

011
023

D
00000059602

00000059602

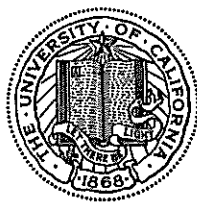
The Variable Interpretation Convention:

*A Condition on Variables
in Syntactic Transformations*

Wendy K. Wilkins

ifornia
onal
ity

UIC



THE LIBRARY
OF
THE UNIVERSITY
OF CALIFORNIA
LOS ANGELES

THE VARIABLE INTERPRETATION CONVENTION:
A CONDITION ON VARIABLES IN SYNTACTIC TRANSFORMATIONS

by

Wendy Karen Wilkins

C. E. L. L.

EL COLEGIO DE MEXICO

Reproduced by the
Indiana University Linguistics Club
310 Lindley Hall
Bloomington, Indiana 47405
November, 1979

To Nettie Wilkins, who often knows the
road I will travel long before I realize
I am off on a journey--but who knows me
well enough never to let on too soon.

© by Wendy Karen Wilkins, 1979

TABLE OF CONTENTS

	Page
Acknowledgments	vii
Abstract	ix
Chapter 1. Introduction	1
Chapter 2. A Preliminary Condition on Internal Variables	5
2.1 Passivization: An Initial Formulation	5
2.2 The Lower-S Constraint	12
2.3 A Condition on Non-Variables	13
Footnotes to Chapter 2.....	15
Chapter 3. Raising	17
3.1 Subject Raising	17
3.1.1 A Preliminary Statement of the Subject Raising Rule	17
3.1.2 Grosset Constituent Analysis and the Variable Interpretation Convention	21
3.2 The Questionable Status of Subject Raising to Object	24
3.3 Object Raising	28
3.3.1 Object Raising Predicates and Subcategorization ...	28
3.3.2 c^{\max} and the Raising Rule	31
3.4 Raising, VP', and the Principle of the Transformational Cycle	36
3.5 Raising to Object and the VIC	54
3.6 Alternatives to a Global Constraint on Raising	57
Footnotes to Chapter 3	64
Chapter 4. WH-fronting	71
4.1 The Rule for WH-fronting	71
4.1.1 The Revised Left Branch Condition	71
4.1.2 WH-terms in the Grosset Analysis	75
4.2 <i>That</i> -complements and Relative Clauses	79
4.3 <i>For-to</i> Clauses	89
4.4 Extraction from NP	98
4.4.1 The Revised WH-fronting Rule	98
4.4.2 Gerundive Nominals	99
4.4.3 Presuppositional Anaphoricity	103
4.5 Multiple WH's	106
4.5.1 The Variable Interpretation Convention--Final Version	106
4.5.2 An Alternative to Successive Cyclicity	108
Footnotes to Chapter 4	115

Chapter 5. Passive	121
5.1 Pseudo-Passives	121
5.2 The Revised Account of Passivization	124
5.2.1 Restatement of the Passive Rule	124
5.2.2 An Alternative to A-over-A	127
5.2.3 Passive and Indirect Object Movement	131
5.3 The Transformational Cycle, Raising to Object, and a Rule of Control	132
5.3.1 An Alternative to Raising to Object	132
5.3.2 The Cyclic Nature of Passivization	138
5.3.3 Control Properties and the VIC	144
5.3.3.1 A VIC Formalization of the Minimal Distance Principle	144
5.3.3.2 An Alternative to the Minimal Distance Principle	154
5.4 The Interaction of Passive and Raising	166
Footnotes to Chapter 5	184
Chapter 6. Rightward Versus Leftward and Local Versus Non-local Movements: Some Speculations	187
6.1 Extraposition	187
6.2 The Boundary Attachment Condition and The Fixed Head Principle	196
6.3 Complex NP Shift	198
6.4 Two Hypotheses	202
Footnotes to Chapter 6	209
Bibliography	215

ACKNOWLEDGMENTS

First of all, and above all, I acknowledge the help and influence of Joe Emonds. He is not only my chairman, advisor, and friend--he taught me to think syntax. Joe, thank you.

I can now finally begin to acknowledge and thank Arthur Stickgold. He provided me with boundless emotional support during my impossible evolution from linguistics undergraduate to linguistics department faculty member. He was always properly both sympathetic and skeptical when I claimed that "nothing is working right". Also, Art, thank you for the many "illegal acts"--way below and beyond the call of duty.

To Dollie Meyers first, of all my friends, I give my thanks. She was always "there" for me, was never jealous of my ambitions or achievements, and helped me know that it's okay to "choose RED first". To Leon Jacobson I present the coveted White Stallion Award for the last minute coming-to-the-rescue. I would also like to thank other West Coast linguist/friends. For both support and valuable discussions about syntax I thank Paul Schachter, Peter Culicover, Carlos Otero, Sharon Klein, Randy Hendrick, and Tom Peterson. For friendship and many wonderful (broadening) hours of non-syntax linguistics I thank Cathe Browman, Louis Goldstein, Sharon Sabsay, Pam Munro, and Brent de Chene. For being sometimes a teacher--and always a friend--I thank Peter Ladefoged. To Vicki Fromkin I must give special thanks, for the love she has shown me and the confidence she has had in me. For keeping the UCLA Department of Linguistics functioning just so I could get my degree, I thank Anna Meyer and Teddy Graham. For financial support during my whole life as a graduate student I acknowledge the Regents of the University of California; and I thank the UCLA Linguistics Department faculty, in particular Robert Stockwell who was chairman at the time, for making that financial support possible.

Sharon Barr--thank you for agreeing to do the typing (and the tree drawing). You just don't qualify as "my typist". Thank you for your hand-holding and caring, even during the time you didn't care.

Sheldon Kardener--is there some way of thanking you for me? I know, you merely "helped the rose to grow itself".

Next I would like to acknowledge the help of my East Coast linguist/friends. In particular I thank Joan Bresnan for her friendship, for her help in making me feel so at home so soon, and especially for her support during those times when her confidence in me far surpassed the confidence I had in myself. I thank her also for often reminding me that "it's easy to be clever". I thank Noam Chomsky for his inspiring lectures and discussions, but even more for never "pulling any punches". "It's an interesting paper, BUT THERE ARE NO CONDITIONS ON TRANSFORMATIONS." (Chomsky, September 1976)¹

¹There--does {that
 this} count as a footnote even in the acknowledgments?

I would also like to thank Alberto Rivas both for suffering through so many "clever" but wrong attempts to smooth out the rough spots and, more importantly, for showing me the wonders of the moons of Jupiter. For his friendship and many long hours of talk, both in person and long distance, I thank Joel Rotenberg. I also thank the following Boston and Amherst friends for making the winter of '77 bearable: Jane Grimshaw, Mary-Louise Kean, Ronnie Wilbur, Amy Weinberg, and Lisa Selkirk.

I am indebted to the faculty of the MIT Linguistics Department, in particular Morris Hale who was Acting Chairman, for making possible my year as "visiting scientist". And I thank Ray Jackendoff for initially suggesting that I spend some time at MIT.

Next I would like to acknowledge the support I have gotten from family members--who don't have the good fortune to live on either coast. Thank you Mother, Sandra, Scott, Robin, and Don for letting me harbor the illusion that you can go home again. I also thank my father, Berne, who left me the mixed blessing of once having said "only second!!"

Finally, I thank Breean Stickgold for all the "magic wake-up kisses" and Kira Stickgold for telling me that the way to get to the end is to "go backwards into the future, you know, go ahead to yesterday".

ABSTRACT

This dissertation presents, and argues for, a particular condition on variables in structural descriptions of transformational operations. By use of this condition, the Variable Interpretation Convention (VIC), a theoretical model is constructed in which constraints on the applicability of transformations are determined by the INTERACTION of the FORM OF RULES and the Convention. This model is suggested as an alternative to constraining transformations either by output conditions or conditions on rule functioning alone.

In Chapter 2 a preliminary version of the condition on variables is presented. Additionally, a condition on nonvariables is suggested: namely, that structural descriptions include only crucially affected terms. Chapter 3, basically a discussion of the English raising process, leads to a revision of the preliminary version of the condition on variables. This revision incorporates the definition of GROSSEST CONSTITUENT ANALYSIS, a definition on how phrase markers are analyzed. Grossest constituent analysis assures that for any transformation, with respect to any particular phrase marker, the variable material of the structural description corresponds to the highest (= least embedded) analysis between the nodes of the phrase marker which correspond to the nonvariable terms of the structural description.

The discussion of WH-fronting (Chapter 4) shows that this rule is to be considered one which, in principle, can function over an unbounded domain. The discussion of the WH-fronting process leads to the final version of the Convention:

The Variable Interpretation Convention:

Given $A - B$, where A and B are crucially affected terms of the structural description of a transformational operation T , first T functions where A and B are strictly adjacent; then T may function where A and B are weakly adjacent. A and B are weakly adjacent where, for all non-optional A and B , $A - B \equiv A - X - B$ where X corresponds to the grossest constituent analysis of a phrase marker and X does not contain any A or B or head of an A or B .

Chapter 5 presents a formal account of English passivization and its interaction with other processes. In the discussion in Chapter 5 it is shown that (at least for the movement rules discussed in this dissertation) there is no need for a statement of the principle of the transformational cycle. It is also shown that (for the rules discussed) the A-over-A condition and certain semantic properties of "control" follow from the VIC.

Chapter 6 discusses rightward movements (of S and complex NP's) and shows that upward bounding is predictable from the VIC. In Chapter 6,

additionally, some speculations are made about directions for future research using the VIC. In particular, suggestions are made about a predictable dichotomy between leftward and rightward movements, and a predictable distinction between strictly local movements and movements over a variable.

CHAPTER 1

INTRODUCTION

In the past fifteen years a large proportion of the syntax literature has dealt with the problem of properly constraining transformational operations for extraction, that is, movement and deletion rules. For syntactic transformations to be written in optimally simple form much attention must be paid to the issue of language-independent, general theoretical conditions under which these transformations may be said to apply.

As early as 1962, in "The Logical Basis of Linguistic Theory", in discussing the notion of explanatory adequacy, Chomsky gives a general formulation of perhaps the best-known and most general of all the proposed constraints on transformational operations. This is the constraint which has since become known as the A-over-A condition. The sentences Chomsky discusses are given here as (1) and (2).

(1) a. Whom did Mary see walking to the railroad station?

b. Do you know the boy who(m) Mary saw walking to the railroad station?

(2) Mary saw the boy walking to the railroad station.

Chomsky points out that (2), but not the examples of (1), is ambiguous. The two possible analyses for (2) would be:

(3) a. NP - Verb - NP - Complement

b. NP - Verb - NP

where (3b) would be for the reading of (2) where the object NP has a reduced restrictive relative clause. That the sentences of (1) are not ambiguous in the way that (2) is, is the interesting fact to be accounted for. The general principle which accounts for this, and which explains why the sentences of (1) can only be derived from a structure like (3a), is presented by Chomsky (1962:931) in the following way:

... in the case of (2) with the structural description (3b), this specification is ambiguous, since we must determine whether the second NP--the one to be prefixed [in the derivation of the sentences of (1)]--is "the boy" or "the boy walking to the railroad station", each of which is an NP. Since transformations must be unambiguous, this matter must be resolved in the general theory. The natural way to resolve it is by a general requirement that the dominating, rather than the dominated, element must always be selected in such a case. This general condition, when

appropriately formalized, might then be proposed as a hypothetical linguistic universal. What it asserts is that if the phrase X of category A is embedded within a larger phrase ZXW which is also of category A, then no rule applying to the category A applies to X (but only to ZXW).

With respect to (1) and (2), what this does is to prohibit question formation or relative clause formation from applying to the NP-dominated NP *the boy*. Therefore the sentences of (1) can only be interpreted as having been derived from a structure like (3a) and, hence, are not ambiguous.

Even in this early a paper Chomsky notices that this constraint on the application of transformations is not accurate. He says (1962:933):

Although this account still leaves much unsaid, and several qualifications are necessary, the principle seems well-supported and formally well-motivated, and thus can be proposed as a general hypothesis concerning linguistic structure, to be tested in terms of the consequences to which it leads in various languages.

Ross (1967) explored in great detail the predictions made by the A-over-A condition and the extent to which it was in fact inaccurate, or at least in need of modification. Ross discusses counterexamples to A-over-A which Chomsky himself pointed out (Chomsky 1964:46, fn. 10) where a question word NP is moved out of a bigger NP (as in *Who would you approve of my seeing?*). Ross also gives many cases where a transformation such as relative clause formation can move either "a dominated NP or any one of an unbounded number of NP's which dominate it" (Ross 1967:9). In a very thorough discussion Ross shows many cases where A-over-A is apparently supported but also many ways in which it is too strong and in which a properly weakened version is then not adequate.

The alternative to A-over-A presented by Ross is actually a group of constraints on transformations which reorder elements. These constraints on the functioning of rules (see particularly Chapter 4 of Ross 1967) are The Complex NP Constraint, The Coordinate Structure Constraint, The Pied Piping Convention, The Sentential Subject Constraint, and The Left Branch Condition.

Another very important contribution to the "constraints literature" was made by Bresnan (1972). This constraint is called the Fixed Subject Constraint and was formalized (p. 308) as:

No movement rule T may move a (part of a) subject if it lies next to a complementizer not mentioned in the structural description of T.

What a condition of this sort is able to account for, among other things, is a grammaticality difference such as is illustrated in (4):

- (4) a. What did they believe that he did?
 b. *Who did they believe that did all the work?

In her recent work Bresnan has revised and replaced the Fixed Subject Constraint by the Complementizer Constraint on Variables (Bresnan 1976d) (discussed here in Chapter 4). Bresnan has also suggested a revision of the A-over-A condition (1976a) where its applicability can only be determined relative to the process under consideration. Where Ross's constraints were basically conditions on rule function, Bresnan's conditions make crucial use of the form of the rules.

Chomsky (1973) presents a set of constraints which has come to be known as the CONDITIONS framework. Repeatedly in work since 1973 Chomsky has shown the generality and usefulness of several of the conditions from this work, in particular the Specified Subject, Subjacency, and Tensed-S Conditions. This framework has evolved into a model (Chomsky and Lasnik 1977) in which all rules apply completely freely and where the results can be marked as ill-formed or uninterpretable by virtue of certain output conditions.

These different approaches to constraints on transformations are mentioned here in this brief introduction because the thesis of this dissertation is the proposal that many of the facts currently accounted for in the syntax literature by the various separate conditions can actually be accounted for by a single, simple convention for the interpretation of variables in syntactic transformations. This can be done, I will argue, by recognizing the importance of THE FORM OF STRUCTURAL DESCRIPTIONS AS IT INTERACTS WITH A GENERAL CONVENTION ON RULE APPLICATION.

It is interesting to note here the different direction taken by Chomsky in his recent work where structural descriptions have essentially disappeared from the statement of transformations. By comparison, the proposal presented in this dissertation is formally more similar to work by Bresnan where great attention is paid to the form of particular rules (see particularly Bresnan 1976a, 1976d). I will show in the pages which follow that the apparent need for separate conditions on transformations is a function of overlooking the crucial INTERACTION between the form and functioning of rules (cf. Chomsky 1973: section 1).

The general convention which will be proposed in this thesis is to be considered as an alternative to constraining transformations either by output conditions or separate conditions on rule function. In other words, this dissertation is presented as an alternative to Ross's constraints or the Conditions framework. Additionally, it will be argued that by using the Structure Preserving Hypothesis (Emonds 1970, 1976) and the Revised Left Branch Condition (Emonds 1976), along with the proposal of this dis-

sertation, there will be no need for statements in the general theory of either the A-over-A condition or the condition of cyclic rule application. A-over-A and the cycle are natural consequences of the overall framework presented here.

The general convention which will be proposed here is one which will allow variables to be eliminated in transformations. By "eliminating" I mean that an appropriate general convention can be formulated which will predict in each case what material can intervene between constituents corresponding to nonvariable terms of a structural description. Since this prediction will be based, in each case, on what occurs as the nonvariable terms of the structural description, the conditions under which transformations apply are then predictable from the form of the transformations themselves. This is a strong claim which, if justifiable, is equivalent to saying that, in accord with a theoretical convention, rules constrain their own application. This approach to constraining syntactic processes is similar to recent work in phonology by Vergnaud, Prince, and Halle (this is pointed out in particular in Chapter 4).

In Chapter 2 we begin looking at the process of passivization in English to consider the first approximation of the convention on variable interpretation. In Chapter 3, with a definition on how phrase markers are to be analyzed, the convention is refined, and in Chapter 4 the Variable Interpretation Convention is stated in its final form. Chapter 5 shows how the A-over-A condition and the principle of the transformational cycle (at least as they are relevant to the movement processes considered in this work) follow from the Variable Interpretation Convention. Chapter 6 (and to some extent also Chapter 5) suggests promising directions for future research into the use of the theoretical model presented in this dissertation.

CHAPTER 2

A PRELIMINARY CONDITION ON INTERNAL VARIABLES

2.1. Passivization: An Initial Formulation

A simple statement of the passive transformation for English could well be that given in (1).

$$(1) \text{ NP} - \text{X} - \text{V} - \text{Y} - \text{NP} - \text{Z} \quad \Rightarrow$$

$$5 - 2 - be+3+en - 4 - \emptyset - 6+(by+1)$$

There are of course many questions raised by any statement of a transformation, here for example, concerning the nature of structure-building or morphology-introducing processes, or the theoretical status of a rule which performs more than one operation simultaneously. For now, however, we will work with this formulation of the rule (a more sophisticated account of passivization will be presented in Chapter 5) to begin to consider what can occur contained in the internal variables, that is, variables X and Y.

In the discussion of variables, in both this chapter and what follows, I am making certain assumptions, without explicitly arguing for them, about the nature of the phrase markers to which transformations can apply. I am assuming except where otherwise stated, that the rules apply only to possible English base structures or to the output of structure preserving rules. Making this common sense assumption facilitates exposition since it is then unnecessary to repeatedly exclude possibilities which would never be considered for English.

Considering variable X, rule (1) can obviously apply where X contains an auxiliary (AUX). That X may also contain an adverb is shown by the grammaticality of the passive of (2).

(2) a. Harry could scarcely solve the problem.

b. The problem could scarcely be solved by Harry.

Since the adverb in this sentence is of the type which appears only preverbally, it cannot be plausibly argued that it is moved into preverb position subsequent to the application of the passive rule. It must be that the rule applies where the variable contains an adverb. For a discussion of these adverbs see Emonds 1976, Chapter V.

Example (3) shows that X may contain the negative morpheme (NEG) and the emphatic morpheme (EMP). When EMP, usually *so* or *too*, occurs with NEG it appears as *either*.

- (3) a. Harry didn't either solve that problem.
 b. (I beg your pardon,) that problem wasn't either solved by Harry!

Variable X cannot, however, contain a noun phrase or a noun. This is shown in (4).

- (4) a. The king of England's brother kicked the dog.
 b. *The dog's brother was kicked by the king of England.
 c. The man $\left\{ \begin{array}{l} \text{, who I saw,} \\ \text{(that) I saw} \end{array} \right\}$ caught the ball.
 d. *The ball $\left\{ \begin{array}{l} \text{, who I saw,} \\ \text{(that) I saw} \end{array} \right\}$ was caught by the man.

Example (4d) is the result of passive applying over variable X where it contains, among other things, an NP (I); (4b) is the result of passive applying over an N in X (*brother*). Rule (1) applies properly only where the NP of term 1 is construed to mean the whole first NP, that is, *the king of England's brother* or *the man* $\left\{ \begin{array}{l} \text{, who I saw,} \\ \text{(that) I saw} \end{array} \right\}$. This is the type of example where the A-over-A condition (Chomsky 1962, 1964; Ross 1967) makes exactly the correct stipulation. A rule which moves an NP must be so construed as to move the biggest or highest NP (other things being equal, that is, unless the rule specifies otherwise). With respect to the examples of (4), because of the configuration of NP's in English, saying that X may not contain an NP or an N gives the same result as the A-over-A condition.

Consider additionally now the ungrammatical examples of (5).

- (5) a. *The king of England's $\overbrace{\text{the dog was kicked by brother.}}$
 b. *The man $\left\{ \begin{array}{l} \text{, who} \\ \text{(that)} \end{array} \right\} \overbrace{\left[\begin{array}{l} \text{the ball saw,} \\ \text{the ball saw} \end{array} \right]} \text{ was caught by me.}$
 c. The man from Chicago caught the ball.
 d. *The ball $\overbrace{\text{from Chicago was caught by the}} \text{ man.}$
 e. *The man from $\overbrace{\text{the ball was caught by Chicago.}}$

Examples (5a) and (5d) will be excluded in any account of passive where the rule is written to postpose an NP in the sense of N with the maximal number of bars. (For an indepth discussion of bar theory see the recent works of Jackendoff, particularly 1976). The NP in (1) is meant to indicate

the movement of the phrasal category whose head is N and where N can occur with the full range of specification, modification, and complementation. *Brother* in (4) and (5) is not an NP in this sense, nor is *the man* in (5c). They are not NP's in the sense of N^n where n = the maximal number of bars. For instance, *brother* in (4a) cannot occur in this configuration with a specifier. It is an N internal to an NP which has the possessive NP as its determiner. *Brother*, here, can occur with modifiers and complements so it would be an NP in the sense of N^{n-1} , or N with some less-than-maximal number of bars. With use of bar notation and the usual notion of 'head of a phrase' (where N would be the head of N^n) it is possible to restate the claim made above. The claim that X for rule (1) may not contain a noun phrase or a noun may be rephrased as 'X may not contain any N^n or head of an N^n '.

With respect to examples (5b) and (5e), a condition such as A-over-A accounts for their ungrammaticality just as was the ungrammaticality of (4b) and (4d) accounted for. In (5b) and (5e) the passive rule has violated A-over-A by moving an instance of N^n which was not the highest N^n . For (5e), clearly, a constraint on what can occur in variable X is irrelevant since variable X would have been empty. We will not discuss examples like (5e) in depth now but will return to them in Chapter 5 where more discussion of passivization is presented and this type of example will be accounted for without the use of the A-over-A condition. Notice, however, that (5e) cannot be prevented as was (5a) because *Chicago* here is in fact a full N^n .¹

In considering the ungrammaticality of (5b), where in both cases X contained the verb *saw*, we will be lead to the simple fact that not only may X not contain an NP or an N but neither can it contain a verb. Continuing on, the sentences of (6) are NOT examples of the active-passive relationship.

- (6) a. Mary wanted to hit John.
 b. John wanted to be hit by Mary.

Example (6b) would be derived from a structure where the embedded clause of *John; wanted [Mary hit John;]* undergoes passivization and then subsequently the derived subject, *John*, is deleted under identity to the subject of the matrix clause. It could not be derived directly from (6a).² Since (6a) does however meet the structural description of (1), where *Mary* is term 1, *hit* is term 3, and *John* is term 5, (6b) is prevented as a result of (1) applying to (6a) by saying that X may not contain a verb, here namely the verb *wanted*.

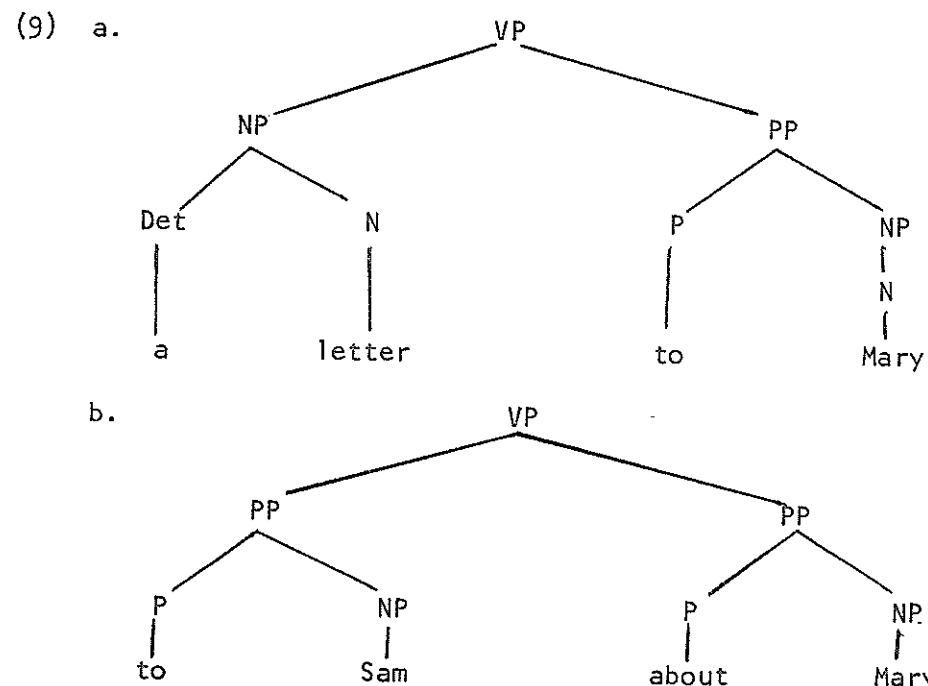
If we now shift attention to variable Y of the passive rule we will see that it cannot contain a verb either.³ That is, example (7), which would result if *wanted* were analyzed as the V of the rule and *hit* were contained in Y, is not acceptable.

- (7) *John was wanted to hit by Mary.

Continuing an examination of variable Y we will see that additionally it cannot contain any NP.

- (8) a. {John talked to Sam about Mary.}
 {John wrote a letter to Mary.}
- b. {*Mary was talked to Sam about by John.}
 {*Mary was written a letter to by John.}
- c. {Mary was talked about by John.}
 {Mary was written to by John.}

In (8b) where passive applied, Y contained {to Sam about} and the result is unacceptable. But in (8c) where Y contained just the preposition, to or about, and no NP, then the result of passive is fine. In this discussion of the examples of (8) I am assuming an analysis of indirect objects like that used by Emonds (1976) where the indirect object prepositional phrase is considered a sister to the direct object NP. In this account the PP about Mary would also be a sister to the PP to Sam. The structures of the VP's of (8a) would therefore be as given in (9).



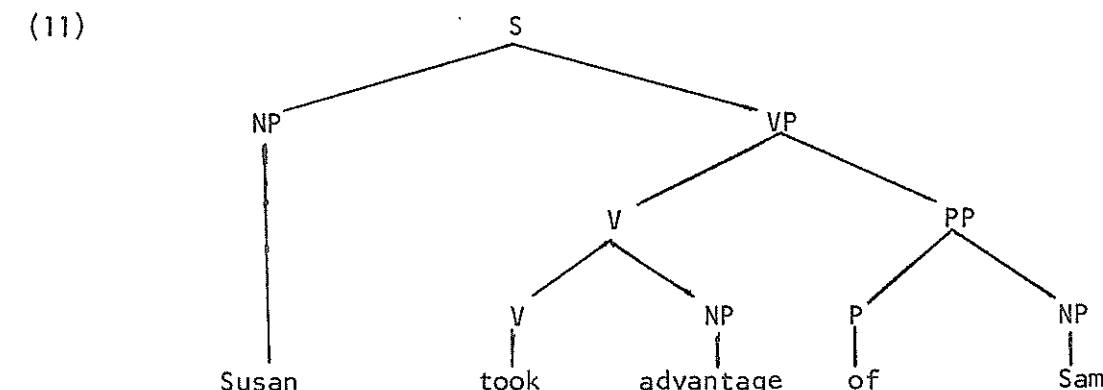
It is clear from this structure that the ungrammaticality of (8b) cannot be attributed to a violation of a principle like A-over-A since the NP dominating Mary is dominated by a PP, not another NP. Any condition to account for the status of (8b) would seem to have to refer to the LINEAR ORDER of elements occurring after the verb and not just hierarchical structure. (Osvaldo Jaeggli

has called to my attention an unpublished paper by Fiengo and Gitterman (n.d.) in which they refer to just such a condition, the A-before-A condition, suggested to them by Howard Lasnik.) It seems to be the case that passive must apply to the first NP following the verb. Or, in other words, Y may not contain an NP.

It might seem at first that an example like (10c) (pointed out to me by Paul Schachter) allows passive to function properly over an NP in variable Y.

- (10) a. Susan took advantage of Sam.
 b. Advantage was taken of Sam by Susan.
 c. Sam was taken advantage of by Susan.

If it were the case that advantage in (10a) were a normal direct object then the grammaticality of (10c) would argue that passive can apply over NP's. I think, however, that the structure of (10a) is actually as given in (11) where take advantage is considered a complex verb.



The passive rule then evidently can analyze either V as the third term of the structural description; where the higher V is so analyzed the result is (10c), otherwise it would be (10b).

Another possible counterexample to this generalization about variable Y not containing an NP (also pointed out to me by Paul Schachter) is provided by the dialect of British English where (12) is well-formed.

- (12) It was given me by Sam.

This might in fact be a counterexample. Alternatively, it might be that the result of indirect object movement for some dialects is an NP encliticized onto the verb. The fact that sentences like (12) seem better when the indirect object is phonologically reducible, that is, when it is a pronoun of some sort, lends support to this alternative.⁴

Another example where Y must be prevented from containing an NP and where an analyzability condition must be sensitive in some way to linear order is illustrated in (13).

- (13) a. The committee elected Joan president.
 b. Joan was elected president by the committee.
 c. *President was elected Joan by the committee.
 d. *Joan president was elected by the committee.

The status of examples (13b) and (13c) shows that Y again may not contain an NP. The ungrammaticality of (13d) shows that, of course, *Joan president* in (13a) is not itself an NP constituent. The ungrammaticality of (13c) cannot be attributed to a condition which just disallows extraction of an NP out of another NP.

It seems then that a generalization about what can be contained in variables X and Y of the passive rule might be "no NP's or V's". The passive transformation therefore could be written including condition (14).

- (14) X, Y does not contain any NP or V (where NP means any NP or head of an NP).

There must, of course, be some special consideration for sentences with verbs showing perfective or progressive aspect where in fact passive does apply over a V in the variable, according to some analyses of *be* and *have* where they are considered verbs in these constructions (e.g. Emonds 1976, cf. Chomsky 1957). For instance, *John had eaten the apple* can become *The apple had been eaten by John*. Verbs which are never auxiliaries, i.e. almost all verbs, may not be contained in the variables. This can be accounted for by including some feature, +F, perhaps [+lexical content], on the V of the passive rule, and hence in (14), so that examples like (7) would be excluded (because *hit* would be +F) but not those sentences with progressive *be* or perfective *have*. That is, a -F verb would be allowed in the variable. There is independent justification for including the feature +F on the V of the passive rule because any treatment of passive must be able to exclude from the domain of the rule sentences with *be* plus a predicate nominative or *have* followed by an NP. If *be* and *have* were said to share a particular feature, or rather together not contain a particular feature, it would then not appear to be accidental that they behave in ways very similar to each other. It would not be treated as coincidental that *be* and *have* do not allow passivization whether they are functioning as main verbs or as auxiliaries. Lacking the proper feature on the verb, sentences like *Susan is a doctor* or *Sara has a book* would not meet the structural description of the passive rule.⁵

In an attempt to make broader use of condition (14) we might wonder whether it is always the case that variables in transformations may not contain NP's or V's. We can see immediately however that this is not the case, since other rules, WH-fronting for instance, can freely move terms over NP's or verbs. What is an interesting fact, though, about WH-fronting is that a WH term may not be fronted over another WH term (Chomsky 1964).

- (15) a. *Who does John like the man who saw Ø?
 b. *What does John forget when to do Ø?
 c. *What did John remember why he bought Ø?
 d. *Who did the man who saw Ø talked to John?
 e. *What did the man who ate Ø ordered dessert?

I will return to a discussion of examples like these in Chapter 4, the chapter on WH-movement, but for now it is enough to notice that a WH does not move over another WH in a variable. It is also a fact that just as NP and V are mentioned in the structural description of the passive rule, WH is mentioned in the structural description of the WH-fronting rule. The generalization I suggest is that a variable may not contain an item that is the same as a term in the structural description, or in other words:

- (16) In the structural description of a transformation
 $A - B \equiv A - X - B$ where X does not contain any A
 or B or any head of an A or B.

(This notion of "contain" will be refined below in Chapter 3 because, of course, generalization (16) must not exclude perfectly good sentences such as *What did the man who ordered the big dinner have for dessert?* where without a precise notion of "contain", the variable in WH-fronting would seem to contain a WH.)

Up to this point principle (16) effectively accounts for the acceptability facts presented. It is basically this principle, with certain modifications, which provides the backbone, the main thesis, of this study.⁶ Its level of generality and power will be shown repeatedly in subsequent discussion, but already there is an interesting result of this approach.

Let us return for a moment to the examples of (8) and add to them the examples of (17).

- (17) a. Mary was written a letter by John.
 b. *A letter was written Mary by John.

What the examples of (8) and (17) together show is that if an indirect object is moved in front of the direct object it can then become the passive subject. Once this movement of the indirect object has taken place (and I take the non-occurrence of the preposition *to* to be evidence that indirect object movement has taken place) then the direct object can no longer be affected by passivization (as was pointed out by Fillmore (1965)). Only the first NP following the verb may become the passive subject--the variable *Y* may not contain an NP. The effect of principle (16) therefore is to allow us to avoid extrinsically ordering the rule of indirect object movement to precede the passive rule. An alternative account of these facts would be an ordering statement.⁷

2.2. The Lower-S Constraint

There are certain other conditions under which passive must be prevented from applying where a constraint on variables *X* and *Y* will be irrelevant. Consider the following examples:

- (18) a. John $\left\{ \begin{array}{l} \text{persuaded} \\ \text{promised} \\ \text{expected} \end{array} \right\}$ Mary to hit the ball.
- b. John $\left\{ \begin{array}{l} * \text{persuaded} \\ * \text{promised} \\ \text{expected} \end{array} \right\}$ the ball to be hit by Mary.

In (18a) where the passive rule could analyze *Mary* as term 1, *hit* as the verb, and *the ball* as the second NP, there is nothing contained in *X* or *Y* to account for the ungrammaticality of the first two examples of (18b).

The acceptability facts of (18) might be accounted for by some version of a condition suggested by Chomsky (1965), which has been referred to both as the Lower-S Constraint and as the Insertion Prohibition. This condition, suggested originally to account for certain facts about reflexivization, prohibits the introduction of morphological material into a configuration dominated by *S* after the cyclic application of rules to that configuration has been completed. In an analysis including the assumption that the VP is the head of the *S* (following Jackendoff 1974) and a convention about pruning stating that a node does not prune if it still dominates its head (following Kuroda 1965) then the infinitives in (18) would still be dominated by *S*. The Insertion Prohibition then would prevent any material from the higher matrix *S* (such as the object of *persuade* or *promise* or the agentive preposition *by*) from being inserted into the embedded *S*. The acceptability of the *expected* example of (18b) is then accounted for because passivization applies on the lower cycle and *the ball* is only subsequently raised to object in the matrix *S*.

This account, using the Insertion Prohibition, clearly makes crucial use of the assumption that the agent *by*-phrase gets lowered into the VP by passive. The *by*-phrase could as well be attached as a sister to the infinitive VP.⁸ Regardless of where the agent phrase would be properly attached in the examples of (18), it would certainly be the case that the *be-en*, which in this account of passive is added by the rule, would have to be inserted into the lower domain. For now the first two cases of (18b) are excluded by the Insertion Prohibition. When we return to passive and raising and their interaction, in Chapter 5, we will see that there is no longer any need for this additional constraint.

2.3. A Condition on Non-variable Terms

To return now to the main point of this chapter, the passive transformation can be written as (19) where principle (16) will account for what can intervene between the terms of the structural description.

$$(19) \text{ NP} - \text{V} - \text{NP} = 3 - \text{be}+2+\text{en} - \emptyset - (\text{by}+1) \\ +F$$

It is now the terms of the structural description along with principle (16) which determine what can occur as internal variable material. It is these terms which actually constrain the applicability of the transformation. Therefore there should be some way to prohibit the addition of unnecessary terms in the structural description just to limit applicability. In other words, the inclusion of context terms in the rule just to condition where the rule may or may not apply must be prevented. This must be prevented or this approach to handling constraints on transformations would amount to no substantial improvement over a listing of conditioning environments on the rules. The strongest prohibition then on terms of a structural description would be to exclude any one which was strictly contextual. It may turn out that this complete prohibition of context terms is too strong; however, for now, structural descriptions will be constrained to contain only terms which are crucially affected by the rule. "Crucially affected" here is meant to indicate "actually involved in the rule" or, in other words, any term which is moved, which provides the target site for a movement,⁹ or which exhibits a change in morphology.¹⁰

Principle (16) can now be revised in accordance with this restriction on non-variable terms.

- (20) In the structural description of a transformation, only terms which are crucially affected may be mentioned and for all terms mentioned $A - B \equiv A - X - B$ where *X* does not contain any *A* or *B* or head of an *A* or *B*.

In the next chapter we will discuss the movement processes of English which

raise an NP from an embedded S into a higher S. In the course of the discussion it will become clear that the notion "contain" in (20) must be carefully considered. A precise definition will be presented and after (20) is revised and given in its final form we will return, in Chapter 5, to a further discussion of passivization using the revised principle.

FOOTNOTES TO CHAPTER 2

¹Actually it will be shown at the beginning of Chapter 6 that by following a certain suggestion made by Arthur Schwartz (1972) passive can, interestingly enough, be written as just a movement of N, rather than of NP. This, of course will mean that (5a) is accounted for differently.

²In Chapter 5 examples like (6b) will not be handled by an equi-deletion process but rather by the interaction of passivization and the semantic property of "control".

³Often in statements of the passive rule no variable Y is provided for between the verb and the NP to be preposed. Variable Y must in general be empty. In the model discussed in this dissertation, where no variable will be written into any rule, there does not exist the option of just excluding Y from this particular rule.

⁴This possibility was suggested to me by Peter Culicover based on the probability that a sentence like *It was given Mary by Sam* would be less acceptable than (12).

⁵We will return to the question of a feature included in the structural description of passive and an alternative formulation of the passive rule(s) in Chapter 5. In Chapter 5 not only will an account be given where passive is prevented from applying to sentences with the verbs *be* and *have* but also where it is prevented from applying in sentences with verbs like *resemble* that just never have passive counterparts to their active constructions.

⁶Already, particularly in the work of Halle, Prince, and Vergnaud (1976), it has been shown that some very interesting results can be obtained where a principle similar to that in (16) is used in phonology. While it is certainly not necessary that phonological and syntactic processes be shown to be similar it is always gratifying to find that any proposed condition is not only of use in one aspect of linguistic theory but that it had wider ramifications as well. The work of Vergnaud and others on Q-variables in phonology represents a distinct but parallel line of research to the work of this dissertation.

Colette Dubuisson, in work I have not been able to obtain, reportedly has also begun to consider a condition similar to (16) in her research on adverbs in French.

⁷Another proposed account of the facts of (8) and (17) is that given within the Relational Grammar framework where once indirect objects become direct objects then they can go on to become subjects. The old direct object may no longer become a subject. This is an accurate description of what happens but in Relational Grammar this description is just stated as a law. By the use of principle (16) this description is not postulated but, rather, the generalization is derived from a broader convention.

⁸This was pointed out to me by Peter Culicover.

⁹It has been pointed out by Emonds (1976) and by Chomsky (class lectures 1976) that for the class of structure preserving transformations there is probably no need to indicate the target site in the structural description. It would be predictable from the structure preserving hypothesis and the category of the term which moved.

¹⁰This notion of crucially affected can, of course, very naturally be extended to include terms which are deleted and terms which control the deletion (or interpretation). There is some evidence from deletion processes that a complete prohibition against context terms is too strong. Examples (i) and (ii) (from Bresnan 1976a), with respect to VP deletion, were pointed out to me in this regard by Joan Bresnan.

- (i) *First people began pouring out of the building and then smoke began.
- (ii) First people began to pour out of the building and then smoke began to.

In order to properly account for the material which can be deleted it might be necessary to mention the context term AUX, where here *to* would have to count as AUX.

CHAPTER 3

RAISING

3.1. Subject Raising

3.1.1. A Preliminary Statement of the Subject Raising Rule

We will now turn, in this chapter, to a consideration of the processes which move NP's out of embedded clauses into higher sentences. It will become clear that principle (20) given in Chapter 2 is not accurate as it stands. It will also become clear, however, that the basic motivation, the spirit of it, is in fact accurate and that what is needed is a rigorous method for the analysis of phrase markers. In other words, we will be led, in this chapter, to consider precisely what it means for a variable term to "contain" material.

Let us begin now with a discussion of the process of subject raising in English. This rule might normally be formalized as something like:

$$(1) \text{ it} - X - V - Y - S[\text{NP} - \text{VP}] = \\ 5 - 2 - 3 - 4 - \emptyset - 6^1$$

Variable X is part of the structural description of the rule because some provision must be made for variable material which can occur between terms 1 and 3, such as an AUX or an adverb. Between terms 3 and 5, that is, variable Y, there is the possibility, for instance, of a prepositional phrase and an adverb. So from a deep structure similar to (2a), where X is *would certainly* and Y is *to them sometimes*, we can derive sentence (2b):

- (2) a. it would certainly appear to them sometimes
[giving parties be useless]
- b. Giving parties would certainly appear to them some-
times to be useless.

In a statement of subject raising such as (1) a reason for including the verb of the matrix clause is that raising is a process which is sensitive to what the main verb of the sentence is. At times the raising is sensitive to a particular adjective of the main clause.

- (3) a. it is $\left\{ \begin{smallmatrix} \text{probable} \\ \text{certain} \end{smallmatrix} \right\}$ [the girl win the game]
- b. The girl is $\left\{ \begin{smallmatrix} * \text{probable} \\ \text{certain} \end{smallmatrix} \right\}$ to win the game.

It is clearly the adjective *certain*, and not the verb *be*, which controls the raising in (3).² It seems then that the rule for subject raising needs some

provision for a feature on the V to indicate that it only operates where there is a raising verb. Additionally, it needs some way to include in its structural description an appropriately marked adjective. Since abbreviatory braces are not usually incorporated into the formalism for syntactic transformations it might be that a separate rule would be necessary for subject raising where it is conditioned by a particular adjective. Alternatively, the raising rule could refer to just a particular feature which raising verbs and adjectives would have in common.

In the formalism for syntactic rules being discussed here this question does not even arise, however. Since neither a verb nor an adjective of the matrix clause is affected by the raising process, by principle (20) of Chapter 2 neither one could be mentioned in the structural description anyway. The mentioning of V or Adj would be including a strictly contextual term.

There must be some mechanism, naturally, for specifying where raising may or may not take place. A viable alternative to the use of a feature on the verb or adjective could be a particular use of subcategorization. Lexical items (verbs or adjectives) could be subcategorized for some element, say a semantically empty *it* in the proper configuration. A verb like *appear*, which allows subject raising, would be subcategorized for a preceding *it*. A verb like *believe* which allows subject raising to object (assuming for now the existence of this controversial process) would be subcategorized for a following *it*.

Then if raising is written as a replacement of this semantically null *it*, it will only appear where such an *it* is provided for. Some provision must be made, of course, for the fact that not all non-referring *it*'s can be replaced by NP's. For instance, *It was surprising that John left* cannot be transformed into **John was surprising to leave*. There must be a way to distinguish between this *it* and the *it* that is replaced in raising. Below we will consider in detail the interaction of raising and extraposition. For now it is enough to point out that this distinguishing of elements, as to whether or not they can be replaced, is no more of a complication than marking each verb with a feature to indicate whether or not it allows raising and, if so, which kind.

There is actually some precedent for this type of alternative to the marking of verbs as to whether or not they allow certain processes. Richard Kayne (1969) uses a similar analysis in his treatment of the exceptions to Clitic Placement in French. To account for some verbs which can have à-NP complements but unexpectedly may not be preceded by the corresponding personal clitics (that is, *Je pense à lui* but **Je lui pense* where the general case would be like: **Je parle à lui*, *Je lui parle*), Kayne says that he will consider two possibilities:

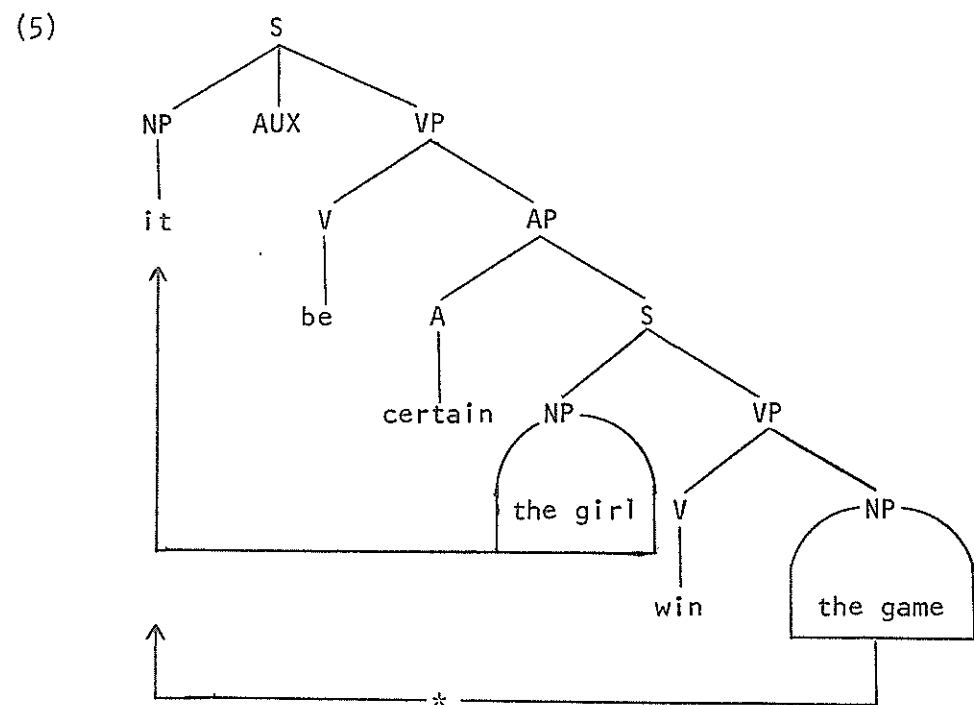
One, verbs like 'penser' are marked as exceptions to the rule of clitic-placement. This is the rule-feature hypothesis [footnote omitted]. Second, the two classes of verbs are to be distinguished in the lexicon either by postulating two distinct 'à's, or by saying that verbs like 'parler' take 'dative' complements, whereas the complements of verbs like 'penser' are true prepositional phrases. For the purposes of exposition, we shall assume two distinct 'à's, but the argument would hold equally well if 'dative' were shown to be the relevant feature. (1969:86)

Kayne is then able to go on to show not only two arguments where the rule feature hypothesis would be observationally inadequate but also that verbs like *penser* function as a natural class in the grammar, not only with respect to clitic placement. He says that "this type of generalization is of the kind predicted by the subcategorization hypothesis, but not by the rule feature hypothesis" (p. 87).

For the cases of raising verbs and adjectives in English it doesn't seem to be possible to find places in the grammar where they function as a natural class and where the subcategorization for a particular pro-form captures a generalization aside from the raising facts. The motivation here for using this mechanism is that with this approach, the marking of the *it* and the appropriate subcategorization of the verbs and adjectives, term 3 of the subject raising rule in (1) is not necessary. The rule is then revised as in (4):

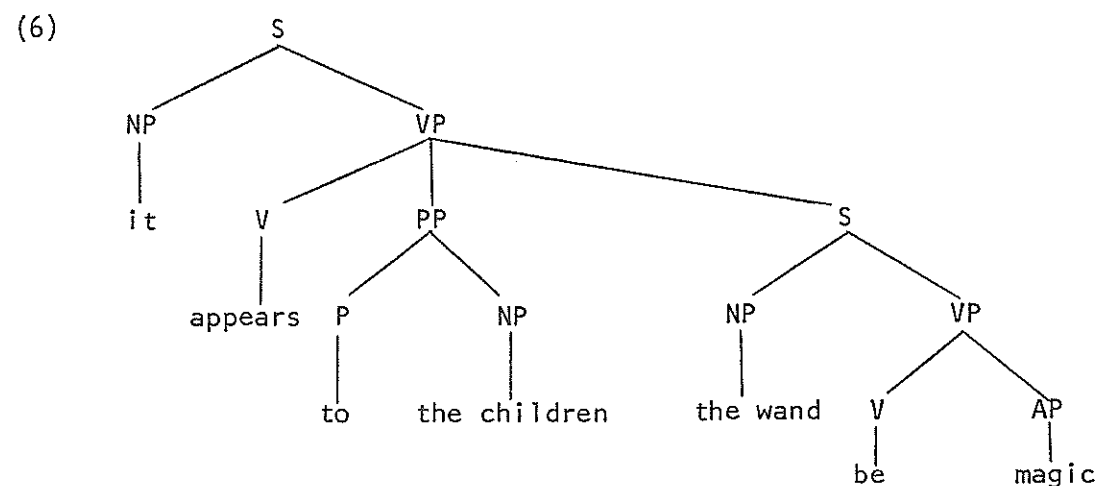
$$(4) \quad it - X - {}_S[NP - VP] = 3 - 2 - \emptyset - 4$$

We must now consider term 4 of (4), the embedded VP. It is included in the formal statement of the subject raising rule to make sure that the NP which moves is the subject NP and not an object contained in the VP. It is therefore another example of a context term which, by the principle we are following, should not be included in the rule. To see that in fact it is not necessary to include this VP in the structural description consider phrase marker (5).



Principle (20), from Chapter 2, effectively excludes the possibility of the object NP raising. Since an NP, term 3 of (4), must be mentioned in the structural description (it is what moves), no NP can occur contained in the variable. The object NP (*the game* in (5)) cannot be moved across the subject NP (*the girl* in (5)). The fact that it is a subject NP which blocks the movement of the object here is of course immediately recognizable as a case where the Specified Subject Condition (Chomsky 1973) makes exactly the right prediction. We will be returning repeatedly to discussion of the Specified Subject Condition because it turns out to be a special case of the principle argued for in this thesis, where the principle applies to noun phrase movements.

But consider now phrase marker (6).



From an underlying structure like that in (6) can readily be derived the well-formed sentence (7).

(7) The wand appears to the children to be magic.

Here the subject NP of the embedded S (*the wand*) can be raised over the NP *the children*. Thus far this constitutes a violation of principle (20).

3.1.2. Grossest Constituent Analysis and the Variable Interpretation Convention

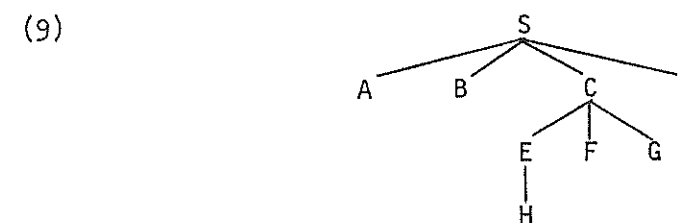
If we now compare carefully the structure of phrase marker (5) with that of phrase marker (6), we see that the NP which interferes with the raising process (the embedded subject in (5)) is, in a sense, in the 'highest direct line' between the *it* of the matrix clause and the object NP of the embedded clause. In (6) the NP *the children*, which occurs within a prepositional phrase, does not seem to lie in this type of 'highest direct line' between the *it* and the NP *the wand* which raises.

