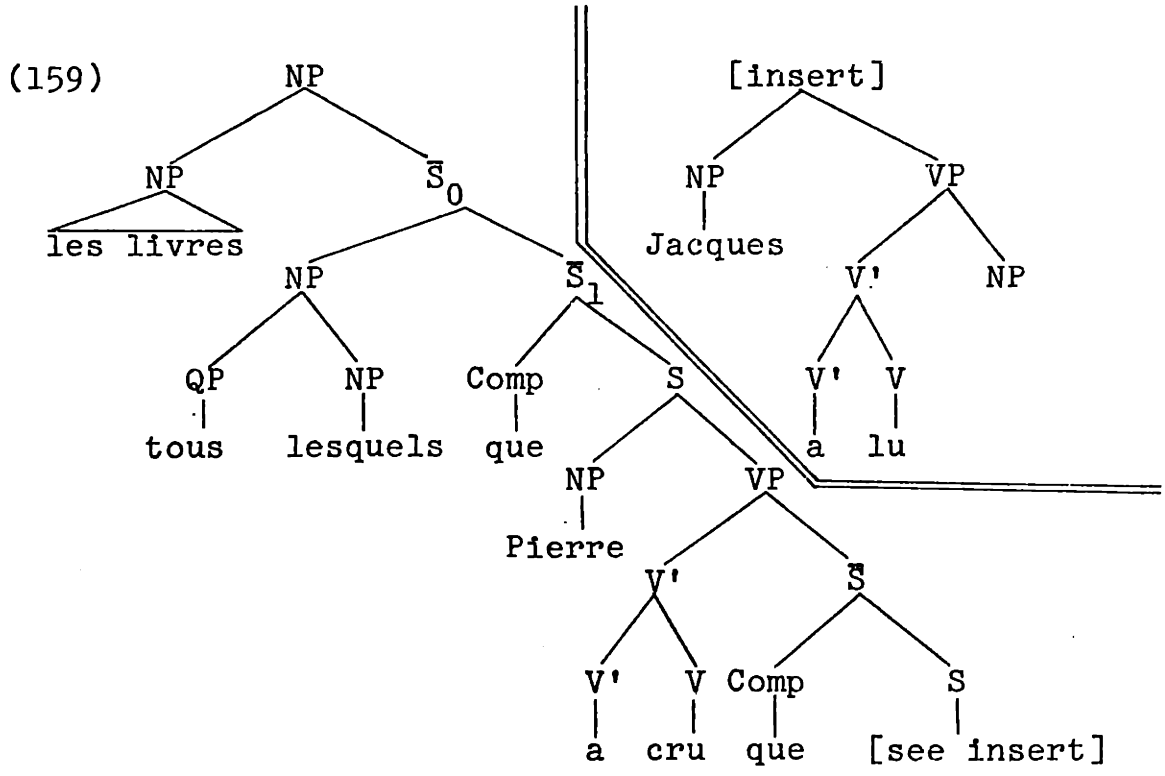


if the tous, which started out in the object of the complement, were to be floated into the matrix, it would violate the Tensed-S and Specified Subject Conditions. However, nothing in Quicoli's paper says anything about moving tous lesquels into $Comp_1$, then $Comp_0$, and then floating tous from $Comp_0$ to a position in S_0. He seems to be presupposing the Comp-to-Comp Condition of Chomsky (1973), which says that once an element moves into Comp, the only position it can move to is another Comp. Therefore, Q-Float must apply before wh-movement.

As things stand, however, the Comp-to-Comp Condition is another stipulation in the theory. I will now attempt to derive the empirical effects of the Comp-to-Comp Con-

dition from Subjacency. I propose that wh-movement is really an adjunction to $\bar{S}$, rather than a Comp-substitution (more will be said about this proposal in the next chapter). Therefore, suppose the input to the ungrammatical (157) would be (159):

(159)

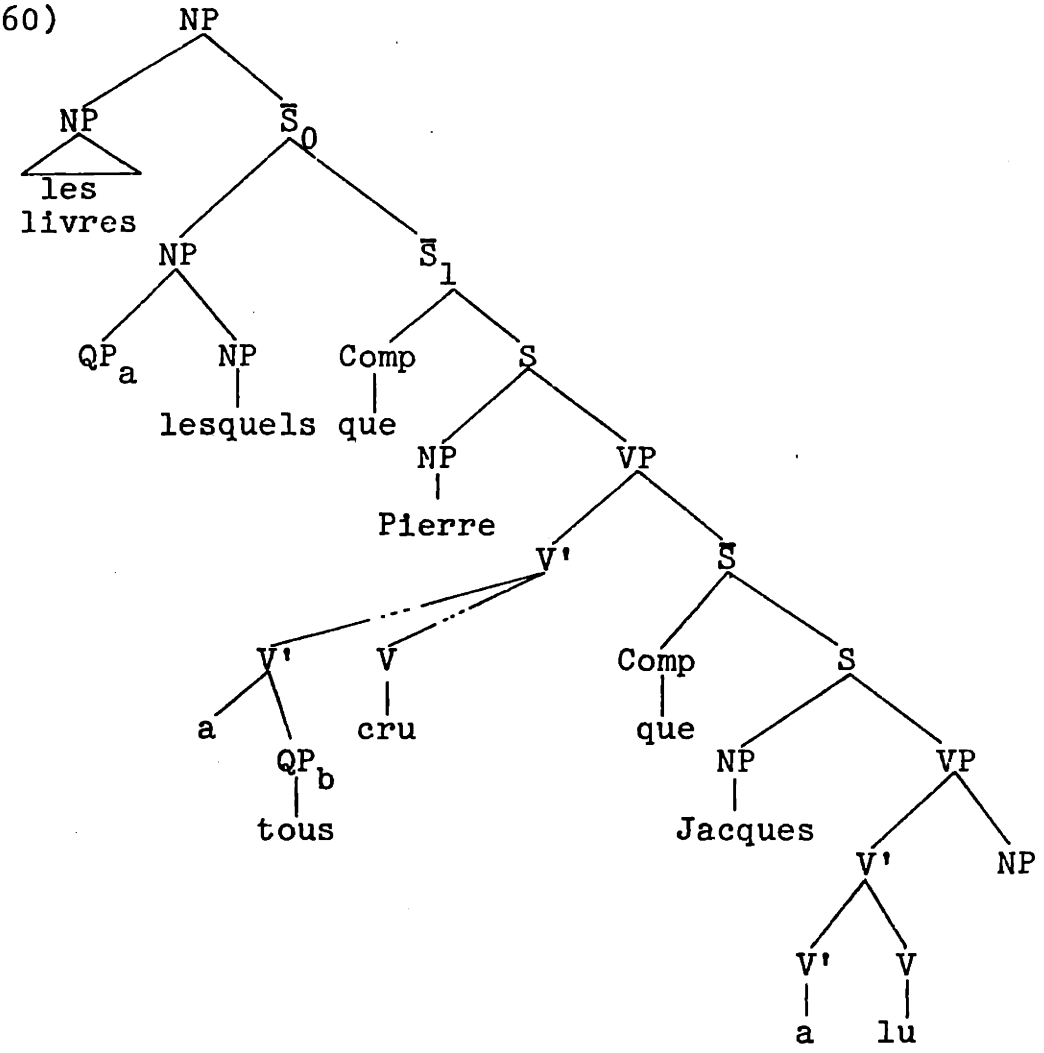


As was mentioned at the beginning of this chapter, $\bar{S}$ and not S seems to be the bounding node for French. With this in mind, since NP is also a bounding node, any movement out of an NP under $\bar{S}_0$, into $\bar{S}_1$ would be violating Subjacency. In this case, then, we can derive the effects of the Comp-to-Comp Condition from Subjacency, an independently needed principle in the theory of grammar.

Let us therefore consider the formulation of Subjacency which is most appropriate here. Suppose we said that Subjacency was a condition of trace-binding, rather than on

the application of movement rules directly. In that case, the structure of (157) would, more precisely, be the following:

(160)

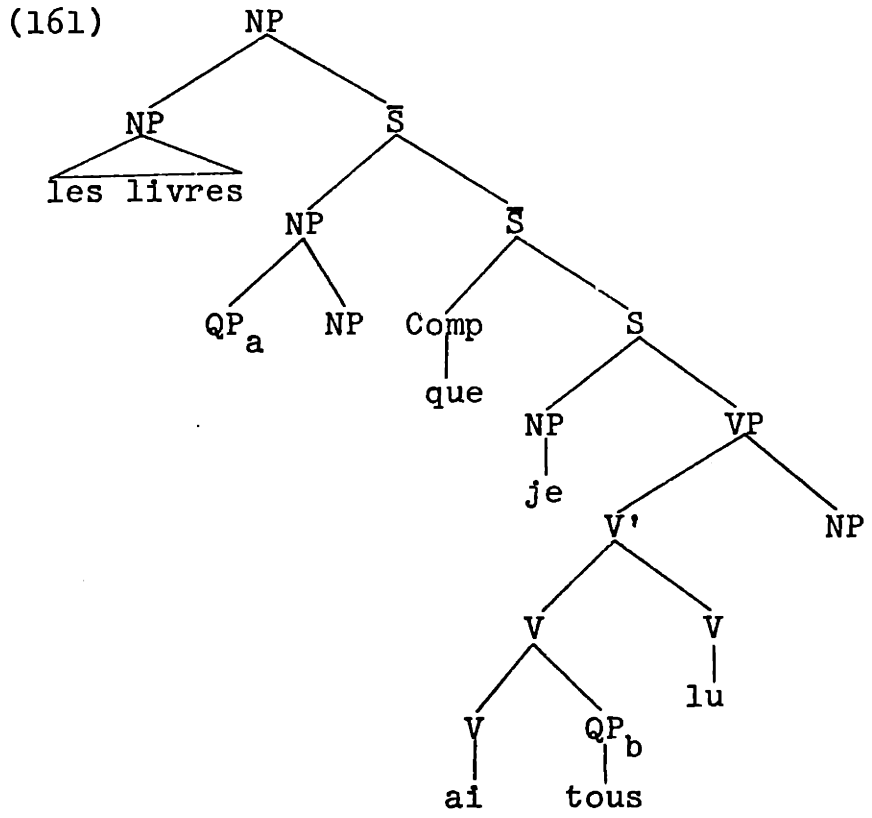


Under this conception of Subjacency, which allows movement rules to apply freely, but constrains the operation of trace-binding, it would be the co-indexing between QP_a, the trace, or empty category left behind, and QP_b, the category which has been moved, since QP_a and QP_b are in non-subjacent domains.

If we view this as the correct conception of Sub-

jacency, however, we must ask why (88) would not be blocked along with (147). Since Subjacency would be a constraint on trace-binding, holding at the level of logical form, the structure of (156) would be (161):

(161)



As we can see, QP_a and QP_b are separated by just as many bounding nodes in (161) as in (160), and so trace-binding should not be able to apply to (161); yet it obviously can, since (151) is grammatical.

On the other hand, Subjacency viewed as a condition on analyzability for syntactic rules seems to fare much better. To derive (151), we apply Q-Float before wh-movement, thus not violating any conditions, either Subjacency, Specified Subject, or Tensed-S. On the other hand, there would be no way to generate (157); Q-Float could not apply

out of the complement of croire, since it would violate Tensed-S and Specified Subject Conditions, and if it applied from $\bar{S}_0$ in (158), it would violate Subjacency. Therefore, there would be no possible derivation for (157).

This case is particularly interesting for deciding on the proper formulation of Subjacency, since we have a case where the structures are identical in the relevant respects; the sentences, with respect to Q-Float, could differ only with respect to the derivations, and yet one sentence is grammatical, and the other ungrammatical.

Furthermore, given the arguments of Chapter 2, we could not claim that the ungrammaticality of (157) is due to a construal mechanism in which the QP in (161) must be construed with NP_b rather than NP_a. If this alternative, which was suggested by Joan Bresnan and Noam Chomsky (personal communications) were viable, the construal operation in (150) would violate the Tensed-S and Specified Subject Conditions. As we saw in Chapter 2, however, construal does not operate in that way. Such a construal mechanism would be totally incompatible with the construal schema developed in that chapter.

Therefore, Subjacency is best viewed as a condition on analyzability. One consequence of this is that Q-Float must be a syntactic movement rule in French, since, as we have seen, the condition which distinguishes (157) from (151) is a condition operating within the syntax.

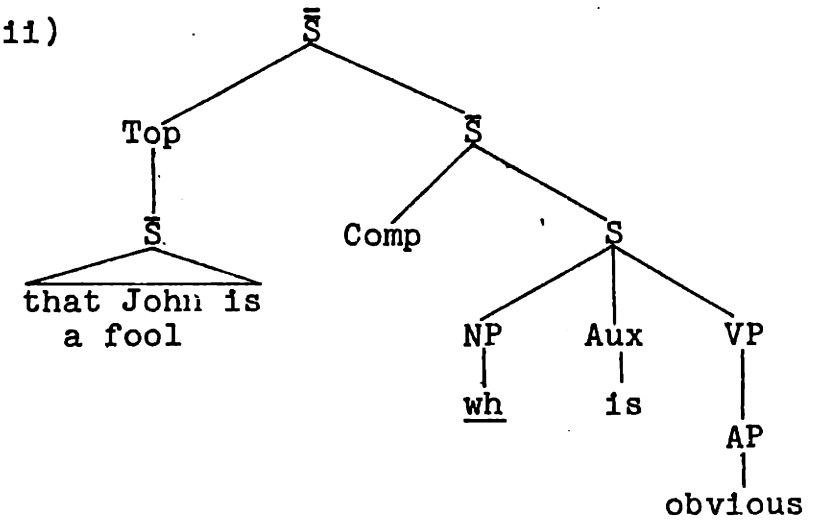
FOOTNOTES TO CHAPTER 3

1. Noam Chomsky (personal communication) has suggested that (91) may be ungrammatical because a PP in topic position cannot cause an NP in Comp to delete under identity, since the two elements are of different categorial types. However, Chomsky accepts Koster's (to appear) analysis of sentential complements, in which a that-clause in initial position, as in (i):

(i) That John is a fool is obvious.

is really in a Topic position, and binds a wh, which deletes in Comp. Thus, he accepts an analysis in which the underlying structure of (i) is (ii):

(ii)



The wh, under this account, which is an NP, moves into Comp, and then deletes. Thus, a proponent of this alternative account of the ungrammaticality of (91) would have to allow deletion of a category type distinct from that of the controller in some cases, but not in others. I regard this as an intolerable consequence.

Furthermore, non-NPs must clearly bind NP wh's in a wide variety of cases, as noted in Jackendoff (1977):

(iii) John came home late last night, which surprised me.