This difference, illustrated by the NP's in (5) and (6) turns out to be of critical importance. This important difference can be captured with the definition of a condition on how phrase markers are to be analyzed.

(8) Definition:

A GROSS CONSTITUENT ANALYSIS of X , with respect to $A - X - B$ but not including A or B , $= C_1 C_2 \dots C_n$ where C_i is a constituent and for every other analysis $C'_1 C'_2 \dots C'_k$, then $k \geq n$. If for each C_j in a gross constituent analysis $X_1 = C_1 \dots C_{j-1} C_j C_{j+1} \dots C_n$, there is no gross analysis $C_1 \dots C_{j-1} C'_j C_{j+1} \dots C_n$ such that C'_j dominates C_j , then $C_1 \dots C_j \dots C_n$ is the GROSSEST CONSTITUENT ANALYSIS of X .

For an illustration of this consider (9).



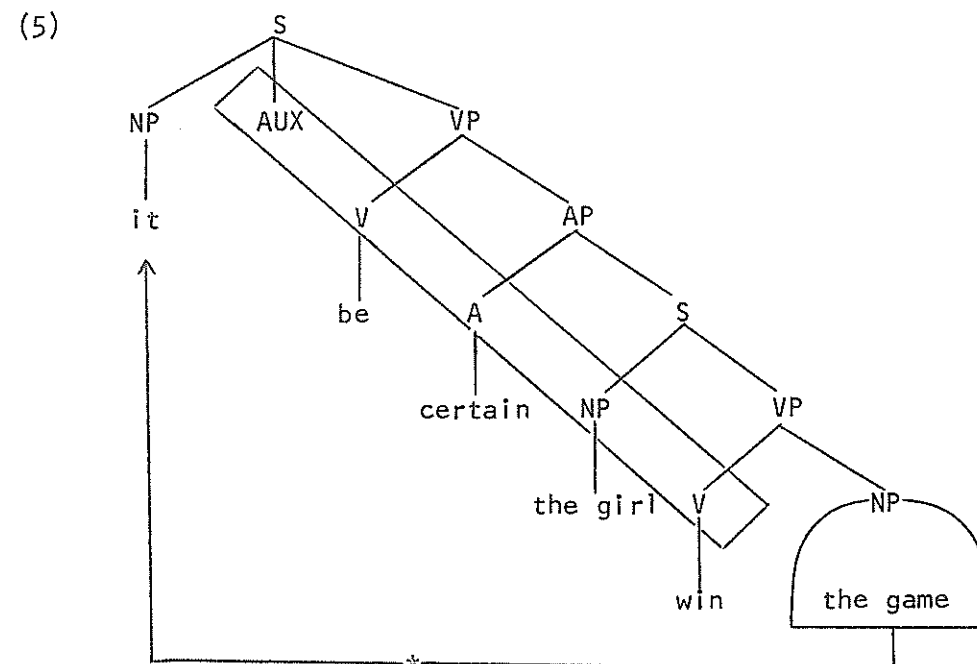
If the structural description of a rule were to mention A - D then a gross constituent analysis of X (that is, for A - X - D) would be B - C, and not B - E - F - G, since for the analysis B - E - F - G there would be another analysis, namely B - C, with fewer constituents. If the structural description of a rule were to mention B - G, then there would be two gross analyses of X, namely E - F and H - F, but E - F would be the GROSSEST constituent analysis since there is no other gross analysis higher in the tree.³

With this definition of grossest constituent analysis principle (20) of Chapter 2 is now revised and presented as the Variable Interpretation Convention.

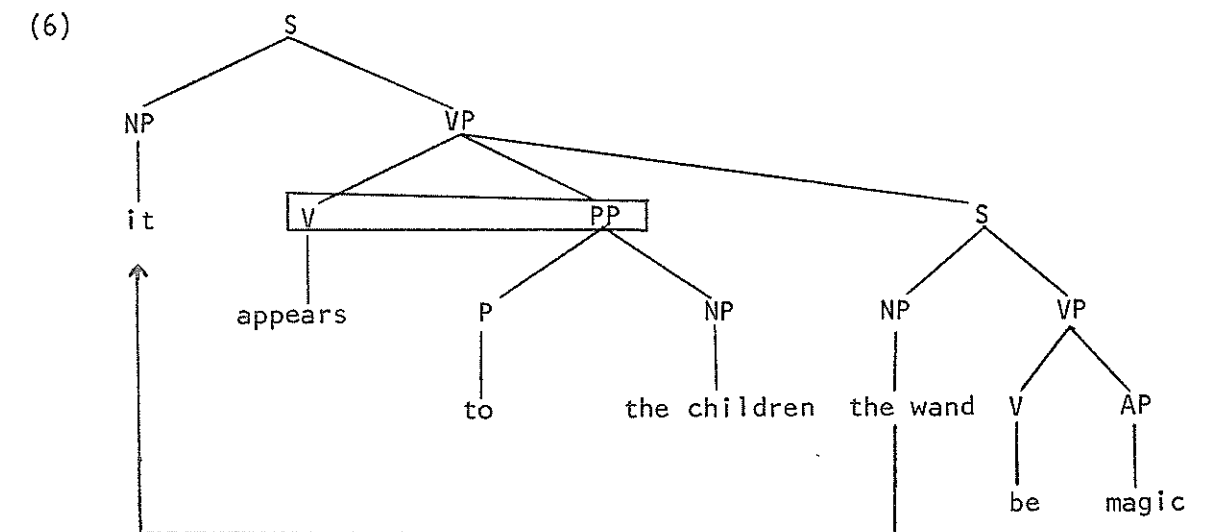
(10) Variable Interpretation Convention:

In the structural description of a transformation only terms which are crucially affected may be mentioned and for all terms mentioned, A - B \equiv A - X - B, where (a) X is a variable in the structural description and corresponds to the grossest constituent analysis of a phrase marker, and (b) the grossest constituent analysis does not contain any A or B or head of an A or B.

If we return now to phrase marker (5), we see that the NP *the girl* lies in the grossest constituent analysis (along with AUX, V, A, and V) with respect to the *it* and the NP of the rule in (4) if the object NP *the game* is construed to be term 3 of the rule. Phrase marker (5) is repeated here with the relevant grossest constituent analysis indicated by the enclosure. Throughout this work, where nodes of a phrase marker are enclosed as they are in (5), it is to indicate the grossest constituent analysis.



The object NP in (5) may not be moved. In (6), however, when term 3 of the rule is taken to be the NP *the wand*, the NP *the children* does not lie in the grossest constituent analysis; the PP which dominates it does. The repetition of (6) with the grossest constituent analysis indicated makes this clear.



There being no NP in the grossest constituent analysis of (6) between *it* and the embedded subject, the well-formedness of (7) is properly accounted for.⁴

Naturally in any account of subject raising there must be a way to prevent subjects of *that*-clauses from raising. I am assuming here an analysis of *that*-S complements where the *that* may be inserted late in the derivation (*that*-insertion is discussed in detail in Chapter 4). The insertion of *that* would be sensitive to the existence of an NP environment so that if the embedded S subject has been raised, *that*-insertion will not apply and the unacceptable **The wand appears to the children that \emptyset be magic* is prevented. I am also assuming in this work that verb agreement is a late process, sensitive to the existence of a subject NP, so that **The wand appears to the children \emptyset is magic* is also prevented in a natural way.

Now, with grossest constituent analysis defined and the Variable Interpretation Convention stated, the rule for subject raising can be written without including the context term VP. The rule at this point then is:

$$(11) \quad it - {}_S[NP] = 2 - \emptyset$$

where what occurs between terms 1 and 2 of the structural description is constrained by the Variable Interpretation Convention (VIC).

In (11) there is still in a certain sense a context included, namely the S-boundary. This boundary is necessary in the rule to prevent the *it* from being replaced by an NP in the same clause. For instance, the NP *the children* in (6) must not be analyzed as the NP of the subject raising rule. In the strongest condition on non-variables this boundary should be excluded. It seems however that it cannot be excluded if an NP of the matrix S is to be prevented from replacing this *it*. We will therefore consider the condition on non-variables to be appropriately weakened to allow only a category boundary context, that is, a conjunction of analyzability conditions.⁵

For the Variable Interpretation Convention, then, where it requires that only crucially affected terms be mentioned, it must be remembered that "terms" are defined in the usual sense so as not to include category boundaries, but, rather, only category labels referring to nodes of a phrase marker. As far as the variable material over which rules operate is concerned, this definition of "term" remains accurate. Raising, like any other transformational operation, operates across category boundaries in the variable.⁶

It must be pointed out here that in this framework, where variables are not written into rules but rather are predicted, $S[NP]$ does not mean "an NP which is also an S". This notation should be understood to mean $S[...NP...]$ or, in other words, any NP dominated by S, possibly immediately dominated, but not necessarily so. With the formalism used here, there is no way to express, in a structural description, the notion of "exhaustively dominates".⁷ A rule which could apply ONLY where a term of a structural description corresponded to a node of a phrase marker which exhaustively dominated another node would not be formulable in this model. The prediction is that such rules do not exist. Rule (11) then refers to the movement of ANY NP which is contained in an S just so long as that S does not also contain *it*, term 1 of the rule.⁸

3.2. The Questionable Status of Subject Raising to Object

We will now consider the process known as subject raising to object. This is a rule which has been proposed to account for a transformational relationship between sentence pairs such as those in (12) and (13).

- (12) a. The woman expects that she will win the game.
 b. The woman expects herself to win the game.
- (13) a. We can count on it that her strength will devastate the opposition.
 b. We can count on her strength to devastate the opposition.

The existence of this proposed transformational relationship has been seriously questioned in recent years, most notably perhaps in *Conditions on Transformations* (Chomsky 1973). In this work Chomsky assumes that the process does not exist and that there exists instead, in the base, both tensed and non-tensed sentences. This assumption, among others, allows for the statement of several very general conditions (including the Tensed-S Condition) which have since been refined and repeatedly shown to have significant explanatory power with respect to many aspects of "core grammar" (Chomsky 1976 and class lectures 1976).

In his by now well-known book *On Raising*, Paul Postal (1974) argues for the existence of a rule of raising to object and tries to show, among other things, that the existence of this rule invalidates the *Conditions* (Chomsky 1973) framework.

Joan Bresnan (1976c), without taking a stand on whether or not this rule exists, successfully demonstrates that Postal has presented no new arguments for its existence. She says (p. 485):

..., the book contains a number of empirical and logical pitfalls--arguments based upon false generalizations and conclusions unwarranted by their arguments. These are worth discussing because they are central to an evaluation of the evidence for a rule of Raising in English grammar,...

Later she points out that

On Raising ... contains no formulation of the rule of Raising, which is its subject, no precise descriptions of the structures to which the rule applies, and no systematic analysis of the many constraints, conditions, and rules that are found to interact with Raising. (p. 499)

David Lightfoot (forthcoming) in his review of Postal's book, points out that

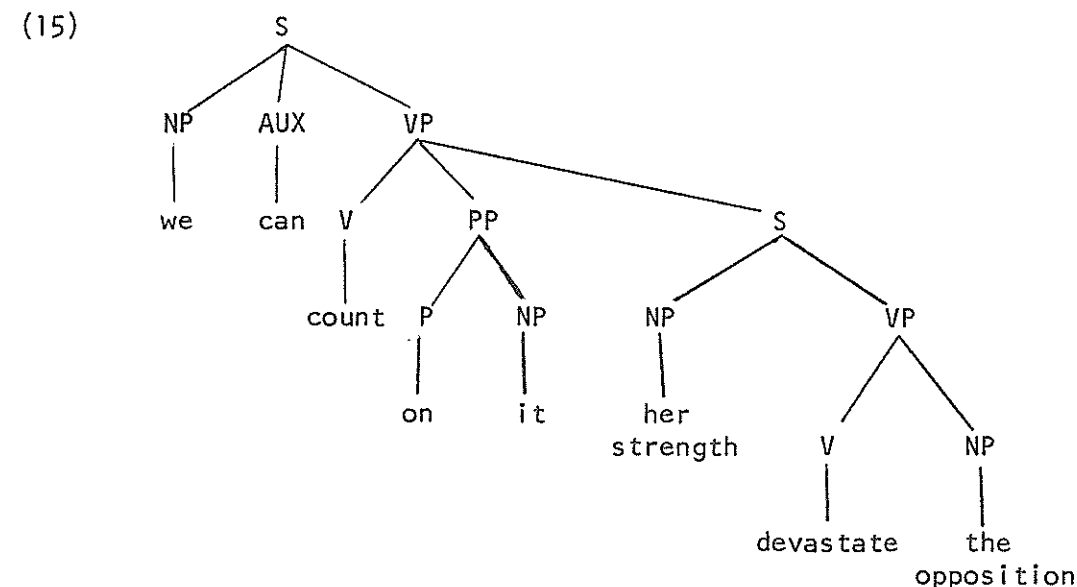
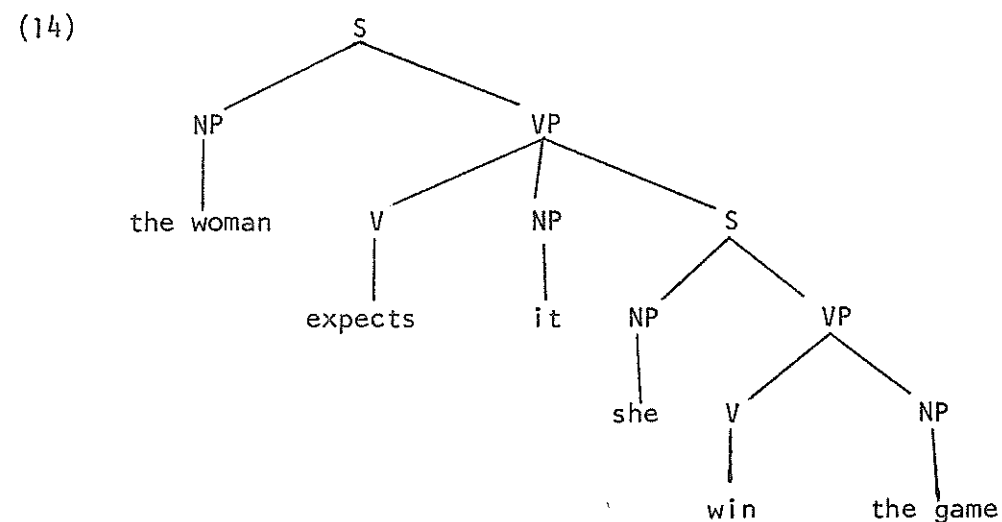
Unlike subject-to-subject raising, subject-to-object is a somewhat unusual rule in that it changes only constituent structure but does not change the order of elements [footnote omitted], and it exists only to feed certain other rules, such as Reflexive, each Movement, Passive and a rule specifying when two NP's cannot intersect in reference. (p. 3-4)

Lightfoot discusses the arguments against raising to object and says:

While I do not think that Postal has formulated a convincing case for Raising, what I want to emerge from this article is that even

if he did build such a case it would not invalidate a 'Conditions' grammar [footnote omitted], if we assume a trace theory of movement rules, nor would it necessarily entail positing a clause mate class of transformations. In this context, which is the context of the book, the existence of Raising has no theoretical consequences and therefore represents a trivial question. Furthermore, as Postal formulates it, his appeal to global devices puts his proposal beyond the reach of empirical falsifiability. (p. 26)

It seems then at this point that while the existence of a rule of subject raising to object has not been satisfactorily demonstrated, it might be the case that it can be. If it is (and we will return below to a discussion of this rule) then whether its existence is of great theoretical importance or not the rule would still have to be stated and its interaction with other processes (rules and conditions) carefully considered. For the purposes of this study it can be shown that rule (11) is an adequate statement of the rule, but that for a discussion of variable material subject raising to object is a process which yields really no new insights. This is so because where the rule would operate felicitously the variable is empty and where it would be blocked the relevant conditions are the same as they are for subject raising to subject position. To see that this is so consider phrase marker (14) for sentence (12a) and phrase marker (15) for sentence (13a).



The basic structures of (14) and (15) are those used by Emonds (1976: 77). In (15) it might be that the embedded S is attached as a daughter of the PP, rather than as a sister to it.⁹ This is a plausible alternative which would mean that the verb *count*, in this sense, subcategorizes just a PP rather than a PP followed by an S. This would entail also that the preposition *on* would subcategorize both an NP and a following S in order for this raising process to be a replacement of *it*. The Emonds version using (15) does not give double subcategorization but rather depends on the analysis of extraposition. In either case, in the derivations of (12b) and (13b), where the embedded subjects replace the formative *it*, nothing occurs in the variable material between the item which moves and its target site.

The examples which must be prevented are given in (16).

- (16) a. *The woman expects the game $\left\{ \begin{smallmatrix} \text{her to} \\ \text{she} \end{smallmatrix} \right\}$ win.
 b. *We can count on the opposition her strength (to) devastate.

For both (16a) and (16b) the subject NP of the embedded clause properly blocks the movement of the object NP by lying in the grossest analysis with respect to rule (11), if the NP of the rule is analyzed as either object, that is, *The game* in (14) or *the opposition* in (15). As with subject-to-subject raising, for subject raising to object the existence of a subject in the embedded clause will always suffice to block the movement of an object NP.

Thus far, rule (11) and the Variable Interpretation Convention properly account for the grammaticality facts not only for the raising of an NP from an embedded clause into higher subject position, but also for the facts where a rule of raising into object position has been proposed. In other words, rule (11) presents a unified account of the conditions under which a subject may be moved out of an embedded S. Next we will look at the process which moves objects out of embedded clauses.

3.3. Object Raising

3.3.1. Object Raising Predicates and Subcategorization

Object raising is the rule of English which moves the object of an embedded clause into subject position of a higher clause. The rule is also known as *tough*-movement because it is a process which is governed by the occurrence of particular adjectives (and an occasional verb), a class of items of which *tough* is an illustrative member:

- (17) a. It is tough to tolerate his pomposity.
b. His pomposity is tough to tolerate.

Since this is a process, like subject raising, which is controlled by particular lexical items we will again use the option of subcategorizing the appropriate lexical items for a semantically null term which can be replaced, rather than including rule features on the governing terms themselves. Before starting the object raising rule we will return momentarily to the way the subject raising process was handled above.

In the statements of subject raising ((1), (4), and (11)) the term which is being replaced is written as *it*. We now must consider in detail the actual nature of this term which is being replaced. Certainly it is not just any *it*, nor is it just any non-referring *it*. The *it* which co-occurs with forms like *surprising* cannot be transformationally replaced. The *it* which co-occurs with verbs like *seem* can be replaced only with subjects of embedded clauses; the *it* of sentences like (17a) can be replaced only by objects of embedded clauses. It might then be that there is an inventory of three semantically empty forms for which lexical items can be subcategorized: one which cannot be transformationally replaced; one which is replaced in subject raising; and the third which is replaced in object raising. This use of subcategorization would provide a way of encoding lexical information into phrase structure and making transformations sensitive to "governing" lexical items without using rule features.

There is a better alternative which we will now adopt.¹⁰ This alternative preserves the idea of the coding of lexical information into phrase structure through subcategorization, but makes unnecessary the positing

of three different *it*'s in the lexicon. The first step in this approach is to realize that the *it* which co-occurs with *surprising* is not in the base but, rather, is the result of extraposition (see Emonds 1976: Chap. IV). Lexical items like *surprising* subcategorize sentential NP subjects and occur on the surface with *it* as subject only where extraposition has taken place. If we assume, then, that this transformationally introduced *it*, the result of extraposing a complement, is NEVER replaceable by a raising process, and if we go on to carefully consider subcategorization of the other relevant verbs and adjectives, we will arrive at an interesting and coherent account of all of the cases where an NP is moved from an embedded domain into a higher one.

First we will look at the verb *seem*. This verb (and others like it) never occurs with a sentential subject so the *it* which appears as its subject cannot be the result of extraposition. Rather, *seem* subcategorizes a semantically empty subject (it places no selectional restrictions on its subject, as has often been pointed out). Additionally, it would subcategorize an optional prepositional phrase and an S complement. Subcategorization for *seem* is as shown as in (18).

- (18) *seem*, Δ — (PP) S¹¹
+F

As indicated in (18) by the feature +F, the delta subject for which *seem* is subcategorized, while semantically (and phonologically) empty is not void of feature specifications. The feature bundle (see Chomsky 1965) for this NP contains only syntactic features. The deep structure subject of *seem* is not empty. Δ is treated in the base as is any other NP which can

undergo lexical insertion. Principle (19) is now postulated and adopted.

- (19) Any item in the lexicon for which another lexical item may be subcategorized is a feature bundle containing at least syntactic features. No terminal element of a phrase marker dominating a lexical item which contains any feature(s) is considered empty.

Principle (19) will be used repeatedly both in the present chapter and in subsequent chapters of this thesis.

Lexical items made up of only syntactic features differ from other lexical items if they occur in surface structure. The result of their occurrence on the surface is ill-formedness:

- (20) Any lexical item which occurs in a surface string empty of both semantic and phonological features causes that string to be marked as ill-formed.

In keeping with principle (20) there are processes which either replace lexical items containing only syntactic features with phonologically full formatives (such as *it* and *there*) or which provide semantic features based on the semantics of other formatives in a string (interpretation and control processes). Additionally, as we will see shortly, lexical items with only syntactic features are replaced in certain movement transformations.

Returning now to subcategorization, some lexical items, *certain* for instance, which in some ways behave like *surprising* and in some ways behave like *seem*, have alternate possible subcategorizations. *Certain* can occur in the base with a sentential subject like *surprising* and undergo extraposition (*That the girl would win was certain* ~ *It was certain that the girl would win*). Alternatively *certain* can occur in deep structure with a delta subject, like *seem*, and then have the possibility of undergoing subject raising (as in example (3), *The girl is certain to win the game*). In contrast, an item like *probable* does not have alternate subcategorizations and as such could never undergo subject raising (**The girl is probable to win the game*). *Probable* only allows for extraposition: *It is probable that the girl will win the game*. The subcategorizations for *certain* and *probable* are given in (21).

- (21) a. *certain*, S ____¹²
 certain, Δ ____ S
 +F
- b. *probable*, S ____

We can now return to the issue of object raising and lexical items like *tough*. *Tough* can occur with sentential subjects (e.g., *For us to tolerate his pomposity is really tough*) and sentences like (17a) are generally considered to be the result of extraposition. Sentence (17b) then does not derive from (17a). Lexical items like *tough*, *hard*, *easy*, etc., have alternative subcategorizations. The two alternatives are as in (22).

- (22) a. *easy*, S ____ (PP)
 b. *easy*, Δ ____ (PP) VP¹³
 +F

Alternative (b) is the one which gives rise to object-raised forms like (17b). As is clear from these proposed subcategorizations, the only time these items like *easy* appear followed by a full S complement is when the complement is extraposed. Otherwise the complement is a VP.

In this account there are then the same two possibilities, that is, (a) and (b) of (22), for verbs like *take*, which also exhibit both extraposi-

tion and raising behavior. From *To fix the car would take three hours* can be derived: *It would take three hours to fix the car*. To a deep structure like Δ takes three hours to fix the car, raising can apply to yield:

+F
The car takes three hours to fix.

3.3.2. C^{\max} and the Raising Rule

We now have the following two possible rules for raising:

- (23) Subject raising: Δ - $S[NP] \Rightarrow 2 - \emptyset$
 +F

- Object raising: Δ - $VP[VP[NP]] \Rightarrow 2 - \emptyset$
 +F

In this statement of object raising it is necessary to include the two VP brackets to assure that the rule does not select an NP of the matrix VP to replace term 1. This formulation of object raising assures, just as does the form of subject raising, that it will be only NP's of embedded domains which meet the structural description.

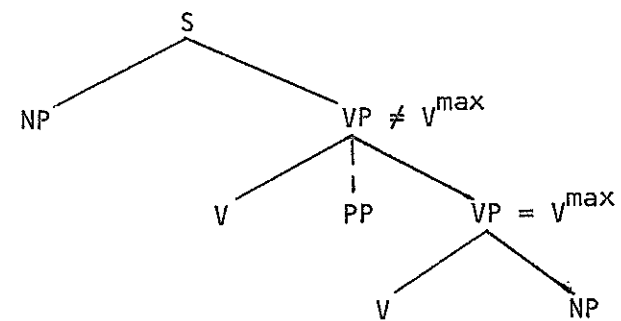
Now we are finally in a position to see how, by use of a simple definition (the idea for which I owe to Joe Emonds), it is possible to collapse subject raising and object raising into one general rule.

- (24) Definition:

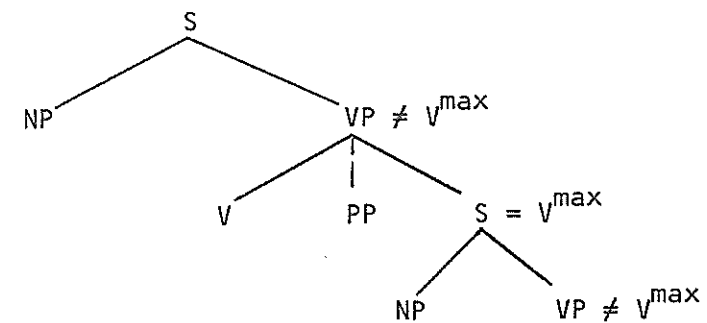
Constituent X, whose head = a projection of a phrasal category C, where X is not the head of constituent Y, is a C^{\max} .

What this means is that the head of a phrase will never qualify as a C^{\max} . In particular, a VP immediately dominated by S, which is the head of the S (following Jackendoff (1974)), will never be a V^{\max} . In contrast, a VP immediately dominated by VP and a sister to V will always be a V^{\max} . Similarly, an S dominated by VP qualifies as a V^{\max} . The phrase markers of (25) illustrate this schematically.

(25) a.



b.

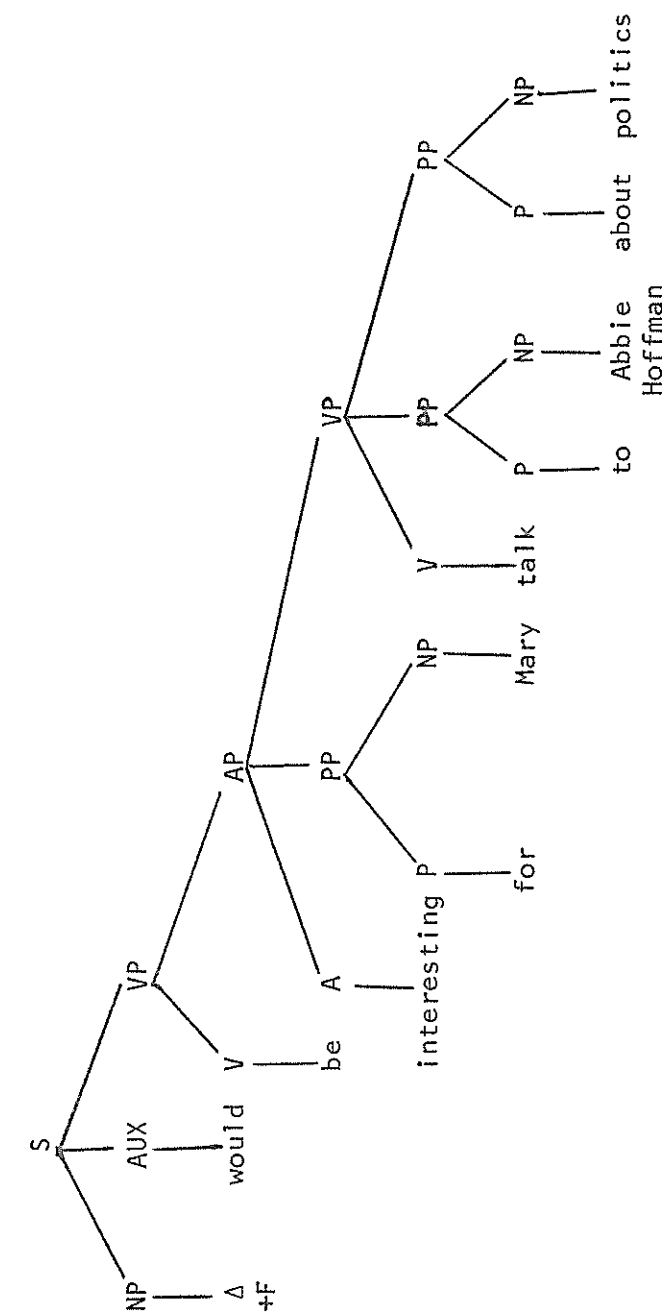


This definition of C^{\max} will again prove to be very useful when we consider (in Chapter 6) the question of rightward movements and upward bounding.¹⁴ For now this notion of C^{\max} provides for the collapsing of the two rules of (23) into one simple rule containing only a single constituent boundary.

(26) Raising: $\Delta -_{+F} V^{\max} [NP] = 2 - \emptyset$

Next we will examine the operation of this raising rule and we will see that by the use of the Variable Interpretation Convention it can in fact properly account for all cases in English where an NP is moved leftward out of an embedded domain.

Consider a deep structure like that of (27a) from which both sentences (27b) and (27c) can be derived.



(27) a.

- b. Abbie Hoffman would be interesting for Mary to talk to about politics.
- c. Politics would be interesting for Mary to talk to Abbie Hoffman about.

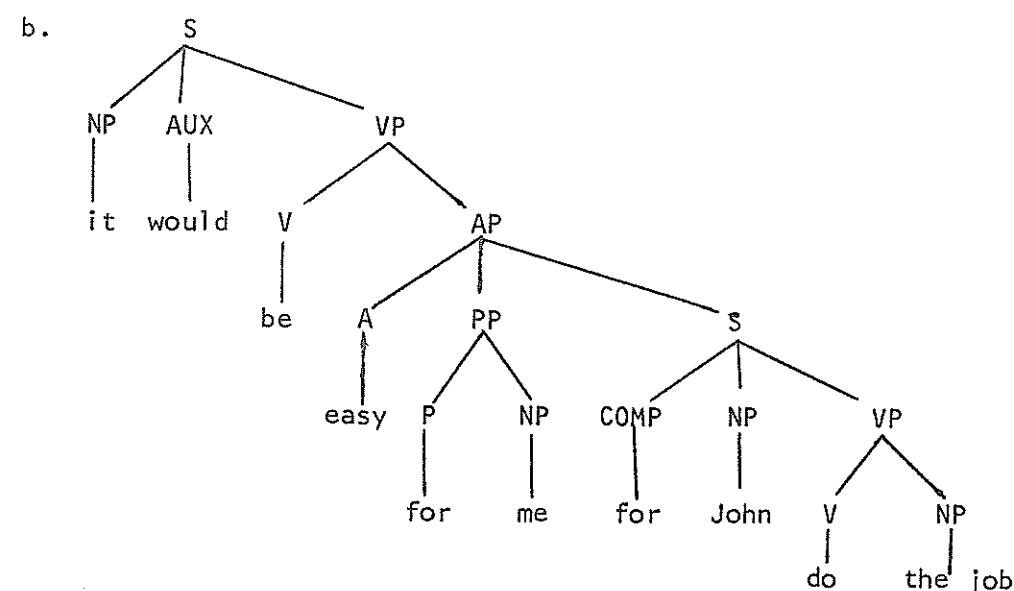
Compare the grammaticality facts of (28), however, with those of (27).

- (28) a. It would be easier for me for John to do the job
than for me to do it myself.
- b. *This job would be easier for me for John to do
(than for me to do myself).
- c. It is a waste of time for me for John to try to
help with this job.
- d. *This job is a waste of time for me for John to try
to help with.

(These examples are from Jackendoff (1972:155).) The difference here has to do with the existence of a subject NP in the embedded clause. For structures like (27a) it has been successfully argued (Bresnan 1971; see also Jackendoff 1972, Chomsky 1973) that the embedded VP is subjectless. Therefore in (27a) the PP *for Mary* is properly attached as a constituent of the matrix sentence. Since the NP *Mary* is contained within a PP it is not in a position to be contained in the grossest constituent analysis where object raising moves an embedded NP.

For a sentence like (28a) the structure is actually more different from (27a) than might at first be realized. Consider (29a) (a simpler string for consideration than (28a)) and the phrase marker (29b).

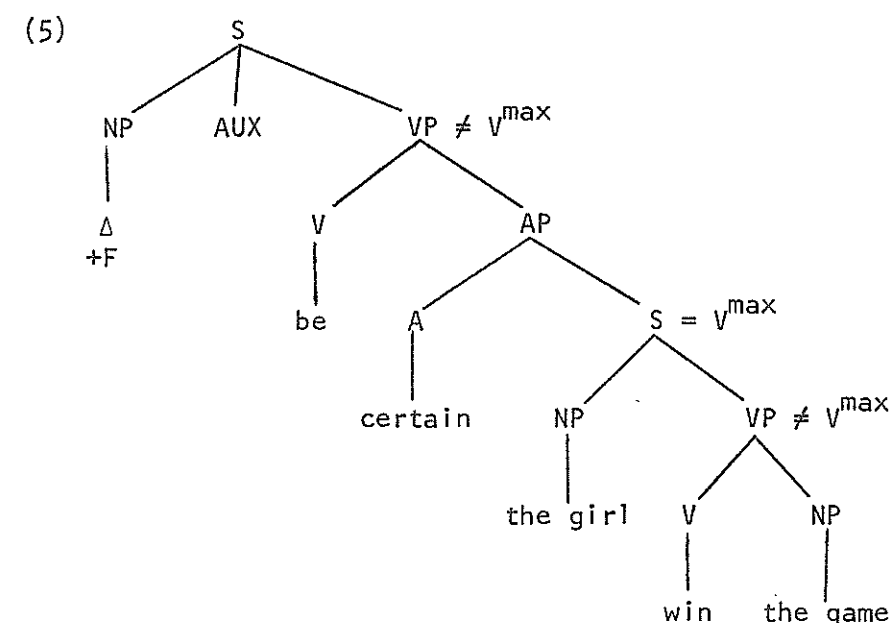
- (29) a. It would be easy for me for John to do the job.



In sentences like (28a) or (29a) the NP *John* must be considered the subject of the embedded clause. Since *easy* therefore has a full S complement, (29b) must be the result of extraposition and not a base phrase marker. There is no possibility, therefore, of raising applying to a structure like (29b).¹⁵ Examples like (28b) and (28d) will never be generated, nor will raising ever apply to derive strings like **The job would be easy for me for John to do* or **John would be easy for me for to do the job* from (29b).

Where the structural description of raising is met and there is a full S complement, that is, where Δ and a VP-final S co-occur, it will always

be the subject NP which raises. An object NP will be prevented from moving, as it was in (5), because the subject NP will lie in the grossest analysis. Phrase marker (5) is repeated here (appropriately modified) for ease of reference.



Notice here that it is the VIC which prevents the raising of the object NP and not just the structural description of raising. Although the NP *the game* is not immediately dominated by a V^{\max} it is CONTAINED in a V^{\max} and as such meets the structural description. In other words, it must be remembered that in this framework V^{\max} [NP] must be understood as V^{\max} [...NP...]. The

only time object raising can occur is when there is a VP with no subject, as in (27), co-occurring with a Δ . The proposed subcategorizations and the

Variable Interpretation Convention make it possible to account for both subject raising and object raising with a single rule.

Next we will see that this treatment of raising processes has some important consequences for the notion of cyclic rule application.

3.4. Raising, VP', and the Principle of the Transformational Cycle

Let us begin this section by looking at the grammaticality facts illustrated in (30).

- (30) a. It seemed to Bill Mary believed John to be sick.
 b. Mary seemed to Bill to believe John to be sick.
 c. *John seemed to Bill Mary believed to be sick.
 d. John seemed to Bill to be believed by Mary to be sick.

Subject raising applies to the subject of *believe* to properly derive (30b). (Although subject raising and object raising are subcases of a single rule in this framework, for ease of exposition I will continue to refer to the processes by different names at times.) The rule cannot apply to move an NP across the NP *Mary* when it is in subject position, (30c). The ungrammaticality of (30c) is not only predicted by the VIC but is also properly accounted for by Chomsky's Specified Subject Condition. The fact that (30c) is not a good sentence might in other accounts be used as evidence that subject raising to object applied in (30a) and that therefore the NP *John* was not available for subject-to-subject raising. In the account argued for here whether *John* is a subject or an object is irrelevant because in neither case could it be moved across an NP in the grossest analysis. In other words, this account of the ill-formedness of (30c) does not rely on the existence of a raising to object rule (neither, of course, would any account using the Specified Subject Condition).

The grammaticality of (30d) shows that (30c) cannot be excluded by an argument based on deep structure depth of embedding. A better argument would be that (30c) is prevented because subject raising is cyclic and on the *believe* cycle there is no place for *John* to move to. The filled subject *Mary* prevents the intermediate cyclic application. The account argued for in this thesis also makes crucial use of the syntactic subjecthood of the NP *Mary* but does not make use of the claim that subject raising must apply cyclically.

If it can be shown that subject raising MUST be cyclic, then, of course, (30) provides no support for the analysis presented here, where it is the NP subject in the grossest analysis which prevents (30c). Obviously a place to look for an adequacy distinction between the VIC account of (30c) and an account crucially relying on cyclic application would be examples where an intermediate cyclic domain was subjectless and where subject raising could not apply on the intermediate domain. If the result

of trying subject raising on the highest cycle were still prohibited it could not be due to an intervening subject. Example (31) provides a relevant example.

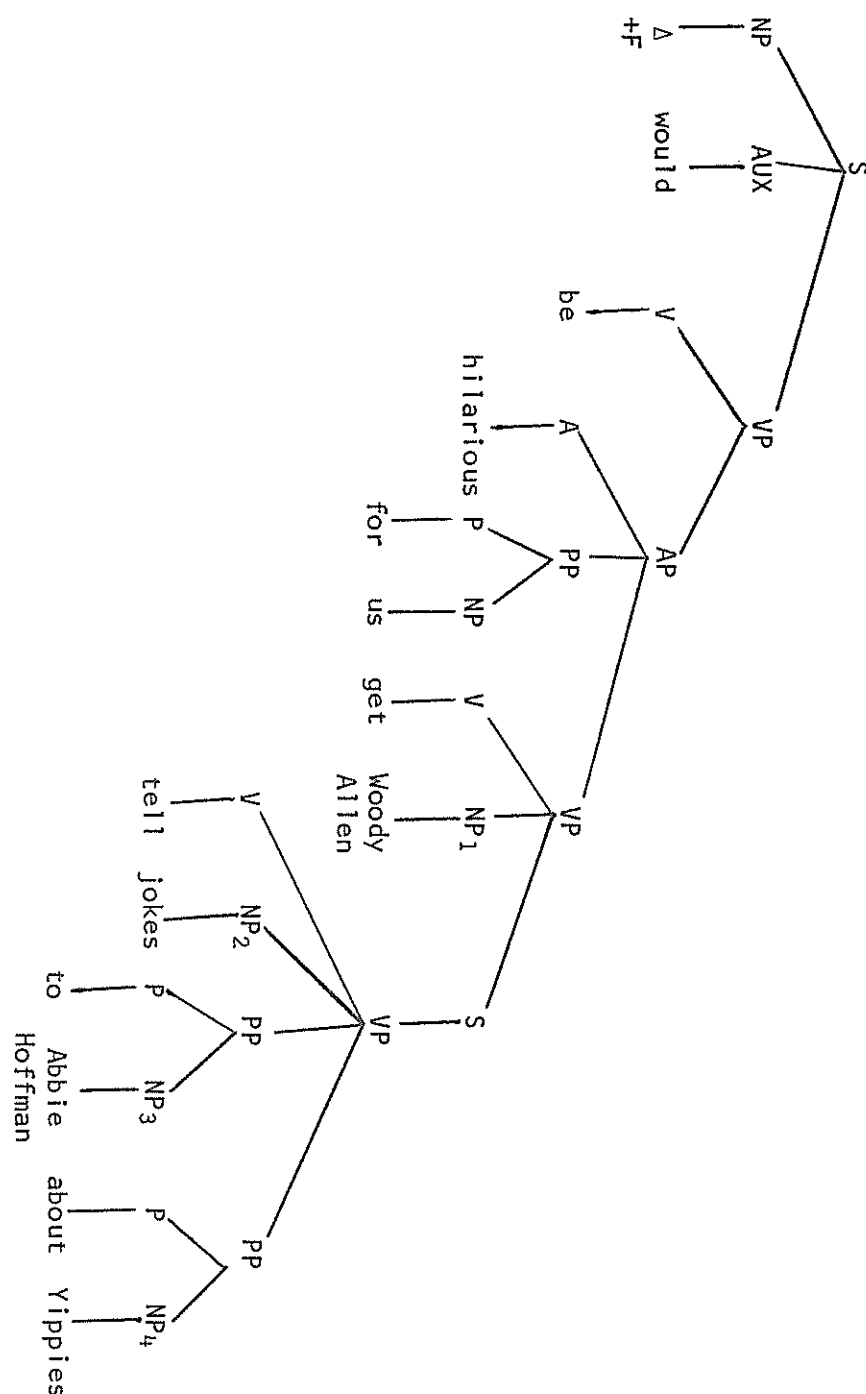
- (31) a. It seemed to be easy to persuade John [Mary be the winner].
 b. *Mary seemed to be easy to persuade John to be the winner.
 c. It seemed John was easy to persuade Mary was the winner.
 d. John seemed to be easy to persuade that Mary was the winner.

If it were the case that subject raising had to be cyclic, (31b) could readily be prevented since *easy* does not allow subject raising (and neither for that matter does *persuade*). Where the raising process affects only adjacent cycles (first object raising, (31c), and then subject raising (31d)) the results are fine. It seems then that the status of (31b) forces the conclusion that subject raising cannot ignore cyclic domains.

Actually, as it turns out, the ungrammaticality of (31b) CAN be accounted for in a VIC analysis. Before seeing that this is in fact the case we will consider further some instances of object raising.

Above where examples of object raising were discussed it was oblique objects which were moved and they were moved from an embedded VP with no direct object (the examples of (27)). Consider now a structure in which there is a direct object NP of the embedded clause.

(32) a.



- b. Woody Allen would be hilarious for us to get to tell jokes to Abbie Hoffman about Yippies.
 c. Jokes would be hilarious for us to get Woody Allen to tell to Abbie Hoffman about Yippies.
 d. Abbie Hoffman would be hilarious for us to get Woody Allen to tell jokes to about Yippies.
 e. Yippies would be hilarious for us to get Woody Allen to tell jokes to Abbie Hoffman about.

As (32b) through (32e) illustrate, any one of the NP's 1-4 may be raised to subject of the matrix clause, even NP₃ and NP₄ where, as things now stand, the direct objects NP₁ and NP₂, would be included in the grossest constituent analysis.

The facts of (32) might at first seem like a serious problem for an account of raising using the VIC. What is actually the case, I believe, is that the problem illustrated here is only an apparent one and the grammaticality of (32c)-(32e) ceases to be problematic if the internal structure of the verb phrase is given proper consideration. I think it can be successfully argued that there must be structure intervening between the main phrasal categories (NP, VP, AP) and the lexical category labels (N, A, V) of AT LEAST ONE node. Here in particular the node VP' will occur between (in the hierarchical sense) the VP and the V. This VP' will then dominate the verb and the objects which are most closely associated with it according to certain syntactic tests.

The arguments for this structure of the VP are three-fold, that is, there are three relevant constituency tests. First of all, a structure including the VP' allows for a simplification of the statement of where adverbs may occur in English. It has often been noted that their occurrence is very free but that one place where they never show up is between a verb and its direct object. They also do not show up between a direct object NP and a PP which is OBLIGATORILY subcategorized by the verb. This is shown in (33).

- (33) a. He suddenly put the book on the table.
 b. *He put suddenly the book on the table.
 c. *He put the book suddenly on the table.
 d. He put the book on the table suddenly.

Notice that if an adverb does occur in order between the direct object and the PP it must be construed as part of the PP. For example, *He put the car immediately behind the house* does not mean *He immediately put the car behind the house*, but rather is close in meaning to *He put the car right behind the house*. These adverbs can occur in the intensifier position within the PP, but not as a sister to PP following a direct object.¹⁶ Thus, adverbs may occur at the beginning or end of VP' but not internal to it.

Secondly, emphatic reflexives can never occur between verbs and their direct objects. This was pointed out by Mark Baltin (1976) who has also noticed that these emphatics may not occur before obligatorily strictly subcategorized PP's.

- (34) a. He himself put the book on the table.
 b. *He put himself the book on the table.
 c. *He put the book himself on the table.
 d. He put the book on the table himself.

Where an emphatic reflexive does occur after the direct object it is actually a rightmost constituent of the NP. The *herself* in *The woman placed Sue herself on the waiting list* can only refer to *Sue*, not to *the woman*.

The generalization made by Baltin is that items move only to constituent boundaries; that the boundaries provide the possible landing sites for terms which move.¹⁷ Therefore, by virtue of where adverbs and emphatic reflexives are attached by the rules which move them, there is evidence of the constituency of VP'.

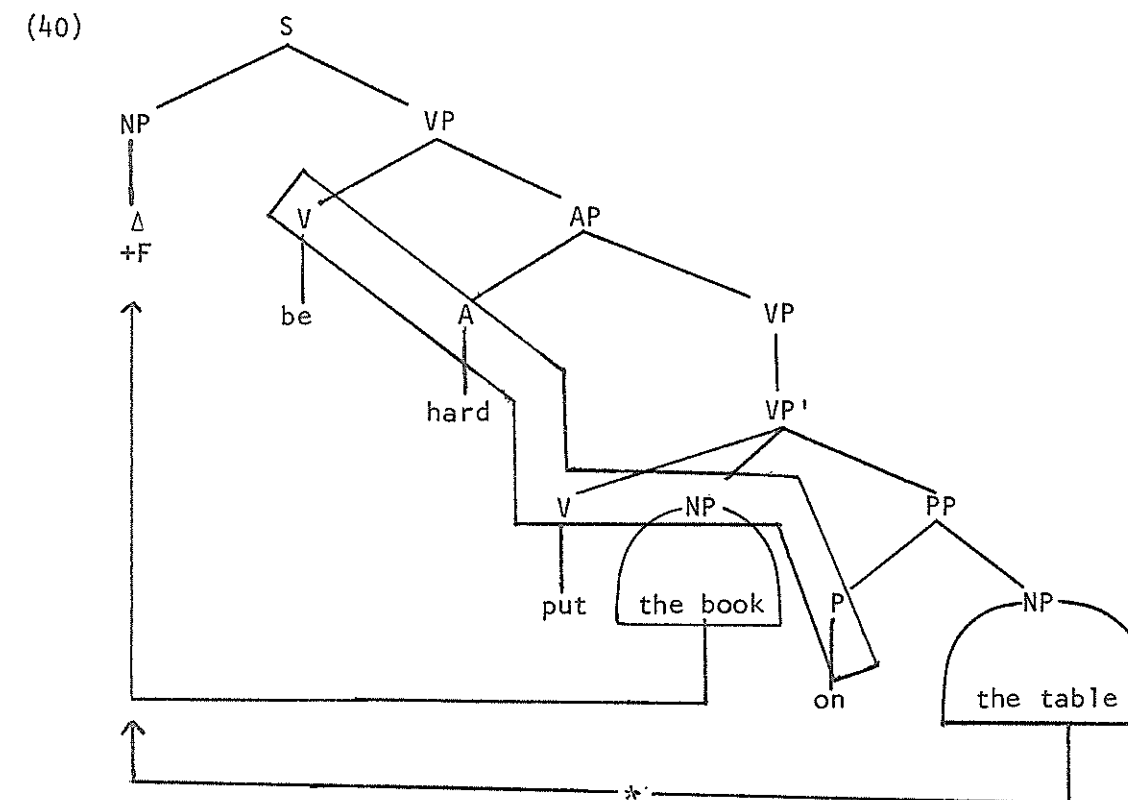
The third argument for the existence of VP' has to do with the behavior of different prepositional objects with respect to the raising process. Regardless of whether or not one accepts the VIC account of raising, there must be a way of preventing NP's contained in obligatorily strictly subcategorized PP's from being moved to subject position of higher clauses.¹⁸

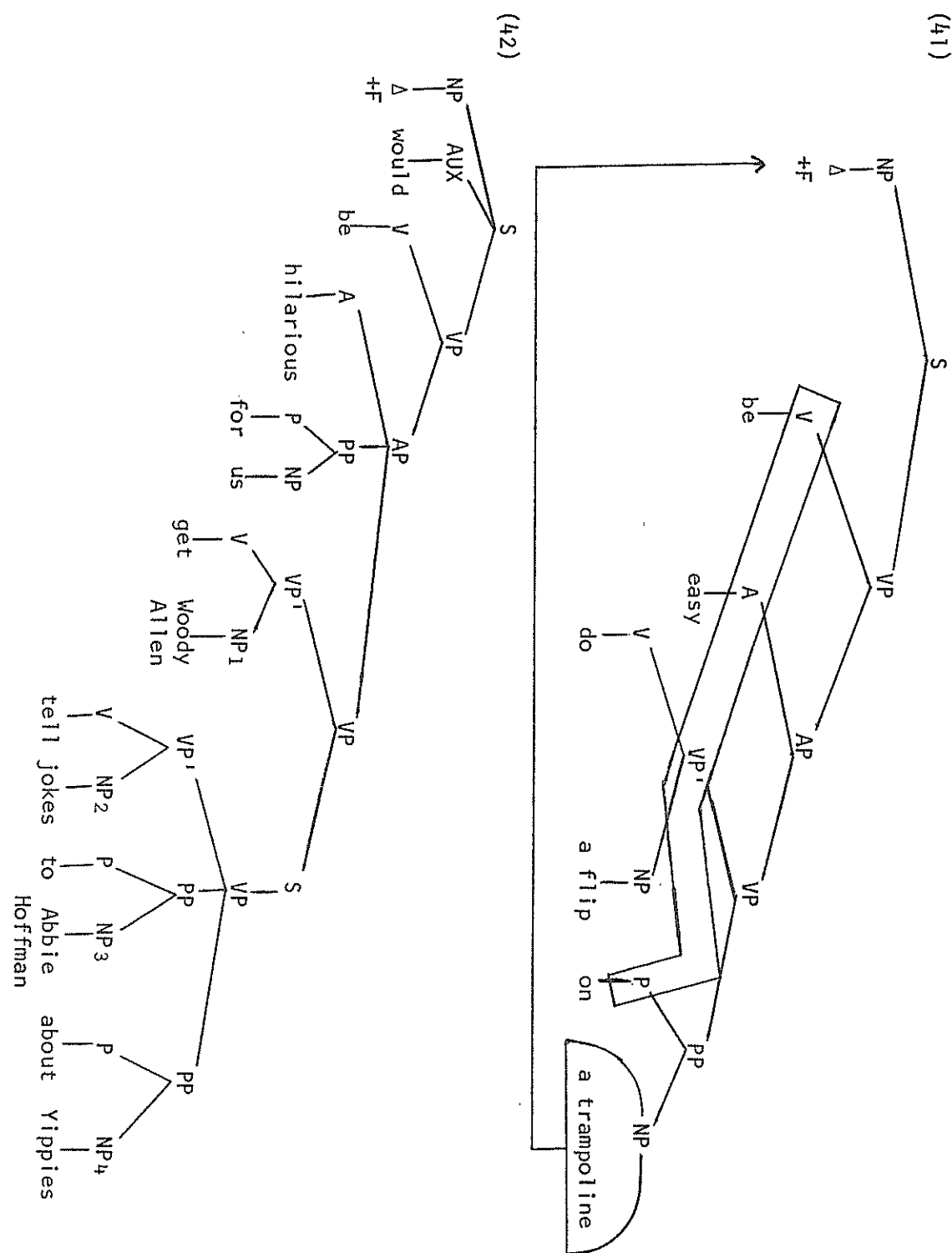
- (35) a. It is hard to put the book on the table.
 b. The book is hard to put on the table.
 c. *The table is hard to put the book on.
- (36) a. It is easy to sit on the fence.
 b. The fence is easy to sit on.
 c. It is easy to set the flowerpot on the fence.
 d. *The fence is easy to set the flowerpot on.
- (37) a. It is tough to place a man on a pedestal.
 b. A man is tough to place on a pedestal.
 c. *A pedestal is tough to place a man on.
- (38) a. It is easy to do a flip on a trampoline.
 b. A flip is easy to do on a trampoline.
 c. A trampoline is easy to do a flip on.

- (39) a. It was hard to perform the dance on the small stage.
 b. The dance was hard to perform on the small stage.
 c. The small stage was hard to perform the dance on.

In any analysis the ungrammaticality of (35c), (36d), and (37c) must be accounted for, as must the difference in status between those examples and (38c) and (39c).¹⁹

Using the Variable Interpretation Convention and assuming the existence of VP', both the facts here in (35) through (39) and the examples of (32) are readily handled. Phrase marker (40) would be the structure underlying sentence (35a); (41) is the phrase marker for (38a); and (42) is the revised underlying structure for (32).





In (40) the NP *the table* is prevented from raising because the NP *the book* occurs in the grossest analysis. Nothing, however, prevents the movement of *the book*. In (41), because the PP is not in the VP', the NP dominating *a flip* does not lie in the grossest analysis if raising applies to the NP *a trampoline*. Sentence (38c) is predictably well-formed. In (42), for none of the object NP's does the grossest constituent analysis for raising contain an NP. Therefore any of the objects may raise, hence the grammaticality of (32b) through (32e). Only where an NP is contained in a PP dominated by VP' will the direct object fall in the grossest analysis for raising. In these cases the VIC will always predict that the result of applying raising will be ill-formed.²⁰

We are now in a position to be able to return to the examples of (31) and to consider further the question of whether raising must be a cyclic process. It seemed that the ungrammaticality of (31b) (repeated here for ease of reference)

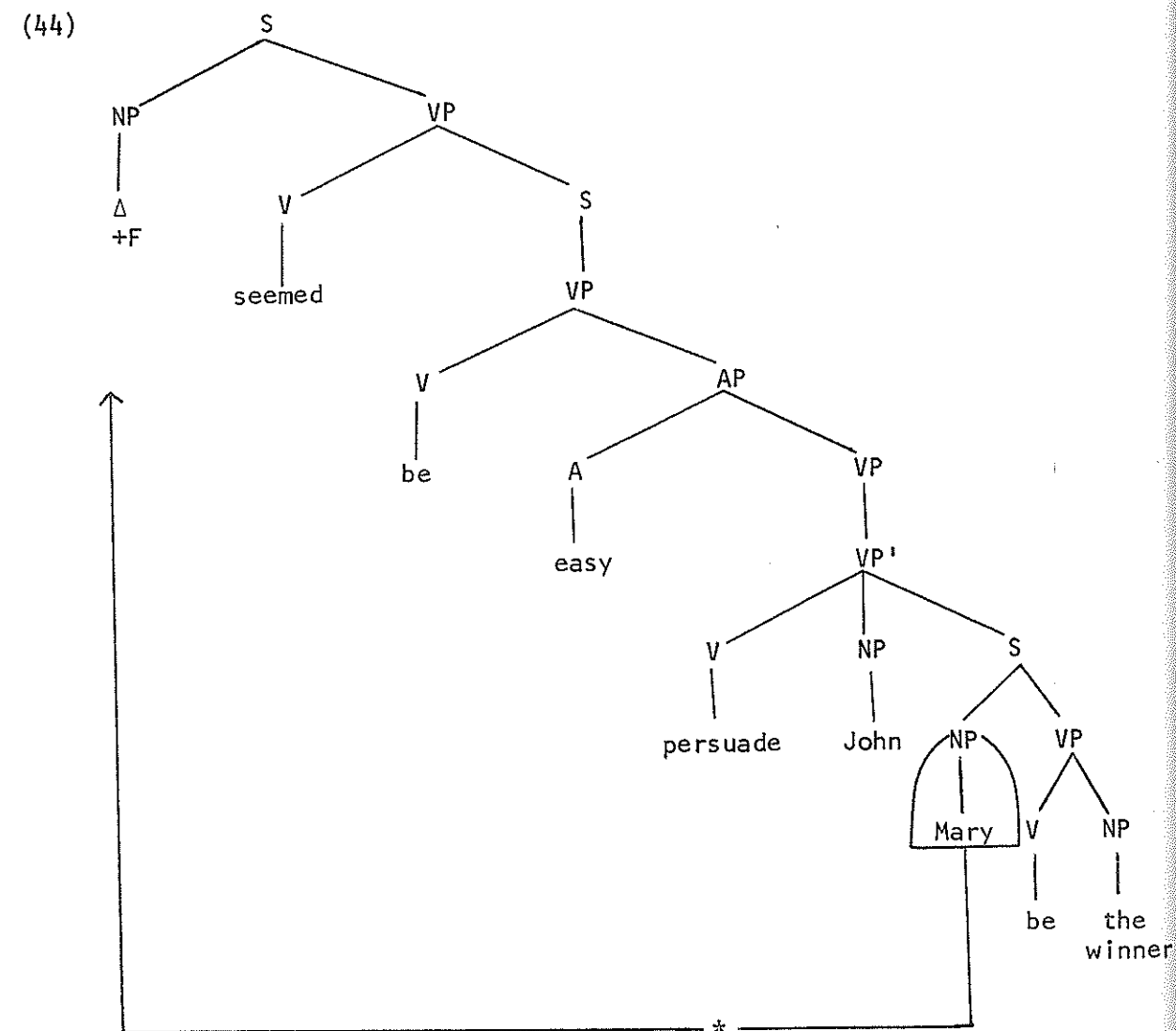
- (31) b. *Mary seemed to be easy to persuade John
to be the winner.

was to be accounted for by the fact that intermediate cyclic domains would not allow subject raising.

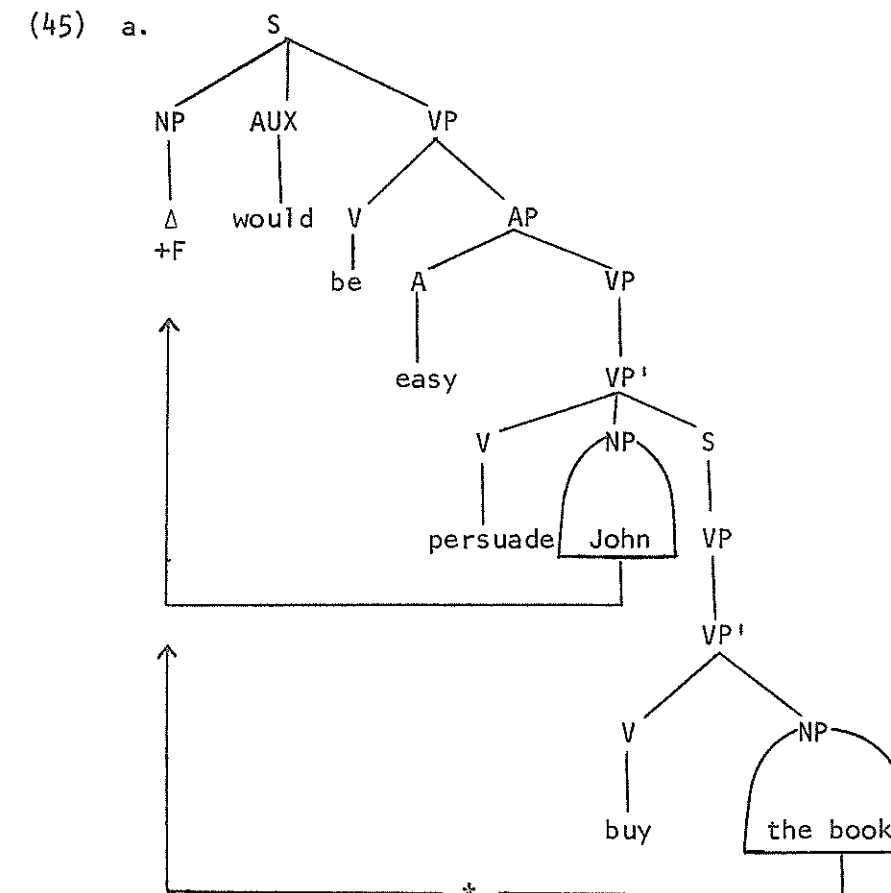
It is possible now though to also account for (31b) by the VIC, again using the existence of VP'. By the adverb and emphatic reflexive placement tests the VP' for a verb like *persuade* contains both the direct object and the complement S:

- (43) a. I {myself
naturally} persuaded John that Mary was the winner.
- b. *I persuaded {myself
naturally} John that Mary was the winner.
- c. *I persuaded John {myself
naturally} that Mary was the winner.²¹
- d. I persuaded John that Mary was the winner {myself
naturally}.

Phrase marker (44) then reveals why, in the VIC account, (31b) is ungrammatical.



The movement of the NP *Mary* is prevented by the occurrence of the NP *John* in the grossest analysis.²² This type of account provides for the right results in examples like (45) as well, where no question of cyclicity arises.²³



b. John would be easy to persuade to buy the book.

c. *The book would be easy to persuade John to buy.²⁴

The fact that the ungrammaticality of (31b) can be accounted for by the VIC shows that that example does not provide the basis for an argument that raising must be constrained by the principle of the cycle. We will see below that actually subject raising can be shown to apply cyclically but that it does so "automatically", that is, without being externally constrained by the cycle. Consider (46).

(46) a. It would be ridiculous for them to expect to win that particular game.

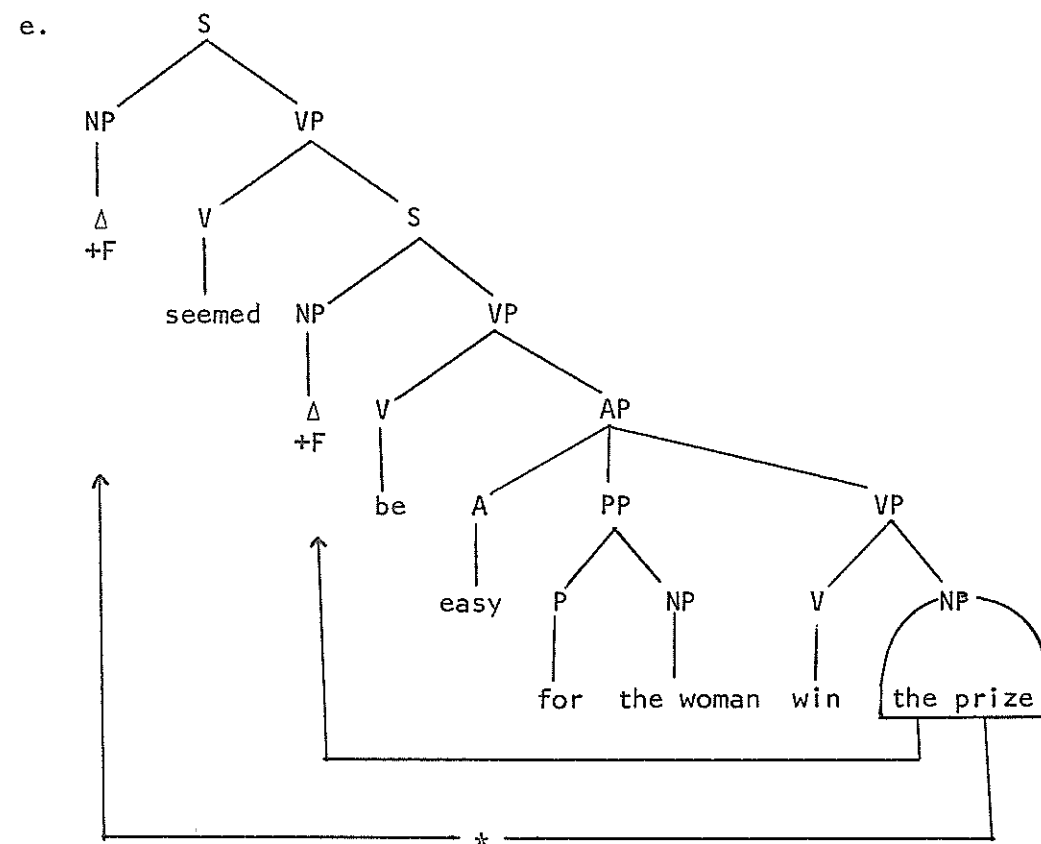
b. That particular game would be ridiculous for them to expect to win.

c. *It would be ridiculous for them that particular game to expect to win.

In this account, where the object NP *that particular game* may become the subject of *would be ridiculous* in a single movement (no NP occurs in the grossest analysis to block it) the ill-formed intermediate form (46c) need never be generated. There would then be no reason to ever say that object raising was obligatory.²⁵

In this section so far I have argued that neither raising process should be constrained to obligatorily apply cyclically. This is not to say that they may not apply within the cycle. What is in fact the case is that the raising rule applies optionally wherever its structural description is met, but in accord with the VIC. Where it does not apply a late process replaces the delta by *it* (see principle (20)). We will see now that this gives the right results. The relevant examples are given in (47).

- (47) a. It seemed it was easy for the woman to win the prize.
 b. It seemed the prize was easy for the woman to win.
 c. The prize seemed to be easy for the woman to win.
 d. *The prize seemed it was easy for the woman to win.

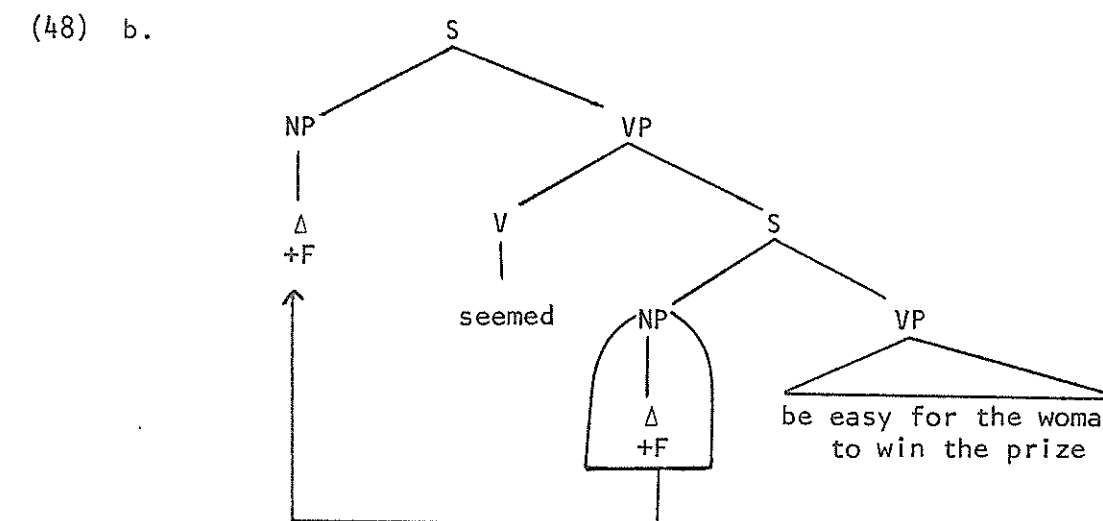


The improper application of raising which would yield (47d) is properly prevented by the VIC and the form of the raising rule. Since the structural description of the rule mentions NP, if the highest delta were analyzed as the first term the structural description would be met where the second term was the NP dominating the object NP *the prize*, but the delta which is the subject of *be easy* would lie in the grossest analysis preventing the movement. If, however, raising were applied and term 1 of the rule was construed to mean the embedded delta, then there would be nothing to block the movement of the object, (47b). Once the object NP replaces the embedded delta a subsequent application of raising properly yields (47c).

For a sentence like (48a),

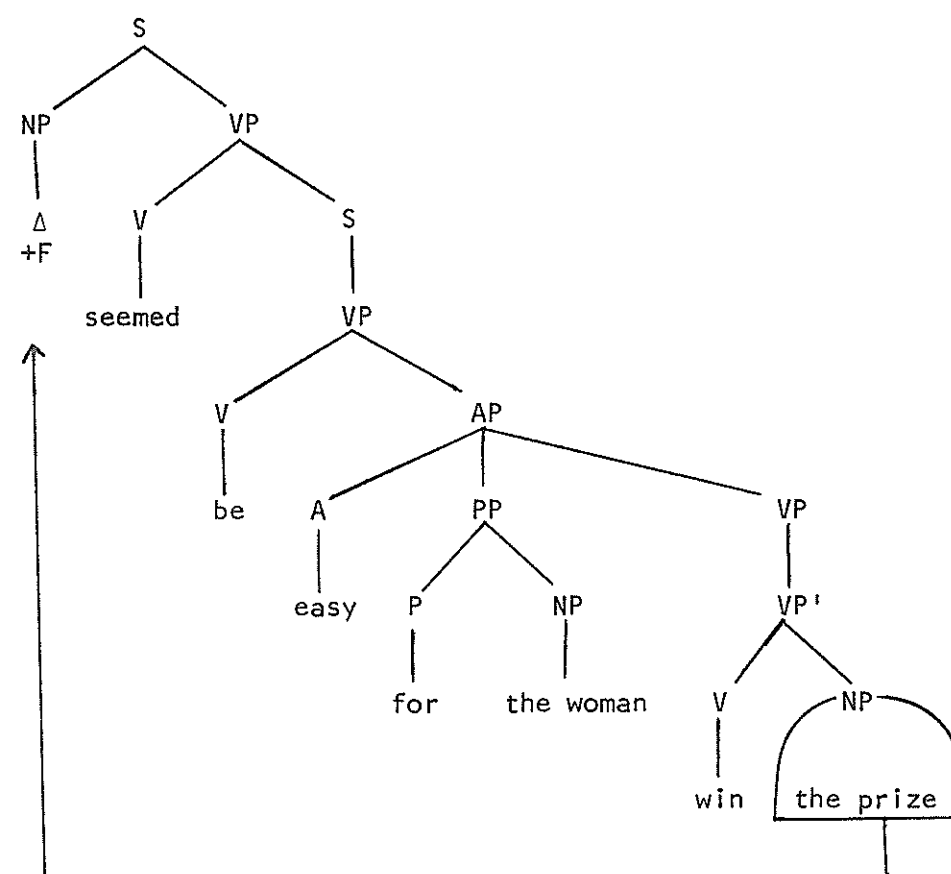
- (48) a. It seemed to be easy for the woman to win the prize.

where the VP *be easy* is subjectless, we must now carefully consider the process by which this subject does not occur on the surface. It might be argued that in a structure like (47e) if no raising takes place to replace the delta subject of *be easy* then the structural description of subject raising is still met on the highest S and subject raising could apply to delta. This is illustrated in (48b):



This would account for the subjectless nature of (48a) since then the process which replaces deltas with *it*'s (where no raising has replaced the deltas) would not apply for the subject of *be easy*. Notice, then, however, that after the process illustrated in (48b) has taken place (and before *it* is inserted) the structural description of raising is again met, as in (48c).

(48) c.



This would then be another possible derivation for the well-formed sentence (47c). Notice, though, that there is a problem with this approach. In (48c) there is nothing to block the improper raising of the NP *the woman*. The structural description of raising is met (there is a delta with the proper feature and the embedded S is a V^{max}), but **The woman seemed to be easy for to win the prize* must be prevented.

This problem does not arise if the process illustrated in (48b) does not take place. And this, in fact, is the solution which I now propose. I suggest that transformations apply to move only categories dominating lexical items which have phonological content--rules do not move deltas. Deltas provide target sites for movements or sites for insertions, but are never construed as the term which is moved by a transformation. This is presented in principle (49).

- (49) No transformational operation T may apply in such a way as to reorder constituent C unless C dominates a feature bundle including phonological feature specifications.

It was suggested to me by Joan Bresnan that an alternative to prohibiting the movement of deltas, that is, an alternative to preventing the process illustrated in (48b), might be to allow it but then in some way to prevent the iteration of a process in the same domain. In other words, for raising, the process could not apply in the domain of the highest NP more than once. This would allow (48b) or (48c) but not both. This very appealing alternative to the use of principle (49), a principle which would disallow the iteration of transformations in the same domain, is not adopted here because I have not explored its ramifications thoroughly. It may, however, prove very useful in future research. Principle (49) will again be shown to be necessary in Chapter 5 where the process of passivization is reconsidered.

The movement indicated in (48b) is prevented because deltas are not moved by transformations and we now return to the question of the derivation of (48a). What I am proposing (following the usual convention) is that where no raising takes place delta can be replaced by *it*. If the embedded delta is replaced by this process and thereby has some phonological shape it can then be raised to yield (48a). If no raising takes place both deltas can surface as *it's*, as in (47a). Example (47d) is properly prevented no matter when *it*-insertion takes place because whether the embedded NP dominates *it* or delta it still lies in the grossest analysis blocking the raising of an NP to replace the higher delta. Once the deltas are replaced by *it's*, no further application of raising is possible because the structural description of the rule is not met. In other words, **The woman seemed to be easy for to win the prize* will never be generated because the only way for *be easy* to occur with no NP subject (which would otherwise be in the grossest analysis blocking this sentence) is where its subject was *it* or some other lexically designated NP (the result of raising) and where this NP was raised to subject of *seem*. In either case no further raising could occur.

We will now look at some cases where there is an interaction between extraposition and raising and see what the implications are for cyclic rule application.²⁶

- (50) a. It seemed to be possible to persuade John Mary was the winner.

b. **John seemed to be possible to persuade Mary was the winner.*

c. **Mary seemed to be possible to persuade John to be the winner.*

d. **The winner seemed to be possible to persuade John Mary was.*

Examples (50c) and (50d) are readily prevented by the Variable Interpretation Convention. For (50c) the object NP *John* lies in the grossest constituent analysis if subject raising applies to the NP *Mary*. For (50d) object raising is blocked because again an NP, in fact two NP's, *John* and *Mary*, lie in the grossest analysis. The potentially problematic case is the ungrammatical (50b). For this example there would seem to be no NP in the grossest analysis to block the movement of the NP *John*. It would appear then that we might be back to relying on constraining raising to be a cyclic process. In that case (50b) would be prevented because on the POSSIBLE cycle no raising could occur.

This actually is not the right conclusion as we will see if we examine carefully the derivation of (50a). The relevant question has to do with why the VP *be possible* shows up here as subjectless. Earlier it was pointed out that the *it* which occurs with predicates like *surprising* is not replaceable by a raising process. *Possible*, like *surprising*, can occur with such an *it* or with a sentential subject.

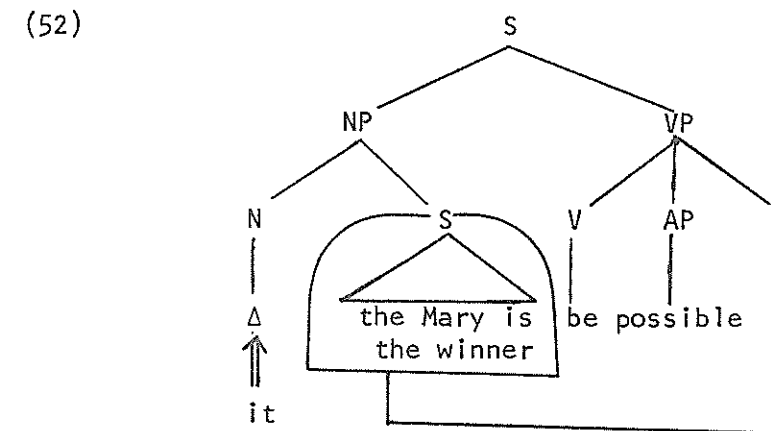
(51) a. It is possible that Mary is the winner.

b. That Mary is the winner is possible.

At the beginning of section 3.3 it was stated that lexical items like this (*possible*, *surprising*) co-occur with a non-referring, semantically empty *it* which is not the same as the *it* which co-occurs with raising verbs and adjectives, but rather which is the result of a rule of extraposition. The question of a transformational relation between sentences like (51a) and (51b) has been given a great deal of attention in the syntax literature. It has been argued (for instance by Rosenbaum 1967, Higgins 1973, Emonds 1976) that there is a rule of extraposition which would derive (51a) from a deep structure similar to (51b). Alternatively, it has been argued (for instance by Emonds 1970, Wexler and Culicover 1973, Culicover and Wexler 1973b) that (51b) is derived from a structure like (51a) by intraposition of the complement clause. Most recently it has been argued, by Jan Koster (1976), that in fact there is no transformational relation between the sentences of (51) but rather examples like (51a) have their complement clauses generated at the end of the VP and examples like (51b) have their complement clauses generated as a satellite under the nonrecursive initial symbol E, as a sister to \bar{S} , and "binding the (phonologically zero) NP-subject of the main sentence" (Koster 1976:7).

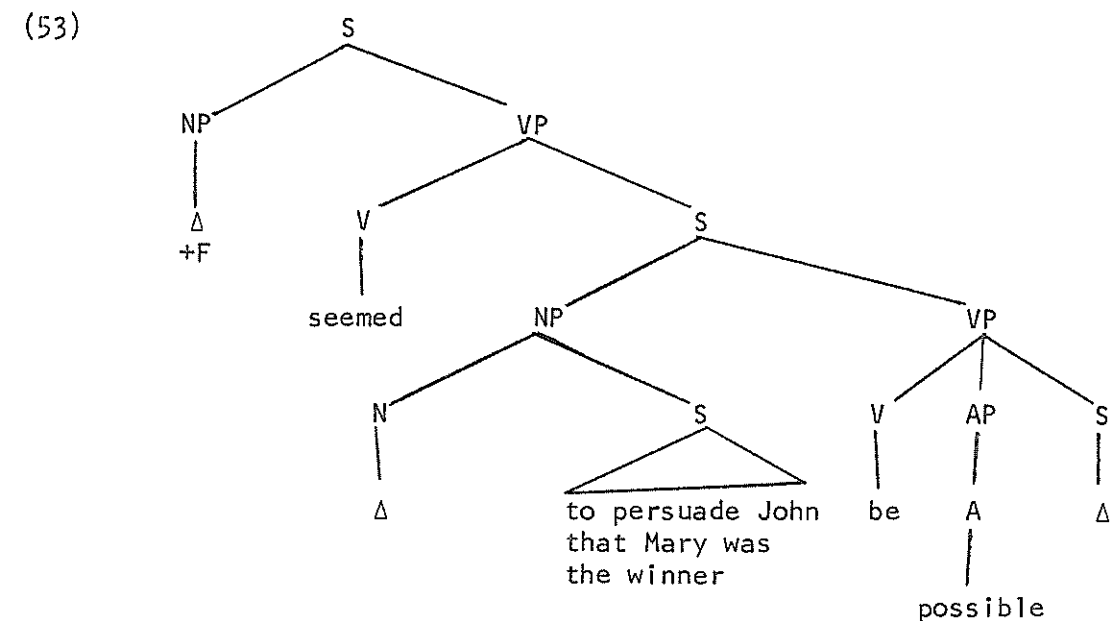
The Variable Interpretation Convention and the analysis of movements presented in this work does not seem to force a choice among these three alternatives. As will be shown in subsequent discussion (Chapter 6) however, the VIC can lead us to an interesting dichotomy between leftward and rightward movements if in fact there is a process of extraposition. Therefore in this work we will assume an extraposition account of the relationship between (51a) and (51b).

Following Emonds (1976:Chp. IV 2.2) the complement subject of (51b) would be generated as $NP[\Delta - S]$ --an NP containing an S which is sister to delta. Sentence (51a) is then the result of applying the rule of extraposition which moves the S to VP-final position and replaces the delta with *it*. This is illustrated in (52).



For Emonds, if extraposition does not apply there is an application of the rule of clausal topicalization which places the $NP[\Delta - S]$ in COMP position and deletes the delta.

With this in mind we can return to example (50a) and consider it as a result of the application of extraposition since the complement of *possible* occurs in VP-final position. The deep structure of (50a) would have been:

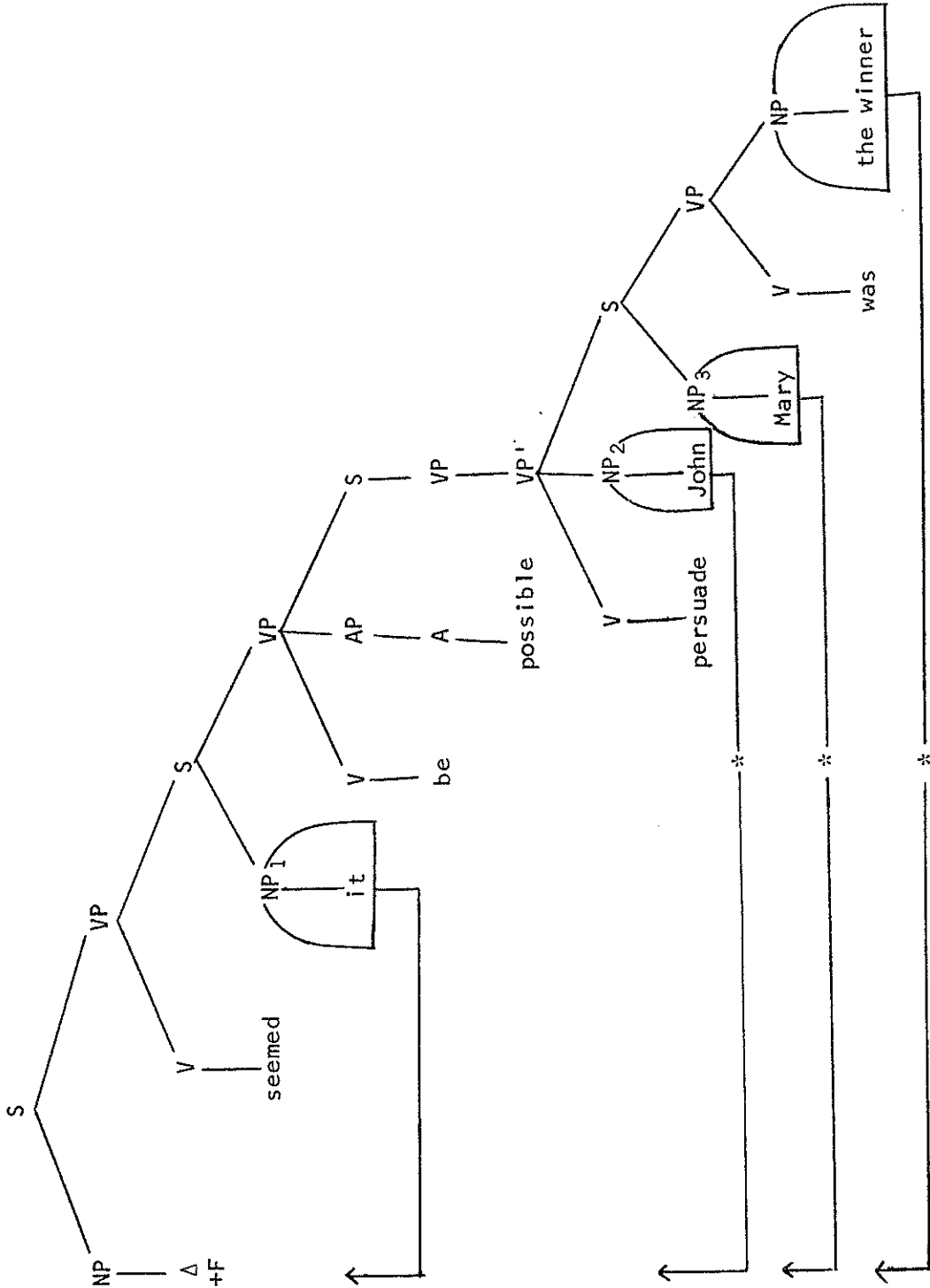


To this deep structure it is obvious that the VIC will properly prevent any NP from the embedded sentence from raising to subject of *seemed*. The N which is left sister to the S will always lie in the grossest analysis. As was pointed out in Chapter 2, where NP, which is N with the maximal number of bars is a term of a rule then any head of an NP may not be contained in the variable.²⁷ There is, however, no grossest analysis to block the raising of the whole sentential subject and, as the VIC would predict, the result of raising is well-formed.²⁸

(54) To persuade John that Mary was the winner seemed to be possible.

We can now see what happens in the case where extraposition applies. The result of applying extraposition to (53) is (55).

(55)



It is now easy to show how the facts of (50), repeated here for ease of reference,

- (50) a. It seemed to be possible to persuade John Mary was the winner.

b. *John seemed to be possible to persuade Mary was the winner.

c. *Mary seemed to be possible to persuade John to be the winner.

d. *The winner seemed to be possible to persuade John Mary was.

can be handled without requiring the raising process to work cyclically. None of the NP's contained in the extraposed S (NP₂, NP₃, or NP₄) is permitted to undergo raising because in each case NP₁ occurs in the gross analysis. NP₁ may undergo raising and it is this application of subject raising which produces the grammatical (50a). No further application of raising is possible since there is no way for the structural description of raising to be met because there is no longer any occurrence of a replaceable delta.

This accounting of (50), along with the other data thus far discussed, is strong evidence that the raising process need not be EXTERNALLY constrained to apply cyclically. It seems, rather, that this rule can be said to apply wherever its structural description is met and the fact that it does apply cyclically (as illustrated in (47)) is a result of the Variable Interpretation Convention.

Additionally it can be pointed out here that nowhere in this discussion of the two rules of raising and extraposition was there a need for an extrinsic ordering statement. It seems thus far that where the rules are constrained by the VIC there is no need for rule ordering.

3.5. Raising to Object and the Variable Interpretation Convention

In section 3.2 it was shown that the status of a rule of subject raising to object was certainly open to question. This rule was briefly discussed before the raising process was actually formalized as (26). A return now to a consideration of raising to object will show that whether or not this rule is assumed to exist there seems to be an unsolved problem, or at least unaccounted for data, in this VIC account of NP movements. When we look again at the process of passivization in light of the VIC, in Chapter 5, and discuss how that process interacts with the raising

phenomenon, we will see that in fact the issue of raising to object leads to no inconsistencies internal to the model. In this section, for now, we will see that there are data which seem to lead us in conflicting directions.

The first relevant examples are given in (56).

- (56) a. *Smith was easy for Jones to expect to recover.²⁹

b. Smith was easy for Jones to believe to be the winner.

The most plausible underlying structure for these examples, which would be consistent with the discussion of subcategorization and object raising (section 3.3.1), would be where *easy* has a VP complement, $_{VP}[to\ expect\ Smith\ to\ recover]$, $_{VP}[to\ believe\ Smith\ to\ be\ the\ winner]$. Whether or not there is a rule of raising to object (56a) and (56b) must surely be derived in the same way. And whether the NP *Smith* was a subject or an object in (56) would be irrelevant since there is only one raising rule which would be applicable in either case. The structural description for raising is met by either (56c) or (56d), either one of which could give rise to (56b).

- (56) c. Δ was easy for Jones $_{VP}[believe\ S[Smith\ to\ be\ the\ winner]]$
+F

d. Δ was easy for Jones $_{VP}[believe\ Smith\ VP[to\ be\ the\ winner]]$
+F

The ungrammaticality of (56a) therefore, at this point, seems problematic for the VIC.

An interesting thing about the examples of (56) is that while for my intuitions the grammaticality facts illustrated are accurate there exists a lot of disagreement among speakers of English about (56b). For all speakers it seems to be the case that (56a) is unacceptable. The status of forms like (56b) however is questionable. Consider also:

- (57) a. It is impossible to expect John to understand that book.

b. *John is impossible to expect to understand that book.

c. It is difficult to believe John to have made such a mistake.

d. *John is difficult to believe to have made such a mistake.

These examples in (57) are taken from Arlene Berman (1973) where she claims (among other things) that "Tough-movement may not apply to any noun phrase that has been Raised into object position" (p. 31). And this generalization

about the data seems in general to be true. NP's which might arguably be derived objects do not generally undergo object raising. With respect to (56), (56a) represents the paradigmatic case. For now we will leave these examples and return to them in Chapter 5 where they will all be accounted for based on the semantic properties of particular lexical items.

Next we will look at some examples which might lead us to conclude that raising to object should not, for the purposes of the model presented in this work, be included in the grammar of English. From Berman's sentences (here (57a) and (57c)), not only are the (b) and (d) versions out, but so are the examples of (58) where object raising applies to the object of the embedded clause.

(58) a. *That book is impossible to expect John to understand.

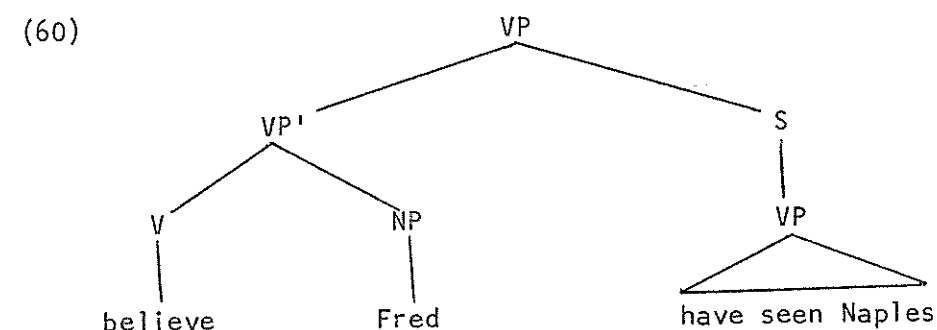
b. *Such a mistake is difficult to believe John to have made.

These facts are readily predicted by the VIC if the NP *John* is considered to be the subject of an embedded tenseless S in each case. For both examples it would then lie in the grossest analysis blocking the movement of the object NP. (This is the same account as would be provided by use of Chomsky's Specified Subject Condition.)

There seems to be some sentences formally very similar (probably identical) to those in (58) but which seem well-formed, or certainly better, such as (59).³⁰

(59) Naples is easy for Mary to believe Fred to have seen.

The grammaticality of this form would seem to argue in favor of a raising to object process where the derived structure of the VP would be:



where the object NP *Naples* could raise across the derived direct object (*Fred*) which would be contained in VP'. In order to account for the status of (59) it would be necessary to give up an account of the ungrammaticality of (58). Since the examples of (58) represent the more usual case it seems that the advantage, within a VIC account of raising, lies with assuming

that there is no rule of raising to object. In that case at least the results of (58) are properly predicted.

In Chapter 5 it will be shown that it is possible to account for the facts of (56), (57), and (58), and to account for the fact that in derived structure verbs like *expect* seem to have object NP's. This will be done, however, with no rule of raising to object. The grammaticality of (59) will remain anomalous.

At this point it should be made clear that within the VIC framework the rule of raising to object is readily formulable (in fact it is just a subcase of the subject raising process as was shown in section 3.2) and its existence or non-existence does not in any way affect the viability of the model as a whole.

3.6. Alternatives to a Global Constraint on Raising

This section will deal with some as yet unaccounted for problems and, where possible, point out some suggestive, if not fully convincing, answers.

In the same paper mentioned above, Berman (1973) makes the generalization that:

...Tough-movement may not move any noun phrase that has been previously moved by any other transformation [footnote omitted], or, otherwise stated, that Tough-movement may move a noun phrase only from its position in underlying structure.

(p. 39)

Berman again seems to be stating a generalization which is basically accurate. It can be seen in the analyses presented in previous sections that where object raising is illustrated it is functioning to move an NP from a base-generated position. Berman seems also, however, to consider her generalization in some way explanatory, thus assuming that a rule can 'know' the derivational history of the structure which meets its structural description. Since this approach is incompatible with the framework presented here, we will look at Berman's examples and consider some alternatives.

We have already discussed the examples (section 3.5) where it seems that object raising does not apply after subject raising to object, and as was said above, we will be returning to these examples later. Another set of examples which Berman presents is illustrated in (61) and (62).

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

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

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

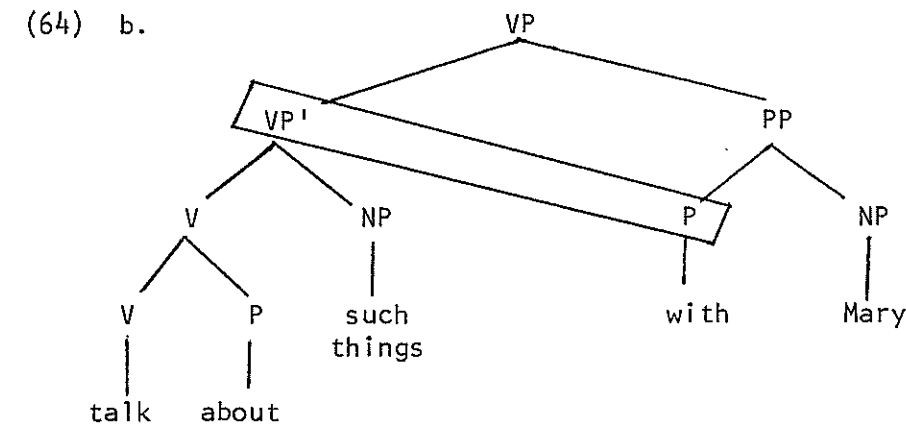
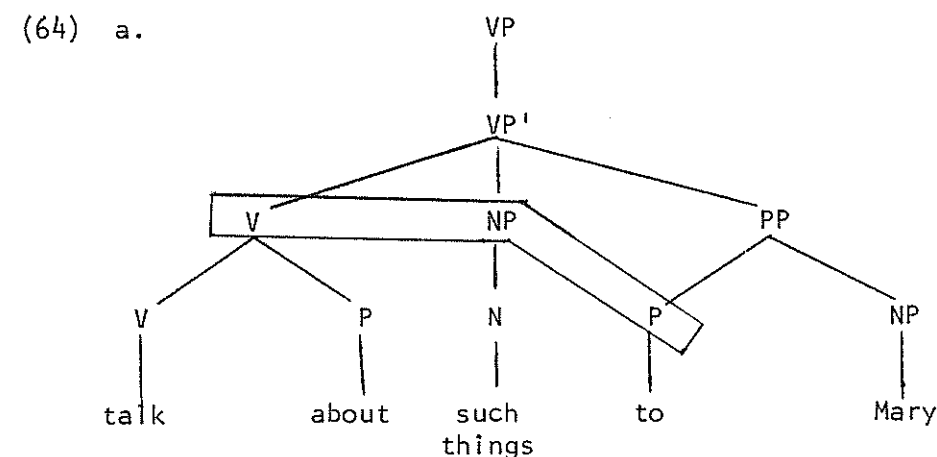
- (62) a. It is difficult to talk about such things to Mary.
 b. *Mary is difficult to talk about such things to.
 c. *Such things are difficult to talk about to Mary.

Granting that the grammaticality facts are as Berman presents them (although to me (62b) and (62c) are not obviously ill-formed), the generalization she makes is that *tough*-movement cannot operate after the order of the NP's has been affected by a transformation she refers to as *About*-movement, which would derive (62a) from (61a).

Regardless of the status of a transformation such as *About*-movement, whatever process creates the order of PP's in (62a) would do so in (63a) as well. But it seems that (63b) and (63c) are certainly well-formed.

- (63) a. It is difficult to talk about such things with Mary.
 b. Mary is difficult to talk about such things with.
 c. Such things are difficult to talk about with Mary.

It would seem then that the ungrammaticality of (62b) and (62c) could not be attributed to a special constraint on object raising. A conceivable alternative would be to consider different structures for the verb phrases of (62a) and (63a). It might be that when the verb *talk* is followed by an *about* prepositional phrase it is analyzed somehow as a complex verb *talk about*. This verb might then have a direct object NP and a closely associated *to*-PP, but a *with*-PP might be considered to be outside the VP'. In other words, compare (64a) and (64b).



The indicated grossest analyses then account for the difference in grammaticality between (62b) and (63b). I think the above-proposed tests for constituency (section 3.4) might have given a certain amount of support for the proposed structural difference, although acceptability judgements are extremely subtle.

- (65) a. *I tried to talk about such things $\left\{ \begin{smallmatrix} \text{naturally} \\ \text{myself} \end{smallmatrix} \right\}$ to Mary.
 b. ?I tried to talk about such things $\left\{ \begin{smallmatrix} \text{naturally} \\ \text{myself} \end{smallmatrix} \right\}$ with Mary.

There doesn't seem to be any structural difference, however, which could account for the difference in status between (62c) and (63c). It must just simply be that for some reason the sequence of prepositions *about to* is prohibited but *about with* is allowed.

Another set of examples presented by Berman has to do with object raising applying after passive has applied in the embedded clause. Her examples are:

- (66) a. It is unpleasant to be kicked by John.
 b. *John is unpleasant to be kicked by.
 c. It is easy to be accepted by that group.
 d. *That group is easy to be accepted by.

Berman's generalization is:

If passive has applied in the embedded sentence of an adjective that triggers *Tough*-movement, *Tough*-movement may not subsequently move the noun phrase that has been made the object of *by*.

Again, as with (62), it is not obvious to me that the examples (b) and (d) of (66) are ill-formed. But granting Berman's grammaticality facts, I would argue that this does not represent a fact about object raising. I think, rather, that object raising should apply to these *by*-phrase NP's (in accord with the VIC of course) and that the possible ungrammaticality of the result is explained by other factors. For instance, consider, along with the examples of (66), those of (67).

- (67) a. *John is unpleasant for me to be kicked by.
b. *That group is easy for me to be accepted by.

For completely independent reasons, that is, reasons not related to the interaction of passive and raising, Jackendoff (1972:154) says that NP's which are raised bear no thematic relation in the higher sentence. For subject to subject raising Jackendoff says:

It is reasonable that a raised NP such as the derived subject of *seem* has no thematic relation at all with respect to its new clause, since thematic relations are related to deep structure grammatical relations....

With respect to object raising, in order to rule out sentences like **Tony is tough for himself to shave*, Jackendoff claims:

... the *for*-phrase is a deep structure constituent of the main clause, but the subject is a deep structure constituent of the complement. If this is the case, the subject has no thematic relation in the main clause, whereas the *for*-phrase does. The modification of the Thematic Hierarchy Condition we have just proposed to handle *It*-replacement³¹ guarantees that whatever the thematic relation of the *for*-phrase, it will be higher on the hierarchy than the subject.

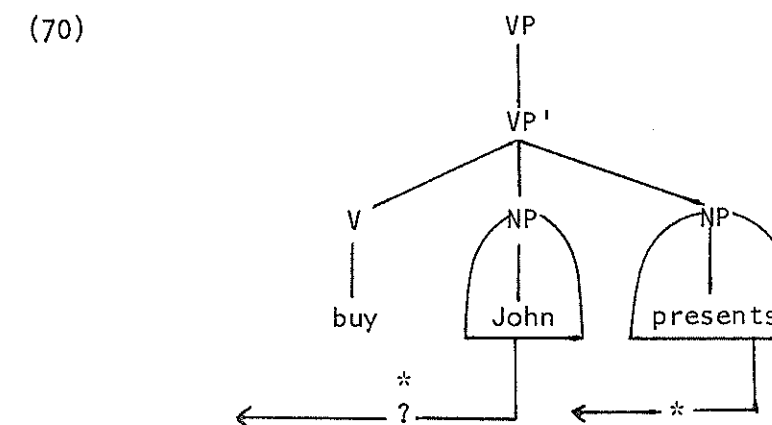
This then provides for a possible account of the ungrammaticality of (66b), (66d) and (67). The occurrence of the *by* in these sentences is an overt indication of the agentiveness of some NP or other. The derived subject NP must be lower on the thematic hierarchy than the *for*-phrase of the main clause in examples like (67). "Agent", however, is the highest relation in the hierarchy. It seems then that an account of (67) might well be that there is a "conflict" in the interpretation of thematic relations. The same basic thing would be true also for (66b) and (66d)--the derived subject of raising constructions cannot be properly interpreted as an agent.³² Where some NP of the embedded clause is raised and originally had some relation on the Thematic Hierarchy which was higher than a *for*-phrase of the matrix, BUT NO OVERT MORPHOLOGICAL SIGN OF THE RELATION REMAINS AFTER RAISING, then ungrammaticality is avoided. So it is not just the underlying agentiveness of *John* in these examples but

that agentiveness along with the occurrence of the *by*, which so strongly indicates "agent", which results in the interpretation problem with respect to thematic relations.

So far there is no reason to assume there must be some special constraint on object raising in order to account for Bresnan's ungrammatical results of applying the rule. There is, however, one set of data which she presents that I, thus far, cannot even suggest an account of. This involves the operation of object raising following indirect object movement:

- (68) a. It is impossible to buy presents for John.
b. Presents are impossible to buy for John.
c. John is impossible to buy presents for.
(69) a. It is impossible to buy John presents.
b. *Presents are impossible to buy John.
c. *John is impossible to buy presents.

The structure of the VP for (69a) would be:



From this it is easy to see how the VIC predicts the ungrammaticality of (69b) but no explanation is provided for (69c). Notice that this would be true regardless of the status of the transformation of indirect object movement (see Oehrle 1976; cf. Emonds 1972). Even if the VP configuration in (69a) is a base structure there is no account of (69c). Berman's account is: "Tough-movement may not move any noun phrase which has been 'displaced' by prior application of Dative Movement" (p. 38). This describes both (b) and (c) of (69). Since (69b) is already explained by the VIC, it would seem natural to hope for an explanation of (69c). Up

to this point it remains unexplained in this framework (see, however, interesting discussion of dative movement in Culicover and Wexler 1973a).³³

In this section we have discussed many apparent counterexamples to the generality of the process of object raising and have accounted in some way for all but one case. Berman's global condition, presented at the beginning of the section, is rejected not only because it provides no explanation of the facts but also now because it would be descriptive of really only one case, that of (69c). The alternative proposal here is that object raising operates freely, with no special constraints. It is conditioned, as are all movement rules, by the VIC.