[sentential relative: S binding NP]

(iv) He went from Boston to California, which is a long way to go.
[PP binding NP]

(v) John was pretty happy, which I was too.
[AP binding NP]

2. Dorel (1978), a recent unpublished paper, responds to an earlier version of this material which was presented at NELS VIII (Baltin 1977). Dorel proposes an alternative account of the ungrammaticality of (97), which purports to render moot all of the arguments for PP as a bounding node. However, her solution is unworkable. She proposes the following filter:

(i) * P - t

However, there are three arguments against the use of a filter like (i) to handle the range of data which motivate PP as a bounding node, one of which is based on an argument in Baltin (1977), the paper which Dorel discusses.

For one thing, such a filter would have nothing to say about the ungrammaticality of (33); if a relative clause were to be extraposed out of the object of a preposition, the trace would not be contiguous to the preposition, and so Dorel would predict, contrary to fact, that (33) would be grammatical. To handle the ungrammaticality of (33), she would be forced to reformulate (i) by inserting a variable between the P and the t; most linguists would agree that the use of variables in the formulation of filters would be a non-trivial increase in power, and should hence be avoided. However, as we see in the text of this chapter, PP as a bounding node makes exactly the right predictions.

Secondly, if (91) were to be ruled out by the filter, one would need a mode of application for filter (i) which would be the opposite of the mode of application of other filters. As was noted in Chomsky and Lasnik (1977), filters obey a sort of A-over-A type condition, in which only the dominating category in a string of categories sub-

mits to the filter's inspection.

For example, based on observations in Bresnan (1977), it is clear that the '<u>that</u> $[_{NP}$ e]' filter, filter (68) of Chomsky & Lasnik (1977), discussed in Chapter 2, must be strengthened to the following:

(ii) *that $[_{\overline{\overline{X}}}$ e]

Bresnan observes that PPs which move to the front of a clause, contiguous to the <u>that</u>-complementizer, cannot be subsequently moved by a rule to create a '<u>that</u> $[_{PP}$ e]' configuration. Her examples are the following:

(iii) It was in Boston that I believe __ were found the remains of Judge Crater.

*(iv) It was in Boston that I believe that __ were found the remains of Judge Crater.

Actually, Bresnan draws different conclusions about these facts, due to the difference between the Complementizer Constraint on Variables and filter (68) of Chomsky & Lasnik, but these differences are immaterial to the point under discussion. Now, in French, <u>que</u> is the equivalent of English <u>that</u>. Given the formulation of the filter in (ii), and assuming its universality, one would have to ask why (v), which my informants accept, is grammatical:

(v) Combien croyez-vous que __ d'enfants sont venus?
'How many do you believe that __ of children came?'

The reason (v) escapes the filter (ii), and is hence acceptable, is clear, under Chomsky & Lasnik's proposed mode of application for filters. The trace of <u>combien</u> is properly contained within the NP contiguous to the <u>que</u>, and the filter inspects only the dominating category contiguous to the <u>que</u>. This mode of application for filters makes exactly the right distinctions, allowing (v) to escape its inspection, but not (iv) or (vi):

*(vi) The girl who I thought that __ liked me was here yesterday.

vs.

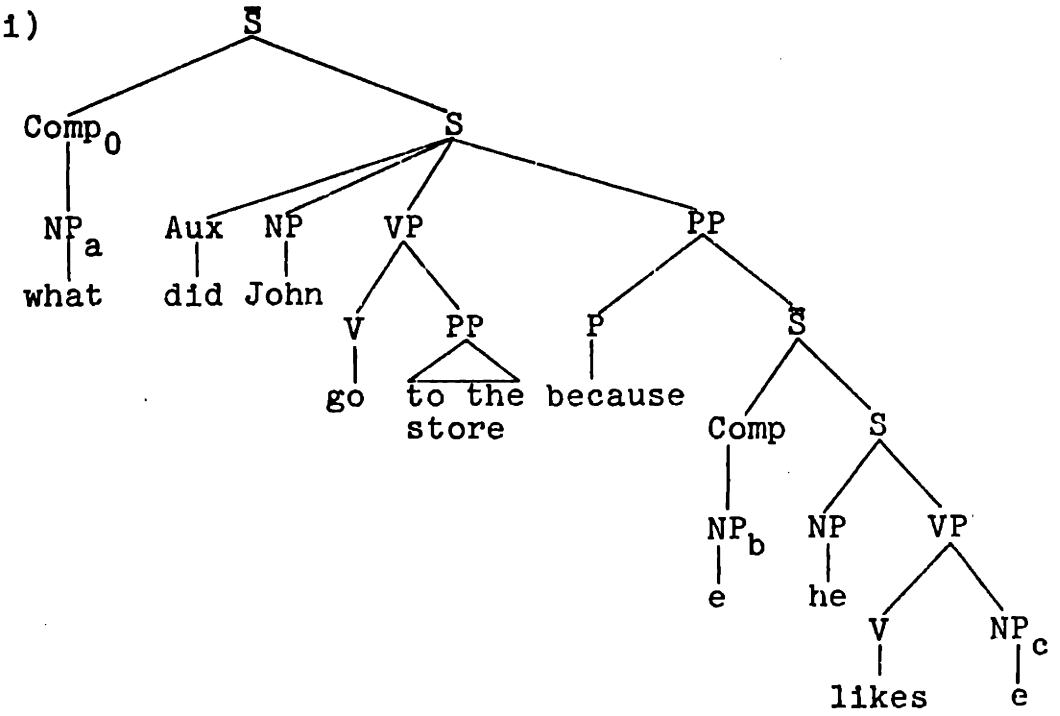
(vii) The girl who I thought __ liked me was here yesterday.

However, if we assume that the mode of application of filters is to check only the highest sequence of categories within a domain, we can see that Dorel's filter could not possible work to rule out (91), since the trace of combien is properly contained within the NP adjacent to the P. Therefore, the trace of combien would not be available for inspection by filter (i), any more than it would for filter (ii).

A third objection to the filter is that it could not rule out extraction out of adverbial subordinate clauses, as in (37). Chomsky & Lasnik (1977), which Dorel accepts, postulate a rule of free deletion in Comp in English, subject to recoverability. (Chomsky (1978) discusses this proposal in more detail.)

Therefore, the trace of what in (37), which would be in $Comp_1$ in (viii):

(viii)



Given that deletion rules apply before filters in Chomsky & Lasnik (1977), and erase the category plus its con-

tents, the free deletion in Comp would be able to apply to NP_b, the trace of <u>what</u>. Thus, (37) would be able to be saved from the clutches of the filter, under this account, and so Dorel's filter would have no way of blocking (37).

For these three reasons, I conclude that Dorel's filter is inadequate by itself; if it is correct at all, it must be supplemented by the postulation of PP as a bounding node.

3. 'np' in this notation, simply means an NP which is unexpanded. It corresponds to what Chomsky (1977a&b) called 'PRO'.

CHAPTER 4: ON THE NOTION "STRUCTURAL CHANGE" IN MOVEMENT RULES

The most prevalent view of transformations today is probably the view sketched in Aspects of the Theory of Syntax (1965) and formalized in Peters and Ritchie (1973). In this view, transformations are of four types: deletions, substitutions, left-adjunctions, and right adjunctions. Restricting ourselves to movement rules, which are the subject of this thesis, we can see that this view of the nature of movement rules is extremely broad; it allows us a great deal of latitude in formulating movement rules. If we restrict ourselves to the adjunction elementaries for the moment, we can see that it allows us to formulate rules like the following:

*(1) QP - X - V - Y
1 2 3 4 ⟶ 2 - 3+1 - 4

*(2) X - Aux - Y - Adv
1 2 3 4 ⟶ 1 - 4+2 - 3 - 0

In other words, the formalization presented in Peters & Ritchie (1973) allows us to adjoin elements to lexical categories like Verb, and non-phrasal categories like Aux. However, as we shall see below, this potentiality is never realized, at least in the grammar of English. Therefore, to the extent that we are permitted this descriptive latitude, our theory predicts a wider class of natural languages than actually exists, and so is too broad.

(To be fair to Peters & Ritchie, their purpose was not to provide a constrained formalism, but rather to formalize the descriptive devices employed in the late 1960's and early 1970's within the framework of generative grammar, so one cannot level this criticism directly at them.) Of course, one could always say that the gaps mentioned above are accidental; the remarks about that move that were made at the beginning of Chapter 2 apply here with equal force.

Within the framework of generative grammar, the first investigator to recognize the nature of this problem, to my knowledge, was Joseph Emonds, in his dissertation (Emonds 1970). Recognizing the methodological wisdom in taking the tack that things are as they are for a reason, and that the positions to which moved elements move are not arbitrary, but rather seem to be selected from a small stock of positions, Emonds proposed the Structure-Preserving Hypothesis. In this theory, movement rules are divided into three types: root rules, that basically (with some qualifications: see Hooper & Thompson (1974), Emonds (1976)) apply only in main clauses; non-root rules, which apply in both main and embedded clauses, but whose output must be a structure independently generable by the phrase structure rules, and local rules, which mention at most two adjacent constituents, one of which must be non-phrasal.

While I agree with Emonds' methodological tack, I would like to propose here an alternative answer to the

question of where moved elements move. If we disregard for the moment the status of the so-called cyclic NP movement rules (i.e., passive, dative, and tough-movement, and raising), which move only NPs, we find an alternative generalization about the remaining movement rules (topicalization, wh-movement, Q-Float, adverb-movement, extraposition of relatives and PPs, and sentential extraposition), all move elements to the periphery of constituents, rather than the center.

If we accept this generalization (arguments will follow below), we must consider a way to incorporate it into the theory of grammar, such that the theory does not allow for logically possible situations to develop in which the generalization is violated. Rather than a formalism which allows rules like (1) and (2) to be expressed, and a stipulation that only allows rules that move elements to the periphery of constituents to be formulated, a more sensible solution would be to devise a notation for movement rules in which the generalization is directly embodied. The proposal I would like to make here is the following; the grammars of natural languages have available to them an inventory of structural positions to which moved elements may move. If we call the inventory of possible positions to which moved elements may move an inventory of <u>landing sites</u>, we can formulate the inventory for English in the following way:

(3) (I) NP (a) Left bracket
(II) VP (b) Right bracket
(III) S
(IV) $\bar{S}$

Instead of mentioning the structural change in the particular movement rule, movement rules will have to select from the inventory. For example, Q-Float, instead of being formulated in the manner in which it is at the beginning of Chapter 2, will be formulated as in (4):

(4) Move QP to IIa.

This notation makes some concrete predictions. For instance, it does not allow one to move elements between a main verb and its object. Therefore, sentences like (5), in which an adverb is niched in that position, are automatically ruled out:

*(5) John ate quickly the steak.

Postal (1974) and Chomsky & Lasnik (1977) have noted this fact but drawn different conclusions about it. They have opted for a filter as in (6):

(6) * V X NP, where X is non-null.

However, there are a number of reasons to reject this filter. For one thing, as Chomsky & Lasnik note, most of their other filters simply refer to constants. Also, in order to allow double-object constructions to pass through

the filter, as in (7):

(7) John gave Mary a book.

one would have to add a proviso to the filter stating "unless X is an NP". However, the filter would then fail to explain why nothing moves between the two NPs in a double-object construction:

*(8) John gave Mary quickly a book.

As we shall see directly, the ungrammaticality of sentences like (8) will fall out of the landing site theory proposed here.

So far, the theory proposed here makes similar predictions to the Structure-Preserving Hypothesis in many cases. The fact that most of the base rules in English create binary branching, in many cases, would lead to situations in which movement rules end up on the periphery, since one way of expressing the Structure-Preserving Hypothesis is to state that the only non-local movement rules that occur in embedded contexts are substitution rules. Therefore, the generalization about periphery is a consequence of other factors for Emonds, where it is applicable, but it is a primitive to the approach to movement rules suggested here.

The way to distinguish the approaches, then, is to see whether there are rules which move elements to the periphery of major constituents, but which do not move them to

positions generable by independently motivated base rules. In this chapter, I will argue that such rules exist. In particular, I will argue that extraposition of relatives not structure-preserving, nor sentential subject extraposition, wh-movement, or topicalization. Furthermore, I shall argue that the positions to which these elements move are structurally distinct.

I. Relative Clause Extraposition

In Emonds (1976), an analysis of relative clause extraposition is postulated in which the extraposed relative moves into the VP or the AP. The rule is structure-preserving, in this analysis, since the base must generate, in any event, S in the VP or AP. The following sentences (Emonds' IV.10) would be instances of this state of affairs:

(9) John was certain that the Mets would win.

(10) Morris whined that night was falling.

(11) A kid shrieked that Baltimore had just gone ahead.

(12) One guest growled to the waiter that an hour was long enough.

(13) Several people mumbled to Harry that he'd better leave.

As Zwicky (1970) has noted, the verbs in (10-13) do not allow passivization:

*(14) That night was falling was whined by Morris.

*(15) That Baltimore had just gone ahead was shrieked by a kid.

*(16) That an hour was long enough was growled to waiter by one guest.

*(17) That he'd better leave was mumbled to Harry by several people.

Therefore, Emonds' analysis, in which a base rule exists of the form,

(18) VP⟶ V $\bar{S}$

is quite plausible. However, there are no base rules of the form in (19):

*(19) VP⟶ V $\bar{S}$ $\bar{S}$

This point is quite crucial to Emonds' account, not only for relative clause extraposition, but for sentential subject extraposition as well, and I will return to this feature of his account in more detail when I discuss that rule. However, Emonds predicts that relative clause extraposition will be blocked when the S position in the VP is filled. Williams (1974) notes that this prediction is falsified, and, in this regard, cites the following sentence:

(20) People claimed that they were sick who weren't sick.

In this case, the complement of claim fills the S, and so, under the Structure-Preserving Hypothsiss, relative clause extraposition should not be able to apply. However,

it obviously can.

In fact, one can find other evidence that extraposed relatives are outside the VP, if one considers VP-deletion a test of verb phrase constituency (Sag 1976, Williams 1977). As it turns out, one can VP-delete and leave an extraposed relative to the right of the deletion site:

(21) Although nobody would ride with Fred who knew just him, people would who knew his brother.

This fact would suggest that extraposed relatives are not dominated by VP, a position argued for at more length in Chapter 2.

II. Extraposition

In this section, I discuss Emonds' (1976) analysis of extraposition. Before proceeding to the analysis, a terminological convention must be made. The term 'extraposition', in this thesis, is reserved for extraposition of sentential complements alone. Therefore, the term 'extraposition', when it does not modify the word rule, stands for a different rule than the one which postposes relative clauses.

First, I shall consider some of the key facts that must be accounted for.

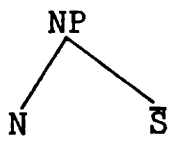
The rule was first discussed by Rosenbaum (1967). It relates sentences like (22) and (23):

(22) That John is a fool is obvious.

(23) It is obvious that John is a fool.