So far then not only is a significant amount of data properly accounted for by this VIC analysis of raising processes, but it also seems plausible that the cyclicity of certain processes is predicted. In Chapter 4 I will show that WH-fronting can be treated as an unbounded movement over a variable (that is, need not apply successive cyclically) and in Chapter 5 I will show that passivization is predictably a "local" process (where "local" here is used in a non-technical sense). In other words, what I hope to show by the end of this dissertation, among other things, is that where the transformational cycle seems necessary it actually follows from the Variable Interpretation Convention.

FOOTNOTES TO CHAPTER 3

¹No particular significance is attached, so far, to the use of the pronoun *it* as term 1 of this rule. The actual nature of term 1 will be discussed in detail below.

²I have been reminded both by Joel Rotenberg and by Peter Culicover that certain nouns can also control raising. For instance, *That girl is a cinch to win the game* is probably the result of subject raising where *cinch* would govern the process. See also footnote 23.

³As far as I am aware there is no process which makes use of gross constituent analysis rather than grossest constituent analysis. The definition of gross analysis is included to facilitate the statement of grossest analysis and also to indicate that for every phrase marker with respect to any structural description there will be at least one appropriate analysis.

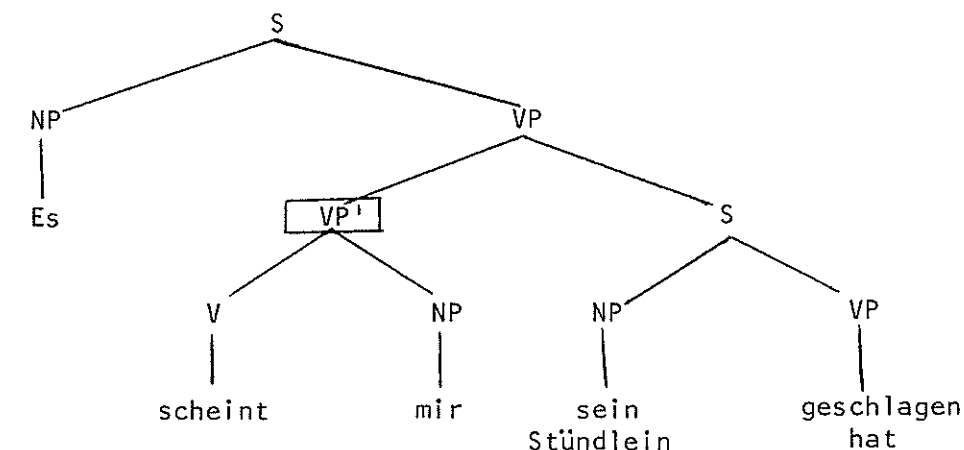
⁴Paul Schachter has pointed out to me that in German, where indirect objects do not take the form of PP's, raising is still possible. For example:

(i) Es scheint mir dass sein Stündlein geschlagen hat.
'It seems to me that his time has come.'

(ii) Sein Stündlein scheint mir geschlagen zu haben.
'His time seems to me to have come.'

This at first seems to present a counterexample to the Variable Interpretation Convention because the NP dominating *mir* would seem to lie in the grossest analysis and as such should block raising. I suggest, however, that this is not a counterexample because the phrase marker underlying (i) would be as in (iii), where the grossest analysis is indicated.

(iii)



The constituent VP' is introduced in this chapter in section 3.4 and used throughout the rest of this work. It seems likely that the internal structure of VP' for German is different from that for English. It seems plausible that the reason for this difference is that in German case may be indicated by morphology on the noun whereas in English it is generally indicated structurally, that is, by a prepositional phrase. As will be shown in section 3.4, where indirect objects are indicated by a prepositional phrase in English they lie outside VP'. After dative movement indirect objects occur, with direct objects, inside VP'.

⁵On the issue of context terms see footnote 10 of Chapter 2 and also a paper by Lasnik and Kupin (1976, unpublished) where it is argued that at most one context term be allowed in structural descriptions. It would be very interesting if it turned out that there was a difference between movements and deletions with respect to necessary terms in a structural description. Perhaps movements allow only category boundary contexts whereas deletions require an actual term to establish the proper context.

⁶Clearly rule (11) could as well be stated without the right bracket of the S-boundary. It is only the left bracket which is actually relevant since raising in English is a leftward movement.

⁷This fact was pointed out to me by Joel Rotenberg.

⁸Peter Culicover called to my attention the following interesting example, which bears not only on the issue of raising but also on the question of what can qualify as the head of a phrase.

(i) it is certain _S [_{NP} [hitting Mary] bother Fred]

(ii) Hitting Mary is certain to bother Fred.

(iii) *Mary is certain hitting to bother Fred.

Although the structural description is met, (iii) must be prevented. It can in fact be prevented by the Variable Interpretation Convention. The gerundive nominal *hitting Mary* is an NP which like any other NP must have a head. The head of this NP must be *hitting* (it is the non-optional element) even though it is a verbal form and not simply an N. An NP may not be moved over the head of an NP even where that head is a V instead of an N.

⁹This suggestion was made to me by Paul Schachter.

¹⁰For the analysis which follows, of subcategorization and extraposition and how they interact with raising, I am indebted to Joe Emonds for many long hours of invaluable discussion.

¹¹Notice that this subcategorization for *seem* means that whenever it occurs with a filled subject, that is, whenever it occurs in a surface string, raising has taken place. The sentences of (i) and (ii) are both the results of raising.

(i) John seems (to be) happy. $\Leftarrow \Delta$ seems [John be happy]
+F

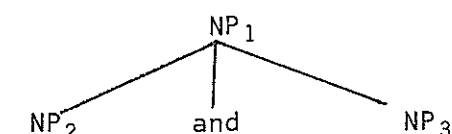
(ii) It seems (to be) a pity (that John is unhappy). $\Leftarrow \Delta$ seems
+F
[it be a pity [John be unhappy]]

There is evidently a late rule which can delete an occurrence of *be* after verbs like *seem*.

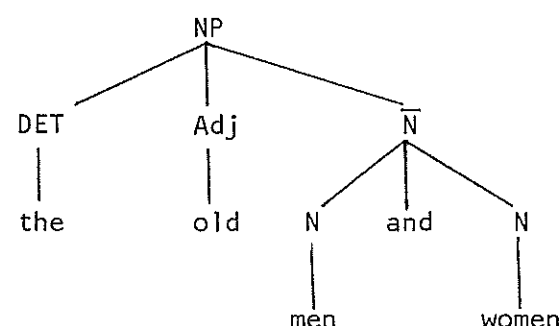
¹²Whether adjectives of this sort subcategorize a verb such as *be*, or whether it is the predicate (*be certain*, *be probable*) which subcategorizes a particular subject is irrelevant here so I will not attempt to choose between these two alternatives.

¹³The fact that *easy* occurs with a VP complement rather than a full S with an NP subject has been argued successfully by Bresnan (1971) and has been used repeatedly by others, such as Chomsky in discussion of the Specified Subject Condition (Chomsky 1973). When *easy* occurs in a string such as *It is easy for the rich to do the work* the NP *the rich* is contained in a PP of the matrix and the embedded clause is subjectless. See also discussion below of examples (27) through (29).

¹⁴Additional motivation for the definition of C^{max} is that it allows for the capturing of an interesting generalization about conjoined structures. In conjoined NP's for instance, where the structure would be:



neither NP₂ nor NP₃ should be considered the head of NP₁. Both would be instances of N^{max} , however, (as would NP₁). The prediction here is that only C^{max} 's occur in conjoined structures. In other words, conjoined structures do not have heads. Notice that C^{max} need not be C^n where n = maximal number of bars. In a structure like:



both instances of N qualify as C^{\max} since neither is the head of any bigger constituent (here \bar{N}), but \bar{N} is not a C^{\max} .

¹⁵Since raising cannot apply in (29b) it is not relevant here whether or not *for*-phrase formation (Emonds 1976) has applied in the embedded S. If this were a structure where an NP could be extracted from the embedded VP it would make crucial difference for the VIC whether the NP *John* were immediately dominated by S or contained in a PP.

¹⁶The sentence *He put the money very slowly into his pocket*, called to my attention by Paul Schachter, seems to be a counterexample to his generalization about where adverbs may occur. The relative obscurity of such examples, together with the fact that even this one is better as *He put the money into his pocket very slowly*, leads me to maintain that the generalization is accurate but that there are a few unexplained anomalies.

¹⁷Baltin's generalization is similar to a constraint on movement transformations suggested by Arthur Schwartz (1972). Schwartz's constraint says: "Boundary attachment: a phrase moved out of its phrase cannot attach to anything but the boundary of the next highest phrase" (p. 37).

¹⁸We are excluding here constructions which might be considered idiomatic, e.g. *It is easy to play tricks on Bill* ~ *Bill is easy to play tricks on*.

¹⁹Notice here that a prohibition against certain types of preposition stranding, an otherwise plausible approach, would be hard to motivate because of the similarity between, for instance, (36b) and (36d). A different alternative might be that verbs like *put* and *set* somehow incorporate following prepositions to form complex verbs and this is blocked where the order is V NP P.

²⁰In a previous discussion of raising in light of the VIC (Wilkins 1976) the facts illustrated in (32) were handled by a definition of what $B[A]$ means in a structural description. The definition was:

$B[A] = B[X A Y]$ where X does not contain an A immediately dominated by B.

This meant that a subject in the grossest analysis could block raising but that a direct object could not. This alternative is abandoned here because, among other things, there would then still be no account of examples like (35c).

²¹Even though this example (43c) does not seem as bad as examples like (33c) and (34c) I will maintain that the S-complement for *persuade* is contained in VP' because of the generality of the account that this allows for.

²²Actually, below, in discussion of examples like (47), the movement indicated in (44) would also be impossible because there would be a subject NP of *be easy* which was either Δ or *it*. *Mary* in (44) would still, however, be prevented from becoming the subject of *be easy* because the NP dominating *John* would be in the grossest analysis:

*It seems *Mary* was easy to persuade *John* to be the winner.

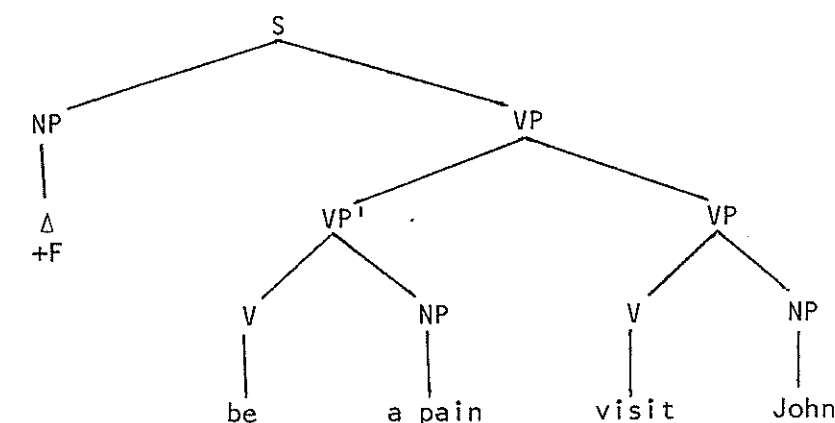
²³There is another type of example where the use of VP' seems indicated. For examples like:

(i) It is a pain to visit John.

(ii) John is a pain to visit.

If a *pain* were contained in the grossest analysis the grammaticality of (ii) would be unaccounted for by the VIC. If, however, the structure underlying (i) is (iii), the right results are predicted.

(iii)



This type of structure makes it seem more plausible that subcategorization for raising predicates needs to take into account not only the noun (*pain*, here) or the adjective (e.g. *easy*) but rather the predicate *be a pain*, *be easy*. See footnote 12.

²⁴Although to my intuitions sentences such as this, (45c), are unacceptable, it seems that for some speakers they are well-formed. Paul Schachter points out that it is better as *That book would be easy to persuade John to buy* where there is some understood contrastive meaning. As such these examples would be unaccounted for by the VIC.

²⁵It is interesting that Chomsky (1976), in order to account for the seemingly unbounded domain over which object raising takes place, argues that it is actually a case of WH-movement which applies successive cyclically.

²⁶The examples of (50) were pointed out to me by Joan Bresnan.

²⁷It will turn out that this instance of N, in order to block movement of an NP out of the embedded S, must dominate something more than a completely empty delta. (The reason for this is discussed in Chapter 4.) It might be that the N-sister to sentential subjects is more properly to be considered a pro-form *it*, or it might be that it is a delta which contains some syntactic feature but, like the delta for raising, no phonological or semantic features. The second alternative is certainly not implausible since this is a delta which is important for the syntactic processes of extraposition and clausal topicalization.

²⁸This is assuming, of course, that clausal topicalization applies (either before or after raising) to erase the delta.

²⁹This sentence is borrowed from Lightfoot (forthcoming) where he, assuming there should be no raising to object, accounts for the ungrammaticality by a violation of the Subjacency Condition (Chomsky 1973). Lightfoot says, referring to his example like (56a) and subject raising, "... but the rule (an extraction rule) will be blocked by the Subjacency Condition since the NP would be moved over two S nodes" (p. 21). Notice, however, that Subjacency would make the wrong prediction about (56b).

³⁰Sentence (59) was pointed out to me by Peter Culicover.

³¹The modification suggested by Jackendoff was to account for certain facts about reflexives. He says (p. 154):

... we will have to modify the Thematic Hierarchy Condition. One possibility that suggests itself is to say that the condition applies only to thematic relations INDUCED BY THE VERB OF THE LOWEST CYCLE in which the reflexive has antecedents.

³²Steve Cushing has brought to my attention the following sentences:

(i) John would be unpleasant to get kicked by.

(ii) That group would be easy to get accepted by.

where there is a sort of forcing of the volitional relation between the derived subject and the rest of the sentence. In these cases the result seems to be grammatical. By establishing a context which forces volition, agentiveness, onto the derived subject of a raising construction a sentence like (66a) can be made much better. For instance:

(iii) Because I am basically a masochist I really think that in a nice friendly soccer game John would be very pleasant to be kicked by.

This seems to suggest that resolving the conflict with respect to the thematic relations leads to grammaticality.

³³As Emonds has pointed out repeatedly (class lectures) dative-moved NP's in general exhibit this behavior of being immune to many transformational processes.

CHAPTER 4

WH-FRONTING

4.1. The Rule for WH-fronting

4.1.1. The Revised Left Branch Condition

In this chapter we will discuss in detail the process of WH-fronting in light of the Variable Interpretation Convention, which relies on the definition of grossest constituent analysis. We will see that there is no need to consider WH-movement as a successive cyclic process (Chomsky 1973, 1976) and that in fact externally constraining it in such a way misses the significant fact that because of its form this rule can, in principle, operate over an unbounded domain. It will be shown that the interaction between the form of the rule of WH-movement and the structure of phrase markers in English results in a process which seems to affect only subjacent domains. It will also be shown, however, that WH-fronting can be permitted to apply across larger domains and needs to be constrained only by the VIC. WH-fronting will be shown to define its own domain of applicability based on the form of its own structural description.

WH-fronting is the process in English which moves a WH-phrase (a question phrase) to the left into complementizer position. The only items in a string which would be affected then would be the WH-phrase and the COMP and therefore by the VIC they are the only terms which can occur in the structural description of the WH-fronting rule. The rule (subject to slight revision shortly) is:

$$(1) \text{ COMP} - \text{WH} \Rightarrow 2 - \emptyset^1$$

This rule must, of course, be construed in such a way that the term that moves is a whole phrase and not just the grammatical formative (or feature) WH. This proper construal is assured, in this framework, by a particular revision of the Left Branch Condition (Ross 1967).

The relevant version of the Left Branch Condition is based on that suggested by Emonds (1976) in his discussion of WH-fronting. Emonds (1976:183) states WH-fronting as:

$$(2) \text{ COMP} - X - \left\{ \begin{array}{c} \text{NP} \\ \text{AP} \\ \text{PP} \end{array} \right\} [(\text{P}) + \text{WH} + \text{Y}] - \text{Z} =$$

$$3 - 2 - \emptyset - 4$$

Emonds then says (p. 185, footnote 10) that this is "somewhat redundant" because it must already be stated in the grammar, in the base, that the

place WH is generated is in either COMP position or in specifier position for NP or AP. The WH-fronting rule then essentially says this again by including the phrasal category labels (and the optional P) and by saying that the rule applies only to WH's in these positions. Emonds then proposes to eliminate this redundancy by restating the Left Branch Condition as:

If a syntactic element C in N'' or A'' is to the left of the head in that phrase, then any major transformational operation that reorders C with respect to elements exterior to N'' or A'' must also reorder N'' or A''.²

Making use of this condition, Emonds says WH-fronting could then be written as:

$$(3) \text{ COMP} - X - (P) + \text{WH} - Z = 3 - 2 - \emptyset - 4$$

Recently Emonds (personal communication) has suggested a slight revision of his version of the Left Branch Condition. It is currently to be stated as given in (4).

(4) Revised Left Branch Condition:

If a syntactic element C in N'' or A'' is to the left of the head in that phrase, then any major transformational operation that reorders C with respect to elements exterior to such N'' or A'' must also reorder all such N'' or A'', and no larger X''.³

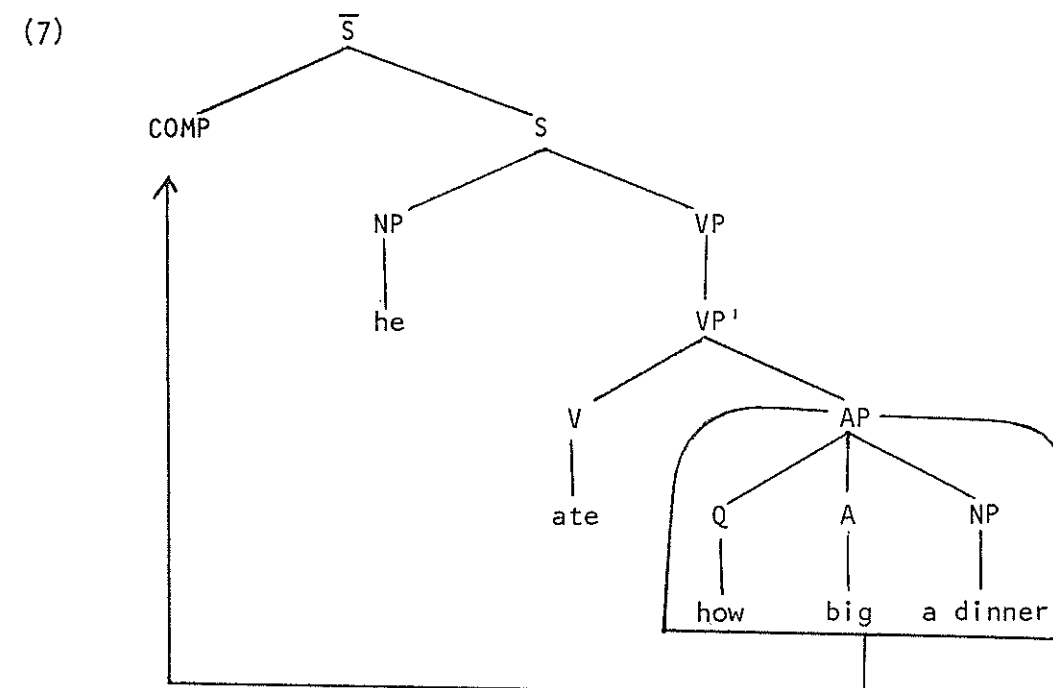
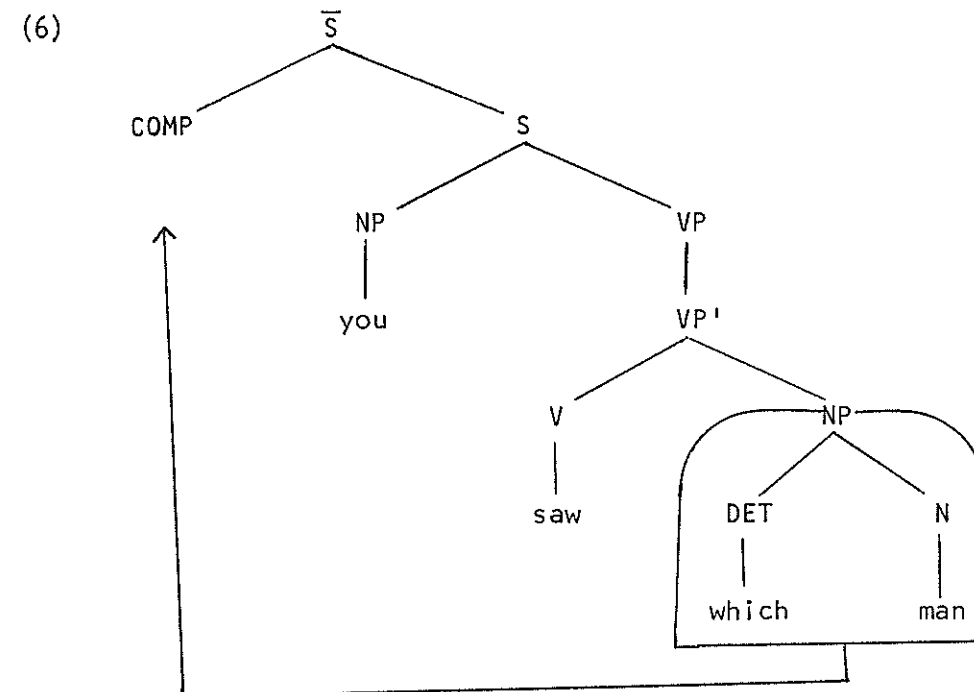
In the VIC framework, assuming condition (4), Emonds' rule (3) can be further revised to:

$$(5) \text{ COMP} - (P) + \text{WH} = 2 - \emptyset$$

This now is the basic form of the WH-fronting rule we will be using in this work and it will be shown to be an adequate formulation of the process.

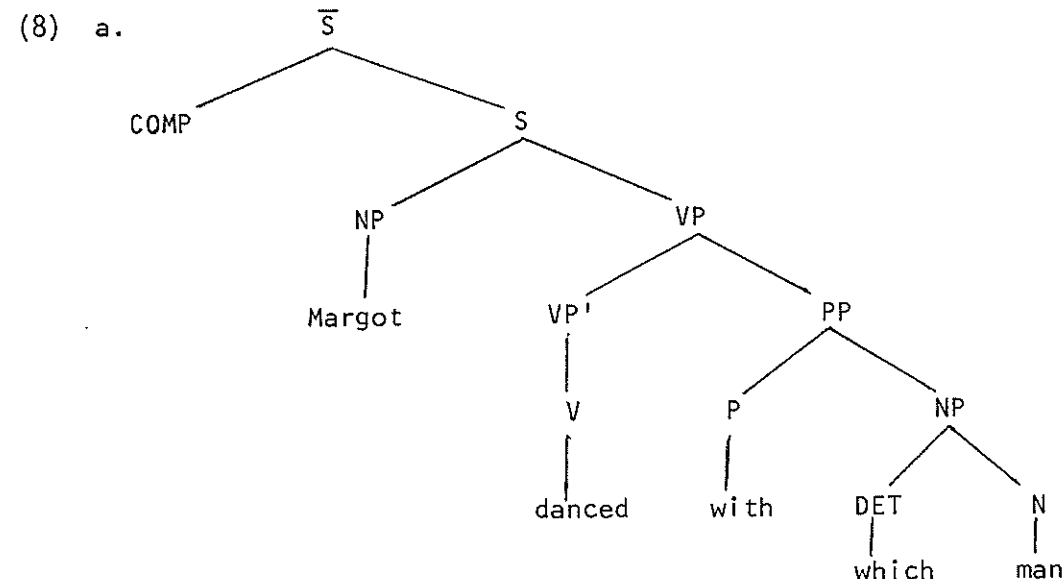
First we will see that (5) functions properly to affect NP, AP, and PP, but not VP. The structural description of (5) includes the optional P (= preposition) because clearly WH-fronting can move prepositional

phrases, (e.g. *[With whom] did you eat dinner?*). Examples (6) and (7) illustrate the functioning of WH-fronting to an NP and an AP, respectively.



In both cases, where rule (5) moves the WH-term, the Revised Left Branch Condition (4) assures that the whole NP or AP moves to COMP position along with the WH.

In (8) the WH occurs in a PP.

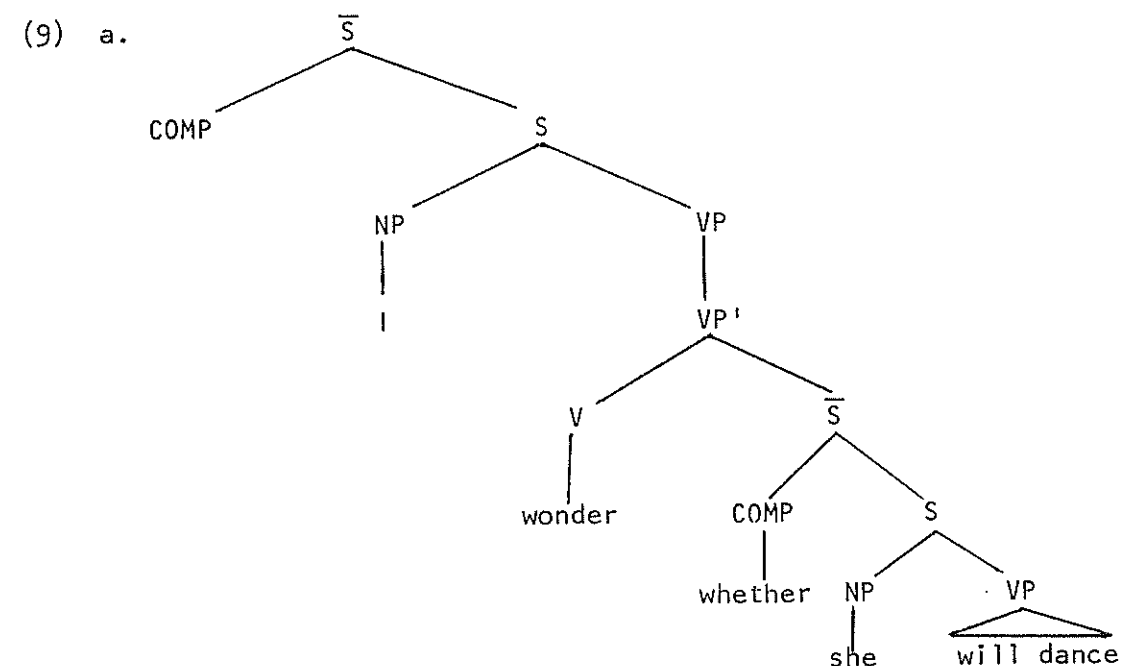


b. Which man did Margot dance with?

c. With which man did Margot dance?

Where the structural description of (5) applies to (8a) without the optional P, the result is (8b). Where the full expansion is used the P, along with the whole NP, is fronted and the result is (8c).

Consider now (9).



b. *Whether did I wonder she will dance?

c. *Whether she will dance did I wonder?.

Both examples (b) and (c) here must be accounted for. (9b) is blocked by a general convention, to which we will return below (section 4.5), which prevents anything from being extracted out of COMP position. (9c) is blocked because, by the statement of the Revised Left Branch Condition, there is no provision for the movement of \bar{S} (or S or VP, for that matter) by virtue of a left branch being reordered. Since the movement of a WH in COMP (*whether* or any other WH-term) will neither invoke condition (4), nor may that WH move by itself, no examples like (9b) or (9c) will ever be generated.⁴

We have now seen that the structural description of (5) can accurately account for the terms which undergo the actual movement of the rule. The VIC, thus far however, does not say anything about OPTIONAL terms. There must be a statement to the effect that structural descriptions can mention optional term(s) which are affected by a rule, but that they are not relevant for determining what can occur in the variable material. In other words, WH-fronting can apply freely across P's or PP's in the variable (*Who did Mary talk to?*, *Where did they search for John?*). The VIC is therefore amended to read as follows:

(10) Variable Interpretation Convention:

In the structural description of a transformation only terms which are crucially affected may be mentioned and for all non-optional terms, $A - B \equiv A - X - B$, where (a) X is a variable in the structural description and corresponds to the grossest constituent analysis of a phrase marker, and, (b) the grossest constituent analysis does not contain any A or B or head of an A or B.

4.1.2. WH-terms in the Grossest Analysis

The VIC now says that for rule (5) anything can occur between the non-variable terms except COMP or WH.⁵ Consider the examples of (11).

(11) a. *Who does John like the man who saw?

b. *What does John forget when to do?

- (11) c. *What did John remember why he bought?
- d. *Who did the man who saw talked to John?
- e. *What did the man who ate ordered dessert?
- f. *Which issue is what they stand for on that we object to?
- g. *Which mazurka did we question her about who danced?

These examples are ungrammatical because in each case a WH occurs in the variable material between the COMP and the WH affected by the rule. The examples of (12) are ungrammatical for the same reason.

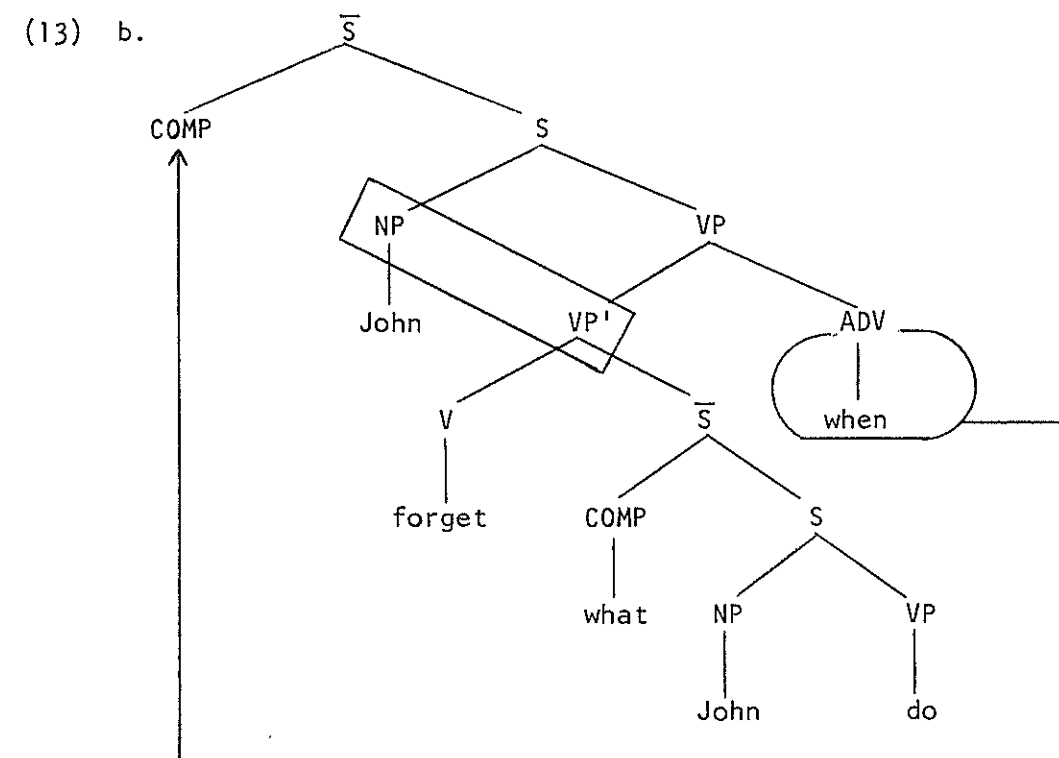
- (12) a. *What did John forget whether he ate?⁶
- b. *Who did John ask whether saw the man?

The WH-term *whether* cannot occur in the variable material of the WH-fronting rule.⁷

The VIC says not only that no WH can occur in the variable of WH-fronting, but also that no COMP can. The examples of (11) and (12) then are in a sense doubly excluded since the WH term is in COMP position and either WH or COMP in the grossest analysis will prevent the rule from applying. We will return below to a refinement of the statement of the rule for WH-movement and it will then be clear that what excludes these examples is a filled COMP in the grossest analysis. In other words, full use will be made of the fact that the WH of relative clauses and indirect questions is in COMP position.

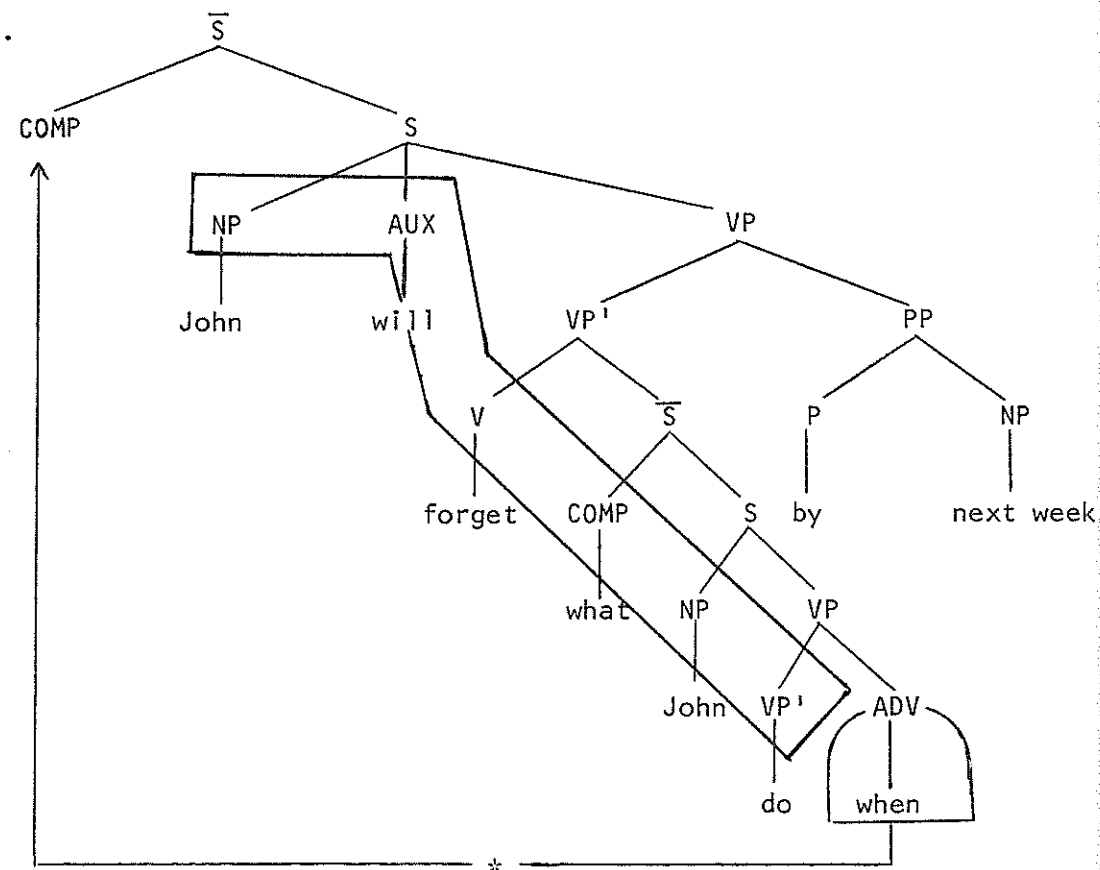
A comparison of examples (11) and (12) with that in (13) shows why again with WH-fronting (just as with the rules previously discussed) it is important to consider only the grossest constituent analysis for variable material.

- (13) a. When did John forget what to do?



Sentence (13a)⁸ represents a case where a WH (the *what* in *John forgot what to do when*) occurred between COMP and a WH to be moved. What is crucial here is that the WH did not occur in the grossest constituent analysis between the COMP and *when*, as the phrase marker (13b) shows.⁹ Compare this with a phrase marker where the adverb must be in the subordinate clause.

(14) a.

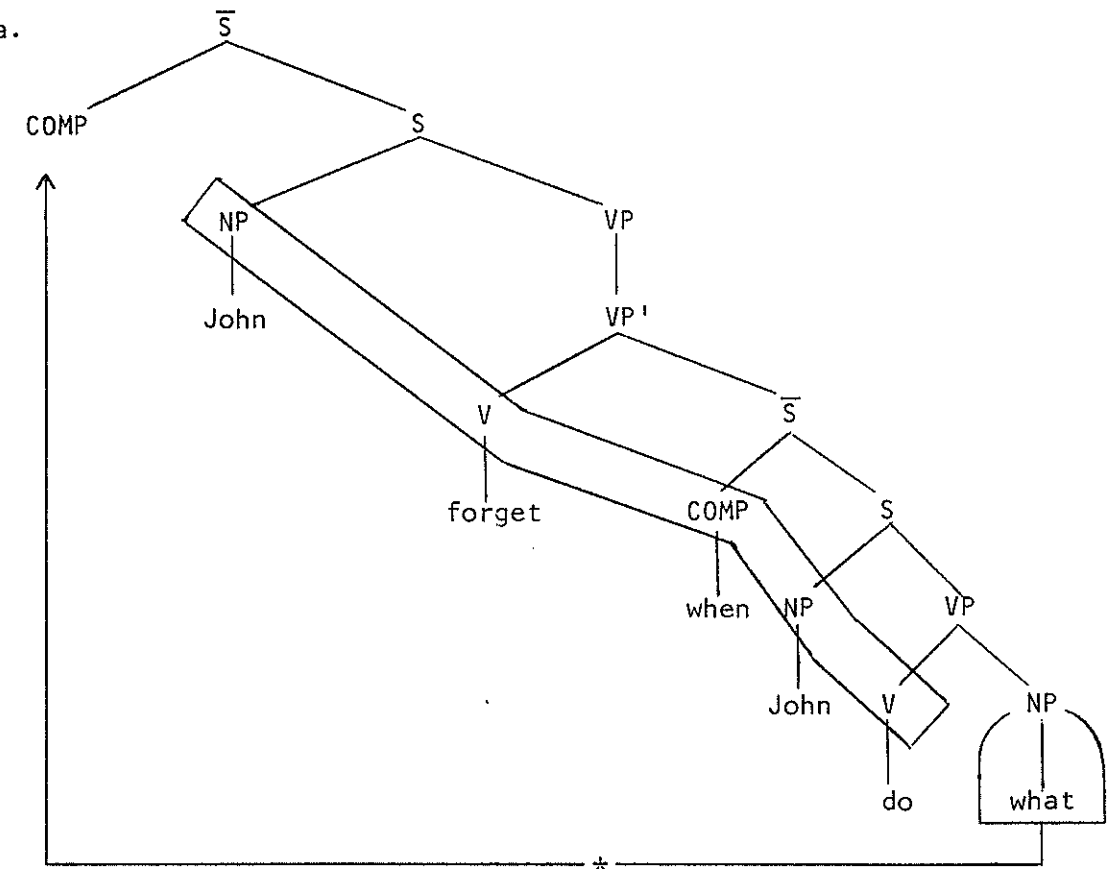


b. John will forget what to do when by next week (= By next week, John will forget what to do when.).

c. *When will John forget what to do by next week?

Also compare to (13b) phrase marker (15) for example (11b).

(15) a.

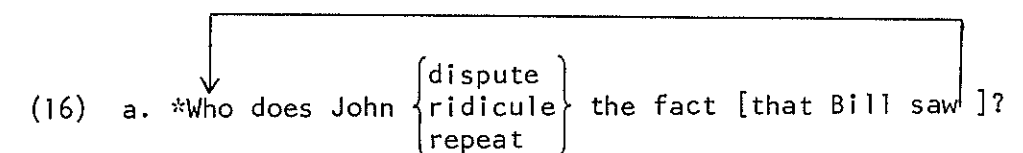


b. *What did John forget when to do? (= (11b))

A COMP contained in the grossest analysis blocks the application of WH-fronting in (14c) and (15b). In (13a) the only interpretation of the sentence is one where what is being questioned is the time of the "forgetting", that is, it can only be derived from a phrase marker like (13b) where the adverb is outside the subordinate S and hence there is no COMP in the grossest analysis.

4.2. That-complements and Relative Clauses

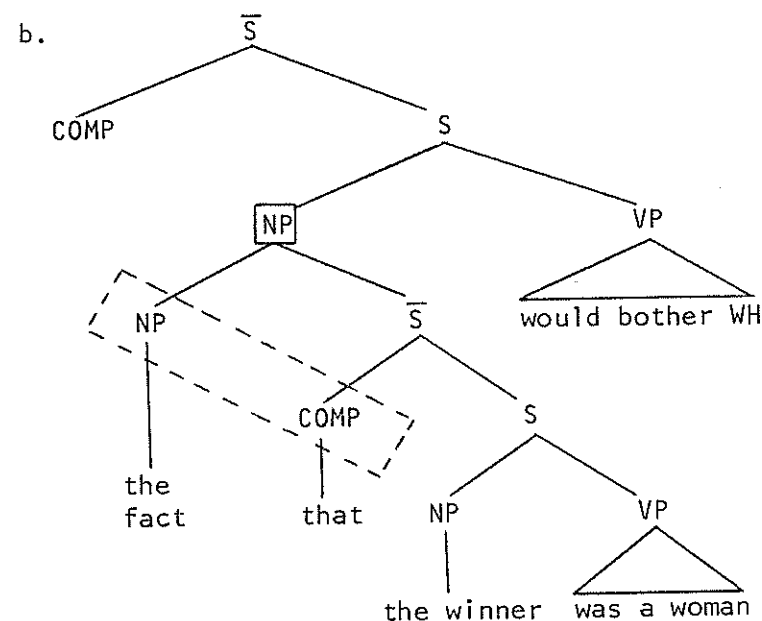
We can now examine cases where the COMP contained in the variable is not a WH.



- (16) b. *Which discovery did those facts undermine the claim
[that Harry made]?
- c. *What did [that John bought] bother Bill?
- d. *Which cookie did the fact [that Mary ate] annoy Tom?
- e. *Which movie did John eat so much [that he could hardly
watch]?

It will always be the case that WH-fronting will produce unacceptable results when an NP is extracted from a sentential subject (e.g. (16c)), a complex NP (e.g. (16b)), or a subordinate result clause (e.g. (16e)).¹⁰ In all these cases there will be a COMP, namely the COMP dominating *that*, in the grossest analysis. A comparison of the examples of (16) with that given in (17) again shows why it is only the material in the grossest analysis which is relevant for the VIC.

- (17) a. Who would the fact that the winner was a woman bother?



Since it is an NP which occurs in the grossest analysis, and not a WH or other COMP, sentence (17a) is well formed. Notice, however, that extraction

from the embedded \bar{S} would be properly blocked since then the COMP dominating *that* would occur in the grossest analysis (this analysis is indicated by the broken line).

Consider next examples (18) through (20).

- (18) Who does Mary believe (that) John said Bill hit?

- (19) Which woman did you {a. think (that)
b. *forget why
c. *answer that} he hit?

- (20) a. *What did John quip that Mary wore?

- b. *What did John complain that he had to do this evening?
(from Chomsky 1973, citing Dean 1967)

For each of these examples there is a COMP in the grossest analysis. But in (18) and (19a) WH-fronting produces acceptable sentences. The important fact illustrated here is that WH-fronting is permitted over a SUPERFICIALLY OPTIONAL complementizer, that is, over a complementizer which is optional in surface structure. According to Chomsky (1973), precisely this fact is pointed out by Dean (1967):¹¹ verbs which require a following *that* do not allow WH-fronting.

In this framework the correlation between the occurrence of an obligatory *that* and the blocking of WH-fronting is not just accidental. Where a complementizer occurs in the grossest analysis for WH-fronting the rule cannot apply. The examples of (18) through (20) show us, as will become clear below, that where a terminal element is empty of lexical material it "does not count" with respect to variable interpretation. The fact that lexically empty nodes¹² are irrelevant for the VIC follows from the definition of a transformational operation. The structural description of a transformation breaks up the TERMINAL STRING into SUB-STRINGS. If some node is empty it has no terminal symbol. Where there is no terminal symbol the VIC does not take a node into account. A more detailed explanation of this is as follows.

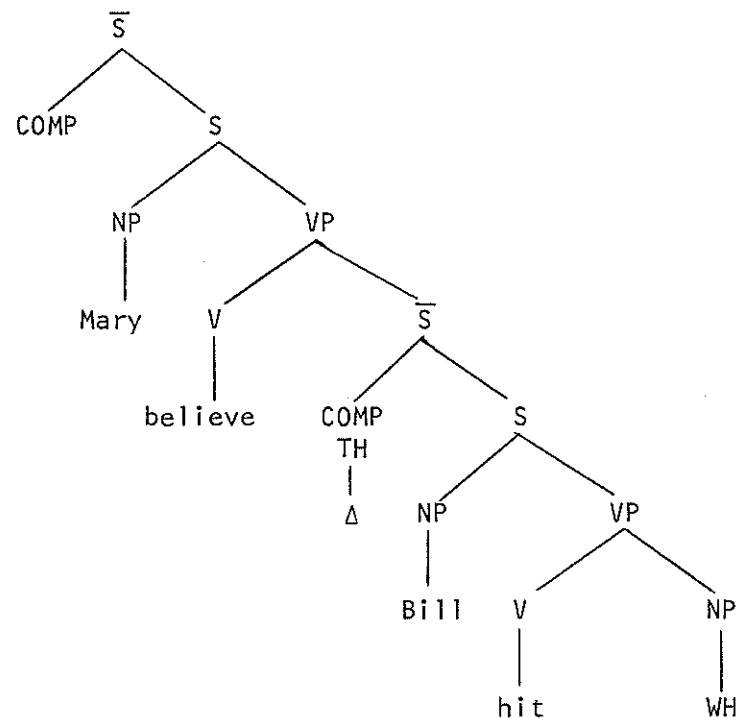
In deep structure the COMP node is obligatory. The phrase structure expansion for \bar{S} is:

- (21) $\bar{S} \rightarrow$ COMP S
 { TH
 { SUBJ
 { WH }

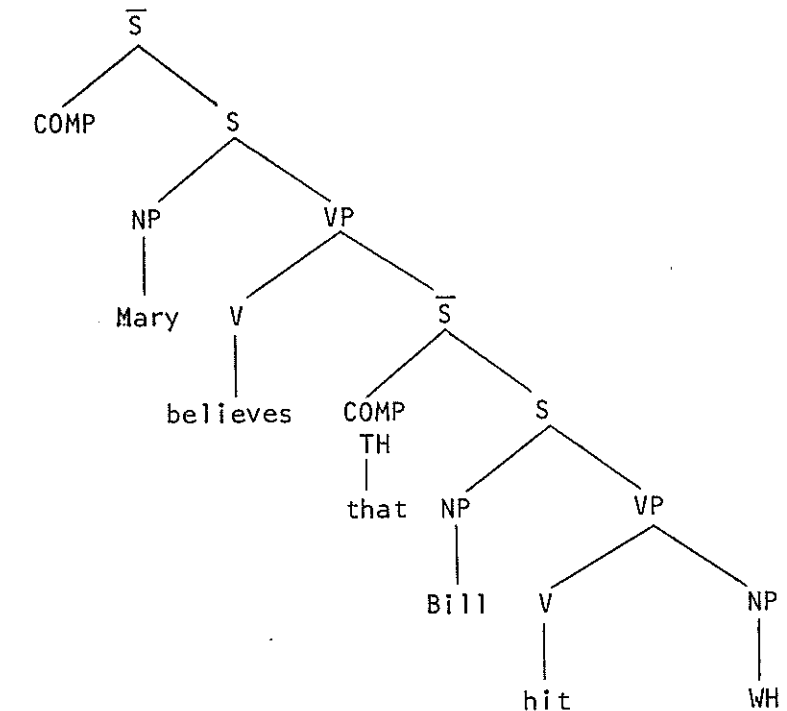
where TH is associated with deep structure *that* (see Pope 1972); SUBJ is associated with *for-to* (see Emonds 1970; Stockwell, Schachter, and Partee

1973; and Emonds 1976); and WH with *whether*. Certain verbs, such as those noticed by Dean (e.g. *quip*, *complain*), which require an obligatory complementizer, would only be inserted into the base phrase marker where the appropriate subcategorized complementizer already appears. Lexical insertion for other verbs, such as *believe* above, would apply freely and would not depend on subcategorization of a complementizer. Where in deep structure the complementizer is lexically designated WH-fronting cannot take place if the COMP node lies in the grossest analysis. WH-fronting IS allowed, however, where the COMP node is empty in deep structure.¹³ Compare the phrase markers of (22).

(22) a.



(22) b.



It is only phrase marker (22a) which can underlie the sentence *Who does Mary believe (that) Bill hit?*¹⁴

What must be accounted for next, of course, are the cases of WH-fronted examples where the *that* occurs in surface structure. There is a rule of *that*-insertion (optional like all others) of the form:

(23) $\text{COMP}_{\text{TH}}[\Delta] \rightarrow \text{that} / \text{--- NP}$

The conditioning NP environment on (23) is necessary to distinguish the acceptability of examples like (24).

(24) a. Who does Mary believe (that) Bill hit?

b. Who does Mary believe (*that) hit Sam?

Since the complementizer is obligatory in deep structure (by rule (21)) and since I am assuming (following the usual convention) that any time a delta occurs in surface structure the string is marked as ill-formed, then the rule for inserting *that* would be, thus far, *de facto* obligatory. Where it does not apply there will be a delta in surface structure. To account for the cases where *that* does not in fact surface there are the two following rules for the deletion of COMP which have been proposed by Joe Emonds (personal communication).

(25) a. NP - COMP[that] - NP ⇒ 1 - ∅ - 3

b. V + (NP) - COMP[Δ]
TH ⇒ 1 - ∅

Rule (25a) is relevant for relative clauses, to which we will return shortly. Rule (25b) is relevant for two different types of cases. Where rule (23) may not apply, because its structural description is not met, as

in *Who does Mary believe COMP[Δ] hit Sam?*, the delta must not cause the string to be marked as ill-formed. Also, where (23) does not apply because

it is optional, e.g. *Who does Mary believe COMP[Δ] Bill hit?*, the string must not be marked ill-formed.

If we can assume that rules which do ONLY insertion or deletion of grammatical formatives, rules such as (23) and (25a,b) operate very close to the surface, or at least after all the movement rules, then there is an adequate account of WH-fronting in the VIC framework which captures the generalization that there is a very high correlation between extractability and the occurrence of surface complementizers. To phrase marker (22a) WH-fronting can apply to give the structure corresponding to (26a).

- (26) a. *Who does Mary believe [Δ Bill hit]?*
b. *Who does Mary believe that Bill hit?*
c. *Who does Mary believe Bill hit?*
d. *Mary believes that Bill hit who?*
e. *Mary believes Bill hit who?*

To (26a) *that*-insertion, rule (23), can apply to give (26b). (26c) is derived by an application of (25b) to (26a). If WH-fronting does not apply to (22a), the result of rule (23) is (26d) and the result of rule (25b) is (26e).

Where relative clauses appear on the surface with no complementizer between the head and the clause, this is the result of deletion rule (25a). Example (11a) (here repeated) shows that extraction is prohibited where a relative clause contains a WH.

- (11) a. **Who does John like the man who saw?*

Extraction is also to be prohibited in examples like (27c) and (27e).

(27) a. *John likes the man that saw who?*

b. **John likes the man saw who?*

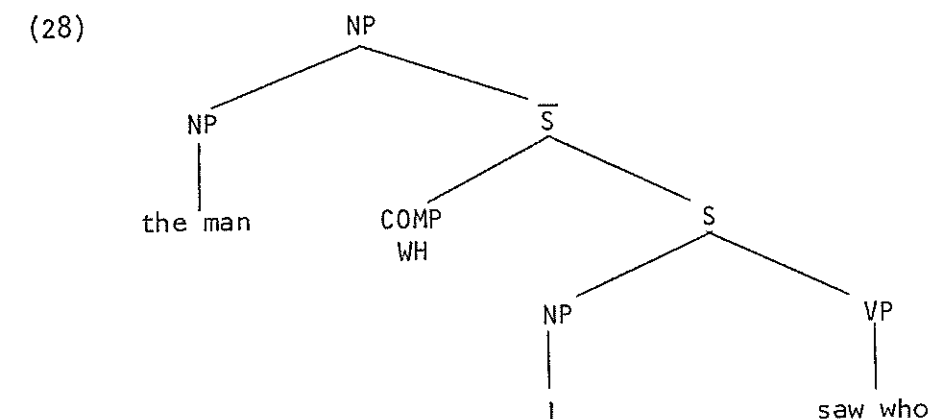
c. **Who does John like the man (that) saw?*

d. *John likes the man (that) Bill gave what to?*

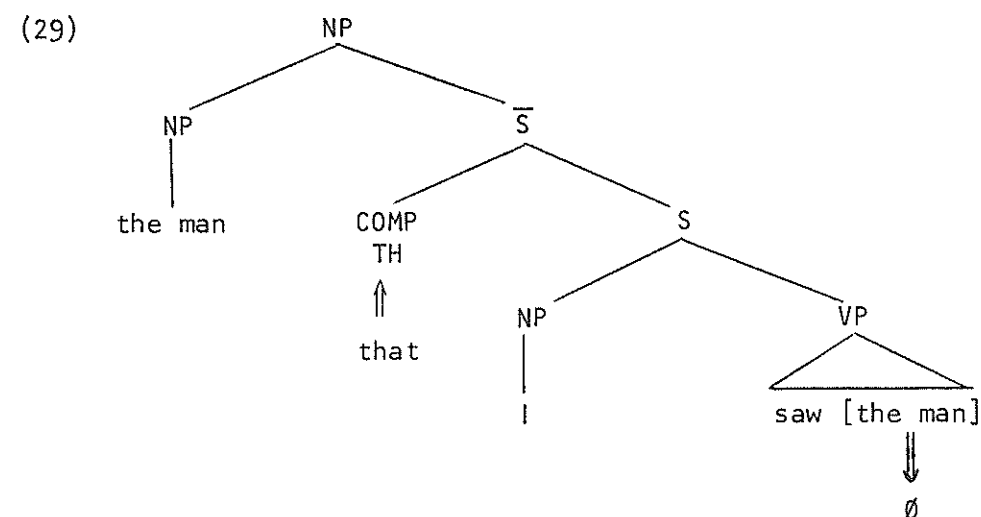
e. **What does John like the man (that) Bill gave to?*

Extraction out of relative clauses is impermissible whether the relative marker is *that* or WH, and additionally is impermissible whether or not the *that* occurs at the surface. If we assume an analysis of relative clause formation based on that given by Emonds (1976) then once WH-fronting has taken place for relative clause formation there is a WH-filled complementizer to block WH-fronting for questioning out of the relative clause. There are certain additional details still to be accounted for, however.

In relative clause formation the pre-relativized structure is an NP dominating an NP with an \bar{S} sister. In the expansion of \bar{S} one of the three possibilities for COMP must be chosen (see (21)). If the one chosen is COMP then relativization by WH-fronting is possible, provided, of course, there is a WH-phrase. This would be as in (28).



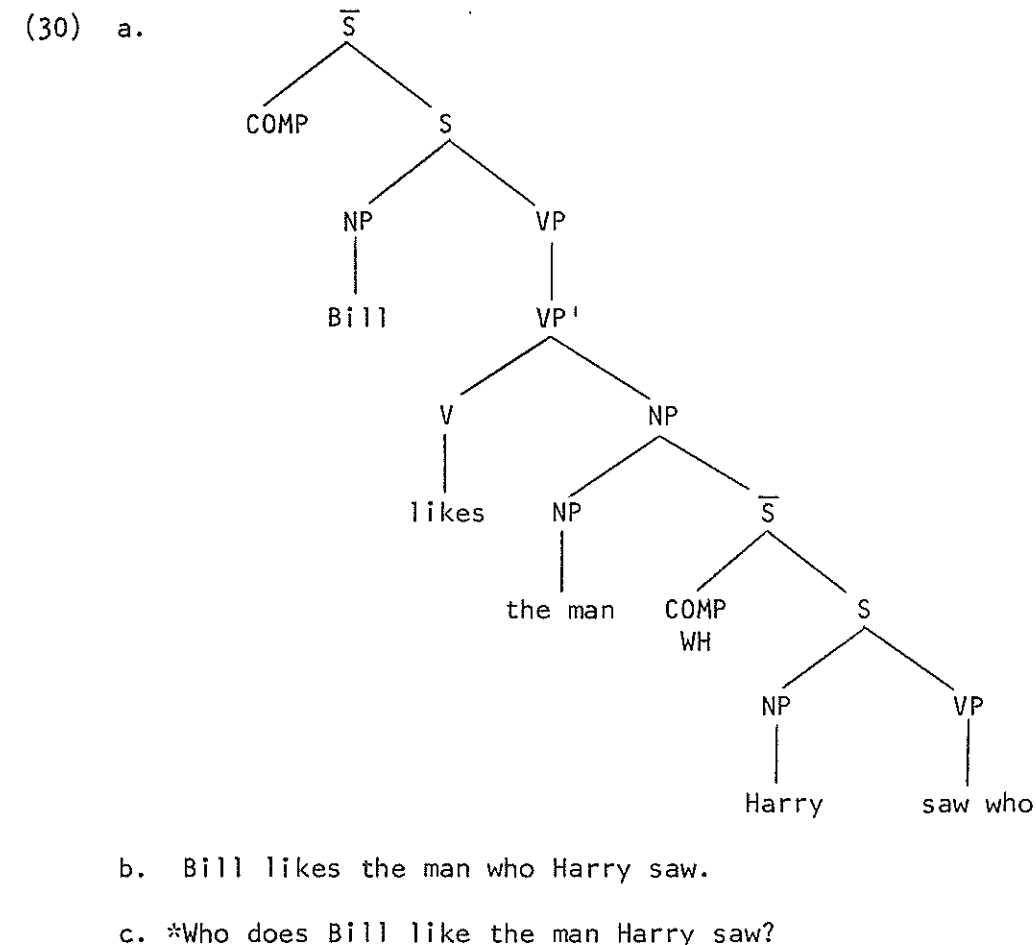
If the COMP chosen is COMP then relativization by WH-fronting is not possible. The alternative is relative deletion.



In this model the deletion process (or reduction process in dialects where a pronoun copy is left) also spells out the TH complementizer *that*.¹⁵ By either relativization process, relative clauses, thus far in the discussion, have non-empty complementizers. As such, extraction out of them by WH-fronting is blocked. This explains the ungrammaticality of (11a) and also (27c) and (27e). (27c) and (27e) will never be derived by WH-fronting so the functioning of rule (25a), with respect to them is irrelevant. To (27d) rule (25a) may apply to optionally delete the *that*. (25a) may not apply to (27a), as the ungrammaticality of (27b) shows, because its structural description is not met.

Before ending this discussion of (25a) it should be pointed out, to avoid confusion, that it only applies where *that* is preceded and followed by a full NP (in the sense of N with the maximal number of bars). In other words the *that* which occurs in other complex NP's will not be subject to deletion since the instance of N which precedes *that* in a sentence like *John ridiculed the claim that Bill saw Mary* is either to be considered N or N' (for further discussion of this see Emonds 1976:149). By contrast, the head of a relative clause is an N' = NP.

Let us return now to some further discussion of the interaction of WH-fronting for relative clause formation and question formation. Consider the structure in (30).



For (30) there are two ways in which the structural description can be met for WH-fronting. One is where term 1 is the embedded COMP and the result of applying the rule is, properly, the grammatical (30b).¹⁶ The ungrammatical (30c) can be accounted for, but it is not a case which is handled by the VIC. There would be no filled COMP or WH-term in the grossest analysis to block the movement represented by the string (30c). It would be marked as ill-formed, however, because there would be no way to remove the delta under the embedded COMP. There could be no further application of WH-fronting and the structural description of neither (23) nor (25b) is met. The only well-formed derivation is where WH-fronting applies in the lower domain first.¹⁷

In summary, let us look again briefly at examples (19c) and (20) (we will discuss further examples like (19b) below where the occurrence of more than one WH is considered). The claim being made here is that WH-fronting applies freely, conditioned only by the VIC. At the time that WH-fronting applies, for these sentences, there would have been a filled complementizer in the grossest analysis which would block the rule. The

that which occurs with verbs like *guip* is not inserted by rule (23), but, since it is strictly subcategorized by the verb, must be present in deep structure.

We will now turn to a class of verbs which seem to present counter-examples to the VIC. Since these verbs represent a certain semantic class the account I suggest will make use of that fact. Consider the examples of (31) which were pointed out by Erteschik (1973:101):

- (31) a. *He regretted you did it.
 b. What did he regret that you did?
 c. *It alarmed me she liked it.
 d. What did it alarm you that she liked?

For a verb like *regret*, since WH-fronting is permissible, in the VIC account it must be that it can occur in deep structure with a lexically empty complementizer. Sentence (31b) is the result of WH-fronting followed by rule (23), *that*-insertion. The ungrammaticality of (31a), and similarly (31c), must be because rule (25b) may not apply to remove the delta. This is confirmed by the ungrammaticality of **What did he regret you did?*. If *that*-insertion does not apply to replace the delta in these examples the result is ill-formed.

Erteschik presents the examples given here in (31) to show that the correlation which she has pointed out, namely that "deletion [of *that*] is possible in approximately the same contexts in which extraction is possible" (1973:101) is not precise. What Erteschik found was basically that factives do not allow *that*-deletion and that nonfactives do allow *that*-deletion (both with some exceptions). Erteschik concludes that

that-deletion is much more dependent on factivity than extraction. The exceptions are too few to base a comprehensive discussion on, but they do indicate that the same factors that determine extraction are operative here [for *that*-deletion], although to a much less extent. (p. 102)

In relation to factivity and extraction, there is also the case of subjunctives. Subjunctives allow extraction from complements and also require a *that*.

- (32) a. She demanded that he think of a clever sentence.
 b. *She demanded he think of a clever sentence.
 c. What did she demand that he think of?
 d. *What did she demand he think of?

As in the *regret* cases, verbs which require subjunctive *that*-complements must not allow (25b) to apply even where (23) has failed to. (32b) and (32d) are evidently surface strings which still contain unfilled deltas. It seems reasonable to assume, along with Erteschik, that the obligatory occurrence of *that* is conditioned in some way by factivity.¹⁸ It might be that the occurrence of *that* is semantically necessary to "signal" something about factivity (either that the clause is factive or that there is a switch to non-factive in subjunctive clauses). Where (23) does provide the necessary *that* the string is not considered ill-formed.

In summary again, the account of WH-fronting presented in this work allows the rule to apply wherever its structural description is met as long as no WH or lexically designated (= having undergone lexical insertion) COMP occurs in the grossest analysis to block it. The fact that only lexically designated items are relevant follows from the usual notion of terminal string. To avoid any possible misinterpretation, however, the VIC is amended slightly to make this fact obvious.

(33) Variable Interpretation Convention:

In the structural description of a transformation only terms which are crucially affected may be mentioned and for all non-optional terms mentioned, $A - B \equiv A - X - B$, where (a) X is a variable in the structural description and corresponds to the grossest constituent analysis of a phrase marker, and (b) the grossest constituent analysis does not contain any lexically designated A or B or head of an A or B.

4.3. For-To Clauses

So far in this chapter we have considered the WH-term complementizers and *that*. We will now look at extraction out of *for-to* clauses. Consider examples (34) and (35).

- (34) It is annoying to me for children to watch television.
 (35) a. It is easy for children to watch television.
 b. It is easy for parents for children to watch television.

In our discussion of (34) and (35) we will be led naturally in two directions. We will be led to compare extraction by WH-fronting with the extraction of NP by the raising process (discussed in detail in Chapter 3) and we will also be led to compare the VIC account to be proposed below with that provided for by Bresnan's Complementizer Constraint on Variables (Bresnan 1976d).

First it should be remembered that in the VIC account of raising presented in Chapter 3 the structural description of the rule ($\Delta - \text{v}_{\text{max}}[\text{NP}]$)

is met only where adjectives like *easy* have VP complements. Where *easy* appears with an S-complement, as in (35b), it is the result of extraposition and the output of extraposition can never serve as the input to raising. Neither (34) nor (35b) allows extraction out of the *for-to* clause by raising. To an example like (35a), prior to the insertion of *it*, raising readily applies to produce *Television is easy for children to watch*.

This brief reconsideration of raising is presented here because for raising, by the VIC, the variable environment which blocks the application of the rule is an NP in the grossest analysis. For example:

(36) a. It is easy for you to see (that) Tom needs $\left(\left\{\begin{smallmatrix} \text{some} \\ \text{the} \end{smallmatrix}\right\}\right)$ help.

b. $\ast\left(\left\{\begin{smallmatrix} \text{some} \\ \text{the} \end{smallmatrix}\right\}\right)$ help is easy for you to see (that) Tom needs.

(36b) is blocked because the NP *Tom* lies in the grossest analysis. In the statement of WH-fronting there is no NP mentioned. The prediction, by the VIC, then would be that an NP in the grossest analysis would have no effect. And this is in fact correct, as (37) shows.

(37) What is it easy for you to see (that) Tom needs?

What does affect the functioning of WH-fronting is a complementizer in the grossest analysis. To example (35a), just as raising can apply, so can WH-fronting.

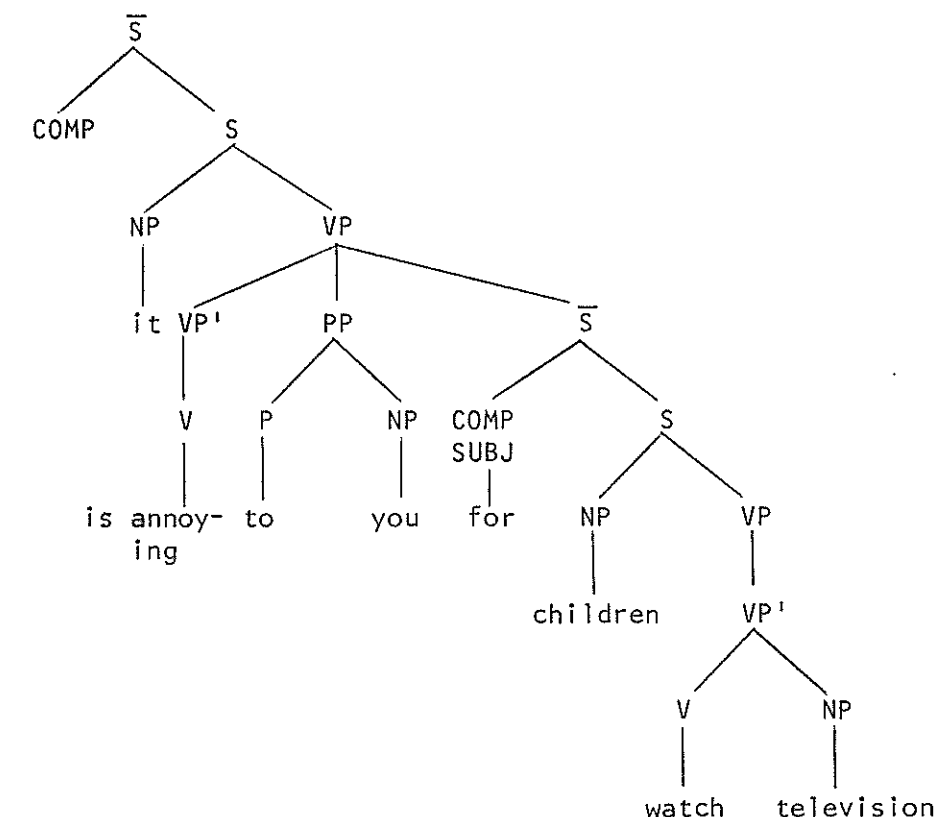
(38) What is it easy for children to watch?

These two processes can apply to (35a) because the *for* + NP (*for children*) is a prepositional phrase of the matrix clause and not a complementizer and subject of an embedded clause.

In both (34) and (35b) there is an S-complement which is the result of extraposition (see discussion of this rule in Chapter 6). In each of these embedded clauses there is an NP subject and a lexically designated deep structure complementizer *for*. The question of extraction by raising does not arise since the structural description is not met. What we must consider is extraction by WH-fronting, and in particular the effect of the complementizer *for*.

The phrase marker underlying a sentence like (34) (postponing for the moment the issue of the rule of *for*-phrase formation (Emonds 1976:195-200)) is given in (39a).

(39) a.



b. Who(m) is it annoying to for children to watch television?

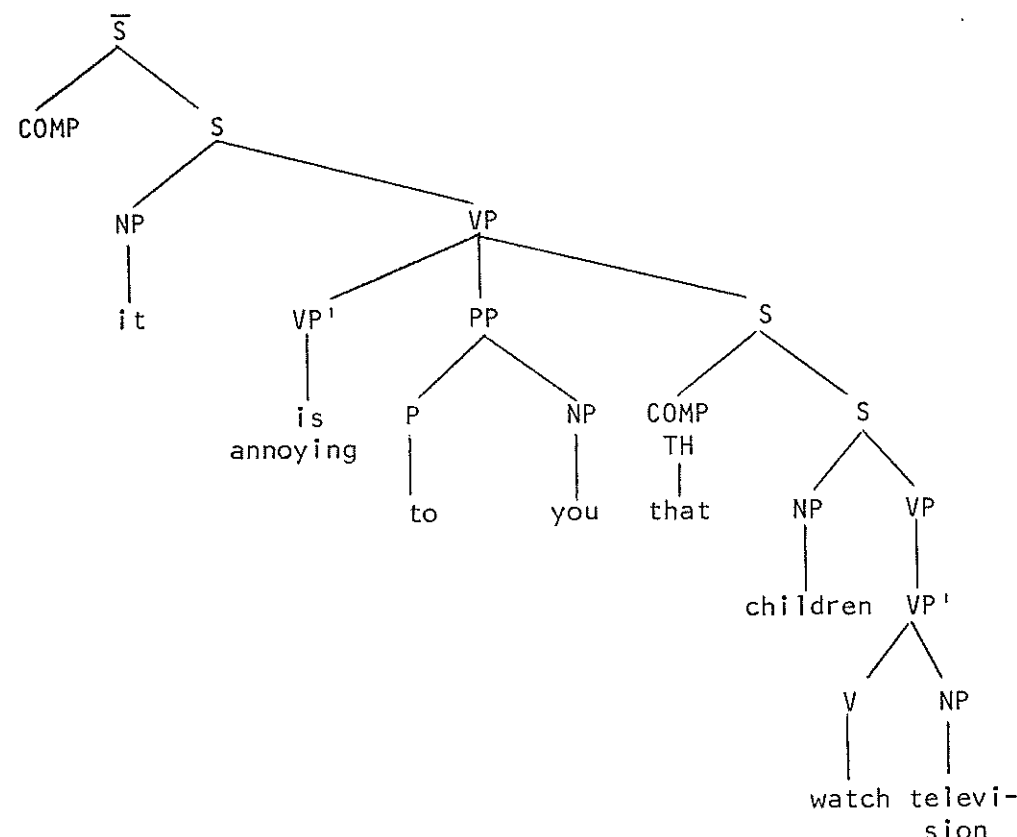
c. To who(m) is it annoying for children to watch television?

d. \ast Who is it annoying to you for to watch television?

e. What is it annoying to you for children to watch?

The difference in grammaticality between (39d) and (39e) will be explained shortly. Before that is taken up, consider also the examples of (40).

(40) a.



b. Who(m) is it annoying to that children watch television?

c. To who(m) is it annoying that children watch television?

d. *Who is it annoying to you that watch television?

e. ?What is it annoying to you that children watch?

We will be returning to a discussion of (40e) in particular, but first consider additionally Bresnan's Complementizer Constraint on Variables (Bresnan 1976d:34):

(41) The Complementizer Constraint on Variables

For any proper analysis (... , X, A, Y, ...) such that X and Y are variable factors and A is a constant factor to be deleted, if X = ---COMP, then --- must be empty (of terminals).

Since Bresnan here follows Ross (1967) in characterizing question movement as a "chopping" rule, and since "chopping" involves both copying and

deleting, this constraint in (41) is applicable not only for deletions *per se*, but also for what we have been considering extractions. What this constraint means then, is that a rule may apply to A either, (a) where X contains COMP and nothing else, or (b) where COMP is not the rightmost term of X. In other words, subject NP's of an S may be moved where X is the left end variable COMP of S, but otherwise they may not be moved.

By (41) the ungrammaticality of (40d) is properly predicted as shown in (42).

(42) It is annoying to you that children watch television
 --- COMP
 X A Y

The ungrammaticality of (39d) is also correctly predicted:

(43) It is annoying to you for children to watch television
 --- COMP
 X A Y

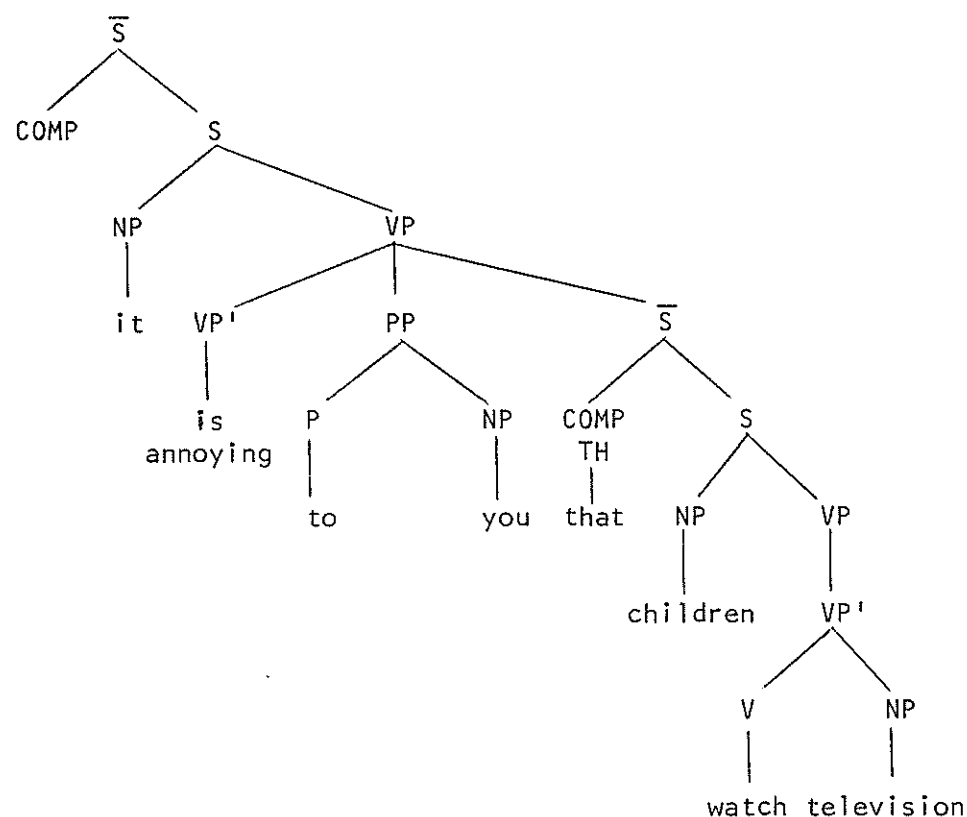
The Complementizer Constraint on Variables would predict that both (40e) and (39e) are well-formed. This is because for neither one would COMP be the rightmost constituent of X.

(44) a. It is annoying to you that children watch television.
 --- COMP
 X A Y
 b. It is annoying to you for children to watch television.
 --- COMP
 X A Y

To my intuitions, example (40e) is ungrammatical or at least worse than (39e). If this is the case, then (41) makes the wrong prediction. Example (39e) is clearly well-formed, as is predicted by Bresnan's constraint.

Let us return now to the account which is provided by the VIC. To phrase marker (40a) (here repeated for ease of reference)

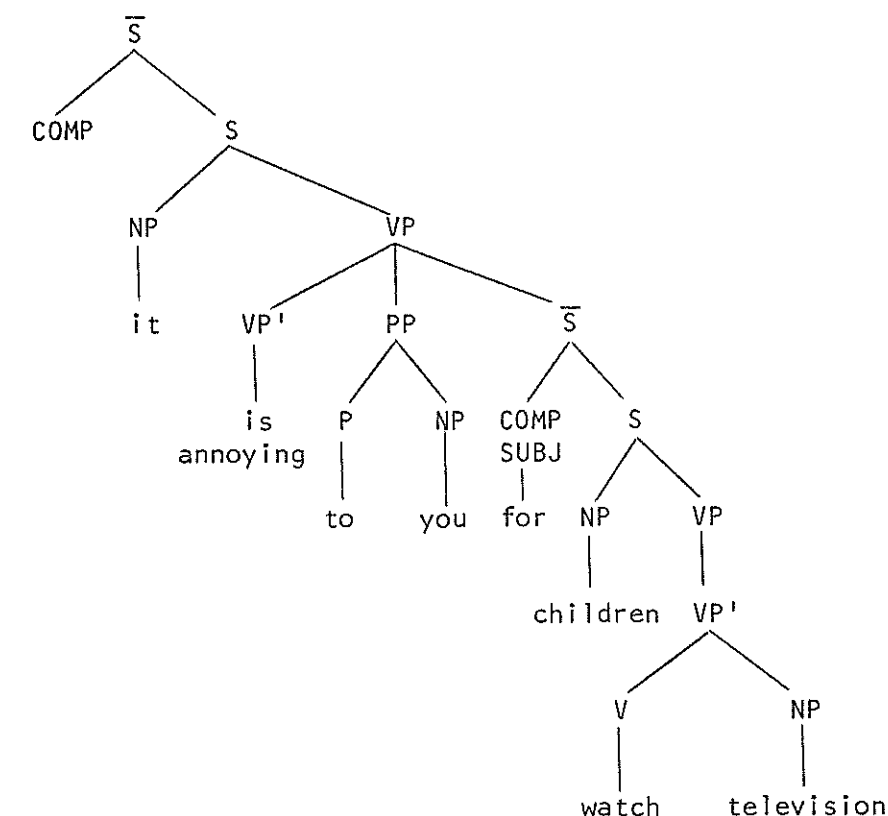
(40) a.



WH-fronting is prevented from applying to either NP of the embedded S because the COMP *that* lies in the grossest analysis. The ungrammaticality of both (40d) and (40e) is predicted. What is not predicted is the fact that (40d) is certainly worse than (40e).¹⁹

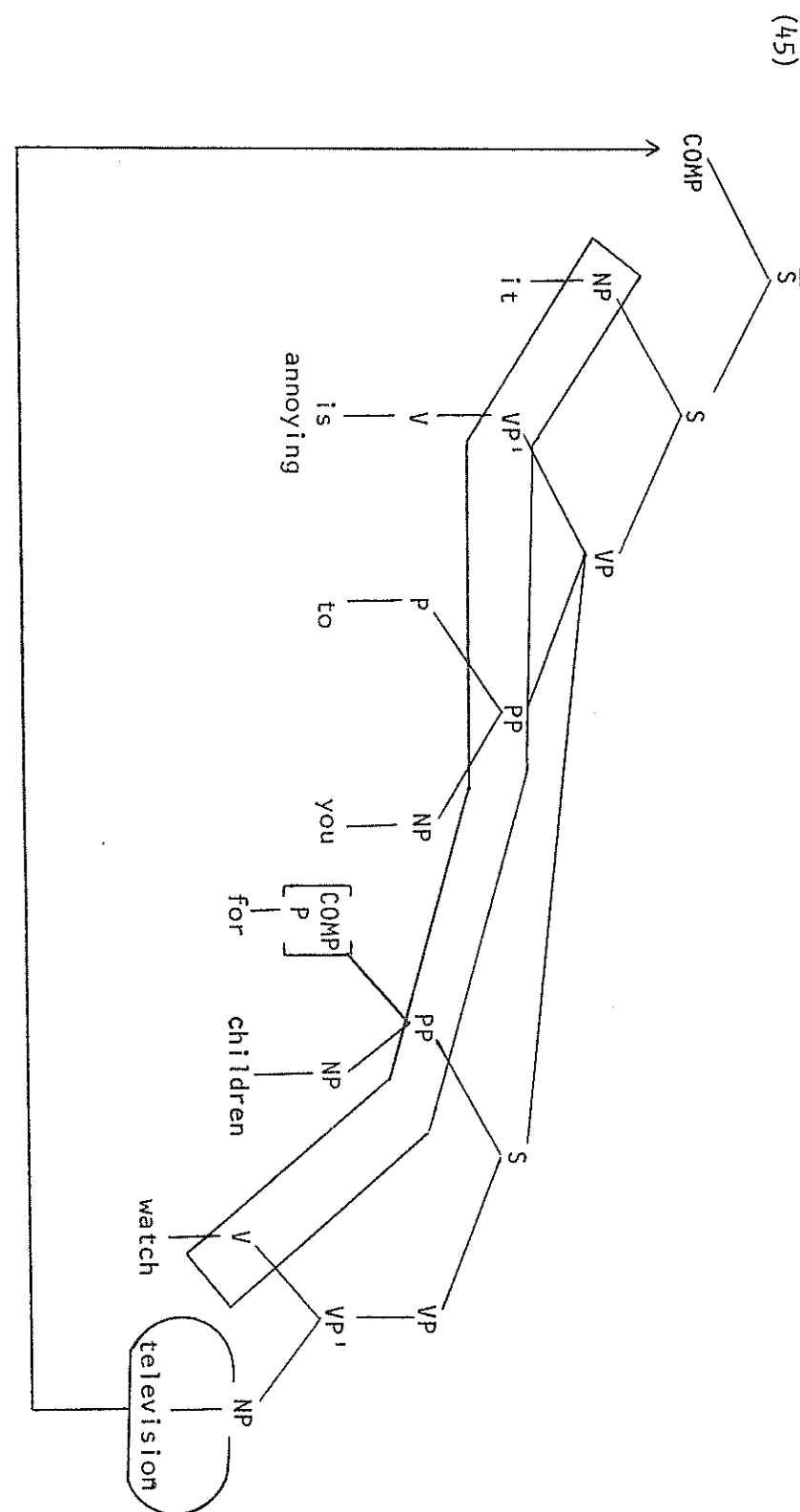
Consider again phrase marker (39a):

(39) a.



For this phrase marker the VIC would predict that any extraction by WH-fronting out of the embedded S should be blocked because the COMP *for* would be in the grossest analysis. As was just pointed out, however, sentence (39e), where the embedded object is questioned (*What is it annoying to you for children to watch?*) is grammatical. This is accounted for in the VIC framework as follows.

It has been argued, by Emonds for instance, that the *for* and the subject NP of a *for-to* clause form a constituent. Emonds argues that there is a rule of *for*-phrase formation which applies in such cases (Emonds 1976:195-200). Alternatively, it might be the case that the *for*+NP constituent is a base configuration (for some discussion of this alternative see Schachter 1976). In either case, at the time WH-fronting takes place the structure underlying the examples of (39) would be as given in (45).



In (45) the structure of the node labelled PP in the embedded S might be considered controversial. If, however, we take seriously the claim (following Chomsky 1965) that all category labels are actually feature bundles then the node immediately dominating *for* would be more properly represented as something like

$$\begin{bmatrix} +\text{COMP} \\ +\text{PREP} \\ \vdots \end{bmatrix}$$

For the process of WH-fronting and the VIC it is the COMPness of the node which is relevant. The NP sister to the *for* may not be fronted because the $\begin{bmatrix} \text{COMP} \\ P \end{bmatrix}$ lies in the grossest analysis, hence **Who is it annoying to you for to watch television?* (= (39d)). As is indicated in (45), however, the embedded object may be questioned, *What is it annoying to you for children to watch?* (= (39e)).

Additionally, with respect to (45), not only is there a prohibition against the movement of the NP immediately preceded by the COMP, but there is also a prohibition against the movement of the whole *for* + NP phrase

(**For whom is it annoying to you to watch television?*²⁰). As was mentioned above, (section 4.1, example (9b)) nothing is ever extracted from COMP position. We will return to this below in the section on multiple occurrences of WH.

That the ungrammaticality of examples like (39d) necessarily involves the fact that the *for* is a COMP can be easily seen by the grammaticality of the examples of (46), particularly (46c).

- (46) a. It would be interesting for us to attend seances.
 b. What would it be interesting for us to attend?
 c. Who would it be interesting for to attend seances?
 d. For whom would it be interesting to attend seances?

The *for*-phrases here are PP's of the matrix, as has been repeatedly pointed out. The ungrammaticality of (39d) cannot be accounted for simply by saying that there is a prohibition against the sequence *for-to*. Rather, the disallowed sequence is COMP[*for*] - *to*. The complementizer *for*, like the complementizer *that* or WH-complementizers, when it occurs in the grossest constituent analysis, will block the application of WH-fronting.

In this section we have seen how, with respect to embedded *for-to* clauses, the correct distinction can be made between where raising applies and where WH-fronting applies (cf. (36) and (37)). For WH-fronting, which mentions no NP in its structural description, NP's in the grossest analysis are irrelevant. Additionally we have seen where the Complementizer Constraint on Variables and the VIC make different predictions. Where the predictions are the same for *for-to* clauses it seems that Bresnan's constraint is a subcase of the VIC.

For *that*-clauses where the *that* is optional (see the discussion above in 4.2), Bresnan's model and the one proposed in this thesis can account for the same facts (notice that the difference in predictions above was where the *that* was obligatory in surface structure and necessarily, therefore, lexically designated in deep structure). The following examples are from Bresnan (1976d:29, 30), where '___' indicates the extraction site.

- (47) a. Jack claimed (that) one of his cats had eaten one of his birds.
 b. *Which one of his cats did Jack claim that ___ had eaten one of his birds?
 c. Which one of his cats did Jack claim ___ had eaten one of his birds?
 d. Which one of his birds did Jack claim (that) one of his cats had eaten ___?

For Bresnan the (b) example here is explained by the fact that WH-fronting may not apply because X of (41) is ---COMP and --- is not empty of terminals. In the VIC framework (b) is accounted for by the fact that *that*-insertion (rule (23) of 4.2) cannot apply where there is no NP environment. *That* can be inserted only where, as in (d), there would be an immediately following NP. For both Bresnan's model and the VIC framework (where \emptyset indicates an extracted element) the sequence COMP - \emptyset - VP will never be well-formed, unless the extraction has moved the term to the adjacent COMP (as in *Who came?*).

4.4. Extraction from NP

4.4.1 The Revised WH-fronting Rule

We will now consider the question of extraction from gerunds and other nominal constructions and we will be led to an important reformulation of the WH-fronting rule. First of all, for gerundive nominals, we notice that there is a difference in extractability depending on whether or not there is a possessive NP occurring with the gerund.

- (48) a. Which movie did you anticipate seeing?

- b. *Which movie did you anticipate Harry's seeing?

The grammaticality difference between (48a) and (48b) might be accounted for by the Specified Subject Condition (Chomsky 1973) where the possessive NP is considered the subject of the larger NP that it occurs in. WH-fronting cannot take place because elements cannot move over specified subjects except through a COMP node, which an NP does not have. There is, however, an alternative account which is consistent with the VIC.

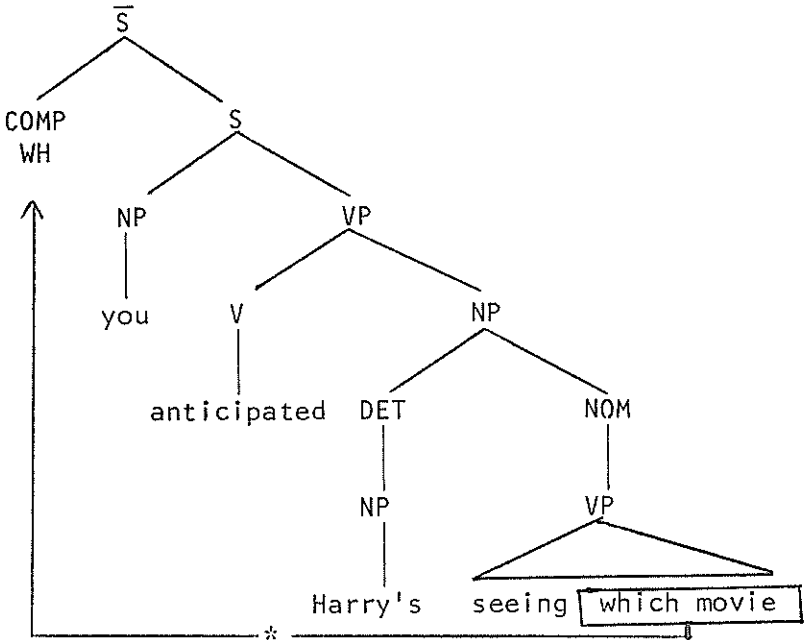
As things stand so far, use of the Variable Interpretation Convention will not account for the difference between sentences (a) and (b) of (48). The reason for this, however, is that the rule of WH-fronting given in (5) is not quite accurate yet. Since it is arguably the case that this process operates on the syntactic element WH and it is also the case that, other than in COMP, WH occurs only in the specifiers of NP or AP, it is possible to write this transformation as a movement of a WH determiner²¹ (for discussion of this treatment of WH-fronting see Emonds 1976:180-95). This alternative is shown in (49).

$$(49) \text{ COMP} - (\text{P}) + \begin{bmatrix} \text{DET} \\ +\text{WH} \end{bmatrix} = 2 - \emptyset$$

4.4.2. Gerundive Nominals

Notice that with a reformulation of the structural description of the WH-fronting rule, the effect of the VIC will be different. Now it will be not only a COMP or a WH in the grossest analysis which will block the application of this rule, but so will a determiner. As will be clear from the examples below, it is not just a WH-determiner which is relevant for the VIC but rather any DET in the grossest analysis. This then is the basis for the grammaticality distinction between (48a) and (48b). I am assuming an analysis of gerundive nominals in which the possessive NP is the determiner of the nominal (for discussion, see Schachter 1976). The VIC, along with the WH-fronting rule as given in (49), would allow extraction from gerunds as long as no DET occurs in the grossest analysis. Consider phrase marker (50).²²

(50)



Extraction will always be blocked where a possessive full NP occurs in the gerund.²³

In his article on gerundive nominals, Schachter (1976) considers DET's other than possessive NP's occurring with gerunds and says that if the analysis is correct other DET's should occur with them since the PS rule for DET would be:

$$(51) \text{ DET} \rightarrow \begin{Bmatrix} \text{ART} \\ \text{QUA} \\ \text{DEM} \\ \text{NP} \\ \vdots \\ \vdots \\ \vdots \end{Bmatrix} \quad (\text{where ART} = \text{article, QUA} = \text{quantifier, DEM} = \text{demonstrative})$$

As evidence that gerunds can occur with determiners other than NP's, Schachter gives two literary examples (here (52a, b)) and says that English speakers can easily coin others (here (53a, b, c)), although perhaps all should be considered marginal.

- (52) a. There is *no* enjoying this world without thee.
 b. I do not like *this* leaving without a word all those to whom she is dear.

- (53) a. I won't tolerate *any more* telling tales out of school.
 b. There's been *too much* telling tales out of school around here lately.
 c. *This* telling tales out of school has got to stop.

If it were the case that (51) in fact were the correct phrase structure expansion for DET, then the VIC would predict that extraction out of gerunds with any of the above-mentioned DET's should be equally ungrammatical. I think, however, that it is not accurate to consider all quantifiers as cases of DET (for discussion of this issue see e.g. Selkirk 1976). A simple argument that they are not all determiners is that some can co-occur with them, as for instance in

- (54) a. The $\begin{Bmatrix} \text{many} \\ \text{seven} \end{Bmatrix}$ books remained on the table for weeks.
 b. The TV commercial really stressed the many extras that came with the bargain.
 c. The whole analysis was based on these too few examples.

This would mean that the category DET would not include all quantifiers. We will return to a discussion of determiners shortly. For now notice that if it is the case that some quantifiers are not determiners, then the VIC would predict that WH-fronting would be blocked whenever an article, a demonstrative, or a possessive NP occurred in the grossest analysis, but not when certain quantifiers did. This seems to be accurate not only for gerunds but for extraction from other nominals as well. First, to continue discussion of gerunds, the examples of (55) are given in place of Schachter's coined examples. They provide a better environment for testing extraction since they don't involve an idiom.

- (55) a. I won't tolerate any more reciting silly limericks in the parlor.
 b. There's been too much reciting silly limericks in the parlor lately.
 c. This reciting silly limericks in the parlor has got to stop.
 d. I can't stand this reciting silly limericks in the parlor any longer.
 (56) a. ?What kind of limericks won't you tolerate any more reciting in the parlor?

- (56) b. ?Which limericks has there been too much reciting in the parlor lately?
- c. *Which limericks has this reciting in the parlor got to stop?
- d. *Which limericks can't you stand this reciting in the parlor any longer?

Whatever the status of Schachter's coined examples (53) or that of the examples of (55) (that is, whether they are fully grammatical or semi-grammatical) there is certainly a difference exhibited in (56). (56c) and (56d) are clearly worse than the others. I have included (56d) which does not parallel an example of Schachter's to have a gerund with a demonstrative in other than subject position since it is always the case that extraction from subjects is disallowed. Even example (56d), where the gerund is not the subject, seems considerably worse than (56a) and (56b). If we now try extraction from the gerunds in Schachter's literary examples we will find similar results. I have tried to simplify sentence (52b) to make WH-fronting seem more natural for (57b) through (57e).

- (57) a. Which world is there no enjoying without thee?
- b. *Without what don't you like this leaving?
- c. *What don't you like this leaving without?
- d. *To whom don't you like this leaving without a word?
- e. *Who don't you like this leaving without a word to?

It certainly seems to be the case that (57a) is more acceptable than the others here. Where there is a quantifier (*any more, too much, no*) in the grossest analysis WH-fronting is permitted. Where what occurs in the grossest analysis is a demonstrative or a possessive NP then WH-fronting is blocked.

We can consider now gerunds with nonpossessive NP's in initial position. For this type of gerund Schachter considers a transformational analysis where the NP is derived from the subject of a sentence and is not a determiner. If this NP in fact is not a determiner, then the VIC analysis would predict that there should be a difference in extraction possibilities where the initial NP is not possessive. This seems to be correct.

- (58) a. Which book did you object to John buying?
- b. *Which book did you object to John's buying?

- (58) c. Which book did you take advantage of John buying?
- d. *Which book did you take advantage of John's buying?

This distinction that we see in gerundive nominals, with respect to whether or not they have determiners, is paralleled in other nominal constructions. For instance, with picture nouns there are the following grammaticality facts:

- (59) a. Who did you see pictures of?
- b. *Who did you see John's pictures of?
- c. Who did you see many pictures of?
- d. *Who did you see $\begin{Bmatrix} \text{those} \\ \text{the} \end{Bmatrix}$ pictures of?
- e. Who did you see a picture of?

4.4.3. Presuppositional Anaphoricity

What seems to be an accurate statement about extraction from nominals is that it is blocked where the specifier of the NP contains a definite determiner, a possessive NP, or a demonstrative. Extraction is allowed over an empty specifier, an indefinite article, and at least some quantifiers. Perhaps the generalization is that no DEFINITE determiner can occur in the grossest analysis if extraction by WH-fronting is to be allowed. There is, in fact, a certain amount of support for this generalization.

Pope (1972) argues that definite NP's, factive S's, and generic NP's (which will not be discussed here) all share a common property, namely, DEFINITENESS. For Pope this property is represented by a TH marker which is generated in the determiner of NP's and in the complementizer of S's. Pope incorporates an argument made by Kuroda (1969) that where an NP is anaphoric it is also definite. She then makes the following statement (p. 12):

... to use a definite NP, one must presuppose the existence of the individual or individuals named; to use a generic NP, one must presuppose the existence of the class named (without necessarily presupposing that the class has members); and to use a factive S, one must presuppose the truth of the proposition advanced.

I think it is possible to reduce these three generalizations to one. That is, I think truth is for propositions what existence is for objects (individuals and classes are all objects, however else they may differ).

On page 13 she adds:

Thus, if we class definites, generics, and factives together as TH categories, existence and truth together as existence, and individuals, classes and situations together as objects, the three generalizations reduce to the following: to use a TH category, one must presuppose the existence of the object referred to.

On page 14 she revises this to:

All I would want to say, then, is (1) categories which have the property of definiteness must be anaphoric in order to be appropriate; (2) if something is anaphoric, it is much more likely (probabilistically) that it exists than that it does not, (3) the relationship between truth and S's is the same as that between existence and NP's.

It is important to remember, in the discussion of features which follows, that the notion of Pope's which I will be using, ANAPHORIC, is one which is INHERENT in certain determiners but not in others. I do not mean to indicate, for instance, that an indefinite NP may not have an anaphor in the real world. This notion of "inherent presupposition" is insightfully discussed by Jackendoff (1972:276-278). He distinguishes between "focal presupposition" and "inherent presupposition". With respect to factive verbs he says (p. 278):

... it seems likely that inherent presuppositions can be formalized like selectional restrictions: just as a verb can presuppose that its object is human or animate or mass, it can presuppose that its complement is true.

And with respect to definite NP's Jackendoff (1972:277) says that another type of inherent presupposition

is that induced by the definite article. Definite noun phrases presuppose that they describe an entity uniquely identifiable within the bounds of the discourse.

Continuing on now, if Pope is right, as I think she basically is in this respect, then it is not surprising that definite determiners and the complementizer *that* in complex NP's function in the same way, to restrict the application of WH-fronting, when they occur in the grossest analysis.

Generally NP's of the form $NP[the\ fact\ that\ S]$ have definite determiners and also obligatory *that* complementizers. Consider the examples of (60) and then those of (61). (The examples of (61) were pointed out to me by Joe Emonds.)

- (60) a. Bill reported the fact that Sara won the game.
 b. *Bill reported the fact Sara won the game.
 c. *Which game did Bill report the fact (that) Sara won?
- (61) a. There exists a possibility (that) Sam might lose the game.
 b. Which game does there exist a possibility (that) Sam might lose?
 c. *Which game does there exist the possibility (that) Sam might lose?

Where the S-complement of NP must be factive, or in Pope's terms, where the complement must be anaphoric (see footnote 5 of Chapter 1 of her dissertation) then the complementizer *that* is obligatory and no extraction may take place. What is crucial here is that in these cases there is a definite determiner. Where the determiner is not definite, as in (61b), there is no presupposition that the complement is anaphoric, it may or may not be, and extraction is permitted--out of a complex NP.

Keeping the notions of definiteness and anaphoricity in mind we now consider the examples of (62). (These sentences were suggested to me by Joel Rotenberg.)

- (62) a. Sometimes you find the odd counterexample to a theory.
 b. Which theory do you sometimes find the odd counterexample to?
 c. He found that counterexample to the theory.
 d. *Which theory did he find that counterexample to?

In (62), even though there is a determiner which in form is definite, extraction is permitted, (62b). What is interesting about (62a) is that it contains a definite NP which is in the relevant sense "non-definite"--it does not presuppose that the NP is anaphoric. In (62c) where the object NP is clearly definite and anaphoric no extraction is possible, hence, (62d).

If we now make use (without going into a detailed account of the semantics of WH-phrases) of the not-implausible claim that all WH-phrases are presuppositionally anaphoric there is a coherent account of all the WH-fronting data thus far presented. Where the WH-fronting rule (49) (presented here again for ease of reference)

$$(49) \text{ COMP} - (\text{P}) + \begin{bmatrix} \text{DET} \\ +\text{WH} \end{bmatrix} = 2 - \emptyset$$

is properly interpreted as the movement of a determiner, it can readily be interpreted not just as the movement of WH but as A MOVEMENT OF THE WH AND ALL THE FEATURES WHICH CO-OCCUR ON THAT DETERMINER WITH THE WH. Pre-suppositional anaphoricity can well be considered a redundant feature predictable from the occurrence of +WH. Then the variable material for (49) could be allowed to contain only determiners which were distinct in their feature specifications from $\begin{bmatrix} \text{DET} \\ +\text{WH} \end{bmatrix}$.

Just as the [+anaphoric] (again in the sense in which Pope uses the term) can be predictable based on the occurrence of the feature +WH, so can a positive specification for the feature [definite] be predicted from the occurrence of [+anaphoric]. If all cases of possessive NP determiners and demonstratives, and most cases of the definite determiner are considered [+anaphoric] then they will be redundantly [+definite]. Where the specification for the feature [anaphoric] is [-], that is, where there is no necessary presupposition of the existence of the NP, then usually the determiner is indefinite. The feature [-definite] is not predictable from [-anaphoric], however, because of examples like (62a).

In summary now, rule (49), the rule for WH-fronting, functions freely unless it is prevented from doing so by the VIC. It will only be blocked by the VIC where either a filled COMP or a filled determiner which is non-distinct from $\begin{bmatrix} \text{DET} \\ +\text{WH} \end{bmatrix}$ occurs in the grossest constituent analysis.

4.5. Multiple WH's

4.5.1. The Variable Interpretation Convention: Final Version

We next consider sentences with multiple occurrences of WH.

- (63) a. What crimes does the FBI know how to solve?
 b. *What crimes does the FBI know whether to solve?
 c. *What books does John know to whom to give?
 d. *To whom does John know what books to give?

- (63) e. John knows what books to give to whom.
 f. John knows to whom to give what books.
 g. *John knows what who saw.
 h. John knows who saw what.
 i. *Who does John know saw what?
 j. *What does John know who saw?

(Examples a-h are from Chomsky 1973:244-45.)