Most linguists have assumed a transformational relationship between (22) and (23) (but see Koster (1976) for an exception), although they have differed on the directionality of the relationship. Emonds (1970) argues that (23) is closer to the underlying structure, while Emonds (1976), responding to Higgins (1973), postulates (22) as the more basic form, following Rosenbaum's analysis. In Emonds (1976), <u>that</u> and <u>for-to</u> complements dominated by NP have the following underlying representation:

(24)



Emonds then proceeds to use the convention that deltas cannot survive in surface structure as a filter. There are two ways, then, that sentential complements which are underlyingly NPs can survive in surface structure:

(a) Some rule deletes the delta.
(b) The delta gets spelled out as <u>it</u> when the sentential complement is extraposed.

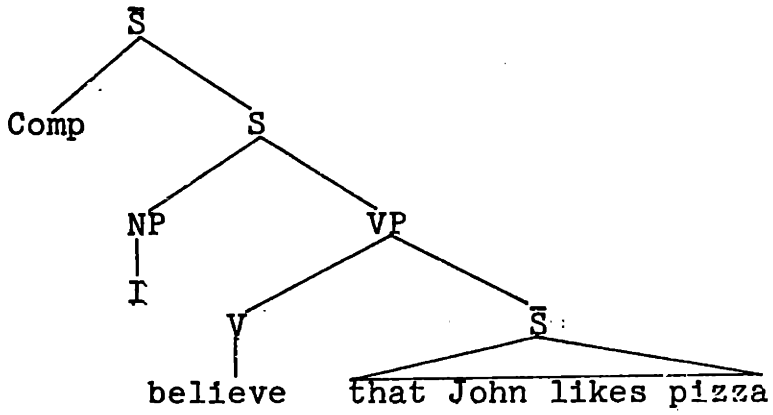
For Emonds, the sentential complement in (22) is not in subject position, but rather in Complementizer position in the matrix S via clausal topicalization (topicalization is viewed as a Comp-substitution rule; more will be said on this view directly), a root transformation. A rule then

deletes the delta in this position. In this way, Emonds captures the Internal-S Constraint of Ross (1967) and Kuno (1973). Emonds has another rule which deletes the <u>it</u> left behind after extraposition from non-subject positions. (Rosenbaum (1967) has substantially the same rule). Thus, sentences like (25),

(25) I believe that John likes pizzas.

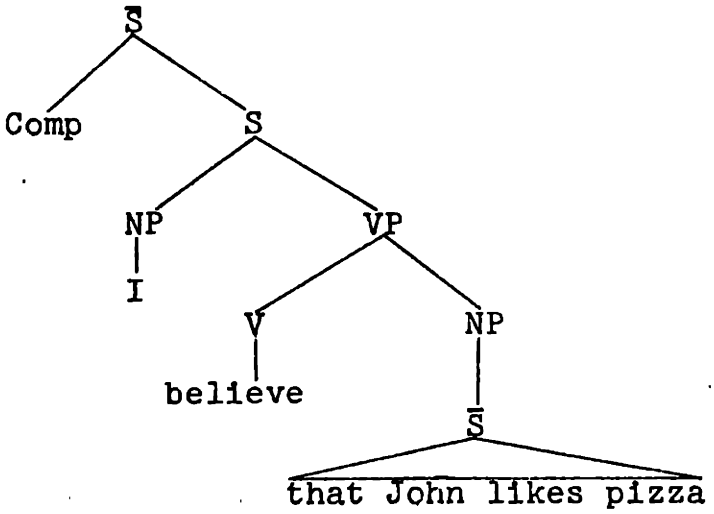
have the surface structure (26):

(26)



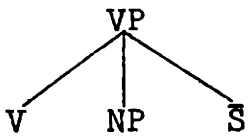
rather than the surface structure (27):

(27)



Given that VP-complementation is well-motivated for some verbs, as well as base-generated configurations like (28):

(28)



(see Emonds (1976) and Rosenbaum (1967) for details), extraposition can be formulated as a structure-preserving rule. Now, as evidence for both the Structure-Preserving Hypothesis and the use of deltas as a filter, Emonds notes the ungrammaticality of sentences like (29):

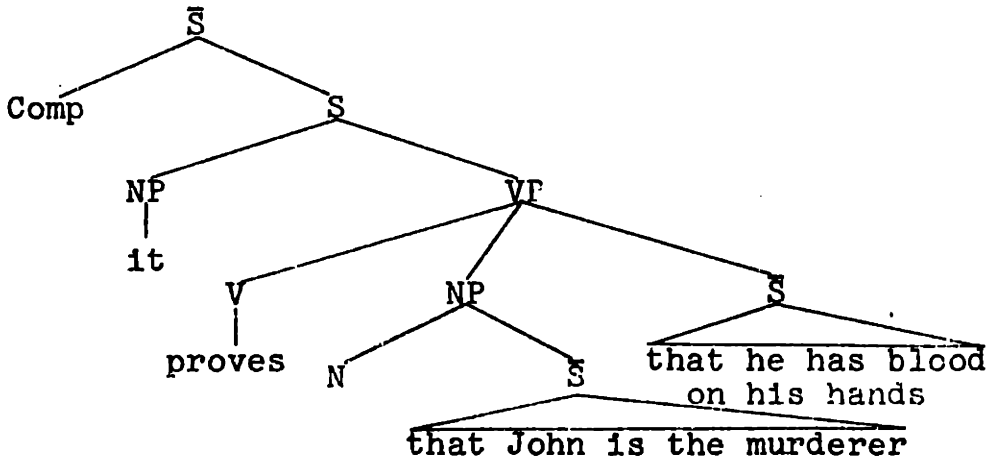
*(29) It proves that John is the murderer that he has blood on his hands.

which would come from (30):

(30) That he has blood on his hands proves that John is the murderer.

In his analysis, the derived constituent structure of (29), after applying extraposition of the subject, would be

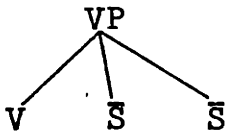
(31)



Given that extraposition is structure-preserving, and given that base rules like (19) do not exist, once extraposition of the subject has taken place, there is no slot for the object to extrapose into, and so the object complement cannot move out from under the domination of NP. Since the delta will therefore remain in surface structure, the sentence is starred for this reason.

I must admit that some speakers with whom I have consulted do not find (29) crashingly bad. In any event, one can see now why the claim that sequences like (32):

(32)



do not arise in non-root sentences is so crucial. However, evne if one accepts the ungrammaticality of (29), one can motivate structures like (32). Consider the following contrast:

(33) It is believed to be obvious by everybody that Fred is crazy.

*(34) John is believed to be certain by everybody that Fred is crazy.

One can show by accepted constituent-structure tests that the complement sequence <u>that Fred is crazy</u> is dominated by VP (in (33)). For one thing, it deletes under VP-deletion:

(35) Although John didn't think that it was believed to be obvious by everybody that Fred was crazy,

(35) (cont.)
it was __.

It would be futile to try to maintain that the it before was in (35) was really anaphoric, and what really underwent VP-deletion was the sequence believed to be obvious by everybody, and that this was the only derivation for (35). Such an analysis would predict that when the two finite complements are distinct, the second could be left behind. Thus, (37), formed from (36), should be good:

(36) Although not everybody thought that it was believed to be obvious by everybody that Fred was crazy, people who thought that it was believed to be obvious by everybody that Fred was stupid felt quite sorry for him.

*(37) Although not everybody thought that it was believed to be obvious by everybody that Fred was crazy, people who thought that it was __ that he was stupid were quite sorry for him.

However, (37) is ungrammatical, according to my informants. Therefore, assuming that VP-deletion deletes constituents, we could explain the ungrammaticality of (37) if we placed the finite complement within the VP.

Now that we have established that the finite complement in (33) is within the VP, we must try to ascertain more precisely whether it is dominated by the infinitival complement or not. The Structure-Preserving Hypothesis would force us to claim that the sentential complements of obvious or certain are both embedded under the infinitive S, rather than the matrix S. However, given that

passives subcategorize for agent phrases, we would have to have a subcategorized element occur in a lower S than the element which subcategorizes for it, definitely an undesirable consequence. Even if we allowed this move for obvious, we would have to block it for certain. Obvious and certain, though, would be required under the Structure-Preserving Hypothesis to have the same structure. Therefore, this non-local subcategorization is untenable.

The same problem arises if we propose that there is a rule which lifts the embedded complement out of the infinitive, and that this rule is stylistic. (For Emonds, stylistic rules apply quite late in the grammar, separated from the basic syntactic transformations in a separate component.) It is true that one can probably motivate a stylistic rule which detaches complements. To do so, one would have to show the rule applying after a rule which operates in a post-transformational component; if one assumes that there is a level of logical form which is post-transformational, and a rule of non-coreference (Lasnik 1976, Reinhart 1976) in some form (Chomsky 1978), one can make the argument.

Consider the following contrast (from Reinhart 1976):

(38) Everyone loved him who knew Fred.

(39) It was obvious to him that Fred would lose the race.

One can construe the two underlined noun phrases in (38), as coreferential, but not (39). To account for this

contrast, Reinhart proposes to attach extraposed sentential complements under VP or AP, but extraposed relatives under S, and then use her C-Command Constraint on anaphora. If we accept Reinhart's account, however, we must explain why the two noun phrases in (40) are perceived as non-coreferential:

(40) Just how obvious to <u>him</u> was it that <u>Fred</u> would lose the race?

In this case, non-coreference should not apply, since neither of the two noun phrases in (40) C-Commands the other. Therefore, if we wish to keep Reinhart's account, the simplest solution would be to say that non-coreference applies to the phrase-marker in which the two NPs are in the appropriate domain relationship, and a late rule detaches the complement. Three alternatives, with the data being considered up to this point, are equally viable: The ordering in (41), the one in (42), or the one in (43):

(41) (a) Non-coreference
(b) Detachment
(c) Wh-movement

(42) (a) Non-coreference
(b) Wh-movement
(c) Detachment

(43) (a) Wh-movement
(b) Non-coreference
(c) Detachment

All three of these orderings would place the two NPs in the appropriate domain relationships, so that non-coreference could apply at the appropriate point. However, the ordering in (41) and (42) would have a rule of logical form applying before a syntactic rule, contrary to our assumptions about the organization of the grammar. Moreover, there is an empirical argument against (41) and (42). Postal (1970) has shown that the domain relationships for anaphora must be established after wh-movement. Therefore, in the following sentence, one can establish coreference between an NP in a questioned wh-phrase and the subject:

(44) I wonder which of the boys that Sally dated she hated.

as opposed to (45):

(45) I know that she hated some of the boys that Sally dated.

in which coreference is impossible, as predicted by Reinhart's constraint.

Therefore, the ordering in (43) is the best-supported one. Note also that the ordering in (43) is the only ordering of the three which forces us to view detachment as a stylistic rule. In (41), detachment would apply before wh-movement, assumed here to be syntactic. In (42), detachment applies after a syntactic rule, so it could be stylistic, but does not have to be. However, in (43), we have detachment separated from a syntactic rule by a rule

of logical form, and therefore, assuming separability of components, it must be of a different character from the syntactic rule (wh-movement).

Notice that detachment must differ crucially from the rule we are calling here extraposition by the argument presented. While detachment must follow non-coreference and wh-movement, extraposition crucially precedes these rules. Reinhart (1976) notes this fact, by citing the following contrast:

(46) That nobody liked John amazed him.

(47) It amazed him that nobody liked John.

While the underlined NPs in (46) can be coreferential, the ones in (47) cannot be. Therefore, Reinhart argues, the sentential complement in (47) must be in final position in the VP, and the rule which places it there, if (47) is not base-generated, must apply before non-coreference. Although there seems to exist a stylistic rule of detachment, then, it cannot operate to directly derive (47) from (46), since claiming that extraposition and detachment are identical would predict that the coreference possibilities in (47) would be reducible to those of (46).

We can thus make a good case for detachment as a stylistic rule, but this does not help us to maintain the Structure-Preserving Hypothesis, for even if we do accept a detachment rule which is stylistic and hence not

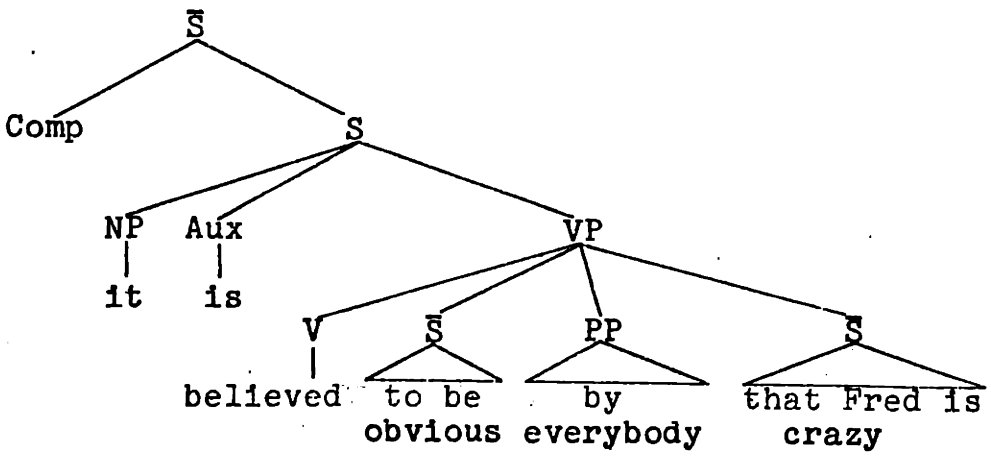
within the purview of the Structure-Preserving Hypothesis, and which postposes complements to the ends of clauses, we are at a loss in explaining why it applies differentially to (33) and (34), since both are assumed to have exactly the same input structure in Emonds' analysis. The same objections to placing the agent phrase in the infinitive apply here with equal force. In this case, we must conclude that detachment is clause-bound.

It seems that we are trying to avoid the inevitable. These previous alternatives have no natural way to differentiate between (33) and (34), and no way to correlate the difference with the fact that (33) has an optional variant:

(48) That Fred is crazy is believed to be obvious by everybody.