In order to account for these sentences with multiple WH's we will be led to a very important reformulation of the VIC; a reformulation motivated, in part, by the apparent need for some version of the Superiority Condition (Chomsky 1973).

Further, in accounting for the facts in examples like (63), we will be using the already-alluded-to necessity of banning extraction from COMP. Not only must examples like (9b) (here repeated)

- (9) b. *Whether did I wonder she will dance?

be prevented, but so much "prepositional dangling" in COMP position. Consider, for instance,

- (64) a. I wonder whom she will dance with.
 b. I wonder with whom she will dance.

- c. *Whom (do) I wonder _{COMP} [with] she will dance?

This generalization about no extraction from COMP seems accurate, at least at the descriptive level, since the only case I know of where an analysis includes a process extracting items from COMP position is the successive cyclic treatment of WH-fronting (Chomsky 1973, 1976). Since this framework presents a viable alternative to the claim that movement of WH is subject to Subjacency (Chomsky 1973) and hence a successive cyclic rule, there seems to be no case where extraction from COMP takes place. This generalization will not only be useful in accounting for the examples of (63) but it has already been of use in allowing a very simple statement of the WH-fronting rule (see section 4.1) and, as has just been shown, accounts for the fact that prepositions don't get stranded in COMP.

The condition which must be incorporated into the VIC, the reformulation alluded to above, is one which accounts for the fact that in a sentence like (63h) the first possible application of WH-fronting moves the WH-subject into COMP position. To assure that this movement of an embedded WH-subject into an embedded COMP position is the first application of WH-fronting to take place (that this is necessary will be clear below) it is merely necessary to require rules to apply first where the variable is \emptyset . In other words, rules apply first to adjacent terms if they can and then only subsequently across variable material. This amendment of the VIC provides for an interesting formal account of a suggestion made informally to me by Morris Halle, based on recent work by Jean-Roger Vergnaud. The suggestion was that all rules (phonological as well as syntactic) can be said to affect only adjacent terms--as long as there exists the relevant definition of "adjacent".

In this VIC model, where rules are already written without including any variables, it is intuitively natural that a structural description of the form A - B is taken, wherever possible, to mean that the A and B are right next to one another. We are now in a position to make the strong statement, as a postulate, that:

(65) Rules affect only adjacent terms.²⁴

"Adjacent terms" in (65) is then construed as "strictly" adjacent (i.e., where NO variable material intervenes between A and B) wherever this is possible. If an operation can in no way apply in accord with the requirement of STRICT ADJACENCY, then adjacent terms in (65) is construed to mean "weakly" adjacent. Terms A and B of a structural description are WEAKLY ADJACENT where no A or B lies in the grossest analysis between them.

The VIC is now properly reformulated as (66):

(66) Variable Interpretation Convention:

Given A - B, where A and B are crucially affected terms of the structural description of a transformational operation T, first T functions where A and B are strictly adjacent, then T may function where A and B are weakly adjacent. A and B are weakly adjacent where, for all non-optional A and B, $A - B \equiv A - X - B$ where X corresponds to the grossest constituent analysis of a phrase marker and X does not contain any lexically designated A or B or head of an A or B.

4.5.2. An Alternative to Successive Cyclicity

Now, with this final formulation of the VIC, we can see how the examples of (63) are handled. Sentence (h) is well-formed because the only movement

- (63) e. John knows what books to give to whom.
 f. John knows to whom to give what books.
 g. *John knows what who saw.
 h. John knows who saw what.
 i. *Who does John know saw what?
 j. *What does John know who saw?

(Examples a-h are from Chomsky 1973:244-45.)

In order to account for these sentences with multiple WH's we will be led to a very important reformulation of the VIC; a reformulation motivated, in part, by the apparent need for some version of the Superiority Condition (Chomsky 1973).

Further, in accounting for the facts in examples like (63), we will be using the already-alluded-to necessity of banning extraction from COMP. Not only must examples like (9b) (here repeated)

- (9) b. *Whether did I wonder she will dance?

be prevented, but so much "prepositional dangling" in COMP position. Consider, for instance,

- (64) a. I wonder whom she will dance with.
 b. I wonder with whom she will dance.

- c. *Whom (do) I wonder COMP [with] she will dance?

This generalization about no extraction from COMP seems accurate, at least at the descriptive level, since the only case I know of where an analysis includes a process extracting items from COMP position is the successive cyclic treatment of WH-fronting (Chomsky 1973, 1976). Since this framework presents a viable alternative to the claim that movement of WH is subject to Subjacency (Chomsky 1973) and hence a successive cyclic rule, there seems to be no case where extraction from COMP takes place. This generalization will not only be useful in accounting for the examples of (63) but it has already been of use in allowing a very simple statement of the WH-fronting rule (see section 4.1) and, as has just been shown, accounts for the fact that prepositions don't get stranded in COMP.

The condition which must be incorporated into the VIC, the reformulation alluded to above, is one which accounts for the fact that in a sentence like (63h) the first possible application of WH-fronting moves the WH-subject into COMP position. To assure that this movement of an embedded WH-subject into an embedded COMP position is the first application of WH-fronting to take place (that this is necessary will be clear below) it is merely necessary to require rules to apply first where the variable is \emptyset . In other words, rules apply first to adjacent terms if they can and then only subsequently across variable material. This amendment of the VIC provides for an interesting formal account of a suggestion made informally to me by Morris Halle, based on recent work by Jean-Roger Vergnaud. The suggestion was that all rules (phonological as well as syntactic) can be said to affect only adjacent terms--as long as there exists the relevant definition of "adjacent".

In this VIC model, where rules are already written without including any variables, it is intuitively natural that a structural description of the form A - B is taken, wherever possible, to mean that the A and B are right next to one another. We are now in a position to make the strong statement, as a postulate, that:

(65) Rules affect only adjacent terms.²⁴

"Adjacent terms" in (65) is then construed as "strictly" adjacent (i.e., where NO variable material intervenes between A and B) wherever this is possible. If an operation can in no way apply in accord with the requirement of STRICT ADJACENCY, then adjacent terms in (65) is construed to mean "weakly" adjacent. Terms A and B of a structural description are WEAKLY ADJACENT where no A or B lies in the grossest analysis between them.

The VIC is now properly reformulated as (66):

(66) Variable Interpretation Convention:

Given A - B, where A and B are crucially affected terms of the structural description of a transformational operation T, first T functions where A and B are strictly adjacent, then T may function where A and B are weakly adjacent. A and B are weakly adjacent where, for all non-optional A and B, $A - B \equiv A - X - B$ where X corresponds to the grossest constituent analysis of a phrase marker and X does not contain any lexically designated A or B or head of an A or B.

4.5.2. An Alternative to Successive Cyclicity

Now, with this final formulation of the VIC, we can see how the examples of (63) are handled. Sentence (h) is well-formed because the only movement

- (63) e. John knows what books to give to whom.
 f. John knows to whom to give what books.
 g. *John knows what who saw.
 h. John knows who saw what.
 i. *Who does John know saw what?
 j. *What does John know who saw?

(Examples a-h are from Chomsky 1973:244-45.)

In order to account for these sentences with multiple WH's we will be led to a very important reformulation of the VIC; a reformulation motivated, in part, by the apparent need for some version of the Superiority Condition (Chomsky 1973).

Further, in accounting for the facts in examples like (63), we will be using the already-alluded-to necessity of banning extraction from COMP. Not only must examples like (9b) (here repeated)

- (9) b. *Whether did I wonder she will dance?

be prevented, but so much "prepositional dangling" in COMP position. Consider, for instance,

- (64) a. I wonder whom she will dance with.
 b. I wonder with whom she will dance.

- c. *Whom (do) I wonder COMP [with] she will dance?

This generalization about no extraction from COMP seems accurate, at least at the descriptive level, since the only case I know of where an analysis includes a process extracting items from COMP position is the successive cyclic treatment of WH-fronting (Chomsky 1973, 1976). Since this framework presents a viable alternative to the claim that movement of WH is subject to Subjacency (Chomsky 1973) and hence a successive cyclic rule, there seems to be no case where extraction from COMP takes place. This generalization will not only be useful in accounting for the examples of (63) but it has already been of use in allowing a very simple statement of the WH-fronting rule (see section 4.1) and, as has just been shown, accounts for the fact that prepositions don't get stranded in COMP.

The condition which must be incorporated into the VIC, the reformulation alluded to above, is one which accounts for the fact that in a sentence like (63h) the first possible application of WH-fronting moves the WH-subject into COMP position. To assure that this movement of an embedded WH-subject into an embedded COMP position is the first application of WH-fronting to take place (that this is necessary will be clear below) it is merely necessary to require rules to apply first where the variable is \emptyset . In other words, rules apply first to adjacent terms if they can and then only subsequently across variable material. This amendment of the VIC provides for an interesting formal account of a suggestion made informally to me by Morris Halle, based on recent work by Jean-Roger Vergnaud. The suggestion was that all rules (phonological as well as syntactic) can be said to affect only adjacent terms--as long as there exists the relevant definition of "adjacent".

In this VIC model, where rules are already written without including any variables, it is intuitively natural that a structural description of the form A - B is taken, wherever possible, to mean that the A and B are right next to one another. We are now in a position to make the strong statement, as a postulate, that:

(65) Rules affect only adjacent terms.²⁴

"Adjacent terms" in (65) is then construed as "strictly" adjacent (i.e., where NO variable material intervenes between A and B) wherever this is possible. If an operation can in no way apply in accord with the requirement of STRICT ADJACENCY, then adjacent terms in (65) is construed to mean "weakly" adjacent. Terms A and B of a structural description are WEAKLY ADJACENT where no A or B lies in the grossest analysis between them.

The VIC is now properly reformulated as (66):

(66) Variable Interpretation Convention:

Given A - B, where A and B are crucially affected terms of the structural description of a transformational operation T, first T functions where A and B are strictly adjacent, then T may function where A and B are weakly adjacent. A and B are weakly adjacent where, for all non-optional A and B, $A - B \equiv A - X - B$ where X corresponds to the grossest constituent analysis of a phrase marker and X does not contain any lexically designated A or B or head of an A or B.

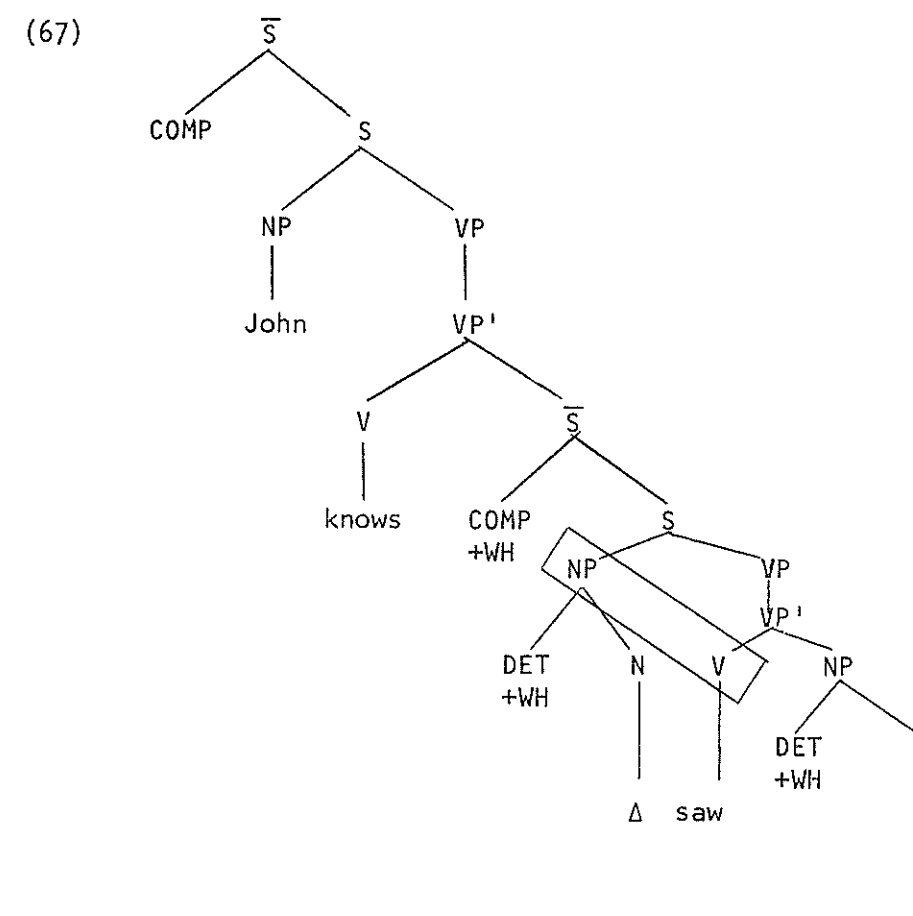
4.5.2. An Alternative to Successive Cyclicity

Now, with this final formulation of the VIC, we can see how the examples of (63) are handled. Sentence (h) is well-formed because the only movement

which has taken place is where the embedded subject has moved into embedded COMP position--an application of WH-fronting where COMP and WH were strictly adjacent. Sentence (e) is well-formed because the movement which took place was from embedded object to embedded COMP position. Since there was no way for the rule to apply to strictly adjacent terms it functioned properly where the COMP and the WH were weakly adjacent.

Example (63b) is excluded just as are other examples where the COMP *whether* occurs in the grossest analysis. Examples (c), (d), and (j) are excluded by the VIC because in each case a filled COMP occurs in the grossest analysis.

It is to exclude strings like (g) and (j) that it is important that in examples like (h) the first WH-term is in COMP position. If the *who* in (h) were still in subject position the phrase marker underlying it would be that in (67).

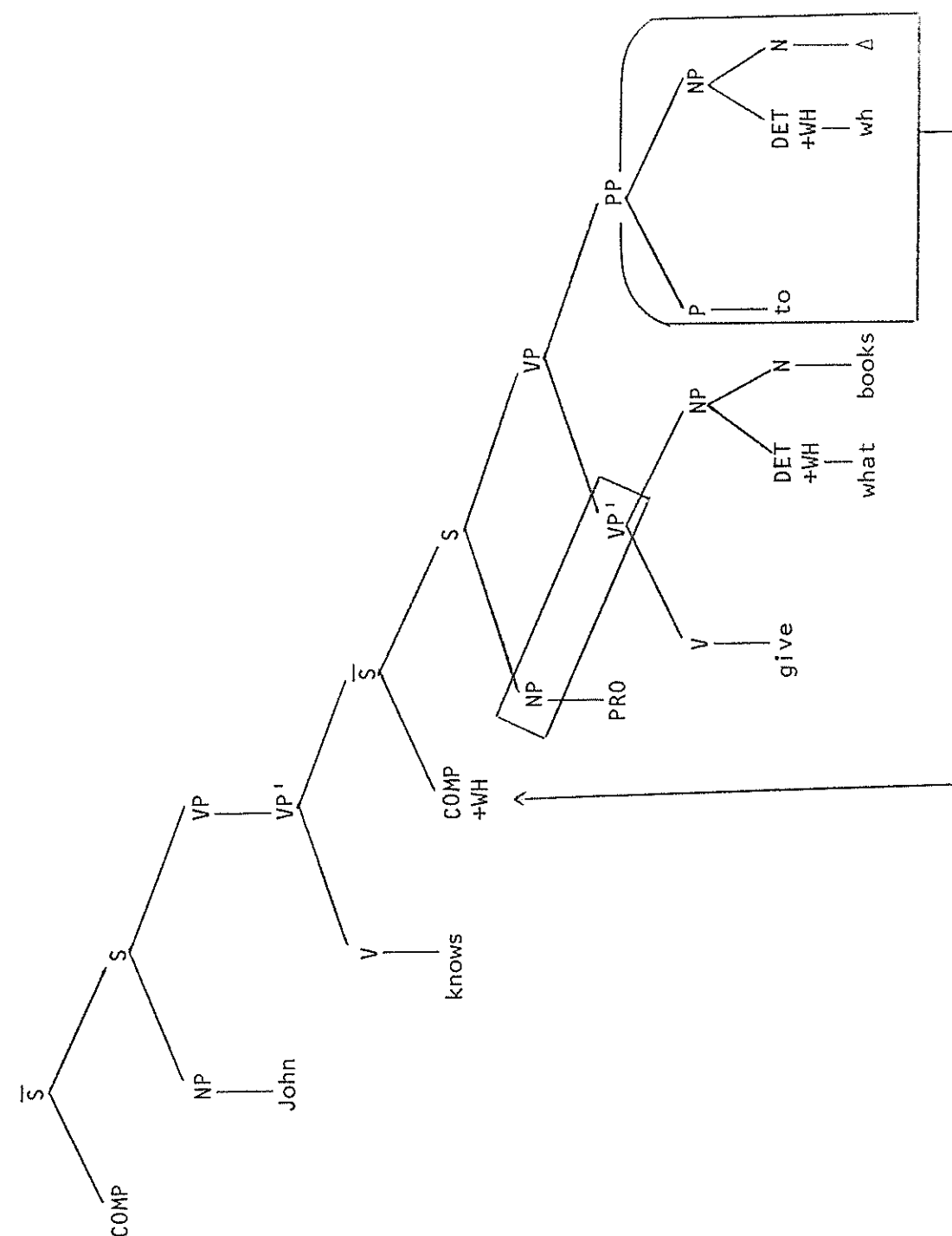


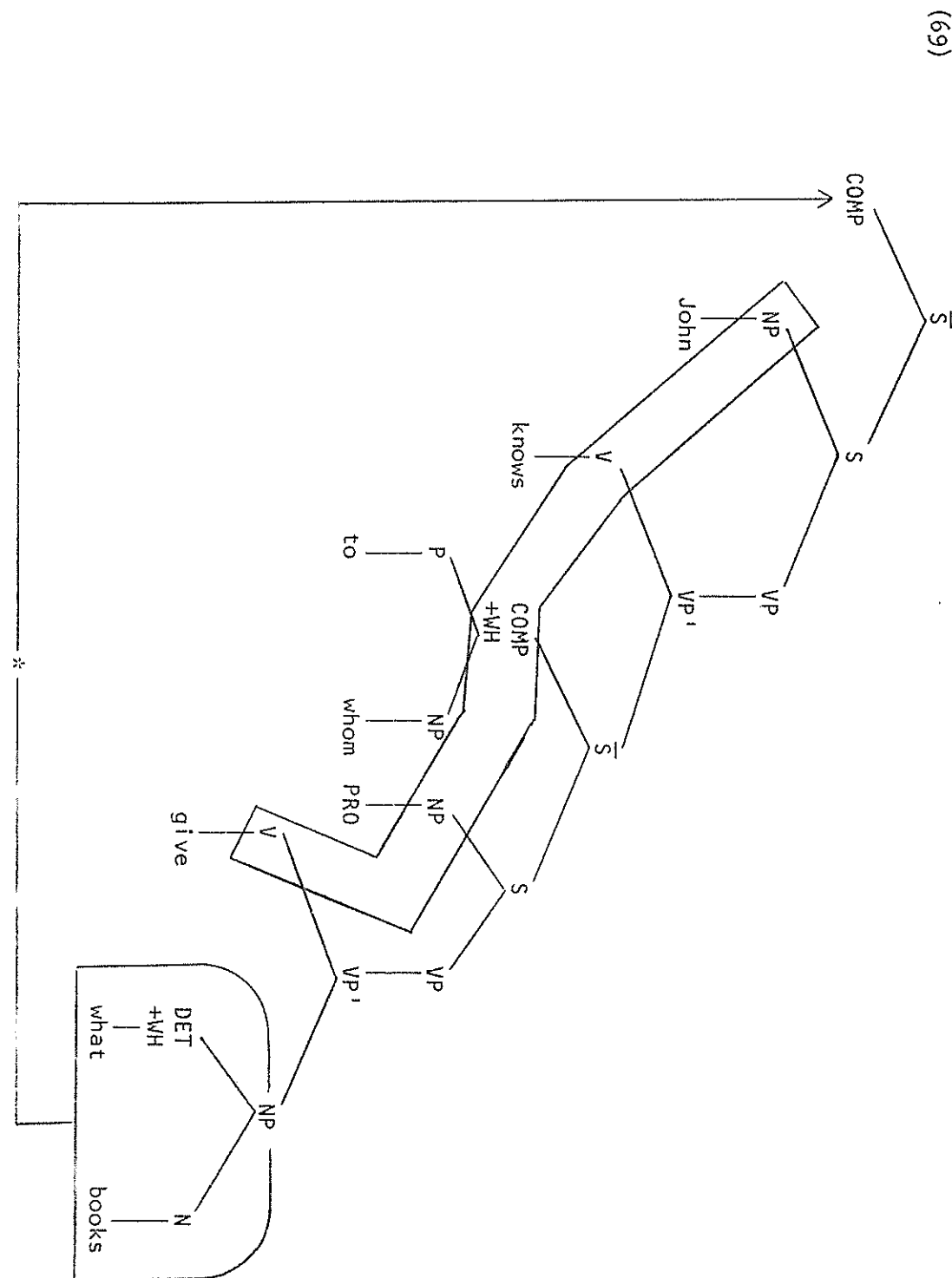
This account of WH-fronting is still following that of Emonds (1976) where WH in the base is generated in the specifier of NP or AP. Where there

is no lexical material as sister to $\begin{bmatrix} \text{DET} \\ +\text{WH} \end{bmatrix}$ in NP there are appropriate spelling rules for *who*, *what*, etc. What is important to notice in (67) is that there would be nothing to block the movement of the object NP over the subject NP. No COMP and no DET occurs in the grossest analysis. To block sentences (63g) and (63j) it must be that WH-fronting first moves the WH-subject, as in (63h), in local fashion, into COMP position. Example (g) is then blocked because the embedded COMP cannot be doubly filled, and example (j) is blocked because the filled embedded COMP occurs in the grossest analysis.

Consider now phrase marker (68) for the well-formed example (63f), as compared to phrase marker (69) for the ungrammatical example (63c).

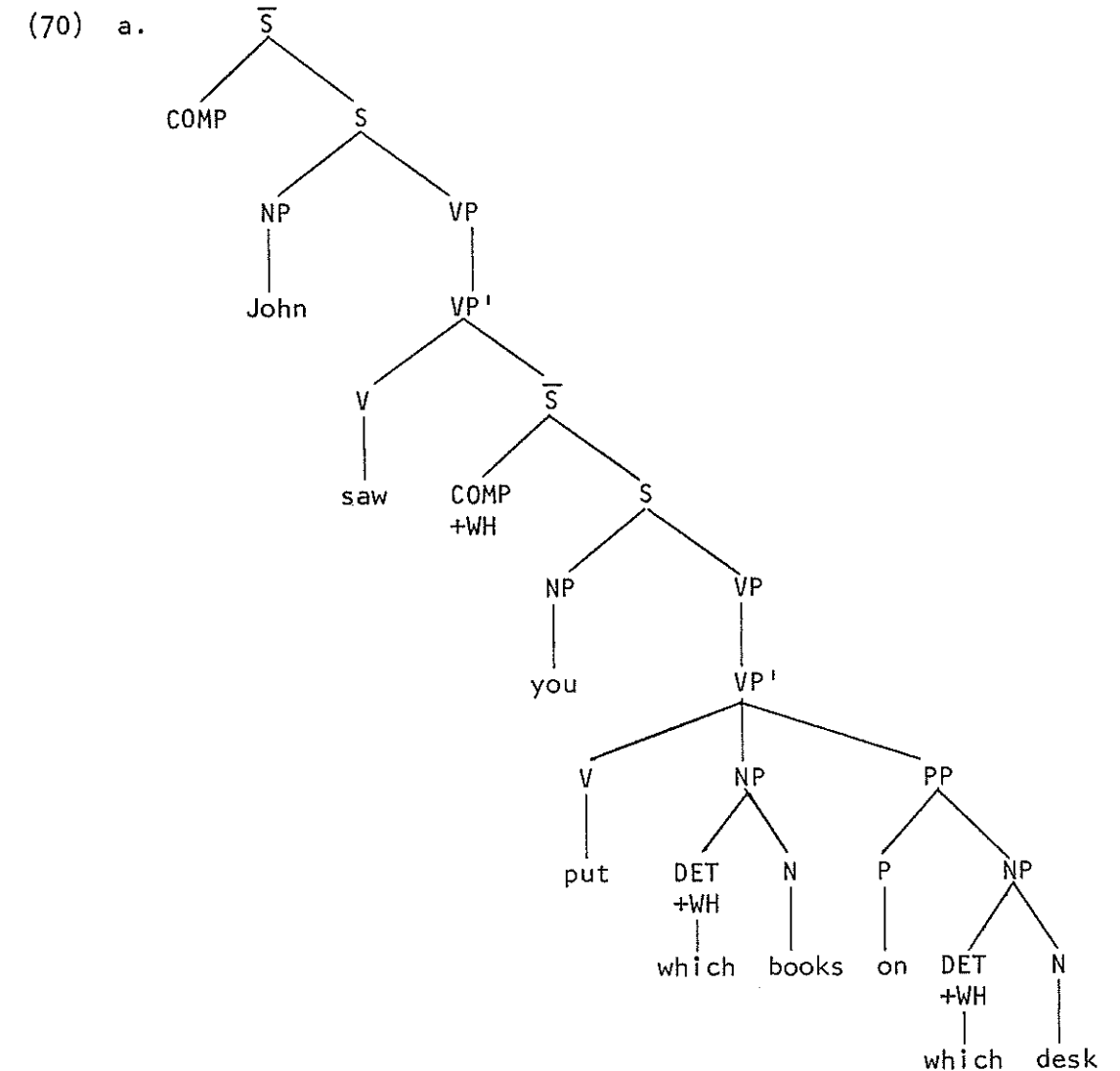
(68)





In (69), but not in (68), a filled COMP occurs in the grossest analysis and blocks the application of WH-fronting. In (68) the movement can take place because the WH is not in the grossest analysis, the VP' which dominates it is.

Even in a case where both WH-terms must be considered to be inside the VP' the VIC makes the right predictions for WH-fronting.



From (70a) can be derived all of the following sentences:

- (70) b. Which desk did John see you put which books on?
 c. Which books did John see you put on which desk?

(70) d. John saw which books you put on which desk.

e. John saw which desk you put which books on.

The WH-determiners in (70a) never occur in the grossest analysis so they never block the application of WH-fronting.

Returning now to (63), we account for the grammatical sentence (a) as follows. It has been claimed, by Chomsky and probably others, that *know how* is a single lexical item, say subcategorized for a following infinitive. Since *how* therefore is not in COMP position there is nothing to stop extraction from the complement of *know how*.

The prohibition against extraction from COMP position would prevent example (i), where once the *who* was in the embedded COMP position it could not move out. This prohibition would also prevent the extraction of the first WH-terms in (e) or (f) as well as any example where the WH-term *whether* of indirect questions is fronted (**Whether did John wonder they could come?*).

In Chomsky's recent discussions of conditions, where WH-fronting must be a successive cyclic rule because it is subject to the Subjacency Condition, there is no way to prevent the application of WH-movement once there is a WH in COMP and a possible COMP for that WH to move into. Instead, WH-moved forms are overgenerated, formed freely, and then unacceptable forms are excluded or given no semantic interpretation at the surface. The analysis presented here is an alternative to the overgeneration and later filtering out of forms at the surface, and as such "lightens the load" on the semantic component. In other words, thus far, no descriptive semantic mechanisms are needed to ensure well-formedness.

FOOTNOTES TO CHAPTER 4

¹The structural change of this rule is possibly more accurately written as $\begin{bmatrix} 1 \\ 2 \end{bmatrix} - \emptyset$, since the COMP node may not be actually replaced but instead filled by the moved constituent.

²A suggestion similar to Emonds' condition was previously made by Chomsky (1973:235, fn. 10) where he mentions that if the specifier of an NP or an AP is extracted then the whole phrase must be extracted.

³This Revised Left Branch Condition will be shown to be very important again in Chapter 5 where, along with the VIC, it can lead to the elimination of the A-over-A condition.

⁴That the Revised Left Branch Condition must be stated so as not to apply to S, or \bar{S} , seems to be some evidence that S is not the maximal projection of V. Clearly when the left branch of S (namely NP) reorders, it does not reorder the whole S. As (9c) shows, the reordering of the left branch of \bar{S} cannot be construed so as to move the \bar{S} . If it can be established that $S \neq V^n$, where n = the maximal number of bars (see Chapter 6 for further discussion of this), then probably the Revised Left Branch Condition can be written in more general form to refer to any X^n where X is any head-of-a-phrase category (N, A, V, or P). For relevant discussion see Emonds (1976:Chapter 1) and Jackendoff (1976).

⁵Chomsky (1964) points out essentially this same fact where he says that either relative clause formation or question formation may apply in a given domain, but not both. He also says that neither of these rules may apply twice in a particular domain.

⁶The fact that for many speakers (12a) is better than (12b) will be discussed below in section 4.3 where the VIC is compared with Bresnan's Complementizer Constraint on Variables (Bresnan 1976d).

⁷That the sentences of (12) are equally ungrammatical if *whether* is replaced by *if* will be accounted for, like other examples below, where any complementizer in the grossest analysis blocks WH-fronting.

⁸This sentence was pointed out to me by Joel Rotenberg.

⁹We are ignoring for now the issue of the coreferentiality between the subjects of the matrix and embedded clauses. We return to this in Chapter 5.

¹⁰There are evidently speakers for whom sentence (16e) is acceptable. The prediction here would be that for these same speakers the declarative counterpart of this sentence would be acceptable without the *that*: *John*

ate so much he could hardly watch the movie. For further discussion of this issue see footnote 14 below.

¹¹I have been unable to obtain a copy of this paper.

¹²A node is empty only where it dominates NO features whatsoever. It is not empty just because it has no phonological features. See principle (19) of Chapter 3.

¹³Lexical insertion, like all other transformations, is assumed to be optional.

¹⁴Since the analysis presented here depends very heavily on a correlation between extractability and the occurrence of a complementizer, the extent to which the correlation is accurate deserves careful investigation. Erteschik (1973) categorizes verbs of saying which take *that*-complements according to whether or not the extraction from the complement is allowed. Her categorization is as follows:

Acceptable:	say, tell, report, announce
Questionable:	grunt, holler, murmur, mumble, mutter, roar, scream, shout, sigh, snort, stammer, wait, whine, tell, exclaim
Bad:	purr, snarl, editorialize, eulogize, coo, jeer, rumble, simper, lisp, quip, croak, dictate, transcribe, ululate, animadvert. (p. 84)

The VIC treatment of extraction constraints would make the prediction that the acceptable verbs would occur grammatically in declarative sentences without the *that*-complementizer. It would say that in declarative sentences the verbs here classified as bad would obligatorily require *that*. About the questionable category of verbs of saying, all that would be predicted would be that speakers who allow the question forms would also allow the declarative forms without the *that*. I think these predictions are borne out.

¹⁵Rather than an identity-deletion process this might instead be some type of indexing of a PRO form, or some other type of identity "checking" process, to assure the semantic well-formedness of the relative clause.

Schachter (1973) has shown that there are certain conditions on the grammaticality of relative clauses that this type of account, thus far, does not handle. For instance, there are cases where the antecedent (the head) is selected by the verb of the relative clause, as in (i), suggested to me by Schachter (personal communication), and (ii), from Schachter (1973:32).

(i) a. *I am pleased by the headway.

b. I am pleased by the headway that you are {making }
*cooking}.

(ii) a. Lip service was paid to civil liberties at the trial.

b. *I was offended by (the) lip service.

c. I was offended by the lip service that was paid to civil liberties at the trial.

A process which just assures the identity of the relevant NP's will not account for the facts of (i) and (ii). Nor would it be able to account for the facts of (iii) and (iv) (also from Schachter 1973:32-33) where reflexivization and the use of *each other* must be "backwards" and may not be "forwards".

(iii) a. The portrait of himself that John painted is extremely flattering.

b. *The portrait of John₁ that himself/he₁ painted is extremely flattering.

(iv) a. The interest in each other that John and Mary showed was fleeting.

b. *The interest in John and Mary that each other showed was fleeting.

A detailed account of relativization should, of course, be able to account for all of these facts. Acknowledging the inadequacy of a simple relativization by deletion or PRO indexing process, what is important here is that however relative clauses are formed they include either a WH- or TH-complementizer.

¹⁶This, of course, is following the usual assumption that unfilled COMP's in main clauses are deleted.

¹⁷In a structure like (30a), if the embedded COMP was TH and the NP was a WH-term, then no derivation could be well-formed because neither relativization process could apply to form the relative clause.

¹⁸Actually, the conditioning environment for the occurrence of *that* for these verbs is probably more complicated. It seems that there is also some relevant phonological conditioning. The following sentence was pointed out to me by Joel Rotenberg: *I regret you've done it*. This sentence seems,

if not fully acceptable, at least better than its counterpart with no contraction. Another relevant example which seems even better is: *I really regretted you'd done that.*

¹⁹Notice that the VIC would also predict the ungrammaticality of both

- (i) *Who do you wonder whether watch television?
(ii) ?What do you wonder whether children watch?

It is not clear to me whether (ii) should be considered ungrammatical or not. If it is well-formed then clearly the Complementizer Constraint on Variables makes the right prediction but the VIC makes the wrong one.

²⁰This was pointed out to me by Peter Culicover.

²¹The notion determiner here must be somewhat refined so as to include the prehead position in adjective phrases, where the WH is generated in the base, and of course to include the prehead DET position of NP's.

²²The internal structure of the object NP in (50) is as it would be in Schachter's account of gerunds. Since it is only the determiner status of the possessive NP which is relevant here I am adopting this structure without further discussion.

²³For some reason, which I am as yet unable to explain, possessive pronouns in these constructions behave differently from possessive full NP's. Over some, but not all, possessive pronouns extraction is perfectly acceptable. Compare:

- (i) Who would you approve of my seeing? (from Chomsky 1964:46, fn. 10)
(ii) Who would you approve of her seeing?
(iii) ?Who would you approve of his seeing?
(iv) ?Who would you approve of $\begin{Bmatrix} \text{our} \\ \text{their} \end{Bmatrix}$ seeing?
(v) ?Who would she approve of your seeing?
(vi) *Who would she approve of Mary's seeing?

Emonds (personal communication) suggests that since (i) through (v) are all better than (vi) perhaps these possessive personal pronouns are adjectives, as they are, say, in French.

²⁴There is an unpublished manuscript by Jean-Roger Vergnaud (1976) which I have as yet been unable to obtain, where, I am told, this type of generalization is formalized for phonological rules such as vowel harmony.

CHAPTER 5

PASSIVE

5.1. Pseudo-passives

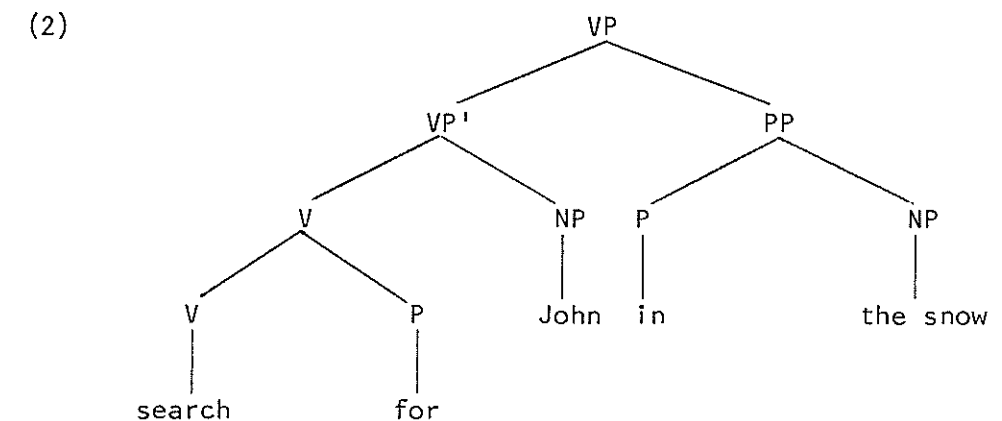
In Chapter 3 we began looking at the process of passivization to start considering what can occur as variable material in a structural description. That discussion concerned a very simple statement of the passive rule and preceded the definition of grossest constituent analysis and the statement of the Variable Interpretation Convention. Here we return to a discussion of passivization in more detail.

Before restating the actual transformation for deriving passives we will look at a new set of examples.

- (1) a. The Mounties searched for John in the snow.
b. John was searched for in the snow by the Mounties.
c. *The snow was searched for John in by the Mounties.

In many cases, where there is no direct object in a sentence, the object of a preposition can become the passive subject, as in (1b). It is only the object of the first PP which may become the subject, however, as (1c) illustrates. With the definition of grossest constituent analysis there is no obvious way, thus far, for the VIC to account for this fact. The grossest analysis between the verb *search* and NP *the snow* in (1a) would contain the PP dominating *for John* and the P dominating *in*. Since it would contain no NP or V (the nonvariable terms of the passive rule) the VIC would predict that (1c) should be grammatical.

One possibility which suggests itself here is that actually the verb is a complex one made up of a V + P sequence and the first NP is the direct object of the complex verb. The structure of the VP for (1a) then would be:



There are three reasons for not adopting this approach. First of all the process illustrated by (1b) is a very general one. This would mean that there were many of these complex verbs and in each case there would be a "simple" verb, identical to part of the complex one, which would require a separate lexical entry. This would seem to be missing a generalization. Secondly, there is evidence that the P NP (here *for John*) is by constituency tests a PP. Processes which front constituents function perfectly well analyzing *for John* in (1a) as a constituent.¹

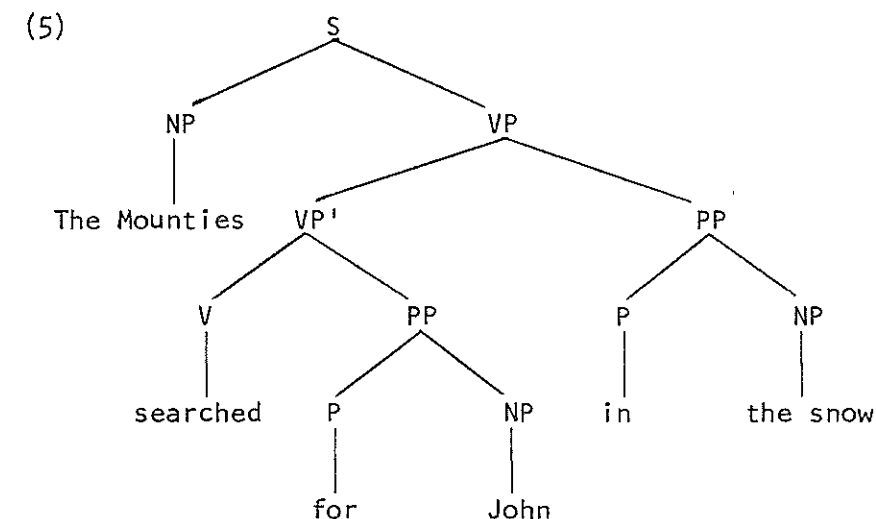
- (3) a. It was for John that the Mounties searched.
 b. For whom did the Mounties search?
 c. For John the Mounties searched (they refused to search for Tom).

The third reason for not adopting a complex-verb analysis of *search for* is one which is internal to the theoretical model presented in this thesis. In stating the passive rule, here repeated for ease of reference,

- (4) NP - V - NP = 3 - be+2+en - Ø - (by+1)
 +F

the motivation for including the nonvariable term V in the structural description was that it is a crucially affected term since it exhibits a change in morphology. Notice then that in a structure like (2) there must be a way to assure that the rule analyzes the lower V as term 2 of the input. This must be so because of the obvious fact that it is just *search* and not *search for* which undergoes the morphological change. To say that a term of a rule must be so construed as to select the LOWER of two possible analyses would seem to go against the quite general observation that originally led to the A-over-A condition. In general the higher analysis is the appropriate one. It might just be, one might argue, that this case is different. Other things being equal, it might be that the higher analysis is chosen, but that here either the structural description must be modified or there must be a special convention to assure that the lower verb is chosen as term 2 of the rule. In this framework this alternative would be particularly undesirable since it is very important that with respect to variable material the highest analysis--the grossest constituent analysis--always be the proper analysis. A special convention with respect to a nonvariable term would seem to be not only a complication, but in this case a complication which intuitively is not in keeping with the basic spirit of the VIC.

What I present as an alternative is as follows. The underlying phrase marker for sentence (1a) would be as given in (5).



The lexical item *search* occurs only once in the lexicon. It has a closely associated PP which occurs with the V in VP'. The PP *for John* in (5) is available for any process which affects PP's, as shown in (3). This V and the following P are also available for a process which can restructure² or reanalyze³ a sequence of terminal elements. This process would create the complex verb *search for*, under the condition that NO material occur

between the V and the P (e.g. **They* _v [*searched for*] *John* *the money*). That this process would have to be strictly local would be accounted for by the definition of restructuring. Restructuring processes would be very highly constrained as compared to transformations.⁴ The fact that the passive morphology is added to just the V *search* is easily accounted for by a revision of the passive rule to:⁵

- (6) NP - V - NP = 3 - be+en+2 - Ø - (by+1)
 +F

The result of the restructuring would be a phrase marker like that shown in (2). The ungrammaticality of (1c) is then readily predicted by the VIC because if the NP *the snow* were preposed the NP dominating *John* would lie in the grossest analysis.⁶

It is interesting to note at this point that passive differs from other movement rules in that it adds morphology, and it is this morphological change on the verb which makes it necessary to mention term 2 in the rule. Compare, for instance, the fact that topicalization, which also moves NP's to the front of the sentence, is very unrestricted with respect to which NP's it can move. It is not just a coincidence that topicalization adds no verb morphology and therefore would have only the NP which moves and the target site in the structural description. Because this is so, topicalization can apply where passivization cannot. The same is true for clefting.

- (7) a. The snow the Mounties searched for John in
(they stopped when they got to the desert).
- b. It was the snow that the Mounties searched for John in.

The mentioning of V in the passive rule greatly restricts its domain of applicability in comparison to other NP movements. It is precisely BECAUSE the verb must be mentioned that the analysis for passive is so restricted. The grossest constituent analysis for passive is lower in the phrase marker than it would be for topicalization. For instance, with respect to (2), a preposing of the NP *the snow* to pre-subject COMP position would have the VP' in the grossest analysis and no NP. As we will see below, when we consider cyclic rule application, this mentioning of the verb in this rule has further interesting implications.

5.2. The Revised Account of Passivization

5.2.1. Restatement of the Passive Rule

The statement of the passive rule given in Chapter 2 and repeated here in (4) contains a verb marked with a feature to indicate that it is some verb other than *be* or *have* (or *become* or *get*, their inchoative counterparts) which cannot undergo passivization. There are numerous other verbs as well which cannot occur in the passive (e.g. **Bill was resembled by Harry*). In Chapter 3, to account for verbs and adjectives which govern raising processes, we found that instead of marking the verbs and adjectives with some sort of feature it was possible to encode the same information on to the particular NP's which the verbs or adjectives could be said to subcategorize. This same approach can also be used for passive. By using it it will be possible, among other things, to account for the data of section 2.2 and much of that in 3.5 which otherwise remains unaccounted for in this framework.

We will begin by discarding the formulation of the passive rule as given in (4) where the passive form is essentially derived from an underlying active. Instead we will consider an approach in which passives are derived from a more abstract deep structure. This approach is motivated by a suggestion made by Emonds as an alternative to an account of passive including a process of agent postposing. Emonds' suggestion (1976:101) is that

The subject NP's of active constructions (both active sentences and active noun phrases) THAT ALTERNATE WITH PASSIVE CONSTRUCTIONS [Emonds' emphasis] are taken to originate in deep structure in a *by* phrase [footnote omitted]. In the domain of sentences ONLY, an agent-preposing rule optionally deletes *by* and replaces an (empty) subject NP with the agent phrase.

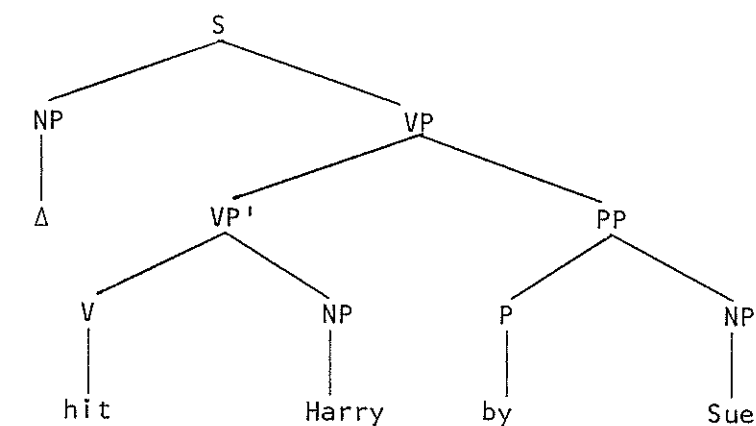
To formalize this approach in the framework of this thesis we would need two rules such as (8) and (9).

$$(8) \Delta - V - NP \Rightarrow 3 - be+en+2 - \emptyset$$

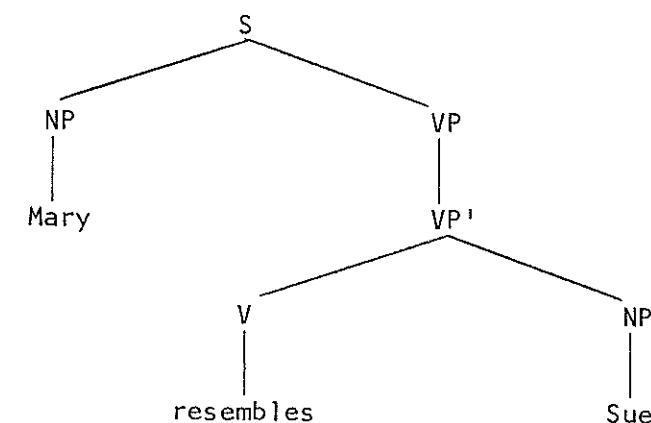
$$(9) \Delta - by - NP \Rightarrow 3 - \emptyset - \emptyset^7$$

From deep structures including empty subject NP's rule (8) would produce passive surface structures and rule (9) would produce actives. Verbs which do not allow for both active and passive constructions would occur in deep structure only with lexically filled NP subjects (for additional recent discussion of base-generated passives see Bresnan 1976b and Wasow 1976). Those which occur only in the active would have no agentive *by*-phrase and those which occur only in the passive would have the optional *by*-phrase and no direct object NP. These three possibilities are illustrated in (10) through (12).

(10)



(11)



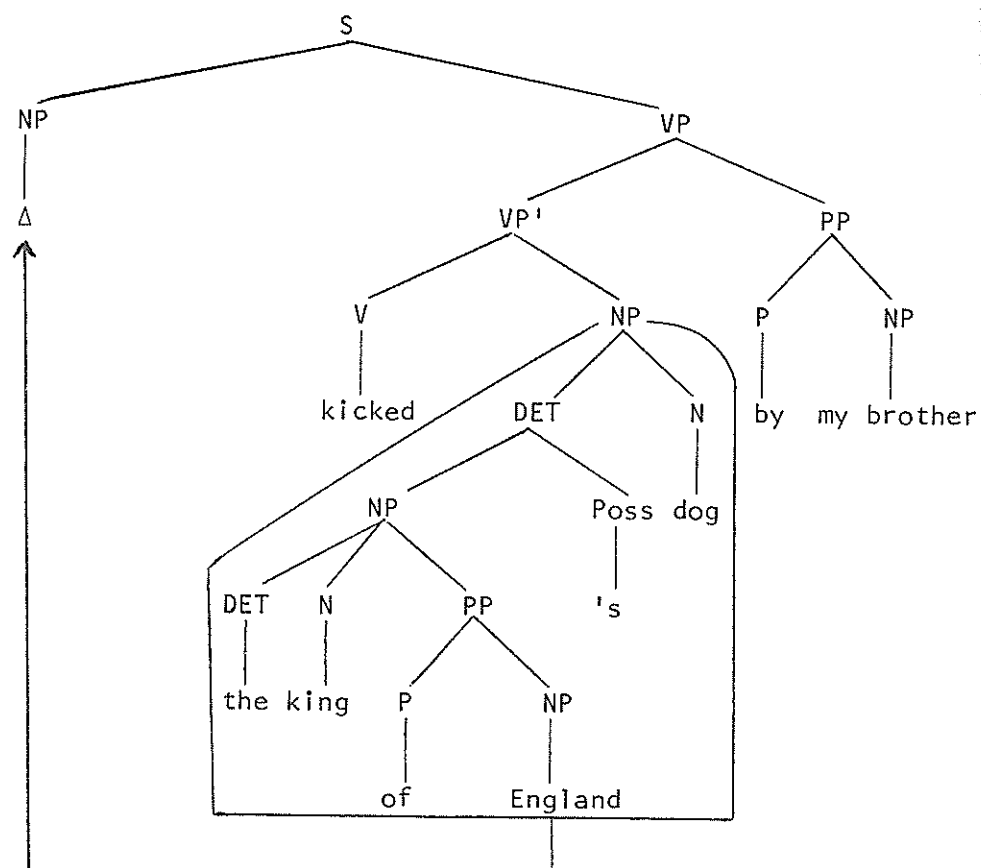
- (14) b. The ball was caught by the man from Chicago.

- c. *The man from Chicago was caught the ball by.

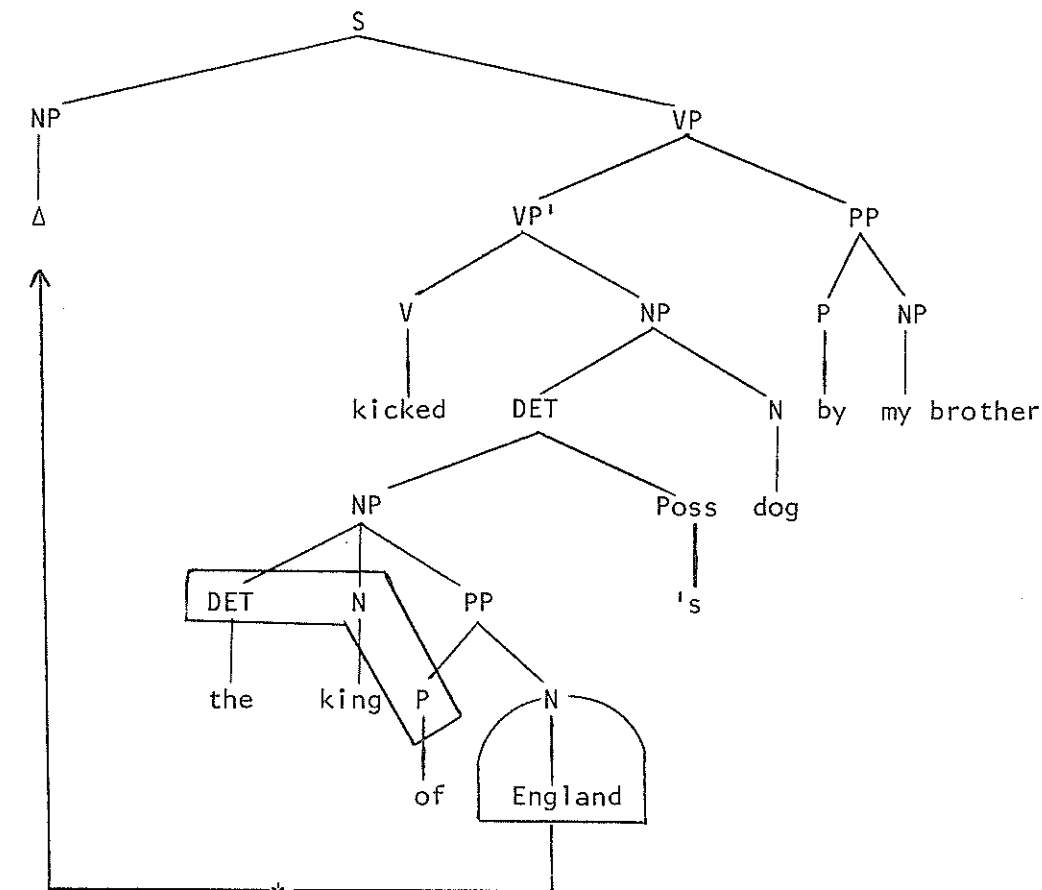
If rule (8) analyzed the NP of the agent phrase as term 3, the NP-sister to the V would occur in the grossest analysis blocking the movement. The VIC will assure that it is always the first NP following the verb which is preposed.

Consider, however, cases where that first NP is a structure which itself contains an NP.

- (15) a.



- (15) b.



- c. The king of England's dog was kicked by my brother.

- d. *England was kicked the king of 's dog by my brother.

- e. *The king[↓] of England was kicked | 's dog by my brother.

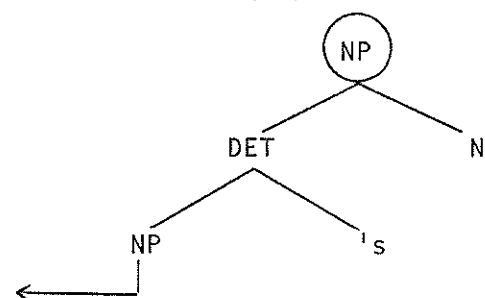
- f. *The king was kicked of England's dog by my brother.

The derivation of (15c) is as indicated in (15a). The ungrammaticality of (15d) can be accounted for in two ways by the VIC. One way is as illustrated in (15b). The NP *England* may not prepose because N, the head of an NP, is in the grossest analysis. Additionally (15d) would be excluded by the provision in the VIC which says that a rule operates first over an empty variable where that is possible (see the final statement of the VIC in Chapter 4). In (15) there are two possible ways for this condition to

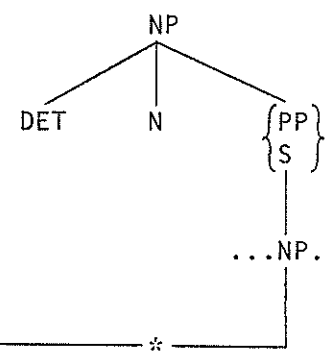
be met. The first is by the application of (8) indicated in (15a). Another way the rule could operate locally would be if the NP to be moved was the NP *the king of England*.⁹ The result of this would be the ungrammatical (15e). This ungrammaticality is accounted for not by the VIC but rather by the Revised Left Branch Condition as given in Chapter 4. The movement of this left branch would require the movement of the whole larger NP, or in other words, would result in the grammatical (15c). The ungrammaticality of (15f) is accounted for simply by the fact that passive only applies to NP's in the sense of N^n (n = maximal number of bars).

Speaking now in more general terms, what was illustrated in (15) is not only that the NP which is preposed by (8) must be the first NP following the verb, but also that it must be the "highest", most inclusive, instance of that first NP. In other words, by the VIC and the Revised Left Branch Condition the same results are obtained as could be achieved by use of the A-over-A condition. As long as the noun phrase movement under consideration is a leftward movement there is no reason to include A-over-A in the inventory of conditions on transformations. Any movement of an NP in a left branch will reorder the larger NP. Any movement of an NP in a right branch will be blocked by the VIC because complements which could contain NP's will always follow the N (or NP) which is the head of the NP (see Chapter 1 of Emonds 1976 for a discussion of base rules). This claim is illustrated in (16).

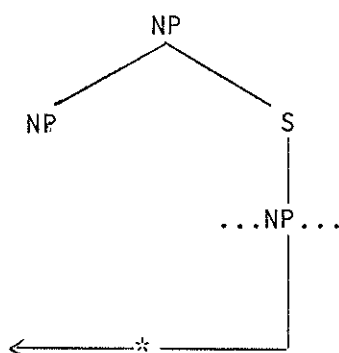
(16) a.



b.



or



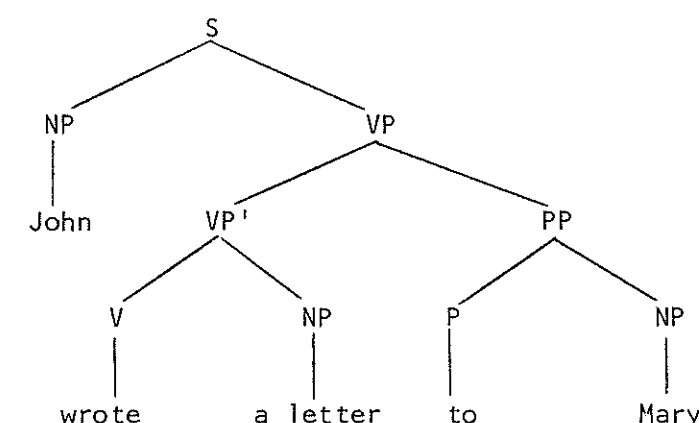
The movement indicated in (16a) will invoke the Revised Left Branch Condition and reorder the whole (circled) NP. This is the only type of structure I am aware of in English where an NP can precede the head of some NP. The

type of movement indicated in (16b) will always be blocked by the VIC. No NP contained in a post-head complement of an NP will be permitted to be re-ordered over the N (or NP) in the grossest analysis. So at least for leftward NP movements, such as passive and raising (and also for WH-fronting if a WH occurs in another WH, e.g. *Whose what did you see?*, **Whose did you see what?*) there is no need for the A-over-A condition.¹⁰ Rather, the results which can be accounted for by A-over-A can be predicted in this model by the independently necessary VIC and Revised Left Branch Condition.

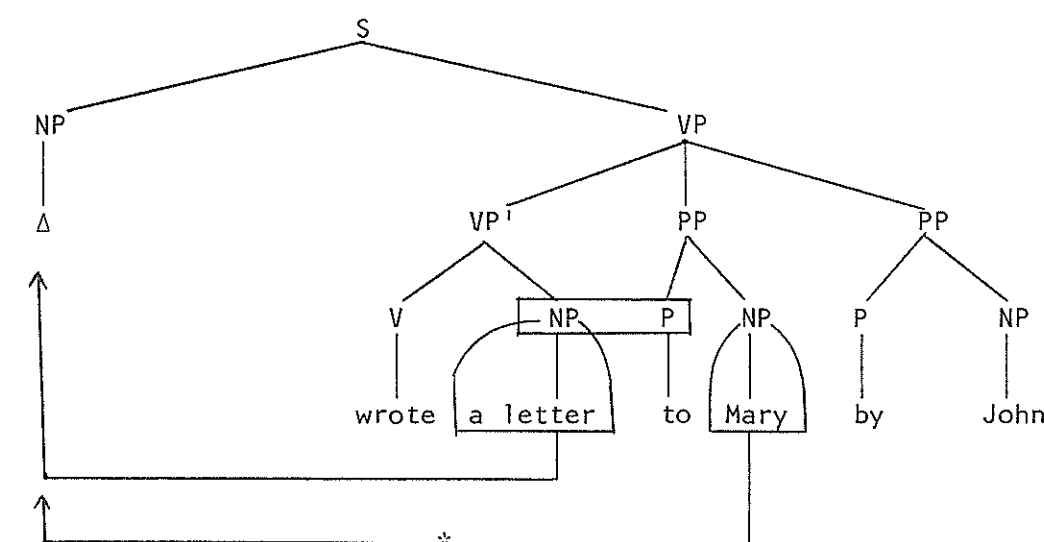
5.2.3. Passive and Indirect Object Movement

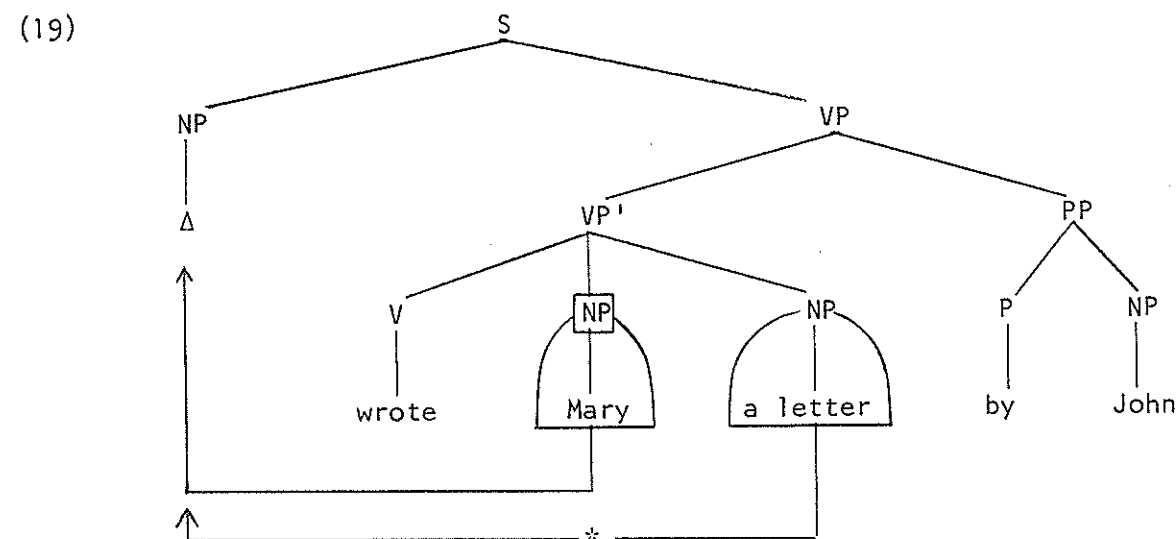
In Chapter 2 it was shown that a VIC account of passivization using a statement of the rule as given in (4) made the proper predictions about the interaction of passive and indirect object movement. It is important here to realize that no generalization is lost by the change to rule (8). Consider the following phrase markers:

(17)



(18)





To (17) rule (8) cannot apply. In (18) and (19) it can apply to move only the first NP following the V and the results are the same as they were in Chapter 2 by the use of rule (4):

- (20) a. A letter was written to Mary by John.
 b. *Mary was written a letter to by John.
 c. Mary was written a letter by John.
 d. *A letter was written Mary by John.

(With respect to the status of example (20d) see example (12) and footnote 4 of Chapter 2.)

5.3. The Transformational Cycle, Raising to Object, and A Rule of Control

5.3.1. An Alternative to Raising to Object

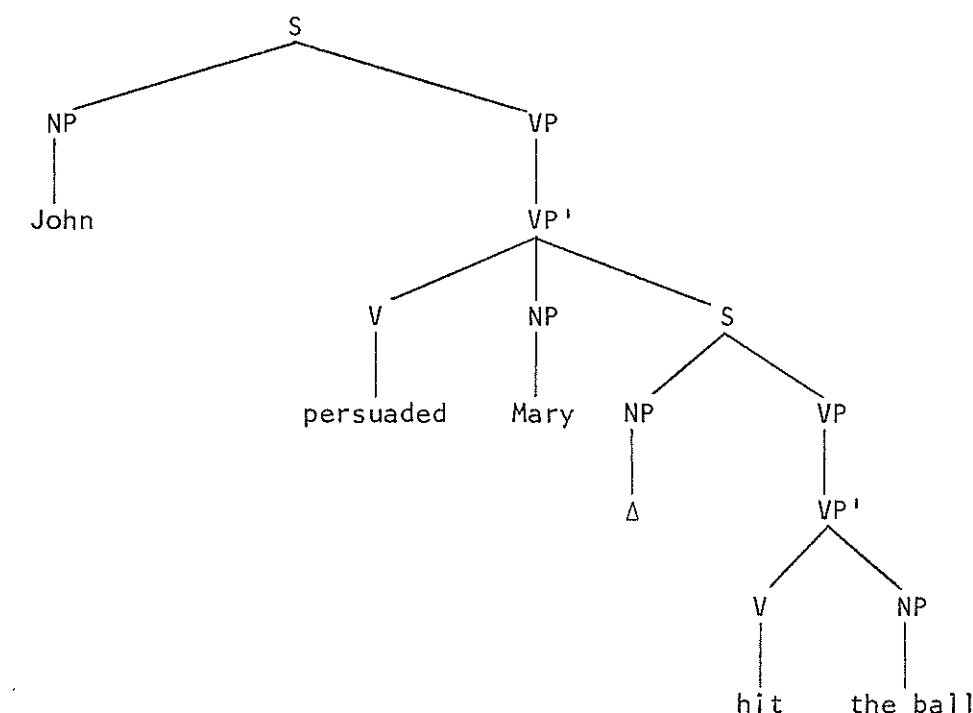
In this section we will reconsider the data presented in Chapter 2, section 2.2. Those examples are repeated here as (21).

- (21) John {a. persuaded
 b. promised
 c. expected} Mary to hit the ball.
 *John {d. persuaded
 e. promised} the ball {to be
 was} hit by Mary.
 f. John expected the ball to be hit by Mary.

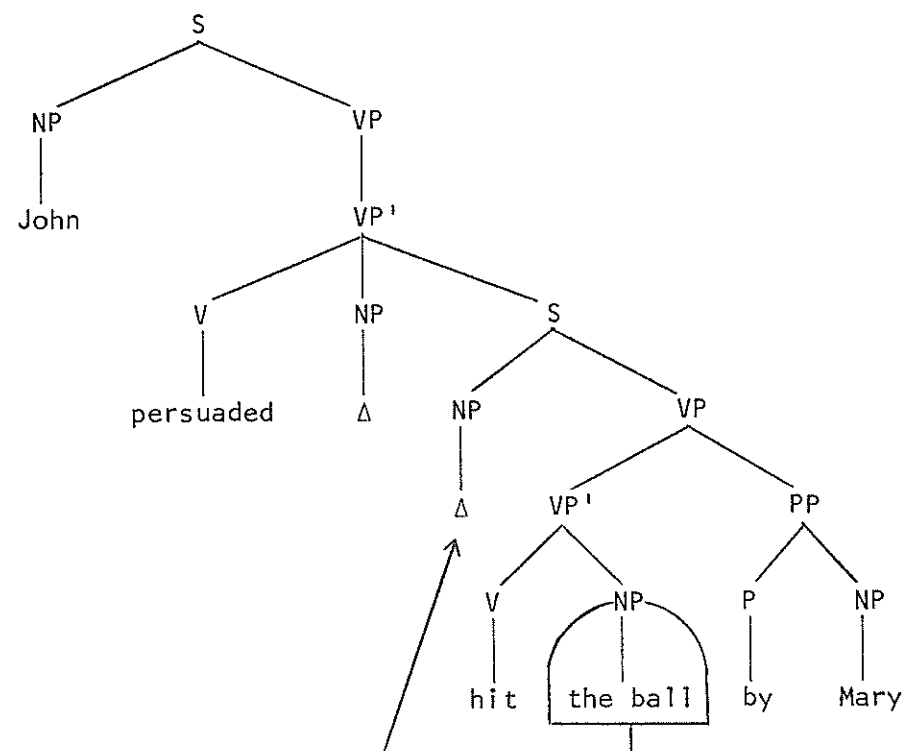
In consideration of these data we will discuss how, at least in these cases, the VIC makes the right predictions with respect to a semantic rule of control (subsection 5.3.3). An in-depth discussion of rules of control is beyond the scope of this thesis, but this consideration will be presented as an account of the facts of (21) and also as an indication of promising directions for further research in the use of the VIC. Additionally this section will show how rule (8) predictably applies in accord with the principle of the transformational cycle (subsection 5.3.2).

To begin, in light of the approach to subcategorization and passivization presented above and the structure associated with *persuade* argued for in Chapter 3 (section 3.4, e.g. example (44)), the deep structure phrase marker for (21a) would be shown in (22a). (We will return below to (22b).)

(22) a.

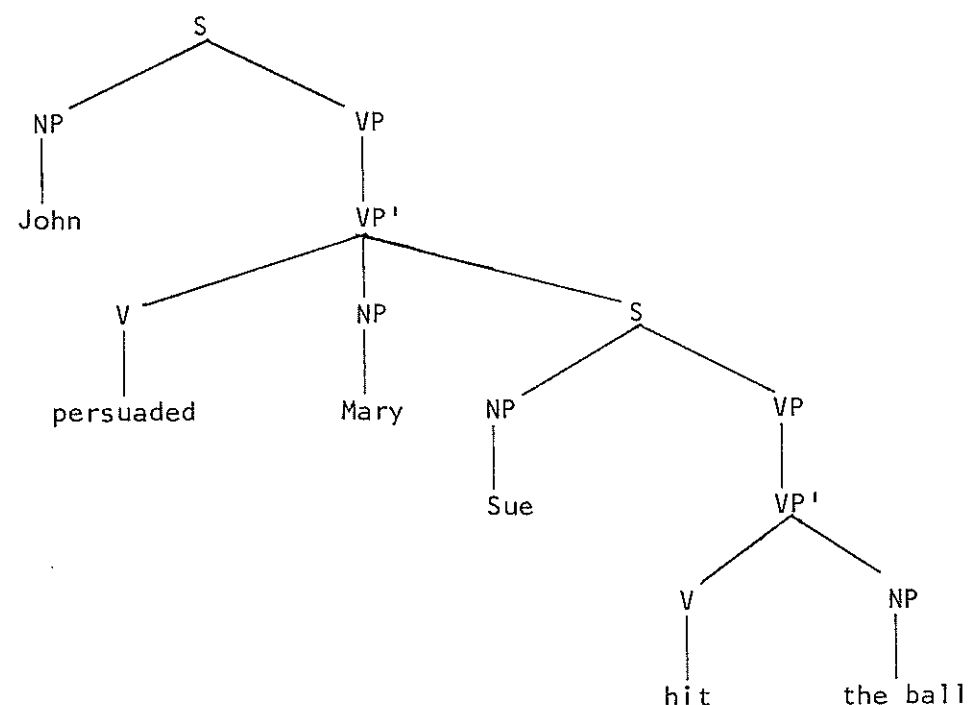


(22) b.



Compare to this a structure like (23a) for the sentence (23b):

(23) a.



b. John persuaded Mary (that) Sue hit the ball.

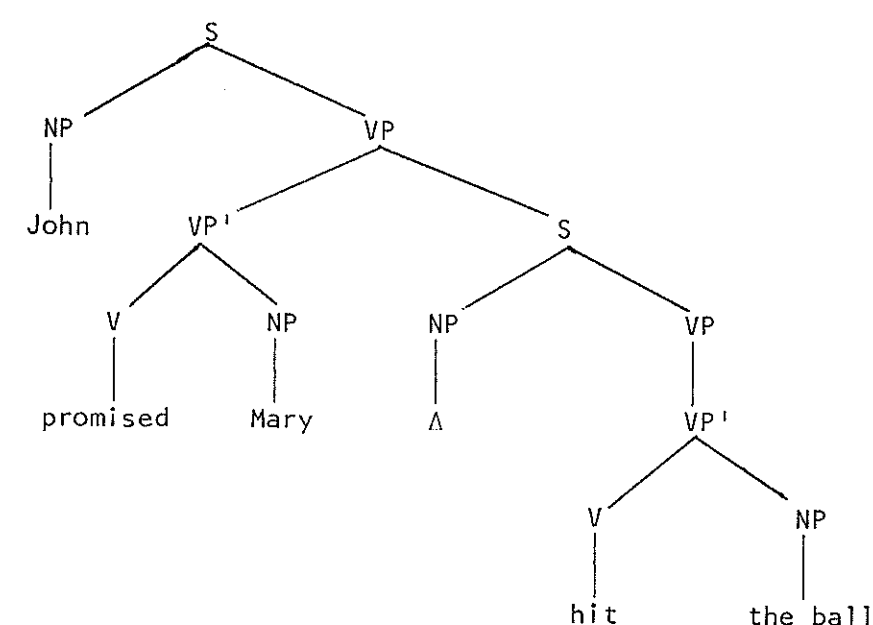
Example (23) provides evidence that *persuade* takes a full S complement. The ungrammaticality of **John persuaded [Sue hit the ball]* is evidence that the object NP in (22) and (23) is obligatory. Notice that if the subject of the embedded clause of (22a) were *Mary* the interpretation for the sentence *John persuaded Mary (that) Mary hit the ball* would be where the two *Mary*'s were distinct in reference. Also, where the subject of the embedded clause is a pronoun, *John persuaded Mary (that) she hit the ball*, there is a possible reading where the subject *she* and *Mary* are distinct in reference. The only case where *Mary* MUST be interpreted as the agent of *hit* is where there is no lexical subject of the embedded clause (for discussion of this issue see Lasnik 1976). We will return to this issue momentarily, but first notice that from (22a) the only possible derived surface structures are (21a) where no transformation has applied, or (24), the result of passive.

(24) John persuaded Mary (that) the ball was hit.

Since there is no longer an agent postposing process in this model example (21d) could not be derived from (22a). Neither can it be derived from (22b). After passive applies, as is illustrated in (22b), the NP object of *persuade* would still be delta. Since there is no way for this delta to be filled it would cause an ill-formed surface string.

Going on now to the (b) example of (21), the deep structure for *promise* sentences would be as in (25).

(25)



Again, as with the *persuade* cases, only when the subject of the embedded clause is lexically empty (as in (21b)) must the interpretation be one of

coreference with an NP of the matrix. For (26a) the two cases of *John* must be distinct; in (26b) *John* and *he* may be distinct in reference.

- (26) a. John promised Mary (that) John would hit the ball.¹¹
 b. John promised Mary (that) he would hit the ball.

The only possible application of passive to (25) will yield (27):

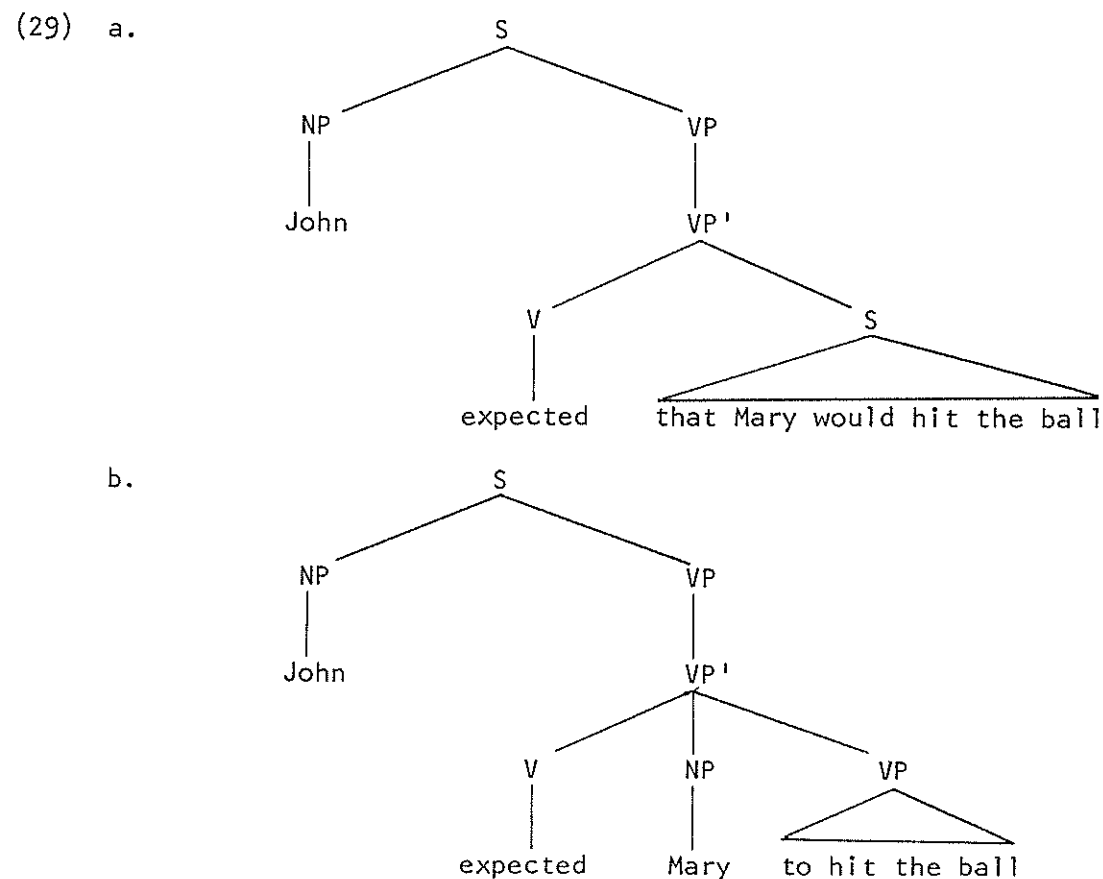
- (27) John promised Mary (that) the ball would be hit.

There is no possible derivation for the ungrammatical (21e).

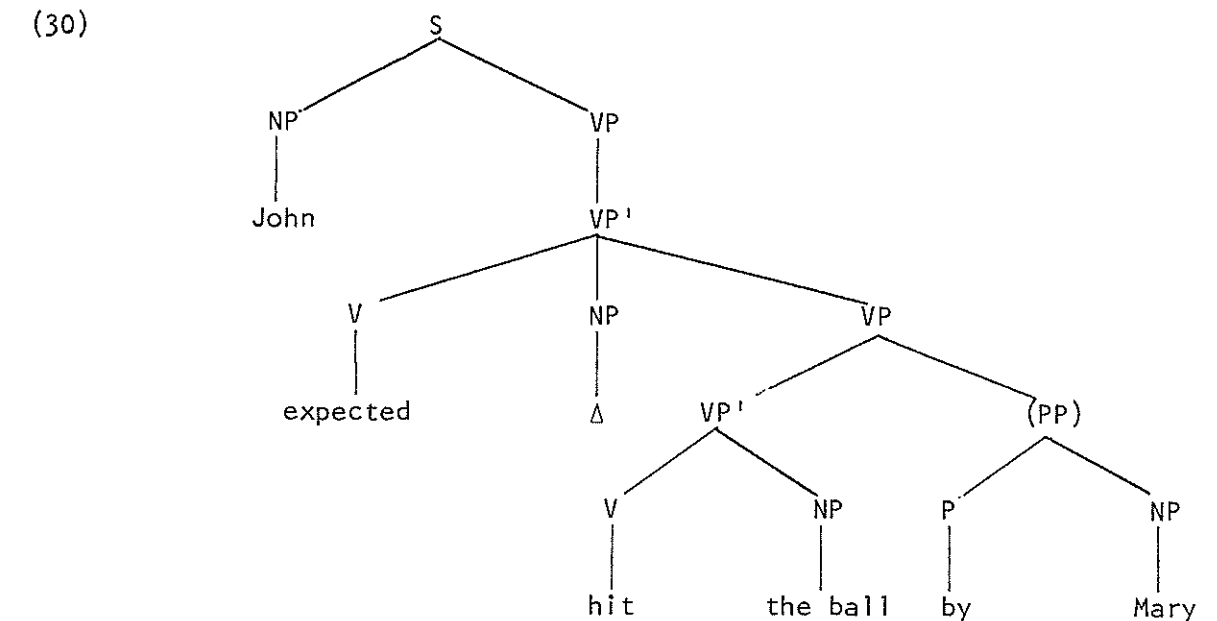
Consider now cases where the verb of the matrix is *expect*. *Expect* can occur on the surface with either a *that*-complement or with a following NP and infinitival VP.

- (28) a. John expected that Mary would hit the ball.
 b. John expected Mary to hit the ball. (= (21c))

The phrase markers associated with the strings of (28) are those of (29).¹²



Where the object NP of *expect* fails to undergo lexical insertion the phrase marker is as shown in (30).



To phrase marker (30) the passive rule (8) can readily apply to replace the delta by the NP *the ball* to yield the grammatical (21f) (here repeated):

- (21) f. John expected the ball to be hit (by Mary).

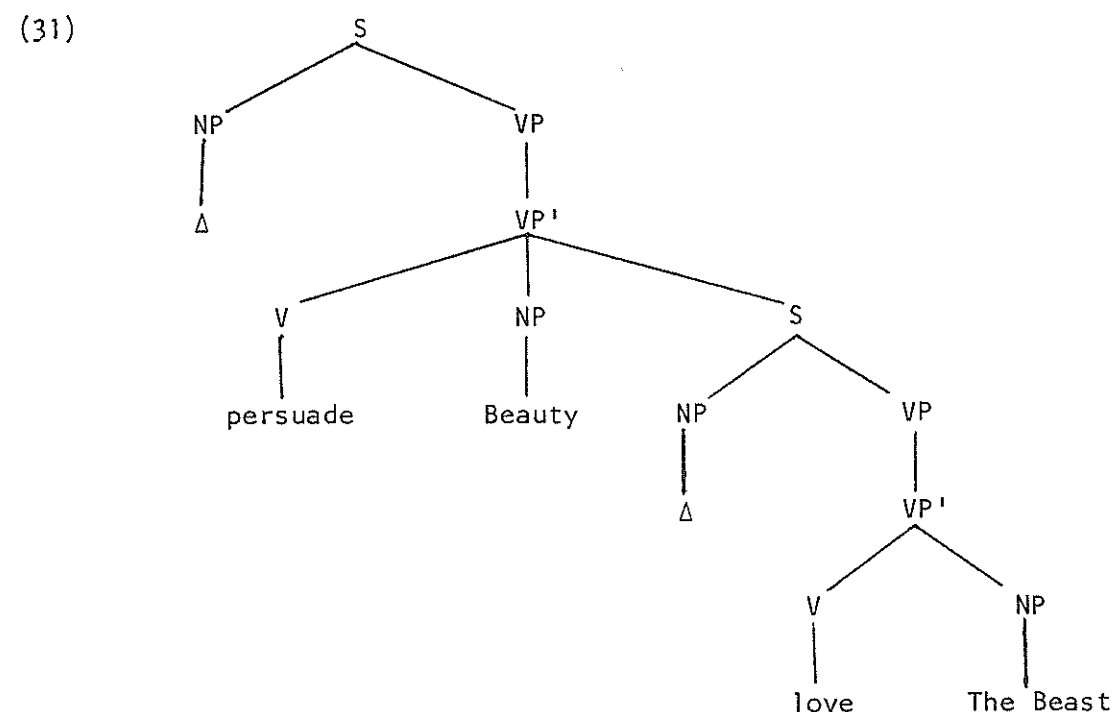
For all of the well-formed strings of (21) the tenseless nature of the embedded predicate is accounted for, in this model, by the fact that verb agreement is a late process which is sensitive to a lexically filled subject NP environment. Tense is spelled out on the verb only where there is a subject NP with some phonological shape; otherwise, the verb remains in infinitival form.

Notice that in this account of the well-formedness of (21c) and (21f) there was no need to assume a process of raising to object position. Neither, however, was it necessary to give up an account of the fact that in derived structure verbs like *expect* have direct object NP's. This is possible here because rule (8), passivization, is not externally constrained to a particular domain. It need not apply only to derive subjects. When (8) applies to phrase marker (30) the result is that an object NP becomes the OBJECT of a higher verb, but by just the normal application of the passivization process. Rule (8) is neither constrained to affect only clausemates nor is it constrained only to create new subjects. It is, however, constrained, as are all movement rules, by the VIC. We will see next how this leads to an account of the cyclic nature of passivization.

5.3.2. The Cyclic Nature of Passivization

Above, in section 3.4, it was possible to show that just by use of the Variable Interpretation Convention, without the additional principle of the transformational cycle, the rule of raising would in effect apply cyclically. The same is true of the passivization process.

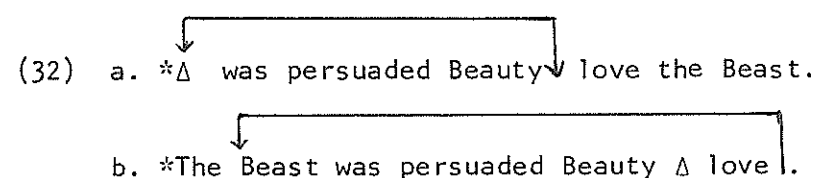
To a phrase marker like (31) there are several possible ways in which it might seem that (8) could apply.



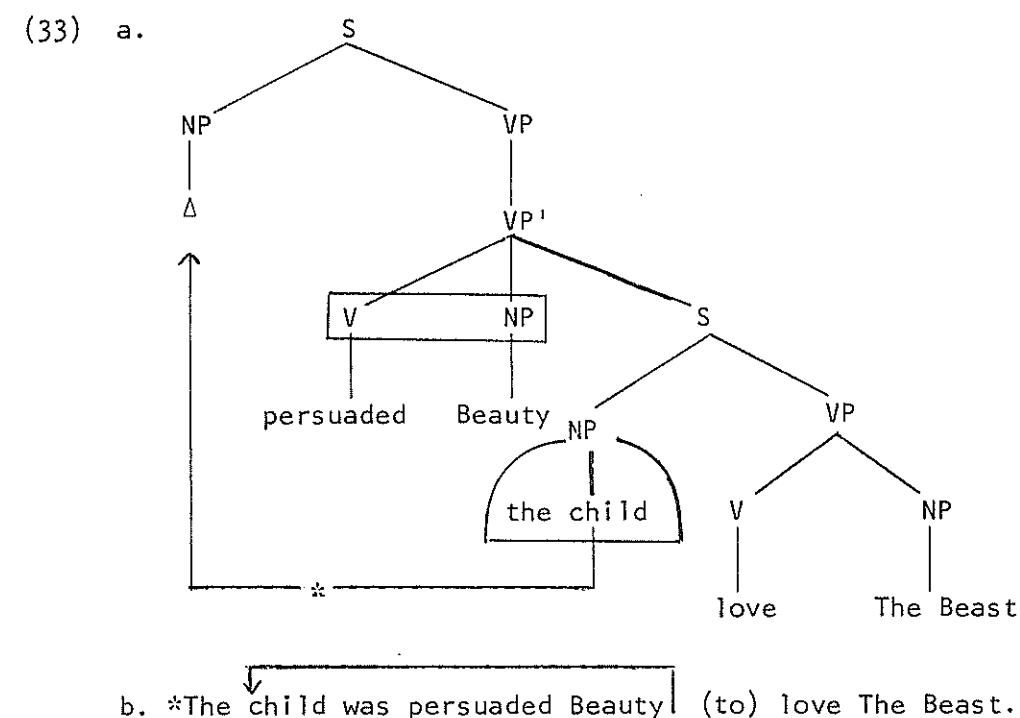
Rule (8) (here repeated for ease of reference)

$$(8) \quad \Delta - V - NP \Rightarrow 3 - be + en + 2 - \emptyset$$

could conceivably analyze the subject of the matrix as term 1, *persuade* as term 2, and then either the subject NP of the embedded S or the NP *The Beast* as term 3. There must, of course, be a way of blocking this since the examples of (32) are ungrammatical.



There are actually many ways these facts could be accounted for. Obviously just the occurrences of delta would be enough to assure that they were marked as ill-formed. This approach is not necessary though, because they could never be derived. (32a) would never arise for two reasons. First of all, transformations never move phonologically empty nodes (see principle (49) of Chapter 3). Additionally (32a) would be blocked because the NP dominating *Beauty* lies in the grossest analysis. This is important because even where the embedded subject is filled this movement must not occur:



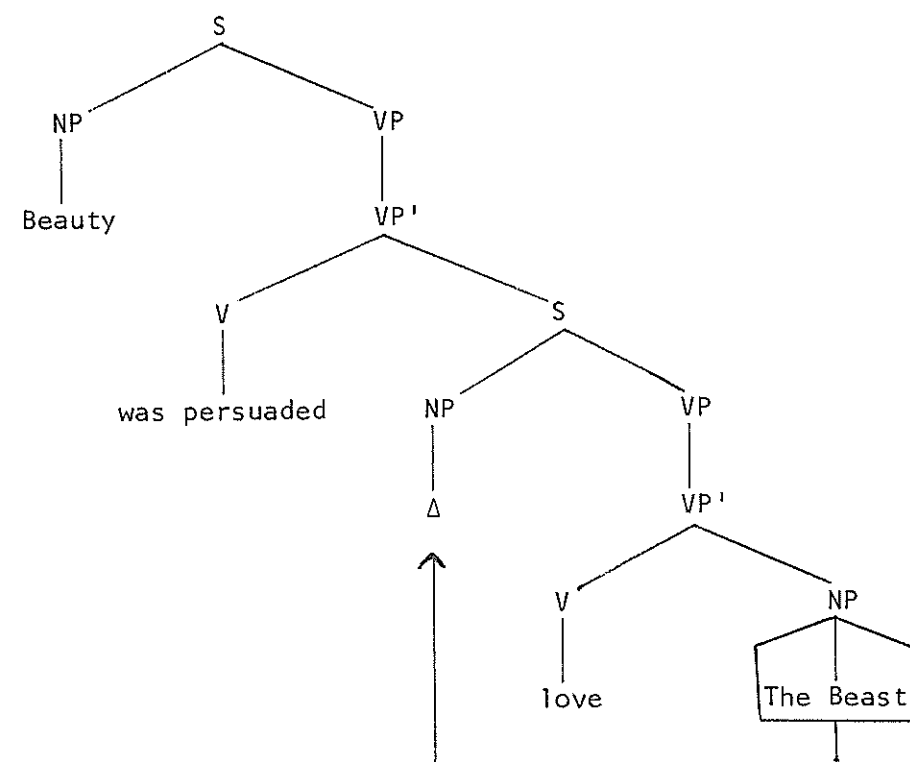
The derivation of (32b) is also properly blocked by the VIC because where the NP *The Beast* is term 3 of (8) the NP *Beauty* and the V *love* lie in the grossest analysis. Either would be sufficient to prevent the rule from applying.

The only application of (8) to (31) which is permitted by the VIC, where term 1 is the matrix subject, is where term 3 is the NP *Beauty*:¹³

(34) Beauty was persuaded to love The Beast.

After passive applies to fill the subject of the matrix clause the structural description of the rule is again met in the lower clause and the rule properly applies.

(35) a.

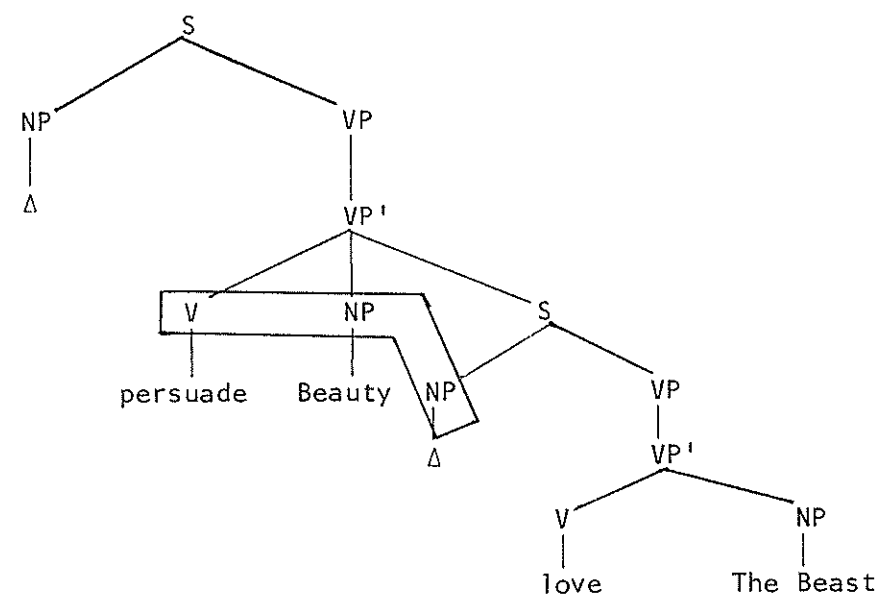


b. Beauty was persuaded (that) The Beast was loved.

The same results are obtained if (8) applies in the lower clause and then in the higher clause.

Let us look again now at (31) (here repeated) to consider the consequences of the two possible analyses for term 2 of rule (8).

(31)

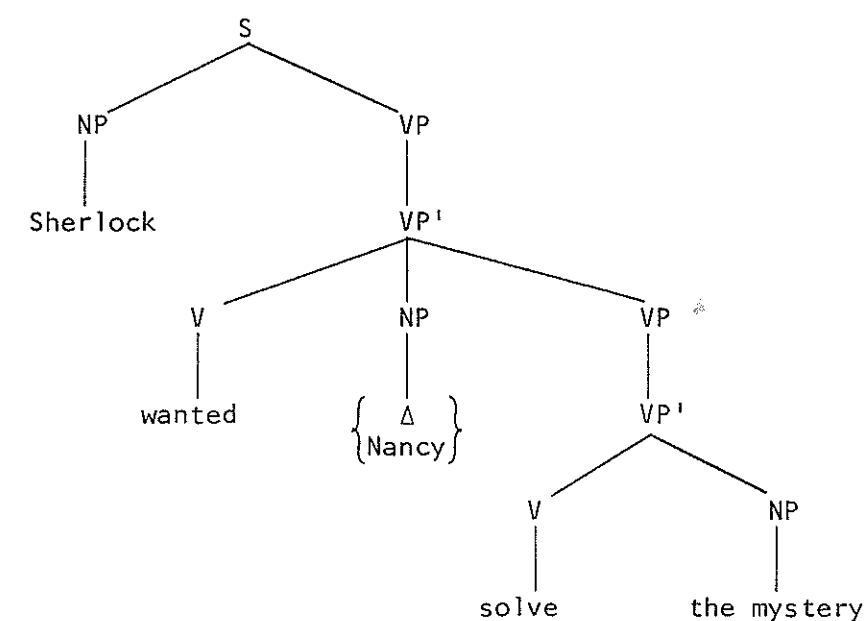


As we just saw, where term 2 is the V *persuade* all the correct results are predicted. The same will be true if (8) analyzes the V *love* as term 2 where term 1 is the matrix subject. No application of (8) is possible where the grossest analysis is as indicated in (31). The analysis between the delta and the V would contain a V and the rule would be blocked.¹⁴ At least for this set of examples rule (8) predictably respects cyclic domains (where "cyclic" here is used for expository ease, since I think there will never be a reason in this model to define the notion cyclic domain).

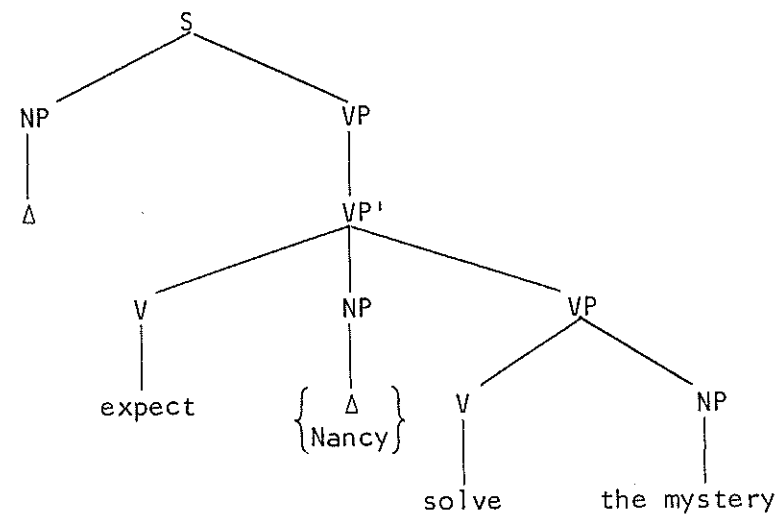
An alternative account for all the delta of (31) through (35) would be the VIC requirement that rules apply to strictly adjacent terms where possible. The fact that rule (8) predictably applies in this fashion shows that, at least in certain cases, this part of the VIC follows from the definition and use of grossest constituent analysis.

For examples like those we have considered so far (using the verb *persuade*) there will always be an NP direct object of the matrix. A verb like *persuade* obligatorily occurs in deep structure with a direct object (*I persuaded that Mary could come, *I persuaded to come). Let us consider examples now where that matrix direct object is optional so that it will not always be the NP available for passivization in the higher clause. The relevant phrase markers are shown in (36) through (38).

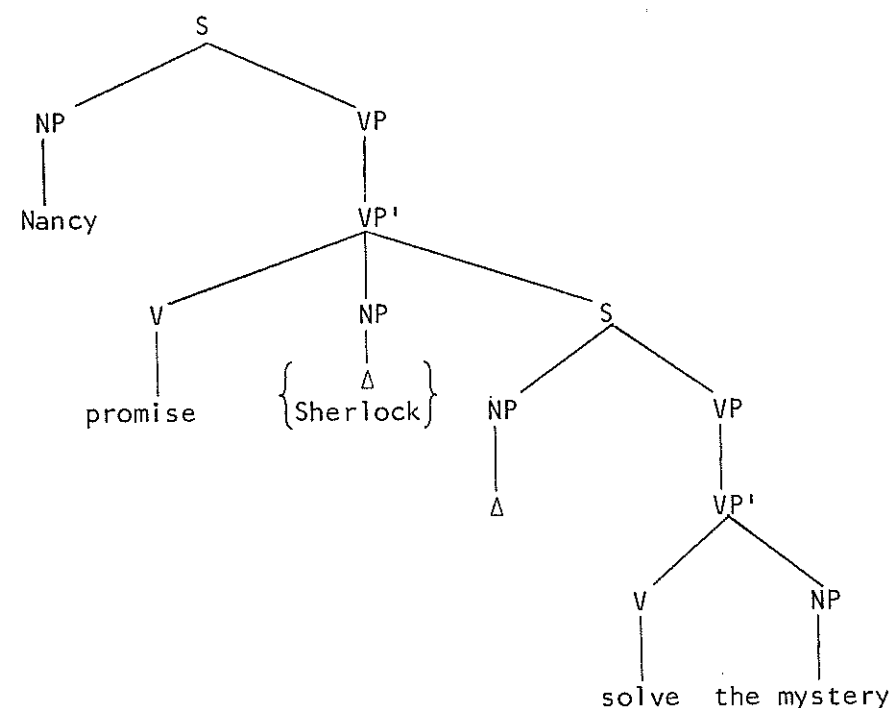
(36)



(37)



(38)

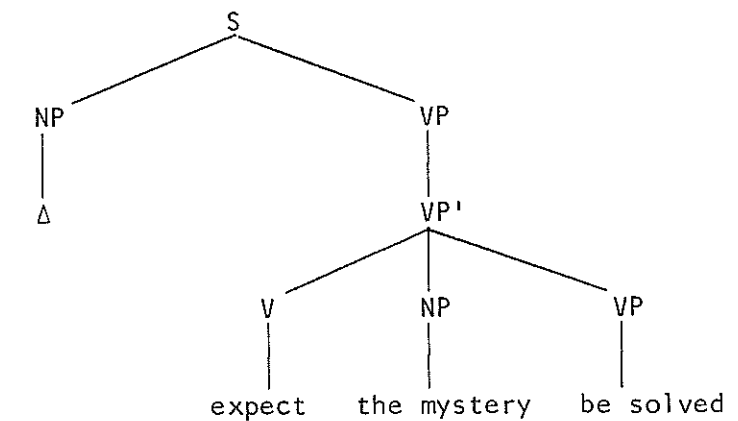


To a structure like (36), where there is an available delta, only one application of rule (8) is possible: *Sherlock wanted the mystery to be solved*. Since *wanted* is not a verb that can undergo passivization it will only occur in deep structure with a lexically filled subject and passive cannot apply in the matrix.¹⁵

Because both *expect* and *solve* can occur in the passive, phrase marker (37) can be used to illustrate that rule (8) predictably will apply first in an embedded domain and then on a higher one, where this type of application is necessary.¹⁶ Where the object of *expect* is empty it cannot be the

NP which is preposed by (8). In that case the only NP which can be analyzed as term 3 of (8) is the NP *the mystery*. If the subject of the higher clause is analyzed as term 1, then whether term 2 is the verb *expect* or the verb *solve*, the VIC will prevent passive from applying. For either analysis for term 2 there would be a verb in the grossest analysis. The only way rule (8) can apply to (37) is to replace the delta object of *expect*. The result of this is (39).

(39)



To this structure passive can again apply and it will properly replace the delta by the NP *the mystery*.

(40) The mystery was expected to be solved.¹⁷

Where it should, passivization properly applies in cyclic fashion.

In (38), where the object of *promise* is not lexically filled there are two deltas which might be analyzed as term 1 of (8). If the direct object delta is so analyzed (8) will be prevented from applying, by the VIC, because there would be a possible local application. The delta subject of the embedded clause must be analyzed as term 1 of (8) by the VIC requirement that rules function first, if they can, where the variable is empty. From (38), (41a) is properly derived and (41b) is properly blocked.

(41) a. Nancy promised (that) the mystery $\left\{ \begin{smallmatrix} \text{was} \\ \text{would be} \end{smallmatrix} \right\}$ solved.

b. *Nancy promised the mystery to be solved.

For passivization, just as for the processes previously discussed, there seems to be no reason to invoke the principle of the transformational cycle.

5.3.3. Control Properties and the VIC

5.3.3.1. A VIC Formalization of the Minimal Distance Principle

We now turn to a brief discussion of a semantic rule of control.¹⁸ Again we will be considering examples with embedded complements where the matrix predicates are verbs like *persuade*, *promise*, and *expect*. As was pointed out above (section 5.3.1) where the complements to these verbs have phonologically designated NP subjects there will always be a possible reading where the referent for the embedded subject is distinct from any previous occurrence of an NP. This is illustrated again in (42).

- (42). a. Freud_i expected (that) he_{i,k} could cure the patient.
 b. Freud_i persuaded Holmes_j (that) he_{j,k} could solve the mystery.
 c. Freud_i promised Holmes_j (that) he_{i,k} would solve the problem.

As the subscripts are meant to indicate, in (42) the pronoun subjects of the embedded clauses can be identical in reference to an NP of the matrix clause (i or j) or their reference can be completely independent (k), that is, determined by any NP, regardless of whether or not it occurred previously in the string. For discussion see Dougherty (1970).

In (43), however, the interpretation of the NP which is represented here as PRO is uniquely determined by a particular previously occurring NP.