It seems, then, that we must derive (33) from (48) by a rule of extraposition, operating in the matrix clause. However, notice the derived constituent structure for (33) that this conclusion forces us to postulate:

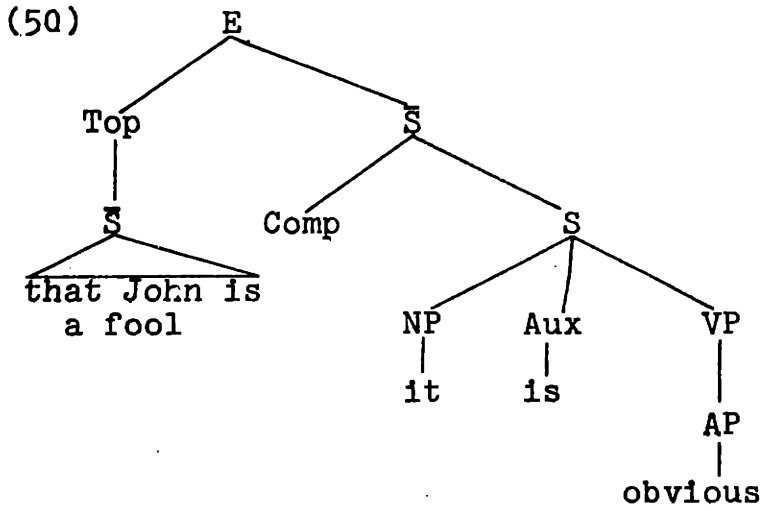
(49)



Notice that we have two S's in the VP, an instance of sequence (32), which is predicted to be impossible by the Structure-Preserving Hypothesis. Therefore, I conclude both that there is a rule of extraposition, and that it is not structure-preserving.

My criticisms of Emonds (1976) apply to Koster (to appear) with equal force, since his proposals are quite similar to Emonds'; clause-initial complements are postulated by both to be topicalized. However, while Emonds views topicalization to be a Comp-substitution movement rule, Koster views clause-initial complements as being generated under an E-node, in the sense of Banfield (1973). Thus, Koster would posit an underlying structure like (50) for (22):

(50)



A rule would then delete the <u>it</u> next to an S which binds it. Crucially, Koster posits no syntactic relationship at all between (22) and (23), and postulates no semantic rule to relate the two structures. Therefore, it is

viewed as a pure accident that any $\bar{S}$ which occurs in topic position can also occur clause-finally.

These two proposals of Koster's are really quite separable. I will not deal with the viability of (50) as an underlying structure, since take-off sites are not my main concern here. One could keep (50) as an underlying structure, and posit an extraposition rule. In fact, Koster would be forced to modify his analysis in this way, since he, like Emonds, would be assuming that the only instances of complements in final position would be those which arose by the phrase-structure rules. Therefore, a structure like (49) would counterexemplify Koster's position to the same extent that it would Emonds'.

It would seem, then, that we have two rules; one which extraposes relatives, and another which extraposes sentential complements; these rules move these constituents to different positions in the sentence. Therefore, a theory which simply said that elements move as far as they can without violating Subjacency would be too strong, since extraposed relatives move to a final position past the verb phrase, while extraposed sentential complements move to a position final within the verb phrase or adjective phrase,[1] and our formulation of movement rules must take account of this (barring an independent account, such as some sort of semantic factor, which could predict the difference, and thus enable us to leave the movement rule more free). The inventory approach manifested in (3) is one way. It strikes me as a

minimal retreat from the position that elements simply move as far as they can from the end.

However, the inventory in (3) must be modified, and I think that the modification which I shall now propose furnishes interesing support for the hypothesis that categories are not unanalyzed primitives, but rather are decomposable into features, much like segments in phonology (Chomsky 1970, 1974; Bresnan 1975, 1976; Jackendoff 1977). In Chomsky's system, the features are ±N, ±V. The categories are then decomposed as follows:

(51) $\begin{bmatrix}+N\\-V\end{bmatrix}$ = Noun, $\begin{bmatrix}-N\\+V\end{bmatrix}$ = Verb, $\begin{bmatrix}+N\\+V\end{bmatrix}$ = Adjective,

$\begin{bmatrix}-N\\-V\end{bmatrix}$ = Preposition

As Bresnan (1975, 1976) points out, this view of syntactic categories is empirical to the extent that rules of grammar make reference to all of the logically possible feature complexes. Thus, some disjunctions are logically possible, such as rules that make reference to -N (i.e., verbs and prepositions), to +V (verbs and adjectives), etc., while other disjunctions are prohibited, such as a rule which referred to noun phrases, adjective phrases, and prepositional phrases, but not verb phrases, since such a disjunction is not expressible within the feature system in (51). Thus, to the extent that disjunctions of three categories do not figure in the formulation of rules, (51) makes a strong and interesting claim which is realized. However,

a feature system of this sort must succeed in two ways: it must rule out disjunctions that do not occur and be able to express the disjunctions that do occur.

With this in mind, suppose we reformulated the inventory to reflect the notion that symbols like NP, VP, etc., are really expository conveniences, just as symbols like /b/, /p/, /t/, /d/, etc. are written as such in phonology papers as nothing more than a shorthand. Thus, position II in the inventory should really be formulated as $\begin{bmatrix} \bar{\bar{X}} \\ -V \end{bmatrix}$ (or whatever the maximal expansion of phrasal categories should be). In this way, we can simplify the inventory of landing sites, and reflect the fact that both AP and VP are the constituents the ends of which extraposed sentential complements move to.

III. On Comp-Substitution

In sections I and II, we looked at the positions to which moved elements move, and saw that even though extraposed relatives and sentential complements both moved to the end of the clause, they had to be distinguished more finely than that. In this section, I wish to motivate an analysis in which wh-movement moves an element to position IVa in the inventory, or left bracket $\bar{S}$, while topicalization moves an element to position IIIa, or left bracket S. I will be making the same point, therefore, with respect to leftward movement rules that I made with respect to

rightward movement rules.

In the past, both topicalization and wh-movement have been viewed as Comp-substitutions by various investigators (Emonds 1976, Higgins 1973, Liberman 1975, Reinhart 1976), as well as other root transformations. The original rationale for this proposal came from Higgins (1973), who wanted to use it to explain the non-interaction between any two such rules, by appealing to the principle that a Comp cannot be doubly filled. An example of the situation which the Comp-Substitution Principle is designed to rule out can be seen in (52), in which both topicalization and wh-movement have applied:

*(52) (a) What these steps did you use to sweep with?

(b) These steps what did you sue to sweep with?

[from Emonds (1976); his II.50]

The source would be roughly (53):

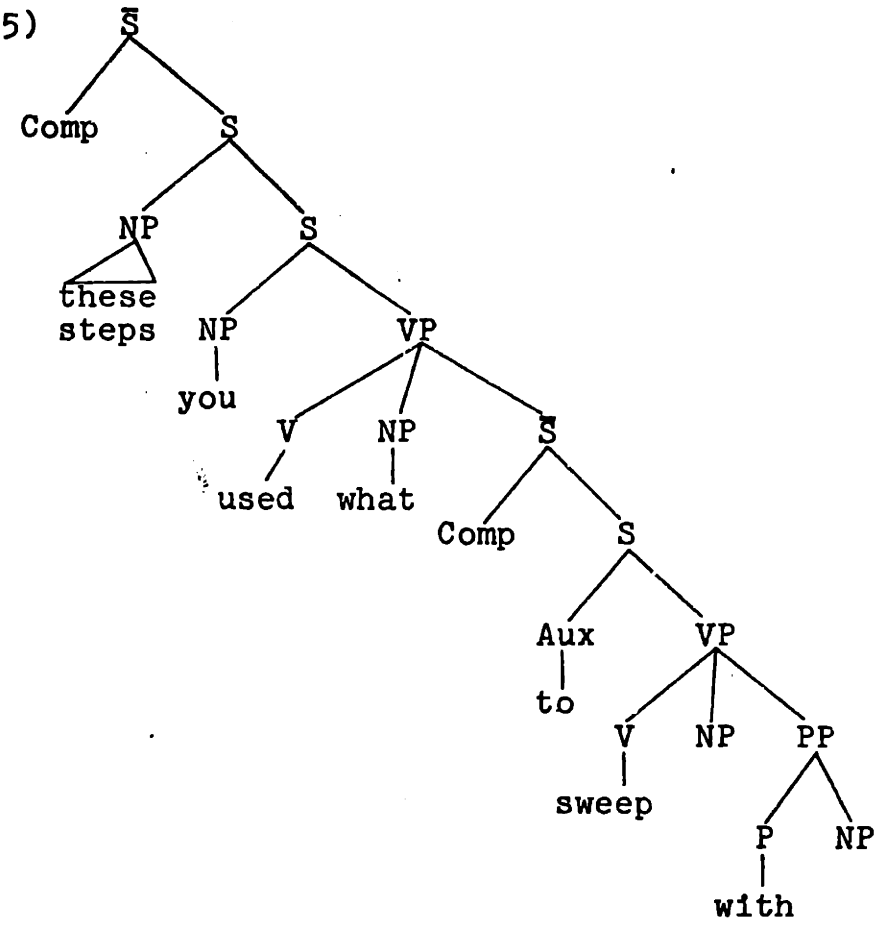
(53) You used what to sweep these steps with.

To rule out this incompatibility, however, one need not appeal to doubly-filled Comps. If one formulated topicalization as in (54):

(54) Move $\begin{bmatrix} \bar{X} \\ -V \end{bmatrix}$ to IIIa.

and formulated topicalization as an adjunction, one would have the following structure after topicalization:

(55)



Given that S is a bounding node for English, one would not be able to perform wh-movement after topicalization without violating Subjacency.

In fact, this analysis makes a prediction that the Comp-substitution analysis does not. It predicts that in a language in which S is not a bounding node, but S̄ is, it will be possible to wh-move over a topicalized constituent.

As we have seen, French is such a language (see Chapter 3 for discussion). As Paul Hirschbuhler and Marie-Thérèse Vinet (personal communications) have informed me, it is possible to topicalize in embedded questions and relative clauses. Thus, they accept the following:

(56) Les hommes à qui les livres j'ai donné.

(57) Je voudrais savoir [quelles lettres][à Jacques] tu as donné.

(58) Le garcon à qui de Jacques j'ai parlé, plutôt que de Marie, c'est Luc.

Evidently, some speakers of French do not accept topicalizations at all, so the grammaticality of (56-58) is dialectal. However, for those who do accept topicalizations in French, one would have to say that doubly filled Comps are possible. Sportiche (1978) shows that doubly filled Comps are not in general possible for wh-movements. Therefore, a problem seems to exist for viewing topicalization as a Comp-substitution, if one believes that wh-movements are such.

Reinhart's constraint on anaphora again has implications, this time for the question of whether topicalized elements and wh-movements have the same landing site. As we saw in the discussion of (44), a subject pronoun can establish a coreference relationship with a noun phrase in a wh-phrase. However, one of Reinhart's main points is that a topicalized PP cannot contain an NP coreferential with a subject pronoun. Therefore, the underlined noun phrases in (59) are not perceived as coreferential, but the ones in (60) are:

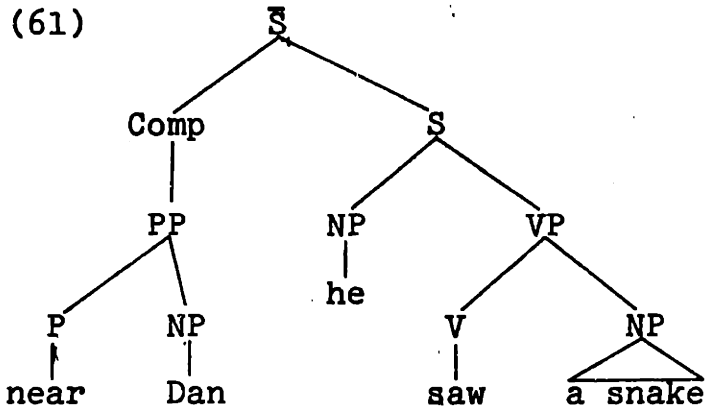
(59) Near Dan, he saw a snake.

(60) Near him, Dan saw a snake.

Reinhart accepts the view of topicalizations as Comp-

substitutions, and she wants to say that the subject NP C-Commands the NP in the topicalized PP. The tree, however, would be the following:

(61)

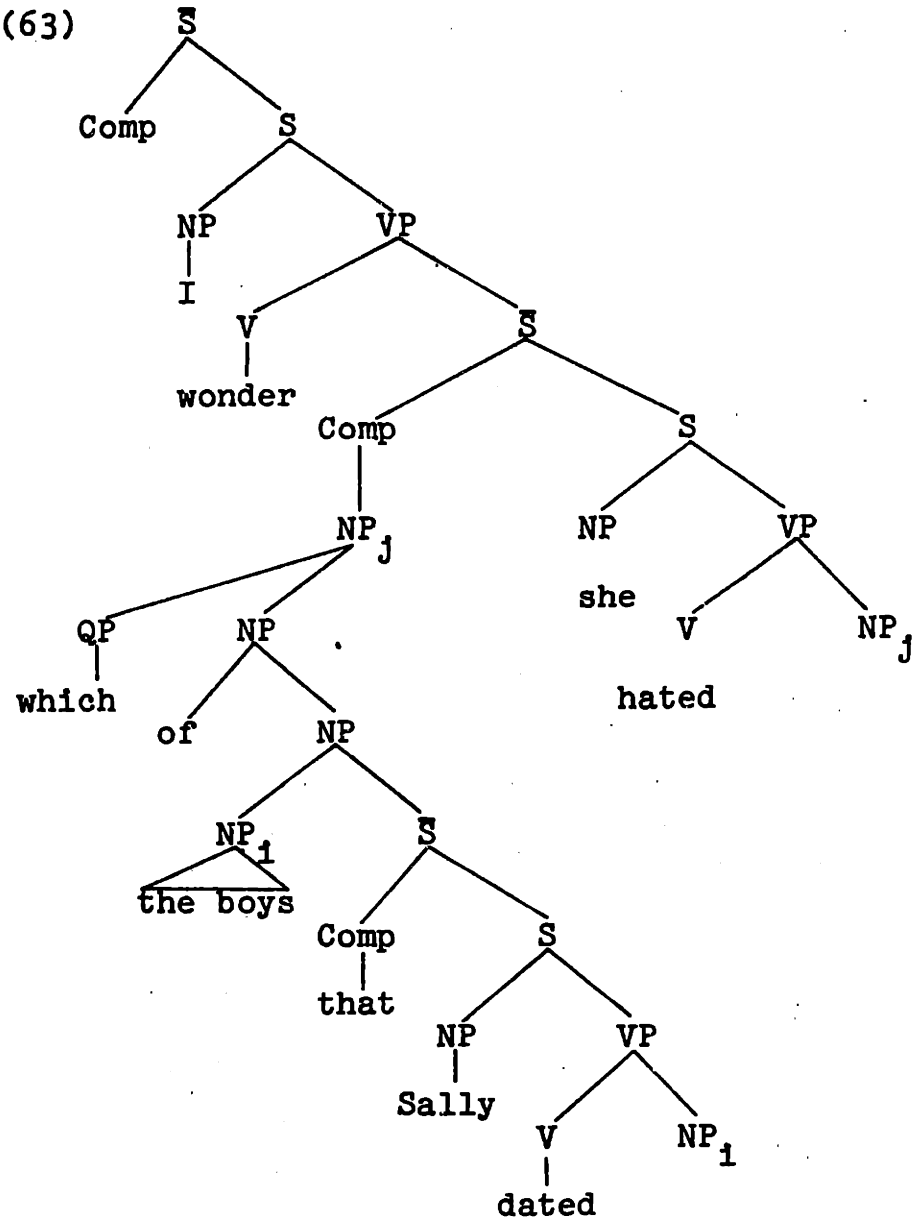


Strictly speaking, the subject NP does not C-Command the object of the fronted preposition, since the first branching node over the subject (S) does not dominate the NP Dan. Therefore, in Chapter 4 of her dissertation, Reinhart modifies the notion of C-Command as follows:

(62) A node A C-Commands a node B if the first branching node α which dominates node A also dominates node B, or if the first node over α is of the same categorial type as α, abstracting away from indexing by bar, and also dominates node B.

Thus, anything in Comp is under S̄, which is of the same categorial type as S, and hence C-Commanded by the subject. Thus, by this modification of C-Command, in sentence (44), the NP Sally will be in the domain of the subject she, since the sentence is hypothesized to have the structure in (63):

(63)



In this instance, then, the theory that both rules move constituents to the same position makes the wrong prediction, since the facts would indicate that the subject C-Commands a topicalized prepositional phrase, but not a fronted wh-phrase. Therefore, if we assume the correctness of the C-Command account of anaphora, we must say that these two rules move elements to different positions in the sentence, such that topicalization moves an element to

a position lower in the tree than wh-movement does. Two alternatives suggest themselves: one could keep wh-movement as a Comp-substitution, formulate topicalization as in (54), and eliminate the rider in the definition of C-Command given in (62) which states abstracting away from indexing by bar. In this way, a stronger definition of C-Command would be in effect, such that adjunctions will be considered to be in the domain of the subject, since both are dominated by S, whereas wh-moved elements will be under $\bar{S}$, and hence will not be C-Commanded by the subject, which is immediately dominated by S. Thus, the distinction can be made in this way.

Another way of making the distinction would be to say that topicalization is a Comp-substitution, while wh-movement is an adjunction to S. One could retain the definition of C-Command in (62), even with the rider about the different number of bars being irrelevant, since the wh-phrase would be too high in the tree to be C-Commanded by the subject.

What is clear from this discussion, however, is that one cannot collapse the landing sites for topicalization and wh-movement. Thus, Maling & Zaenen (1977) have noted this need in Icelandic, which has a verb-second constraint only in main clauses. Therefore, in Icelandic, the verb must be in second position in all clauses. Thus, (64) is good, but (65) is out:

(64) Ég held að smalann muni tröll taka á morgun
I-think-that-the shepherd-will-take tomorrow

*(65) Ég held að smalann tröll muni taka á morgun.
[M&Z's (3b) and (4), resp.]

Now, topicalizations can count as the first constituent in the counting of constituents relevant to the verb-second constraint. Thus, (64) is actually a topicalization, with the meaning of (66):

(66) I think that the trolls will take the shepherd tomorrow.

Wowever, wh-words cannot count as a first constituent. Thus, not only can we get sentences like (67),

(67) Hann spurdi hverjum Olafur hefdi hjalpad.
'He asked who (dat.) Olaf had helped.'

in which the verb is actually the third constituent in the clause, but we cannot get sentences like (68), in which the verb directly follows the wh-word:

*(68) Hann spurdi hverjum hefdi Olafur hjalpad.
'He asked who (dat.) had Olaf helped.'

Again, we see that we must distinguish between topicalization and wh-fronting. Maling & Zaenen propose to account for this distinction by making topicalization an adjunction to S, and counting only constituents within S as relevant to the verb-second constraint. As we have noted, however, all that the data really shows is that a distinction must be made.

There are some considerations, however, which might

tend to show, at least, that wh-movement should be formulated as in (69):

(69) Move $\begin{bmatrix} \bar{\bar{X}} \\ wh \end{bmatrix}$ to IVa.

Of course, once we accept this formulation of wh-movement, we must still make a decision about the derived constituent structure of topicalized elements. We shall return to the question of topicalization after considering wh-movement.

First of all, the following arguments will not go through for English for irrelevant reasons, as we shall see, but there is at least one fairly direct argument from Swedish (due to Andersson 1975) and one slightly less direct argument from French that the formulation of wh-movement in (69) is viable.

First of all, consider the rule of Right Node Raising, discussed by Postal (1974) and Bresnan (1974), who argue for its use as a diagnostic in determining constituent structure. As Bresnan showed, this test would indicate that Complementizers and wh-words are set off from the rest of the sentence. Bresnan's example is (70):

(70) I don't know whether, but I'm fairly sure that, your hypothesis is correct.

Andersson (1975) shows, however, that in Swedish, as in Middle English, wh+that constructions arise in embedded sentences. The following embedded question is an example

of this in Swedish:

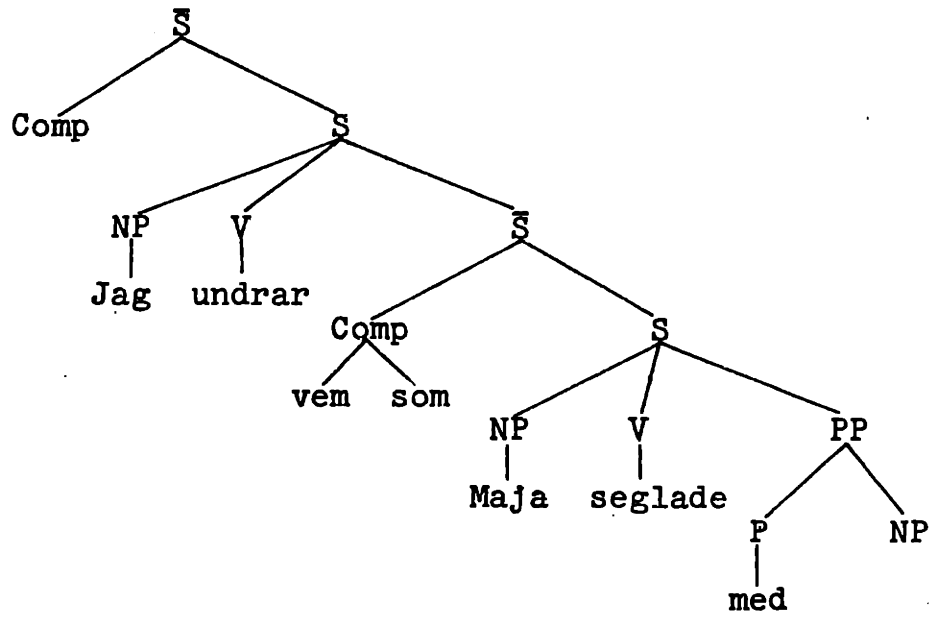
(71) Jag undrar vem som Maja seglade med.
I wonder who that Maja sailed with.
[A.'s III.2 (46)]

Now, Andersson shows that one can strand the wh-word in a right-node raised structure and right-node raise the som plus the rest of the sentence:

(72) Jag vet vem, men du vet nog inte vem, som har varit har.
'I know who, but you probably don't know who, that has been here.
[A.'s III.2 (14)]

(See Andersson, p. 160, for more examples of this.) Consider the two main hypotheses being discussed here. Under the Comp-Substitution Hypothesis, (71) would have the following derived structure:

(73)

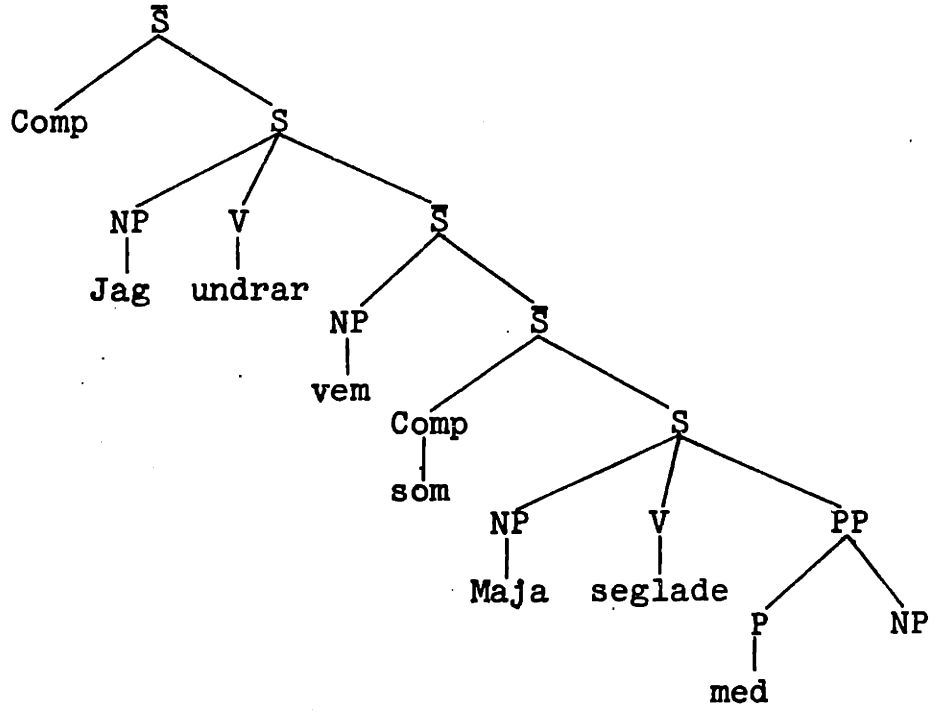


Under this analysis, the som plus S sequence does not

form a constituent, and hence should not be able to right-node raise; therefore, on this account, (72) should be ungrammatical.

If rule (69) were the operative one in Swedish, the derived constituent structure of (71) would be as in (74):

(74)



In this structure, we would be right-node raising S̄, and (72) could therefore be derived. There is fairly good evidence in Swedish that rule (69) is the correct formulation of wh-movement.

One can also provide an argument from French to the extent that one may be able to derive the Comp-to-Comp Condition (Chomsky 1973) from Subjacency if one formulates wh-movement as in (69).

As noted in Chapter 2 of this thesis, sentences in French in which Q-Float and wh-movement of the head have

both applied are both grammatical. Thus, (75) is fine:

(75) Les livres que j'ai tous lu sont bons.
'The books that I have all read are good.'

There are two possible accounts of the derivation of (75). In one account, Q-Float applies first, and then the head is wh-moved, after which it deletes. In the other account, the entire NP first undergoes wh-movement, and then Q-Float then moves the quantifier back into the clause.

There is a way of deciding between thet wo alternatives. Quicoli (1976) has argued that Q-Float in French obeys the Specified Subject and Tensed-S Conditions. Thus, the following relative clause would be out because the quantifier would be extracted from a tensed clause:

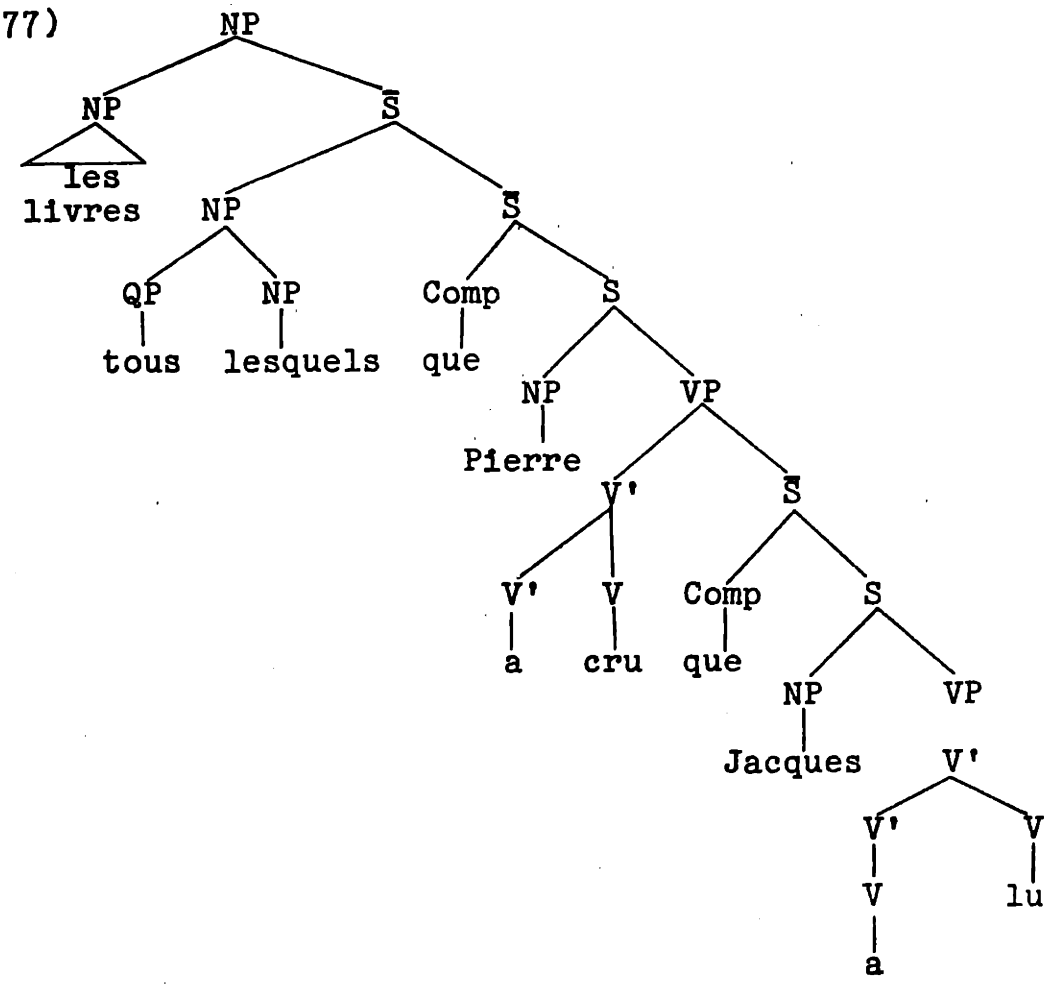
*(76) Les livres que Pierre a tous cru que Jacques a lu.
'The books that Pierred has all believed that Jacques has read.'

However, Quicoli does not discuss the logically possible alternative derivation in which tous lesquels moves to the Comp adjacent to les livres and then tous is floated off, which would give rise to (72) and not violate Tensed-S and Specified Subject. The general constraint that would rule out this derivation is one proposed by Chomsky (1973), which states that once an element is moved into Comp, the only position it can move to is another Comp. Thus, the derivation wh-movement followed by R-tous (Kayne 1975) is precluded.

The Comp-to-Comp Condition, however, as it stands, is another stipulation in the theory. If it could be made to follow as a consequence of other general principles of grammar, it would reduce the number of independent stipulations in the metatheory, and thus lead to a more elegant theory of universal grammar.

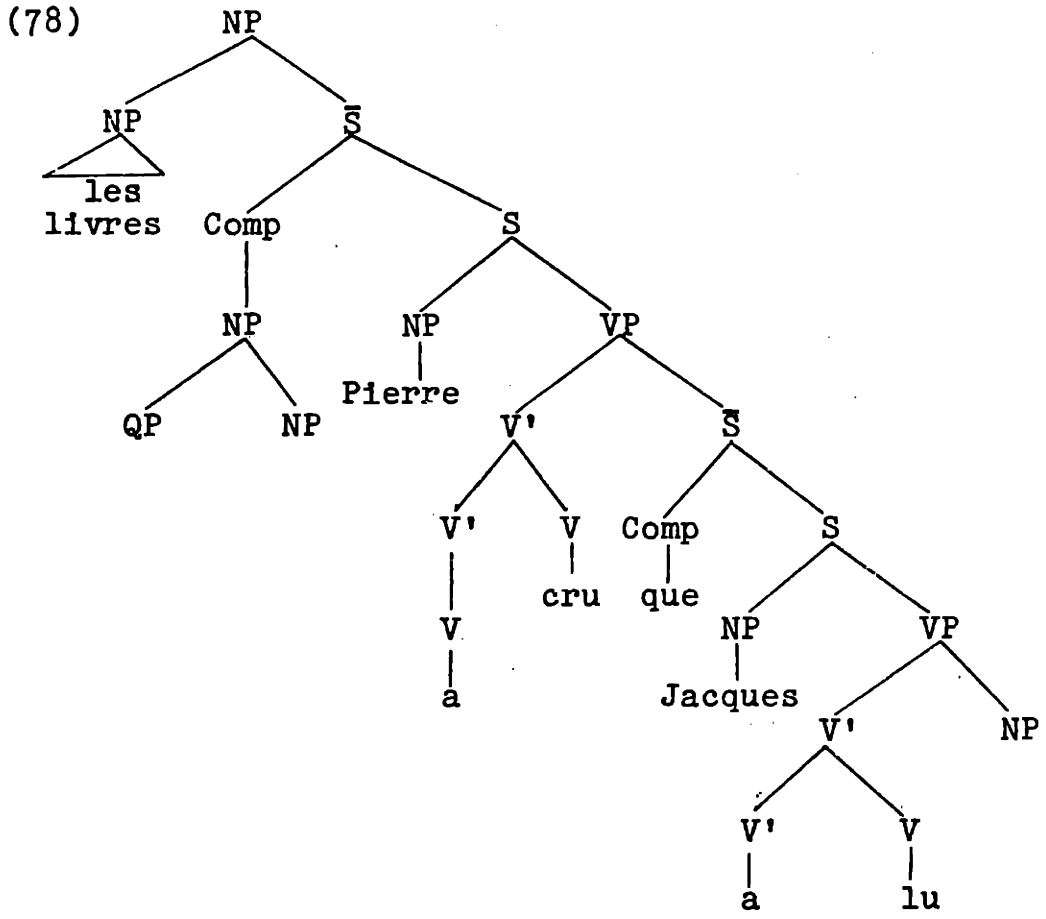
I would now like to suggest a way in which the ungrammaticality of (76) could fall out of Subjacency. Suppose we assumed, as discussed in Chapter 3, that $\bar{S}$ is the bounding node in French. Let us also posit (69) as the rule of wh-movement for French. To derive (76), then, the pre-Q-Float structure would be (77):

(77)



On the constituent V', see Emonds (to appear) for strong arguments. I am accepting this analysis. Now, to derive (76), one would have to float the tous out of the NP and into the $\bar{S}$, two cyclic nodes, and thus violate Subjacency. However, suppose we accepted the Comp-substitution analysis of wh-movement. The pre-Q-Float structure for (76) would be (78):

(78)



In this case, Q-Float out of the NP in Comp would be moving the quantifier out of the NP and the S, which is not a bounding node in French. Therefore, Subjacency, under this account of wh-movement, would not block (76), and we would need the Comp-to-Comp condition to block movement

out of a Comp into an S.

I have not investigated all of the evidence for the Comp-to-Comp condition in that fine a detail, and simply raise this alternative as a suggestion for future research. I would not like to leave the impression that the formulation (69) is totally without problems, however. As Joan Bresnan (personal communication) has pointed out to me, part of the evidence that $\bar{S}$ is a bounding node in a language is that one can violate the wh-island condition if one extracts from theembedded wh-clause itself, but no further. However, if wh-movement is formulated as in (69), any extraction from that clause will violate Subjacency if $\bar{S}$ is a bounding node, assuming that these rules create Chomsky-adjoined structures, since such an extraction would be crossing two $\bar{S}$'s.

Therefore, if one wanted to account for the facts of Italian and French, and keep a universal formulation of wh-movement, one might propose that wh-elements adjoined to $\bar{S}$ undergo a restructuring rule, along the lines of Akmajian, Steele, & Wasow (to appear), which would attract the wh-elements into empty Comps. To ensure the inability of rules like Q-Float to reapply in Comp, and hence to violate the proposed explanation for the ungrammaticality of (76), one might propose that the readjustment rule is a part of the phonology, in the sense of Chomsky & Halle (1968). Bresnan (1971) has proposed that their Nuclear Stress Rule applies after each syntactic cycle, so that phonological

rules and syntactic rules would be interspersed, but not freely (but see Liberman & Prince 1977 for a quite different account of stress). The proposal being made here is identical to hers, except for the nature of the rule.

At this point, it may be instructive to compare Emonds' analysis of wh-fronting with the one presented here. Recognizing that permitting base rules like (79):

(79) $\text{Comp} \longrightarrow \begin{Bmatrix} \text{PP} \\ \text{NP} \\ \text{AP} \end{Bmatrix} \pm\text{Wh}$

would evacuate the Structure-Preserving Hypothesis of empirical content, since these categories would have to be stipulated as being obligatorily empty in the base in this position, he proposes that wh-movement is a substitution not for any syntactic node under Comp, but rather the feature wh, and proposes the following condition:

(80) The Sentence-Boundary Condition

> If A_i is a rightmost or a leftmost constituent of an S, a transformational operation that substitutes B for A_i is structure-preserving if B dominates A_i, provided that there is no S such that
>
> $$B=X[_S\ Y\ A_i\ Z\ _S]\ W.$$
>
> [Emonds 1976, Ch. 5, no. (70)]

Thus, given that wh is a leftmost constituent of an S, any PP, NP, or AP can substitute for it, provided that these constituents dominate a wh-form without any S's intervening.

Criticisms of the feature ±wh are given in Grimshaw (1977 and forthcoming), so I will not repeat them here.

Another noteworthy feature of (80) is that it explicitly mentions the notion rightmost or leftmost constituent. Therefore, we actually have two classes of rules (aside from local rules) under this hypothesis which apply in embedded contexts; the ones which must substitute categories for identical categories (i.e., PPs cannot substitute for NPs), and these apply clause-internally, and rules which can substitute categories for non-identical categories, provided that the categories substituted for are S-peripheral.

It seems that with the Sentence-Boundary Condition, it is recognized that some rules are not structure-preserving in the strict sense, and these rules would correspond to the landing site rules. The other rules, however, which apply freely in embedded sentences, are, strictly speaking, structure-preserving in the sense of Emonds (1970).

However, the claim of landing site theory is that there is only one type of non-NP movement rule, and that the generalization about constituent periphery governs all of these rules. Also, we have shown that extraposition (of sentential complements) does not move the sentential complement to a rightmost position in S, but rather in VP or AP. Furthermore, even if it did move to a rightmost position in S, there would be nothing corresponding to A_j in that position, so the Sentence-Boundary Condition would still not apply. Therefore, it cannot be classed as structure-preserving under either of the two criteria proposed in TAES.

Furthermore, topicalization is viewed in TAES as a

root transformation, and the facts noted by Hooper & Thompson (1974), among others, are accounted for by claiming that the class of root sentences can be expanded to include the complements of verbs of indirect discourse.

However, one would be hard-pressed to make this move in the case of topicalizations which occur in embedded questions and relative clauses, The proposal made here was that the reason relative clauses with embedded topics are out, as in (81) (in English):

*(81) The men to whom the books I have given...

is that S is a bounding node in English, unlike French. Recently, various investigators have suggested (Hooper & Thompson 1974, Green 1976, Bolinger 1977) that the original restrictions noted by Emonds on the inability of certain rules to apply in embedded contexts may be a consequence of the pragmatic effects of the structures resulting from these rules. In any event, the theory of movement rules proposed here makes no formal distinciton between root and non-root rules.

IV. On Movements to VP Boundaries

In this section, I will discuss, in particular, the formulation of Q-Float and certain adverb placements. I have discussed Q-Float in earlier chapters, and argued that the constituent which floats can simply be stated as QP, with no further specification necessary. However, I would

like to look more closely here at the position to which the floated quantifier moves. As far as I can tell, the string positions are the following:

(82) would^have^been
^infl. main verb
^infl. be^

The hatch-marks indicate positions that floated quantifiers can move to. In the speech of me and my informants, therefore, the following pattern of acceptability holds:

(83) The men would all have been working.

(84) The men would have all been working.

*(85) The men would have been all working.

(86) The men are all working.

??(87) The men all would have been working.

(88) The men all are working.

(89) The men all have finished.

Up to now, the movement rules being considered have all moved elements to the periphery of major constituents. If we with to extend this generalization to Q-Float, we must consider what the major constituent would be to account for the pattern in (83-89). What we must do, then, is try to test our theory against the analysis of the Aux and the verb phrase. The analyses considered here will be the following:

(90) S⟶ NP Aux VP
Aux⟶ T (M) (have+en) (be+ing)

(90) (cont.) [Chomsky, 1957]

(91) $S \rightarrow NP \quad Aux \quad V^3$

$Aux \rightarrow \begin{Bmatrix} Tense \\ Modal \end{Bmatrix}$ do

$Tense \rightarrow \begin{Bmatrix} past \\ present \end{Bmatrix}$

$V^3 \rightarrow$ (have) V^2

$V^2 \rightarrow$ (be) V^1

$V^1 \rightarrow$ (be) V (NP) (pp)

[Akmajian, Steele & Wasow (to appear)]

(92) $S \rightarrow NP \quad Aux \quad VP$

$VP \rightarrow \begin{Bmatrix} V \quad VP \\ V \quad (NP) \end{Bmatrix}$

[Sag (to appear)]

Sag and Akmajian, Steele & Wasow both argue for a right-branching structure for all of the elements after the pure modal, and hence dispute (90). Their arguments center on ellipsis possibilities after the modal. For instance, in (93), we see that we can delete either everything after the modal, perfective <u>have</u>, or progressive <u>be</u>:

(93) Those guys would have been working until six, and these guys would $\begin{Bmatrix} \text{a. } \emptyset \\ \text{b. have} \\ \text{c. have been} \end{Bmatrix}$, too.

If we assume that we want rules of grammar to apply to constituents, we could formulate the ellipsis rule as VP-deletion if we accepted (92) (Sag's proposal) or V^n-deletion if we accepted (91) (A.,S.&W.'s proposal). However, in order

to derive (93a or b), the analysis of the Aux in (90) would have the ellipsis rule apply to a non-constituent. Also, subject-aux inversion, under the analysis in (90) would entail movement of a non-constituent, but not under the analysis in (91) or (92).

I am assuming here that the Aux and the VP do not form a constituent. If this is correct, notice that the landing site theory would predict the ungrammaticality of (87), since Aux is not a phrasal category. Therefore, if we accepted the landing site theory as in (3), we would have to formulate Q-Float as in (94):

(94) Move QP to IIa.

At this point, a problem crops up. As we saw in the discussion of (93), we are going to want progressive <u>be</u> to be followed by a verb phrase, since one can delete after it. However, if the string after progressive <u>be</u> is a verb phrase, nothing should prevent (94) from applying to the source to derive (85), which is in fact ungrammatical.

One option we could take is to conclude that this fact is a point in favor of (91), since (91) posits a right-branching structure, but distinguishes the constituents. We might be tempted, then, to expand the inventory to include V^3 and V^2, and modify the statement of Q-Float to the following:

(95) Move QP to Left Bracket V^3 or V^2.

This seems to me to be a theoretically undesirable

move. For one thing, the other movement rules in this chapter tend to move elements to the periphery of maximal projections. In this case, V^3 is a maximal projection, but V^2 is not. A theory which allowed both maximal and non-maximal projections as landing sites would predict the possibility of movement rules in which an element moved from outside the maximal projection of a category to a position inside of that category which is non-maximal, but not to the maximal projection itself. Thus, suppose V^2 were in the inventory in (3), and instead of the formulation of Q-Float in (92), the rule would be (96):

*(96) Move QP to Left Bracket V^2.

Thus, (83) would be ungrammatical, although (84) would be generable. We could say that the fact that the rule is formulated as in (95) rather than (96), is an accident. However, I will make the arbitrary decision that Q-Float does not receive the formulation in (96) for a reason.

The other theoretical problem with (95) is that the landing sites in (95) do not form a natural class, so that braces would be needed in the formulation of the rule. I would like to avoid this, accepting the objections to the brace notation in Ross (1970) and Bresnan (1975, 1976).

So far, I am accepting Akmajian, Steele, and Wasow's analysis of the Aux. In light of the objections to (95) which I have mentioned above, one could keep rule (94)

and rule (96) together. However, one would then be saying that there are two rules of Q-Float, an undesirable consequence, and one would be expanding the inventory in (3).

The essence of our problem here is that we want the sequence after the modal and the sequence after perfective _have_ to both be maximal projections of VP, but we do not want the sequence after progressive _be_, when the modal and/or _have_ precedes this _be_, to be a maximal projection. Therefore, a strong version of landing site theory requires an analysis of the Aux as in (97):

(97) $S \rightarrow NP \quad Aux \quad V^2$

$V^2 \rightarrow \begin{Bmatrix} have & V^2 \\ (be) & V^1 \end{Bmatrix}$

$V^1 \rightarrow \begin{Bmatrix} be \\ V \end{Bmatrix} (NP)...$

As in Sag's (forthcoming) analysis, one could use subcategorization mechanisms to ensure the proper ordering of aspectual _have_ and _be_; this does not seem insurmountable.

This analysis, then, is similar to Akmajian, Steele, & Wasow's, in that the verb phrase is layered, but differentiated. However, the verb phrase under this analysis is only partially differentiated, so that V^2 is recursive and is the maximal projection. Like their system, however, the sequence after the progressive is distinguished from the sequence after the perfective or the modal.

Therefore, it would be worthwhile to consider some of the evidence they give for (91). In particular, we must

focus on the arguments given for distinguishing V^3 from V^2, as well as the arguments for distinguishing V^1 from the others.

Akmajian, Steele, & Wasow note that the rule which has been called VP-preposing (by Emonds 1976, among others) is constrained in that it cannot front VPs with perfective _have_ as the highest verb, or progressive _be_. Thus, the following pattern holds:

(98) Those guys said that they will have finished the job by six, and
*(a) have finished the job they will.
(b) finished the job they will have.

(99) Those guys said that they have been working until six, and
*(a) been working until six they have.
(b) working until six they have been.

Therefore, VP-preposing, they argue, must make finer distinctions than has hitherto been assumed, and they propose to make the distinction by postulating VP-preposing as V^1-fronting.

This argument strikes me as valid. If there were other rules which had to distinguish between V^3 and V^2 such that they only applied to V^2, (91) would be rather well entrenched. The distinctions between V^3 and V^2 rest entirely on distributional evidence, to which I shall now turn.

One bit of distributional evidence given for distinguishing V^3 from V^2 comes from imperatives. Thus, although (100) and (101) are acceptable:

(100) Drink your milk!

[A,S&W's (58)]

(101) Be studying your Spanish when I get home!

[A,S&W's (59)]

one cannot form an imperative from perfective <u>have</u>. Thus, (102) is out:

*(102) Have finished your homework by the time I get home!

[A,S&W's (60)]

According to their analysis, imperatives are generated by the following base rule:

(103) $S \longrightarrow NP\ V^2$

Therefore, they make the strong claim that elements generated outside of V^2 cannot co-occur with imperatives. Thus, (99) is ungrammatical. In this light, consider the following:

(104) <u>Do</u> feel better.

(105) Don't drink that.

As we see in (91), <u>do</u> is generated under the Aux, which is out of V^2. Therefore, Akmajian, Steele, and Wasow have no way of generating (104) and (105). I think that imperatives, therefore, shed no light on any putative distinction between V^3 and V^2. Akmajian, Steele, & Wasow calim that the ungrammaticality of (102) could not be semantic, since they claim that (106) is roughly synonymous as an

imperative, and is acceptable:

(106) Be finished with your homework by the time I get home!

However, it seems to me that (106) and (102) are non synonymous at all, but rather have entailments in common. Thus, if Johnny has done his homework, he is finished with it, but not necessarily vice-versa; if he bribes a classmate to do the assignment, and the classmate completes it, I think that Johnny is finished with it, but hasn't done it. More generally, if we view imperatives as commanding that a state of affaris take place at a set point in time, viewed schematically as in (107):

(107) x _____ y

time of utterance — time of state of affairs

it could very well be that a perfective receives an interpretation which is at some point intermediate between the time of utterance and the time of the state of affairs described. In other words, a perfective does not directly describe a state of affairs, in the same sense that (106), which roughly describes a property (Fiengo 1974) does.

In fact, there may be even more support for the idea that the deviance of (102) is semantic. Carden (1970) discusses a class of verbs which he calls "point-action" verbs, which describe an action which occurs at a split-second point in time; explode is such a verb. Now, such

verbs do not co-occur with adverbails like until, at least with the point action reading:

(108) The tires exploded.

(109) The tires exploded until six.

Of course, predicates which normally describe an on-going process or state can easily co-occur with until:

(110) He was happy until six.

Now, negations of point-action verbs co-occur with until:

(111) The tires didn't explode until six.

I think that the co-occurrence restrictions on until can be explained fairly simply, if we note that the negation of a point-action verb is not itself a point-action; rather (111) describes a state of affairs (see Hindle & Sag (1976) for observations about a dialect which allows positive anymore which point in the same direction). If we make this observation about the interaction of negation and aspect, we note that one can get negative perfective imperatives, as in (112):

(112) Don't have eaten all the bologna by the time I get home!

Again, (91), together with rule (103) for imperatives, would predict ungrammaticality for (112).

This semantic interpretation of perfectives would ac-

count also for their other differences between V^3 and V^2. Thus, some verbs, such as make and let are hypothesized by A,S&W to subcategorize for V^2, rather than V^3 complements:

(113) *(a) We made him have finished his work by the time we were back.

*(b) We let him have eaten supper by 4 o'clock.

(114) (a) We made him be finished with his work by the time we were back.

(b) We let him be done with his supper by 4 o'clock.

If we view the complements of these verbs as designating activities at a later time than the time of the matrix verb's action, we can make the same proposal that we did for the imperatives.

To summarize up to this point, we have seen that a strong version of landing-site theory dictates a particular structure to the Aux and VP. Like Akmajian, Steele, & Wasow, this account assumes a right branching structure within the VP, in which the constituents must be differentiated. However, there are only two branching categories of V within the VP in my analysis, and one is recursive, with subcategorization restriction on have blocking unwanted recursion. To account for the grammaticality of (86), I would accept A,S&W's notion of restructuring rules, with their constraints on them, so that be would be shifted into the Aux when the Aux is unfilled.

I should note here, as have previous investigators (notably Baker 1971) that the positions of floated quanti-

fiers are shared by a class of adverbs which traditional grammarians have called preverbs. Ever and definitely are in this class:

(115) If he would ever have been happy,...

(116) If he would have ever been happy,...

*(117) If he would have been ever happy,...

(118) He would definitely have been happy.

(119) He would have definitely been happy.

*(120) He would have been definitely happy.

Unlike floated quantifiers, it seems to me that these adverbs can occur before the modal:

(121) If he ever would have been happy,...

(122) He definitely would have been happy.

Therefore, following Baker (1977), suppose we generated these adverbs by the following base rule:

(123) S $\longrightarrow$ NP (adv) Aux - V^2

and we can account for the distribution of these preverbs by the following rule:

(124) Move Adv to IIa

Interestingly enough, while (87) is out, (125) is fine:

(125) They all would have enjoyed it.

I think, however, that the quantifier in (125) is not floated, but rather postposed by the rule of Q-Pro flip, proposed by Maling (1976), who argues for a derived constituent structure for sentences like (126) in which the quantifier is actually final within the noun phrase:

(126) She loved them all.

vs.

*(127) She loved the men all.

A few remarks about Maling (1976) are in order. She proposes an analysis in which Q-Floating and Q-Pro flip are distinguished. Q-Pro flip is responsible for (126), and only operates on pronominal heads; therefore, it could not apply to derive (127). Q-Floating, however, operates on both full NPs and pronouns. However, Maling suggests that Q-Floating can float the quantifier not only to a VP, but also to a following NP, AP, or PP. Her examples are the following:

(128) (a) I gave the kids all some candy to keep them quiet.

(b) The tooth fairy promised the kids each a quarter.

(c) Dad bought the twins both bicycles for Christmas.

(d) Mom found the boys all so dirty when she got home that she made them (all) take a bath.

(e) We consider the Joneses both unbearably pompous.

(128) (cont.)

(f) Cinderella's fairy godmother turned the pumpkins all into handsome coaches.

(g) He made his money all in Platypus Platinum.

(h) He looked the twins both in the eye (and said...)

(i) She called the men both bastards.

(j) Hang your coats $\left\{\begin{matrix}\text{all}\\\text{both}\end{matrix}\right\}$ up on hangers.

Again, this fact has interesting implications for the inventory in (3), especially since PP is not otherwise needed in the inventory. Therefore, an analysis of Q-Float as in (129):

(129) $Q - NP - X^n$
$1 \quad 2 \quad 3 \longrightarrow \quad 2 - 1+3$

would be irreconcilable with (3) as it stands. Therefore, a closer look at the derived constituent structure of the floated quantifiers in (128) is in order. If the floated quantifiers were actually moved out of their noun phrases, it should be possible to move the original noun phrase by subsequent movement rules and leave the quantifier behind. Therefore, the following passives (assuming for the moment that there is a rule of NP-preposing) and topicalizations should be acceptable:

(130) *(a) The kids were given all some candy to keep them quiet.

*(b) The kids were promised each a quarter by the tooth fairy.

(130) (cont.)

*(c) The twins were bought both bicycles for Christmas by Dad.

(d) The boys Mom found all so dirty when she got home that she made them (all) take a bath.

(e) The Joneses we consider both unbearably pompous.

*(f) The pumpkins were turned all into handsome coaches by Cinderella's fairy godmother.

*(g) His money was made all in Platypus Platinum.

*(h) The twins he looked both in the eye (and said...)

*(i) The men she called both bastards.

*(j) Your coats you should hang $\left\{\begin{matrix}\text{all}\\\text{both}\end{matrix}\right\}$ up on hangers.

The ungrammaticality of (a-c) and (f-j) would be predicted if the quantifiers in the corresponding examples in (123) were actually final within the NP. The reason that (d-e) are grammatical is the same reason, under the account proposed here, that extraposed sentential complements move to final position in both adjective phrases and verb phrases, i.e. the landing site in the inventory in (3) is simply +V. The examples in (128a-c) and (f-j), I suspect, are derived by some sort of extension of Q-Pro flip. This idea receives some credence when we note that "complex" QPs like <u>all three</u> which, as we have seen in earlier chapters, undergoes Q-Float but does not undergo Q-Pro flip, does not occur in the relevant frames in (131):

*(131) She loved them all three.

(132) *(a) I gave the kids all three some candy to keep them quiet.

*(b) The tooth fairy promised the kids all three a quarter.

*(c) Dad bought the triplets all three bicycles for Christmas.

*(d) Cinderella's fairy godmother turned the pumpkins all three into handsome coaches.

*(e) He looked the triplets all three in the eye (and said...)

*(f) She called the men all three bastards.

Thus, it seems that (129) overgenerates, and the inventory approach to movement rules in (3) can be maintained.

V. On Position I

Compared to the previous sections, the establishment of movement tc the periphery of NPs is an extremely simple task. For example, consider the rule of degree phrase shift, discussed in Woisetschlaeger (1976). This rule operates to create sentences like (133):

(133) I have never seen so intelligent a linguist.

Obviously, degree phrase shift would be formulated in landing site notation in the following fashion:

(134) Move AP to Ia.

and Q-Pro flip, discussed in the last section, would be:

(135) Move Q in env. __ Pro, to Ib.

VI. Landing Sites in Universal Grammar

Until now, the focus of concentration has been evidence from English for the inventory in (3), with evidence from other languages occasionally. However, the extent to which other languages have provided arguments for the inventory, particularly in Section III, has revealed a basic presupposition here that the inventory is universal.

As it stands, this claim is much too strong. As we saw in Chapter 2, for example, the spots to which floated quantifiers move in Persian are the following:

(136) S^O^IO^V

Also, in French, as we have seen, the floated quantifier can end up after the main verb, as in :

(137) Ils veulent tous aller au cinema.
'They all want to go to the movies.'

so that a rule like (94) would be quite difficult to motivate for French.

However, if we scrap the inventory totally, we may be predicting the existence of a wider class of languages than actually exists. For example, extraposed relatives always mvoe to the end of a clause; I have never seen a language which moved elements to, for instance, the positions in (136), or even to the left of a verb phrase (see Baker 1970, Bresnan 1970, Bach 1971). Therefore, the problem here, as in linguistics in general, is to fix the parameters for

language variation, or to construct a metatheory in which one has just enough flexibility to account for the differences between languages and no more.

Notice that the variation cited above centers on Q-Float, a VP rule in English. Suppose we simply restricted the possibility of language variation to position II in the inventory. If this is correct, it may be connected to the claim put forth by Arthur Schwartz, in a few papers (1973, 1975) that the only languages which had VP-constituency are SVO languages. However, it is even possible that not all SVO languages have VP-constituents. In the case of French, however, there is a constituency which has been justified for French which is below the level of VP but above V by Emonds (to appear). Emonds motivates the following phrase-structure rules of French:

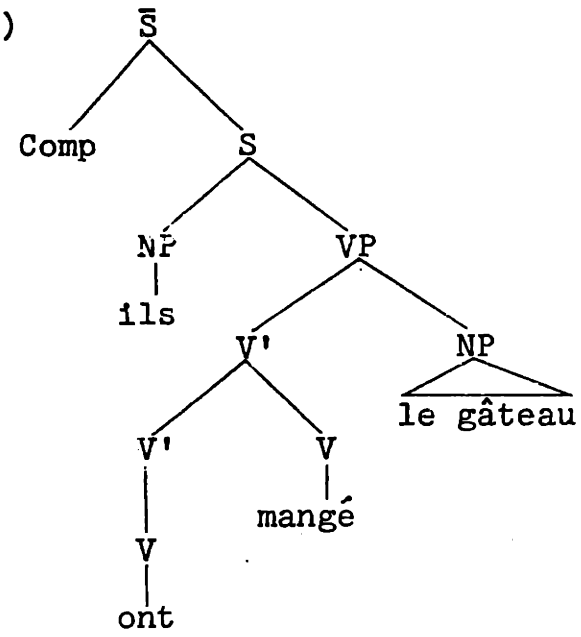
(138) V' $\longrightarrow$ (V') V

and argues that the French "helping verbs" avoir and être are dominated by the first V'. Thus, the constituent structure of (139):

(139) Ils ont mangé le gâteau.
'They've eaten the cake.'

would be (140):

(140)



In this regard, Kayne (1975) notes that floated quantifiers, like adverbs, cannot precede finite verbs or present participles:

*(141) Jean tout lira.
'Jean all will read.'

(142) Jean lira tout.

*(143) En tout buvant...
'In all drinking...'

(144) En buvant tout...

*(145) Jean bientôt partira.
'Jean soon will leave.'

(146) Jean partira bientôt.

*(147) En souvent lisant le journal...
'In often reading the newspaper...'

(148) En lisant souvent le journal...

Thus, if the generalization about movements to maximal projections were to be maintained, one would have to pro-

pose that V' is the maximal projection in French, and that there would be no VP in French which would be a higher projection than V'. As far as I know, French has no rule of VP-deletion. However, Quicoli (1976) proposes an analysis of French causatives in which a rule of VP-preposing figures crucially. Therefore, either Quicoli's analysis must be revised, or else the claims I am making here are too strong.

However, suppose V' were the maximal projection, and Q-Float in French were formulated as in (149);

(149) Move QP to IIb.

Therefore, the inventory would be unchanged, the language-particular parameters being

(a) what the maximal projection would be

and

(b) whether the element moves to left bracket or right bracket.

The universal parameter would be the category to which the element moves.

However, the data in Persian would not allow these moves at all, as we can see from inspection of (136). Also, as Guy Carden (personal communication) has informed me, adverbials niche at the same spots.

One method of retreat, without total surrender of (3), might be to say that universal grammar makes the following provisions for languages without VPs:

(150) Languages without VPs niche the moved elements in VP rules between any two major constituents immediately dominated by S.

In other words, the Persian case would be considered to be the normal alternative. Of course, one may say that it is somewhat rash to make a hypothesis about universal grammar on the basis of one language (Persian in this case). However, it seems to me to be a sounder strategy to try to make the strongest claim about a phenomenon first, and then retreat if one must, rather than formulate a weak hypothesis first, since the weak hypothesis usually does not lead one to make the more testable claim. In this connection, some remarks by Partee (1970), in another connection, are entirely to the point:

> Of course, just being strong doesn't make a hypothesis right; in fact, it increases its chances of being wrong--but it does make it interesting in the sense that it increases our stock of generalizations about the structure of language if it is right.
>
> [Partee (1970), p. 8]

In this connection, let me discuss a view of universal grammar which is somewhat similar to the one which underlies this work, but which differs from the one here in a somewhat crucial respect. In this view, universal grammar would consist of "a fixed, finite list of transformations available for any language" (Bach 1971, p. 154). In this view, universal grammar contains a list of all of the possible rules, and languages will pick from the list; the concept is roughly analogous to a universal inventory of distinc-

tive features in phonology.

Thus, as in my view, one can formulate a universal rule of wh-movement, topicalization, Q-Float, etc. However, I do not believe that this is all there is to universal grammar (at least as far as the transformational component is concerned). If it were, there would be no point to looking for generalizations about the rules in the list; they could vary among themselves without limit. Universal grammar is as far back as the theory of language allows us to go to state a linguistically significant generalization and thus, if universal grammar takes the form of a list and a generalization about the list, this generalization would have the character of a post-hoc explanation about the constructs of a theory, rather than a part of the theory itself. Thus, if the generalizations which I have made in this chapter and Chapter 2 are valid, they should be expressed within the theory directly, by a notation for writing rules from which they follow.

FOOTNOTES TO CHAPTER 4

1. In fact, one can show decisively that extraposed sentential complements are final within an adjective phrase and that extraposition precedes wh-movement if one restricts one's attention to adjective phrases. Sentences like (i) have interesting implications for the range of theories considered here:

(i) Just how obvious that Fred was a fool <u>was</u> it?

The extraposed sentential complement moves with the adjective phrase as a unit by wh-movement; therefore, it must be final within the AP. Also, since it must be part of the AP as the input to wh-movement, it must have gotten to that position prior to wh-movement, again showing that extraposition differs crucially from detachment.

BIBLIOGRAPHY

Aissen, J. (1975). "Presentational *There*- Insertion: A Case for a Cyclic Root Transformation", in CLS XI.

Akmajian, A. (1975). "More Evidence for an NP Cycle", *Linguistic Inquiry* VI, 1.

-----, S. Steele, and T. Wasow (1978). "The Category *Aux* in Universal Grammar", to appear in *Linguistic Inquiry* IX.

Andersson, L. (1975). *The Form and Functioning of Subordinate Clauses in English*, University of Goteborg Dissertation.

Andrews, A. (1975). *Studies in the Syntax of Relative and Comparative Clauses*, unpublished doctoral dissertation, M.I.T.

Baker, C.L. (1970), "Notes on the Description of English Questions: The Role of an Abstract Q Morpheme", *Foundations of Language*.

----- (1971), "Stress Level and Auxiliary Behavior in English", *Linguistic Inquiry*.

----- (1977). *Introduction to Generative Transformational Syntax*; Prentice-Hall.

Baltin, M. (1975), "on the Cyclicity of Extraposition", *Linguistic Inquiry* VI, no. 3.

----- (1977a). "Some Aspects of Cleft Sentences in English", unpublished paper, M.I.T.

----- (1977b). "PP as a Bounding Node", in NELS VIII, M. Stein, ed.

Bach, E. (1971), "Questions", Linguistic Inquiry II.

Banfield, A. (1973), "Narrative Style and the Grammar of Direct and Indirect Speech", Foundations of Language.

Bell, S. (1976). Cebuano Subjects in Two Frameworks, unpublished doctoral dissertation, M.I.T.

Bolinger, D. (1977). "Another Glance at Main Clause Phenomena", Language.

Bresnan, J. (1970). "On Complementizers: Toward a Syntactic Theory of Complement Types", Foundations of Language.

----- (1971). "Sentence Stress and Syntactic Transformations", Language.

----- (1973). "Syntax of the Comparative Clause Construction in English", Linguistic Inquiry IV.

----- (1974). "The Position of Certain Clause-Particles in Phrase Structure", Linguistic Inquiry.

----- (1975). "Comparative Deletion and Constraints onf Transformations", Linguistic Analysis I.

----- (1976), "On the Form and Functioning of Transformations", Linguistic Inquiry VII.

----- (1977a). "Evidence for a Theory of Unbounded Transformations", Linguistic Analysis II.

----- (1977b). "Variables in the Theory of Transformations", in A. Akmajian, P. Culicover, and T. Wasow, eds., Formal Syntax, Academic Press.

----- (1978). "A Realistic Transformational Grammar", in M. Halle, J. Bresnan, and G. Miller, eds., Linguistic Theory and Psychological Reality, M.I.T. Press.

-----, and J. Grimshaw (1978). "The Syntax of Free Relatives in English", to appear in Linguistic Inquiry X.

Carden, G. (1970). "A Note on Conflicting Idiolects", Linguistic Inquiry I.

Carlson, G. (1977). Reference to Kinds in English, unpublished doctoral dissertation, Univ. of Massachusetts, Amherst.

Chomsky, N. (1957). Syntactic Structures, Mouton.

----- (1962). "The Logical Basis of Linguistic Theory", in H. Lunt, ed., Proceedings of the 9th International Congress of Linguists.

----- (1964). "Current Issues in Linguistic Theory", in J. Fodor and J.J. Katz, eds., The Structure of Language: Readings in the Philosophy of Language, Prentice-Hall.

----- (1965). Aspects of the Theory of Syntax, M.I.T. Press

----- (1970). "Remarks on Nominalization", in R. Jacobs and P. Rosenbaum, eds., Readings in English Transformational Grammar, Ginn-Blaisdell.

----- (1972). "Some Empirical Issues in the Theory of Transformational Grammar", in P.S. Peters, ed., Goals of Linguistic Theory, Prentice-Hall.

----- (1973). "Conditions on Transformations", in S. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle Holt, Rinehart, and Winston.

----- (1975). "Questions of Form and Interpretation", Linguistic Analysis I.

----- (1977a). "Conditions on Rules of Grammar", Linguistic Analysis.

----- (1977b). "On Wh-Movement", In Akmajian, Culicover and Wasow.

----- (1978). "On Binding", unpublished M.I.T. paper.

-----, and M. Halle (1968). The Sound Pattern of English, Harper & Row.

-----, and H. Lasnik (1977). "Filters and Control", Linguistic Inquiry VIII, no. 3.

Chung, S. (1976). Case-Marking and Grammatical Relations in Malayo-Polynesian, unpublished doctoral dissertation, Harvard University.

Dorel, M. (1978). "Notes on PP-Extraposition", unpublished paper, Harvard and M.I.T.

Dresher, E., and N. Hornstein (1978). "Trace Theory and NP Movement Rules", to appear in Linguistic Inquiry.

Emonds, J. (1970). Root and Structure-Preserving Transformations, unpublished doctoral dissertation, M.I.T.

----- (1976). A Transformational Approach to English Syntax; Root, Structure-Preserving, and Local Transformations, Academic Press.

----- (1978). "The Verbal Complex V'-V in French", Linguistic Inquiry IX, no. 2.

Fiengo, R. (1974). Semantic Conditions on Surface Structure, unpublished doctoral dissertation, M.I.T.

Fodor, J.A. (1975). The Language of Thought, Thos. Crowell & Sons.

Geis, M. (1969). Adverbial Subordinate Clauses in English, unpublished doctoral dissertation, M.I.T.

Green, G. (1976). "Main Clause Phenomena in Subordinate Clauses", Language.

Grimshaw, J. (1977). English Wh-Constructions and the Theory of Grammar, unpublished doctoral dissertation, U. Mass., Amherst.

----- (forthcoming). "Complement Selection", to appear in Linguistic Inquiry.

Guerón, J. (1978). "The Grammar of PP Extraposition", unpublished paper, University of Paris.

Harris, Z.S. (1964). "Co-Occurrence and Transformation in Linguistic Structure", in Fodor and Katz.

Hendrick, R. (1976). "Prepositions in the X-Bar Theory", in J. Emonds, ed., U.C.L.A. Working Papers in Syntax.

Higgins, F.R. (1973). "On J. Emonds' Analysis of Extraposition", in J. Kimball, ed., Syntax and Semantics, Vol. II, Academic Press.

Hindle, D., and I. Sag (1976). "Some More on Anymore", in R. Shuy and R. Fasold, eds., Analyzing Variation in Language, Georgetown University Press.

Hooper, J., and S. Thompson (1974). "On the Applicability of Root Transformations", Linguistic Inquiry.

Horn, G. (1974). The NP Constraint, unpublished doctoral dissertation, Univ. of Mass., Amherst.

Hornstein, N. (1976). "S and the X-Bar Convention", Linguistic Analysis.

Jackendoff, R. (1972). Semantic Interpretation in Generative Grammar, M.I.T. Press.

----- (1973). "The Base Rules for Prepositional Phrases", in Anderson and Kiparsky.

----- (1974). "Introduction to the X-Bar Convention", distributed by the Indiana University Linguistics Club.

----- (1975). "Morphological and Semantic Regularities in the Lexicon", Language.

----- (1977). $\bar{X}$-Bar Syntax: A Study in Phrase Structure, M.I.T. Press.

Kaufman, E. (1974). "Navajo Spatial Enclitics: A Case for Unbounded Rightward Movement", Linguistic Inquiry IX, no. 4.

Kayne, R. (1975). French Syntax, M.I.T. Press.

-----, and J.-Y. Pollock (1977). "Stylistic Inversion, Successive Cyclicity, Move NP, Move Wh", to appear in Linguistic Inquiry IX, no. 4.

Koster, J. (1977). "Why Subject Sentences Need Not Exist", to appear in S.J. Keyser, ed., Recent Transformational Studies in Europe, M.I.T. Press.

Kuno, S. (1973). "Constraints on Internal Clauses", Linguistic Inquiry IV.

Lakoff, G. (1970). "On Generative Semantics", in D. Steinberg & L. Jakobovits, eds., Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology, Cambridge University Press.

Lasnik, H. (1976). "Remarks on Coreference", Linguistic Analysis II.

Liberman, M. (1974). "On Conditioning the Rule of Subject-Aux Inversion", in E. Kaisse & J. Hankamer, eds., NELS V.

-----, and A. Prince (1977). "On Stress and Linguistic Rhythm", Linguistic Inquiry VIII, no. 2.

Maling, J. (1976). "Notes on Quantifier-Postposing", Linguistic Inquiry.

-----, and A. Zaenen (1977). "TOP-Position", talk presented at the Winter meeting of the Linguistic Society of America, Chicago, Ill.

McCawley, J. (1968). "The Role of Semantics in a Grammar", in E. Bach and R.T. Harms, eds., Universals in Linguistic Theory, Holt, Rinehart, & Winston.

Milsark, G. (1974). Existential Sentences in English, unpublished doctoral dissertation, M.I.T.

Obenauer, H. (1976). "A-Sur-A et les Generalizations Trans-Categorielles", Paper presented at the Colloquium on Romance Linguistics, University of Montreal.

Oehrle, R.T. (1975). The Grammatical Status of the English Dative Alternation, unpublished doctoral dissertation, M.I.T.

Peters, P.S., & R.W. Ritchie (1973). "The Generative Capacity of Transformational Grammars", Information Sciences.

Partee, B. (1970). "On the Requirement that Transformations Preserve Meaning", in C. Fillmore and D.T. Langendoen, eds., Studies in Linguistic Semantics, Holt Rinehart & Winston.

Postal, P. (1970). "On Coreferential Complement Subject Deletion", Linguistic Inquiry.

----- (1972). "The Best Theory", in Peters.

----- (1974). On Raising, M.I.T. Press.

Quicoli, A.C. (1976). "Conditions on Quantifier Movement in French", Linguistic Inquiry

Reinhart, T. (1976). The Syntactic Domain of Anaphora, unpublished doctoral dissertation, M.I.T.

Van Riemsdijk, H. (1977). The Binding Nature of Prepositional Phrases, unpublished manuscript, University of Amsterdam.

Rizzi, L. (1978). "Violations of the Wh-Island Condition in Italian and Subjacency", unpublished paper.

Ross, J.R. (1967). Constraints on Variables in Syntax, unpublished doctoral dissertation, M.I.T.

----- (1970). "Auxiliaries as Main Verbs", in W. Todd, ed., Studies in Philosophical Linguistics, Great Expectations Press.

Sag, I. (1976). Deletion and Logical Form, unpublished doctoral dissertation, M.I.T.

----- (to appear). The Syntax and Semantics of Verb Phrase Deletion, Elsevier Press.

Schwartz, A. (1972). "The VP-Constituency of VO Languages", in J. Kimball, ed., Syntax and Semantics, vol. I, Academic Press.

----- (1975). "Verb Anchoring and Verb Movement", in C. Li, ed., Word Order, Univ. of Texas Press.

Selkirk, E. (1970) "The Determiner Systems of Noun Phrase and Adjective Phrase", unpublished paper, M.I.T.

Sportiche, D. (1978). Syntax Generals Paper, M.I.T.

Taraldsen, K.T. (1978). "Violations of the Wh-Island Condition in Norwegian", unpublished paper, M.I.T.

Weinberg, A. (1978). "On Preposition Stranding", unpublished paper, M.I.T.

Williams, E. (1974). "Small Clauses in English", in J. Kimball, ed., Syntax and Semantics, Vol. IV, Academic Press.

----- (1975). "Comparative Reduction and the Cycle", unpublished paper, Univ. Mass., Amherst.

Woisetschlaeger, E. (1976). "Degree-Phrase Shift", unpublished paper, M.I.T.

Zwaarts, F. (1975). "Some Remarks on the Linear Cycle in Dutch Syntax", unpublished paper, Univ. of Groningen.

BIOGRAPHICAL SKETCH

The author was born April 29, 1950, in Philadelphia, Pa., the youngest of three children. After graduating from Central High School in 1967, he went to McGill University, where he got a B.A. with 1st class honours in linguistics in 1971. After working for William Labov as a research assistant during the academic year 1971-1972, he was a graduate student at University of California, Irvine, in the Cognitive Sciences Program during 1972-73, where he studied linguistics and cognitive psychology, while T.A.-ing the intro syntax course as well as courses in language acquisition and psycholinguistics. The receipt of an N.S.F. graduate fellowship sent him back to the University of Pennsylvania to work with Labov again. After getting his M.A. from Penn in 1975, he enrolled in the Ph.D. program in linguistics at M.I.T., his final graduate school, where he has spent the last three years. In the fall of 1978, he will move to New York City, where he will join the faculty of New York University as an assistant professor of linguistics. His publications include:

"On the Cyclicity of Extraposition", in Linguistic Inquiry 6, no. 3 (1975).

"Quantifier-Negative Interaction" in Studies in Language Variation, R. Shuy and R. Fasold, Eds., Georgetown University Press (1977a).

"PP As a Bounding Node", in M. Stein, Ed., Proceedings of the Eighth Regional Meeting of the Northeastern

Linguistic Society (NELS VIII), U. Mass., Amherst (1977b).