- (43) a. Freud_i expected PRO_i to cure the patient.
 b. Freud_i persuaded Holmes_j PRO_j to solve the mystery.
 c. Freud_i promised Holmes_j PRO_j to solve the problem.

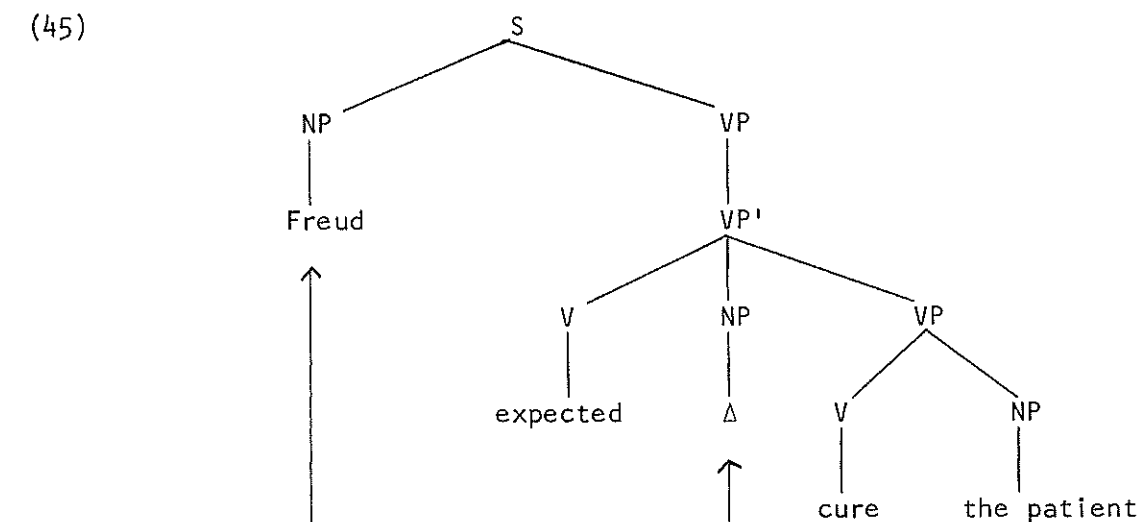
By use of the VIC, at least for the paradigmatic examples, it is possible to account for which NP of the matrix controls the interpretation of the PRO element. In other words there is a VIC account of certain "control properties" of particular lexical items. I will be using the terms "subject control" and "object control" as they have been used recently by Chomsky (class lectures, 1976) where, for instance, *persuade* requires object control (as in (43b)) and *promise* requires subject control (as in (43c)).

We will assume, as is generally taken for granted, that lexically fully designated NP's determine their own reference. This is illustrated in (44) where k is not identical in reference to either i or j.

- (44) Freud_i {persuaded / promised} Holmes_j (that) Watson_k would solve the problem.

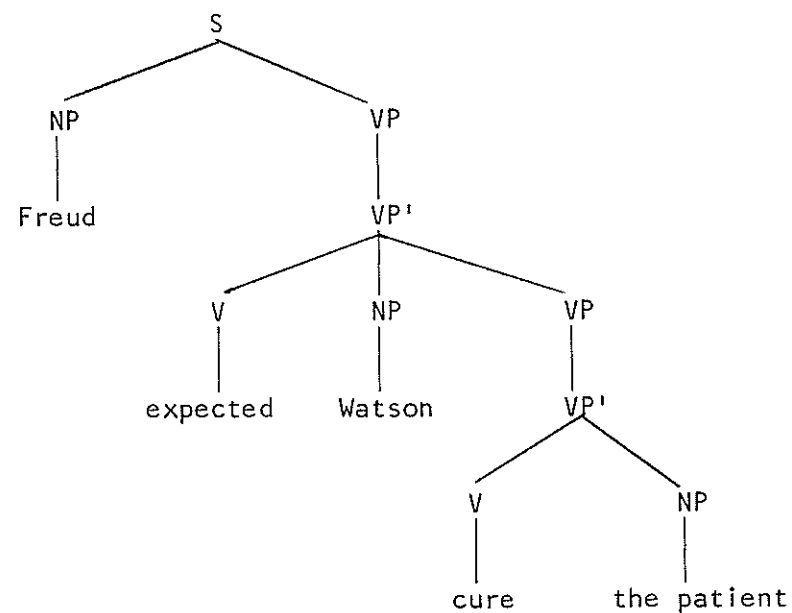
Where there is no NP lacking semantic features the control properties are not invoked. If, however, there is an NP in a terminal string which dominates some item which does not contain semantic features (delta) then semantic features are copied onto that item from some other NP of the string (see principle (20) and the discussion immediately following it in Chapter 3). Which NP that will be depends on the structure associated with that string.

Let us assume for the moment that the general case is that the closest preceding NP determines the interpretation of the delta (this is basically like Rosenbaum's identity erasure process, which is also known as the minimal distance principle (Rosenbaum 1967)). Consider (45):



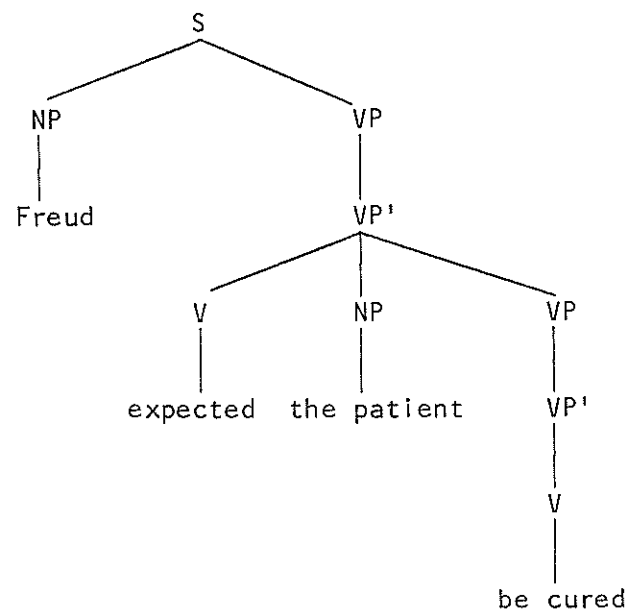
The closest (and here, only) NP which precedes the delta is *Freud* and this provides the proper interpretation as shown in (43a). If it happens that the NP object of *expect* is filled, either by lexical insertion as in (46a) or by an application of passive as in (46c) then no rule of control applies.

(46) a.



b. Freud expected Watson to cure the patient.

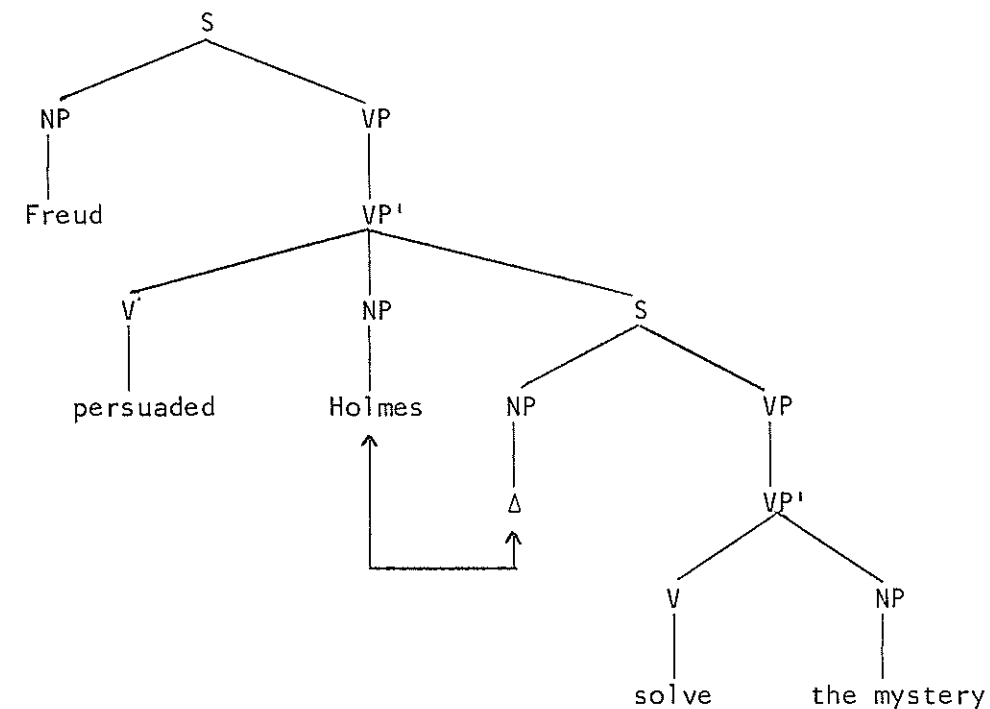
c.



d. Freud expected the patient to be cured.

The structure in (47) is that which would underlie (43b).

(47)



The closest preceding NP to the delta in this case is *Holmes* and it properly controls the interpretation.

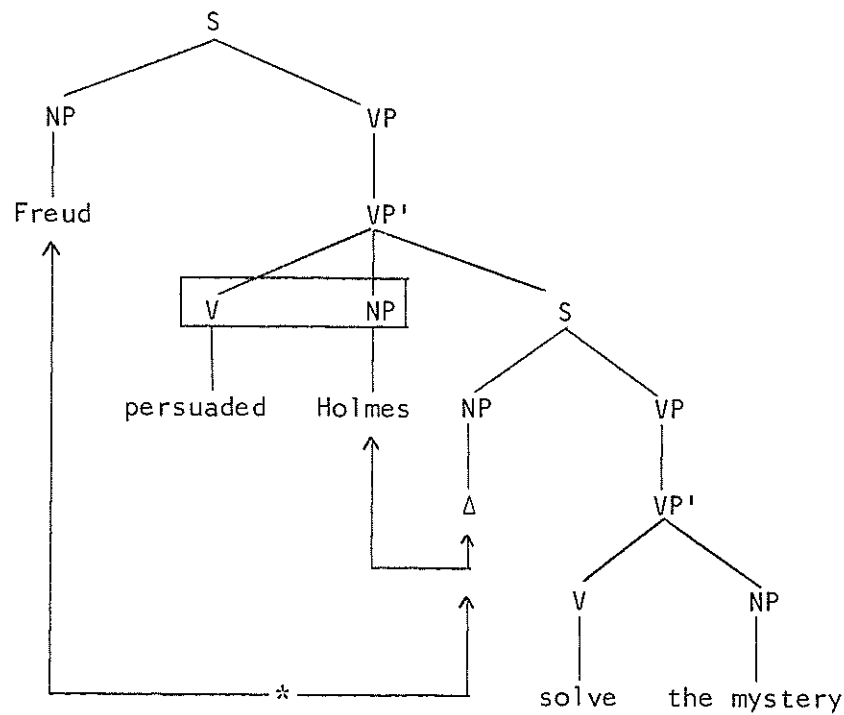
This rule of control can easily be formulated as rule (48):

$$(48) \begin{bmatrix} \text{NP} \\ \alpha F \end{bmatrix} - \Delta \Rightarrow 1 - \begin{bmatrix} 2 \\ \alpha F \end{bmatrix}$$

where $[\alpha F]$ is meant to indicate the entire bundle of semantic features. This rule is then subject to the VIC.

Where only one NP precedes the delta in a given structure, as in (45), it is obvious that (48) will give the right result. (47) (repeated here) is a more interesting case for (48).

(47)



By rule (48) the NP *Holmes* properly controls the interpretation of delta. The indicated grossest constituent analysis, together with the fact that *persuade* has an obligatory NP object (see examples (22) and (23) and the text immediately following them) explains why it is that this verb **REQUIRES** object control.

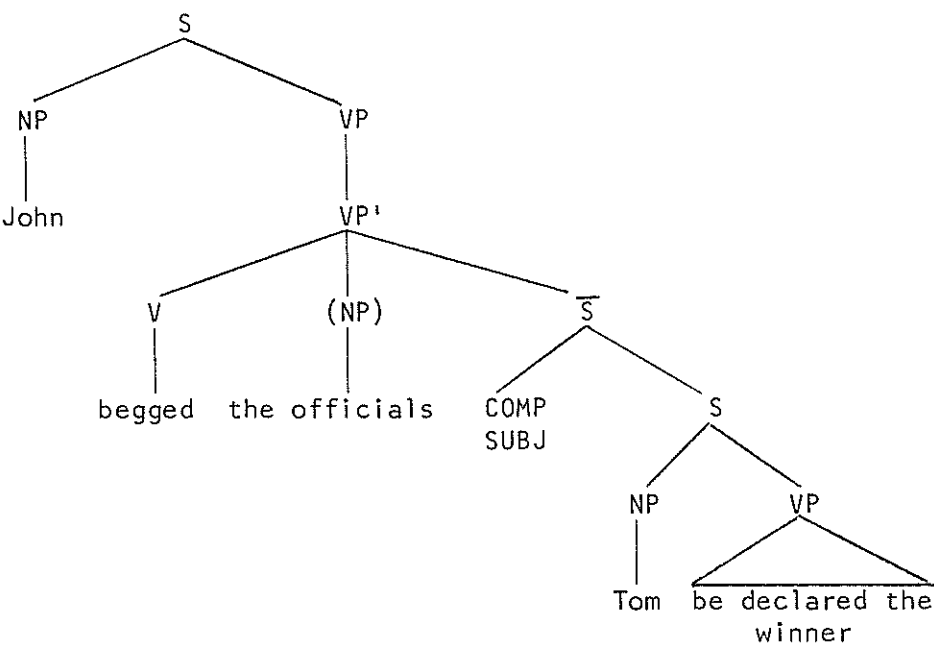
Next consider some verbs which can exhibit both subject control and object control properties. Compare the examples of (49).

- (49) a. Didi {asked
begged} Pozzo to unburden Lucky.
b. Didi {asked
begged} to unburden Lucky.

A verb like *beg* may occur in deep structure with or without a direct object NP, but like *persuade* (and unlike *expect*) it has a full S complement. This is illustrated by (50).

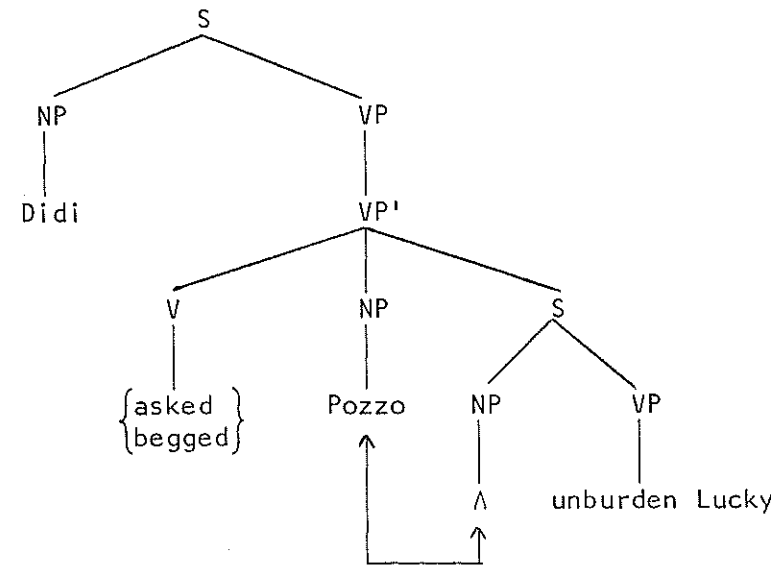
- (50) a. John begged (the officials) {that Tom be declared the winner.
for Tom to be declared the winner.}

(50) b.

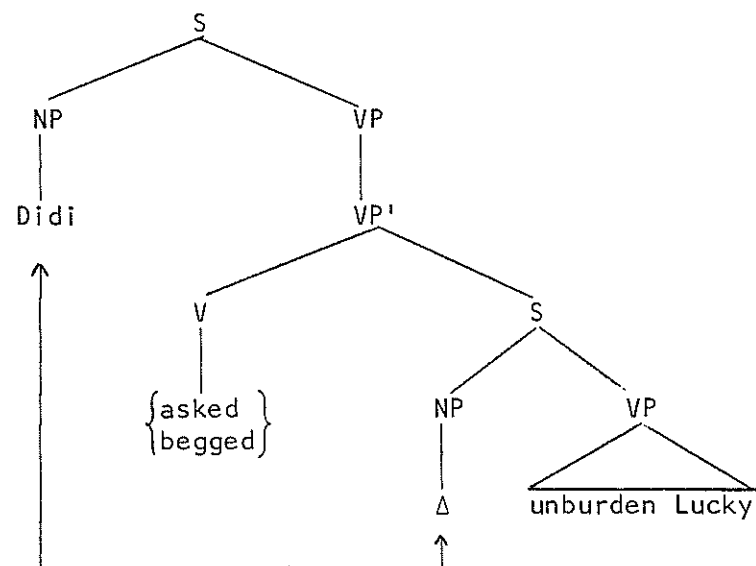


The structures of (51) are now given as the appropriate deep structures for the sentences of (49).

(51) a.



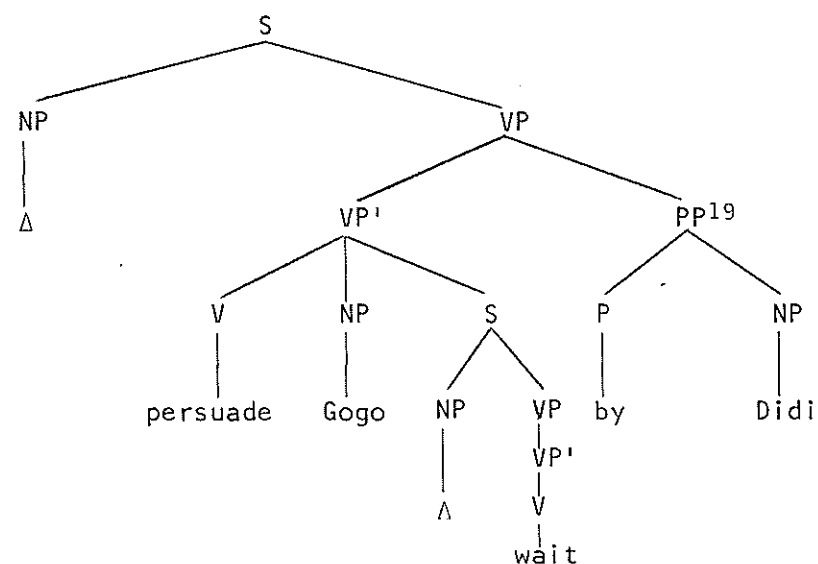
(51) b.



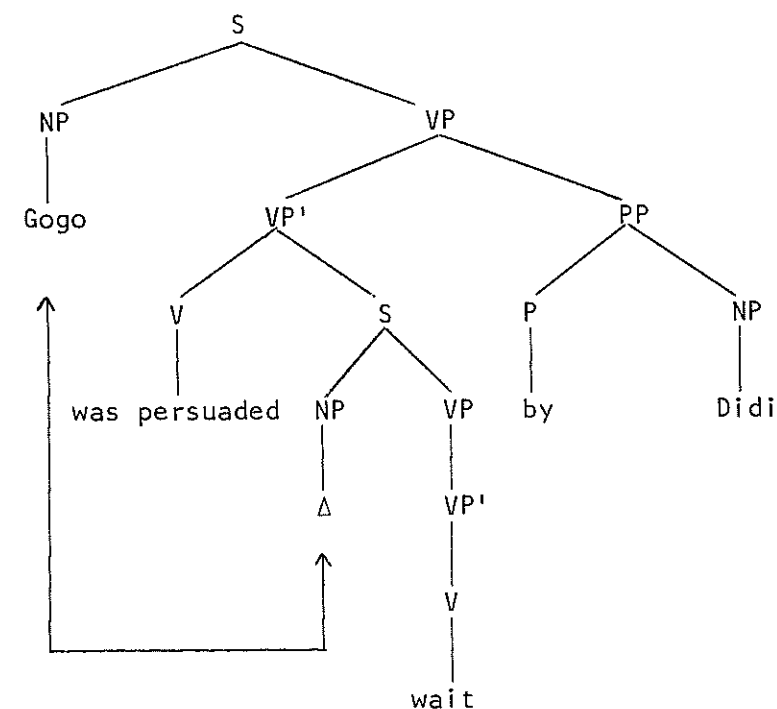
As the arrows indicate in (51a) and (51b), rule (48) makes the right prediction about the control exerted by *beg* and *ask*. Where there is an object NP these verbs may exhibit object control, rather than subject control. Where there is no object NP, as in the (b) example, then the subject of the matrix controls the interpretation of the delta.

Verbs like *beg* do not obligatorily require object control. It is because *persuade* (and other similar verbs, like *convince*, *advise*, *induce*, etc.) has an obligatory direct object that it has obligatory object control. In these cases there is a very direct structural reflex of semantic facts. Because this is so, the VIC and rule (48) provide an interesting account of control. Consider, as further evidence, the result of passivization where the verb of the matrix is *persuade*.

(52) a.



(52) b.

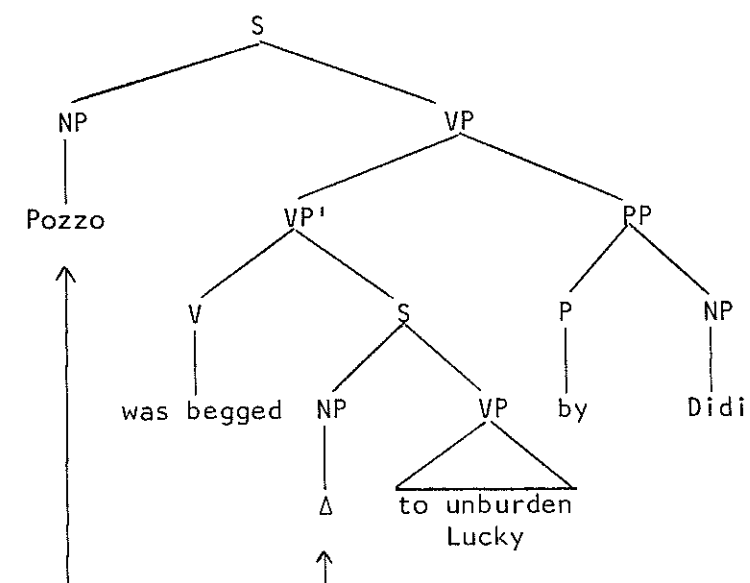


c. Gogo was persuaded to wait by Didi.

(52a) is the pre-passivized structure and (52b) is the result of passive and the rule of control. In (b) it is the grammatical subject of *persuade* which controls the interpretation of the delta, as the meaning of (52c) indicates. Where there is no direct object of *persuade* it allows for subject control.

The same thing is true for passivization and the *beg* case.

(53) a.

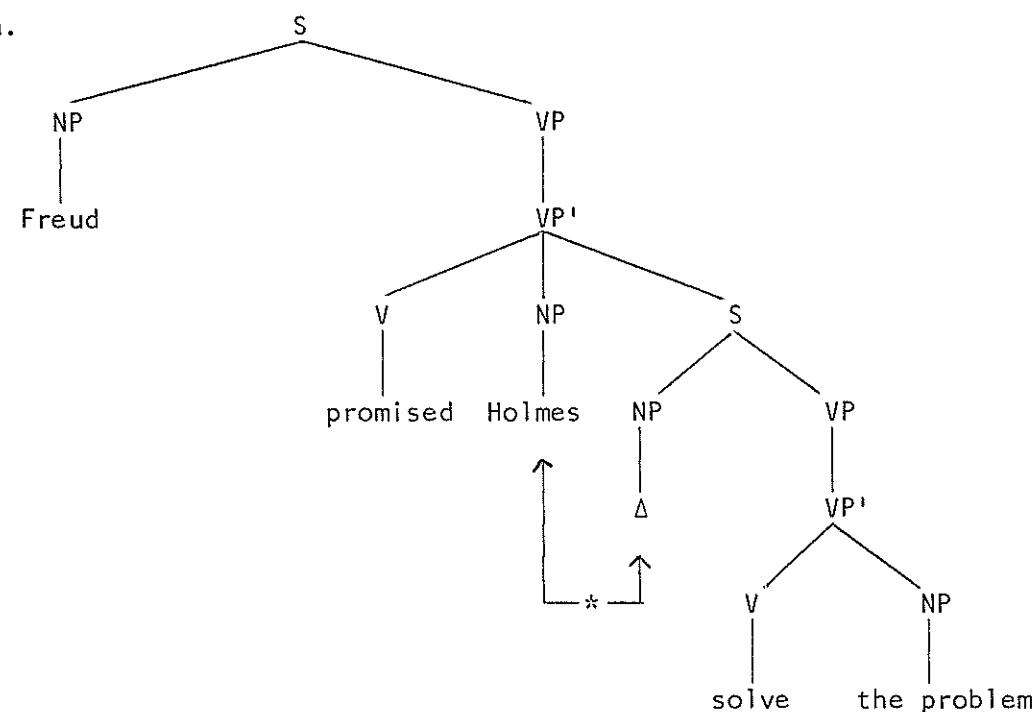


(53) b. Pozzo was begged to unburden Lucky by Didi.

The grammatical subject of (53a) properly controls the interpretation of the delta, as rule (48) would predict.

We now turn to the case of *promise*.

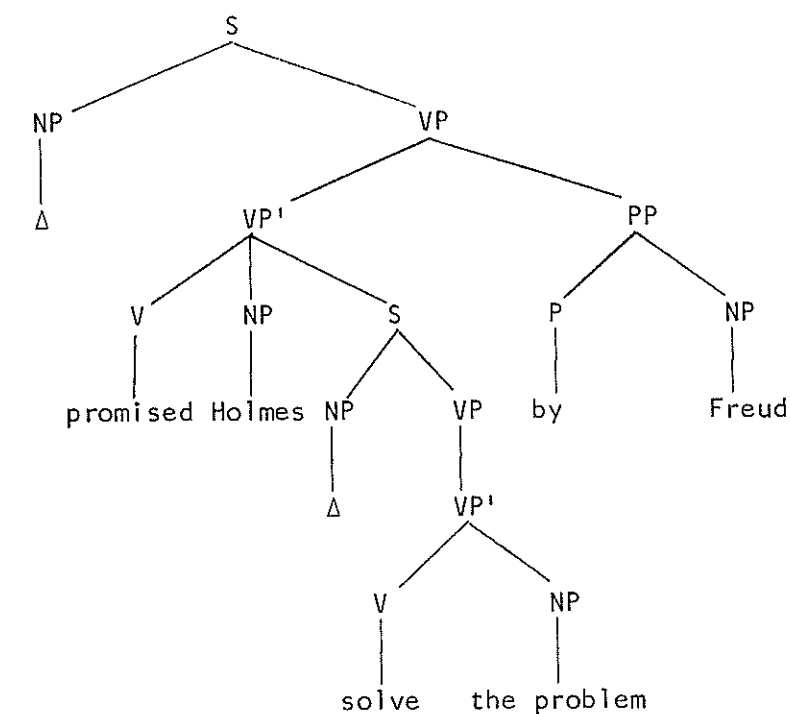
(54) a.



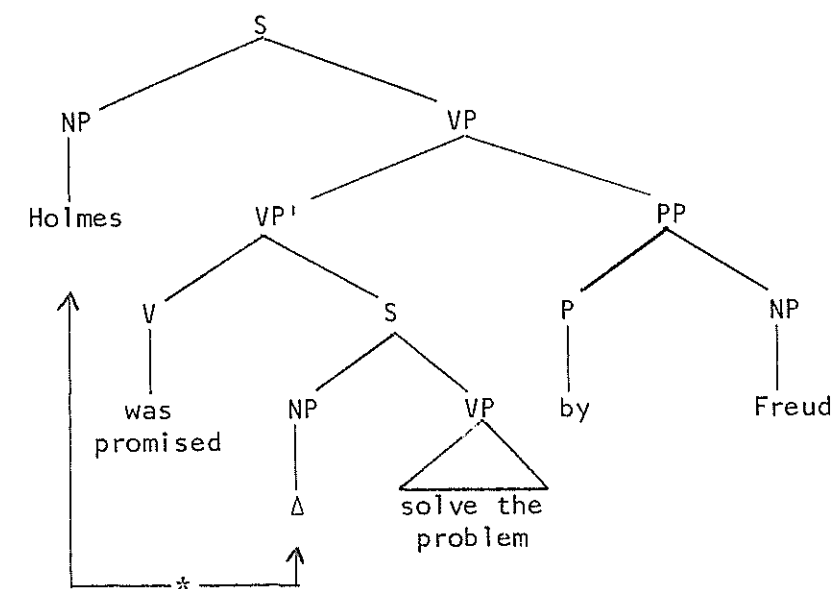
b. Freud promised Holmes to solve the problem. (= (43c))

As the arrows in (54a) indicate, rule (48) gives the wrong results for the interpretation of delta where *promise* is the matrix verb. Here the prediction that the closest preceding NP controls the interpretation is wrong. *Promise* requires subject control whether or not it has an object NP. It seems possible that with the VIC and a certain condition on rule (48) this control property of *promise* can be accounted for. This will be discussed in 5.3.3.2. For now, *promise* is an exception to rule (48). *Promise* is additionally exceptional in that it not only requires subject control, it requires AGENT-subject control. This is shown in (55).

(55) a.



b.



c. *Holmes was promised to solve the problem by Freud.

Where the grammatical subject is not the agent of "promising" no interpretation is possible. The ungrammaticality of (55c) cannot be attributed to the fact that *promise* does not undergo passive because a sentence like

Karen was promised a present by Grandma is perfectly grammatical. *Promise* is just exceptional with respect to rule (48).

5.3.3.2 An Alternative to the Minimal Distance Principle

The basic control property difference between *persuade* and *promise* can in fact be accounted for here if two additional assumptions are incorporated into the VIC framework.

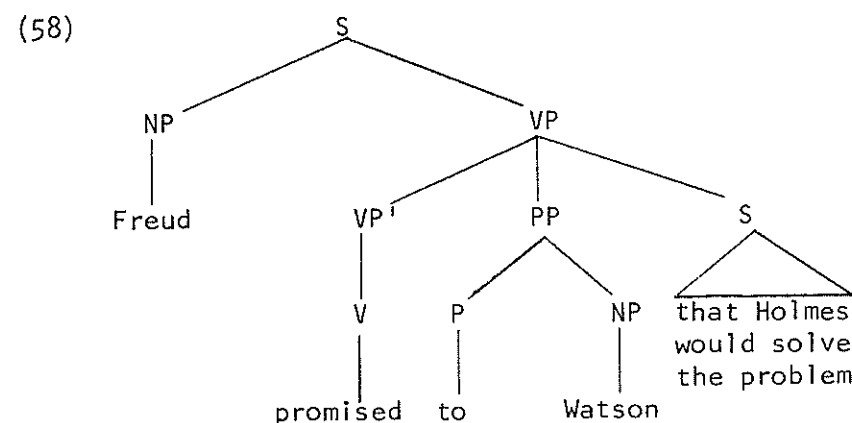
First, it seems plausible that the NP object which follows *promise*, where there is an S complement, is a dative rather than an accusative object. That this is so is confirmed by the synonymy of (56a) and (56b).²⁰

- (56) a. Freud promised to Watson that Holmes would solve the problem.
 b. Freud promised Watson that Holmes would solve the problem.

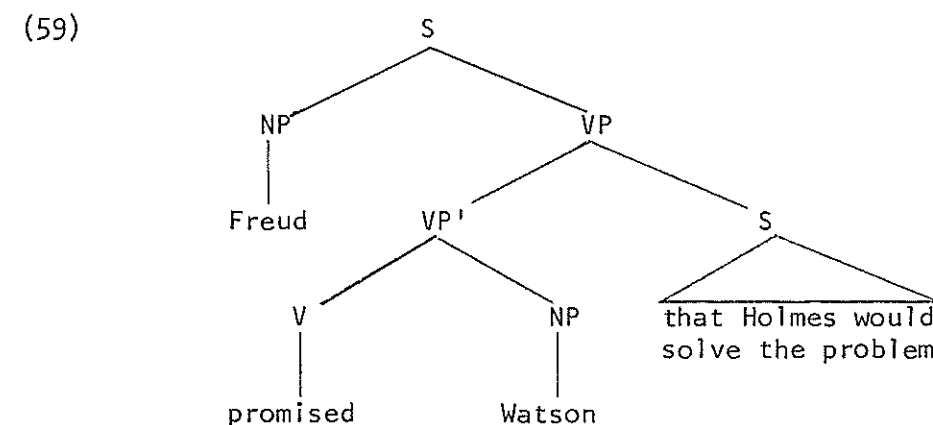
The ungrammaticality of (57a) shows that *persuade* takes only an accusative object.

- (57) a. *Freud persuaded $\begin{Bmatrix} \text{for} \\ \text{to} \end{Bmatrix}$ Watson that Holmes would solve the problem.
 b. Freud persuaded Watson that Holmes would solve the problem.

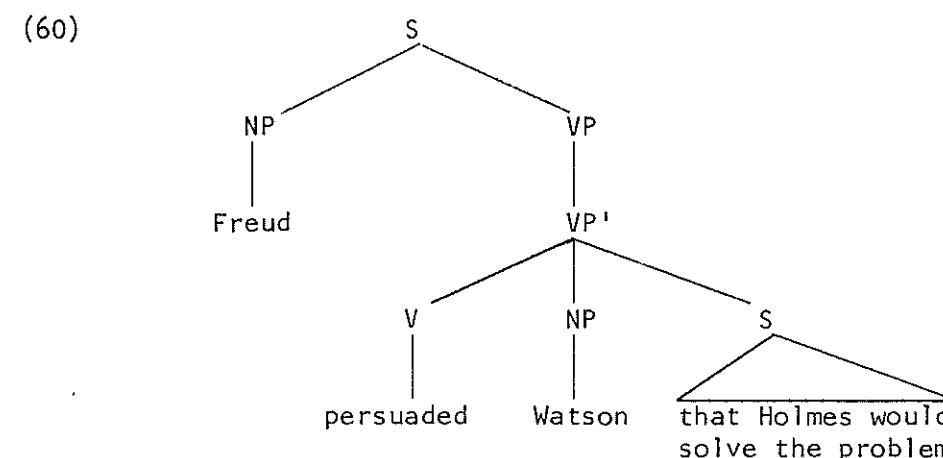
In (58), the phrase marker of (56a), the indirect object prepositional phrase, like all other indirect objects which have not undergone dative movement, occurs outside the VP'. The S complement which follows the PP therefore must also occur outside the VP'.



From (58), by indirect object movement, (56b) can be derived. Indirect object movement may apply whether or not there is a direct object. After indirect object movement, the dative object is attached under VP' (see, for example, (19) in section 5.2.3). The S, however, remains unmoved and outside the VP'. The structure underlying (56b), therefore, would be (59).



The structure underlying (57b) would be (60) where the S, like other S complements, is inside VP'.



The constituency tests provided earlier for VP' (Chapter 3) also suggest that there is a difference between the structure of the VP for *persuade* and for *promise*, but the acceptability judgements are so subtle that not much can be based on them. To my intuitions the (a) examples are better than the (b) examples below. This would indicate that for *promise* the S is outside the VP'.

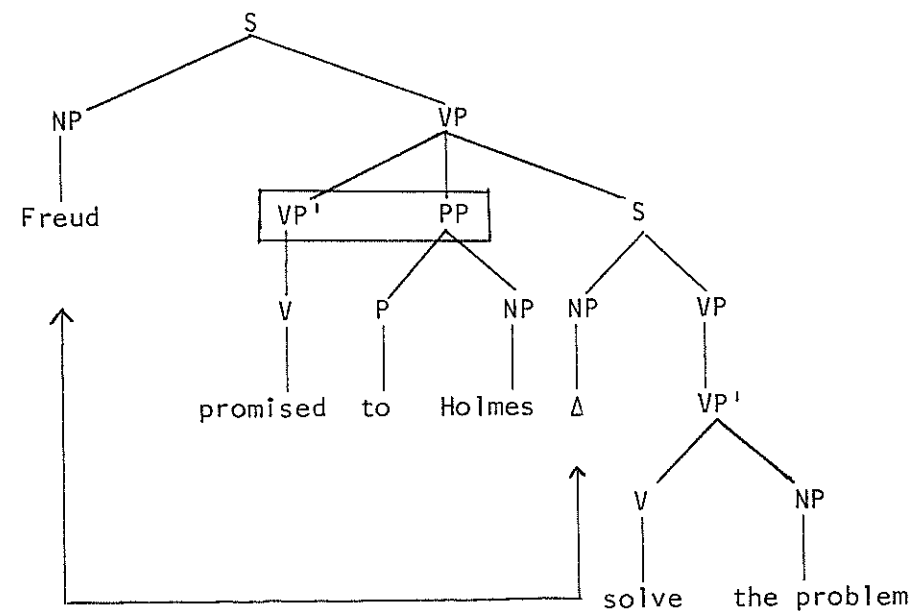
- (61) a. I promised Watson myself that Holmes would be here.
 b. ?I persuaded Watson myself that Holmes would be here.

(62) a. I promised Watson enthusiastically that Holmes would be here.

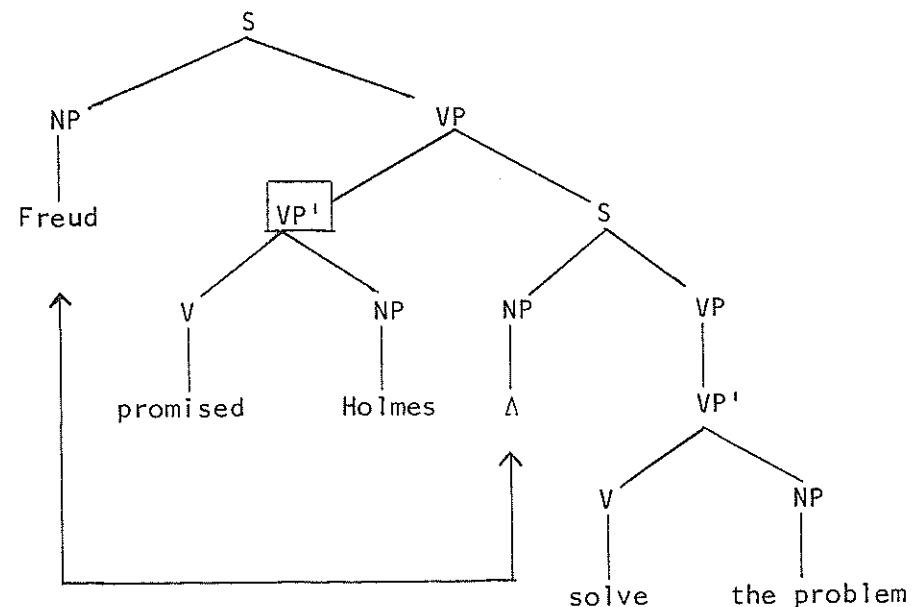
b. ?I persuaded Watson enthusiastically that Holmes would be here.

If we assume that the structures in (59) and (60) are accurate then we can also assume the accuracy of the structures in (63) and (64).

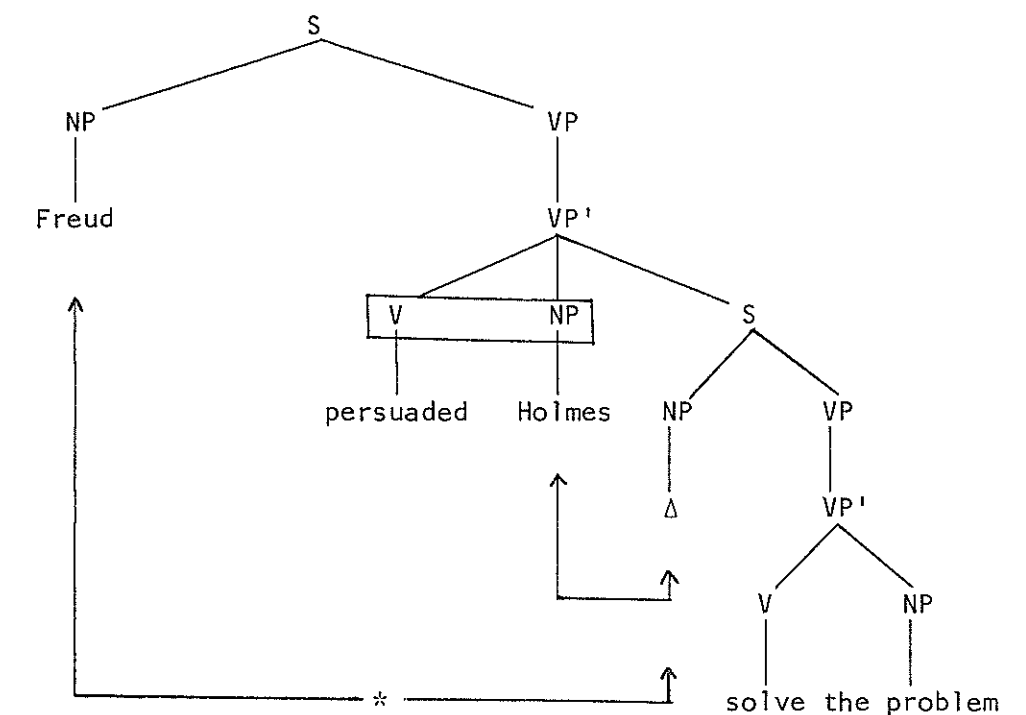
(63) a.



b.



(64)



We can now account for the control facts illustrated by the arrows in (63) and (64) by postulating a condition on rule (48) that it operate from left to right. Rule (48) (repeated here),

$$(48) \begin{bmatrix} \text{NP} \\ \alpha F \end{bmatrix} - \Delta = 1 - \begin{bmatrix} 2 \\ \alpha F \end{bmatrix}$$

is construed to mean that, other things being equal, TERM 1 IS THE LEFTMOST NP IN A STRUCTURE. This leftmost NP controls the interpretation of delta, as in (63), as long as there is no NP in the grossest analysis. Where there is such an NP, as in (64), then control from the leftmost NP is blocked. Rule (48) then applies where term 1 is the next most left NP and assigns control as indicated in (64).

We turn again now to the *beg* class of verbs in English. Reconsider example (49a), repeated here as (65).

(65) Didi {begged} Pozzo to unburden Lucky.

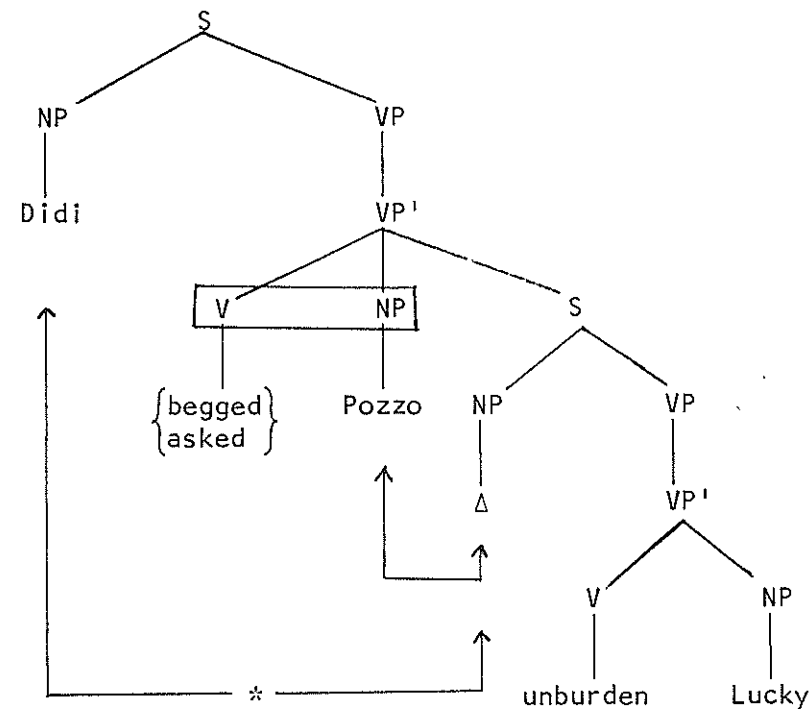
It was shown above that *beg*, like *persuade*, may exhibit object control. It was also shown that unlike *persuade* it may exert subject control (examples (49b) and (51b)) where it has no object in the matrix. *Beg* can also exhibit subject control where there is an NP object of the higher clause. As has been pointed out in the past, examples like (65) are ambiguous.²¹ The reading of (65) where *Didi* is the agent of *unburden* can be paraphrased as (66) or (67).

(66) Didi $\left\{ \begin{smallmatrix} \text{begged} \\ \text{asked} \end{smallmatrix} \right\}$ of Pozzo to unburden Lucky.

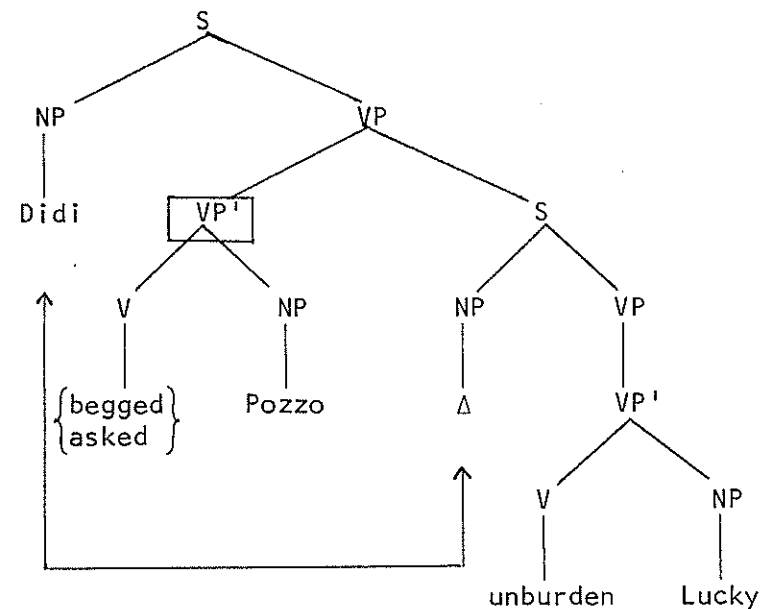
(67) Didi $\left\{ \begin{smallmatrix} \text{begged} \\ \text{asked} \end{smallmatrix} \right\}$ Pozzo $\left\{ \begin{smallmatrix} \text{to be able} \\ \text{to be allowed} \end{smallmatrix} \right\}$ to unburden Lucky.

The relevant phrase markers for the two senses of (65) and for (66) and (67) are given in (68) through (70).

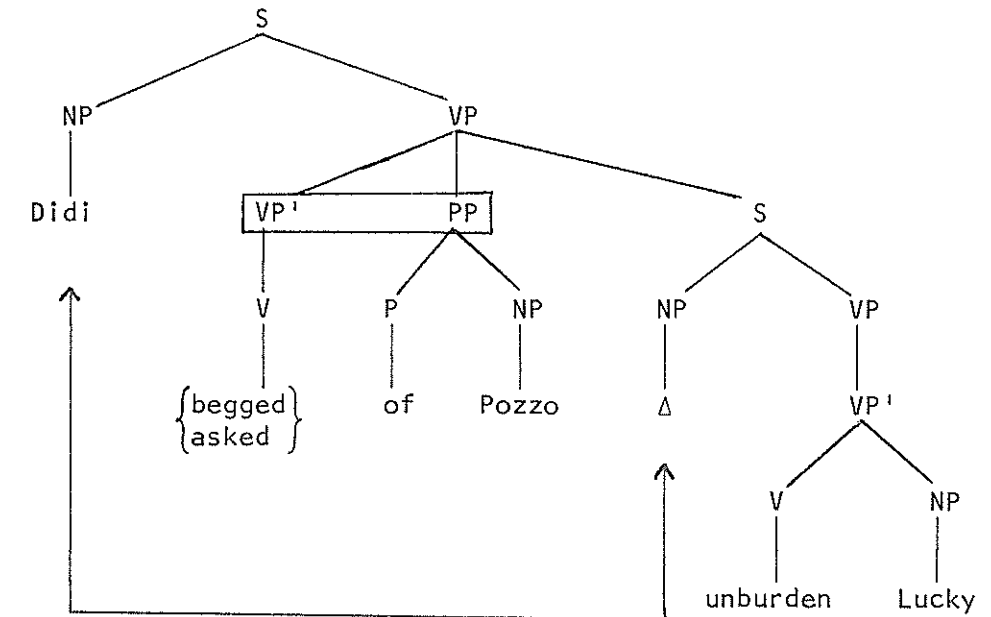
(68) a.



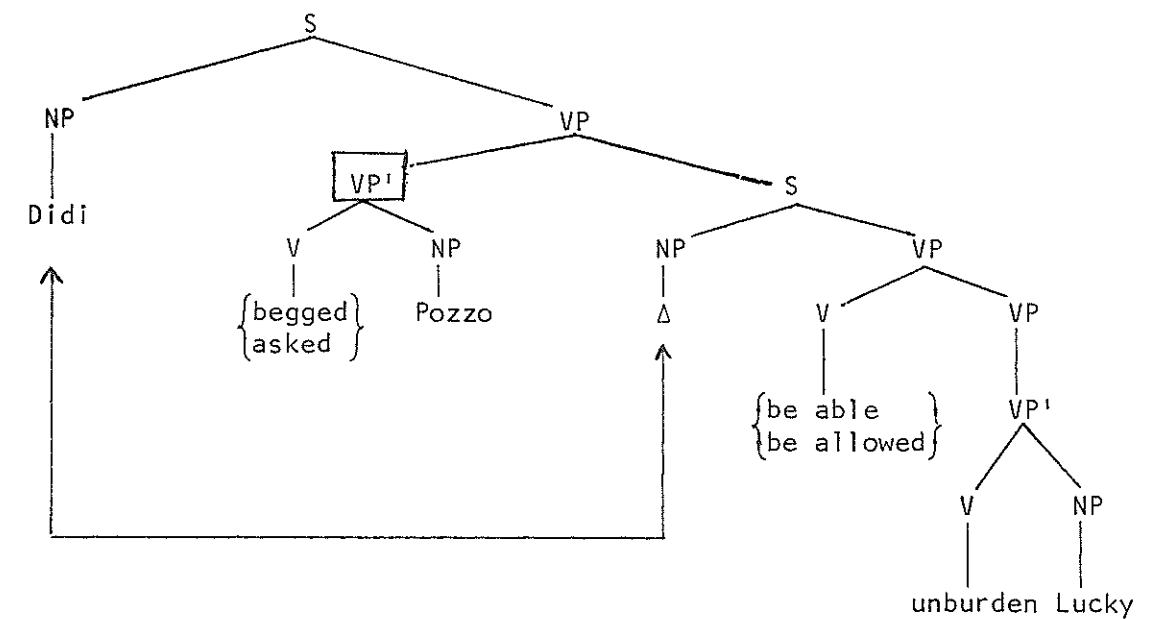
b.



(69)



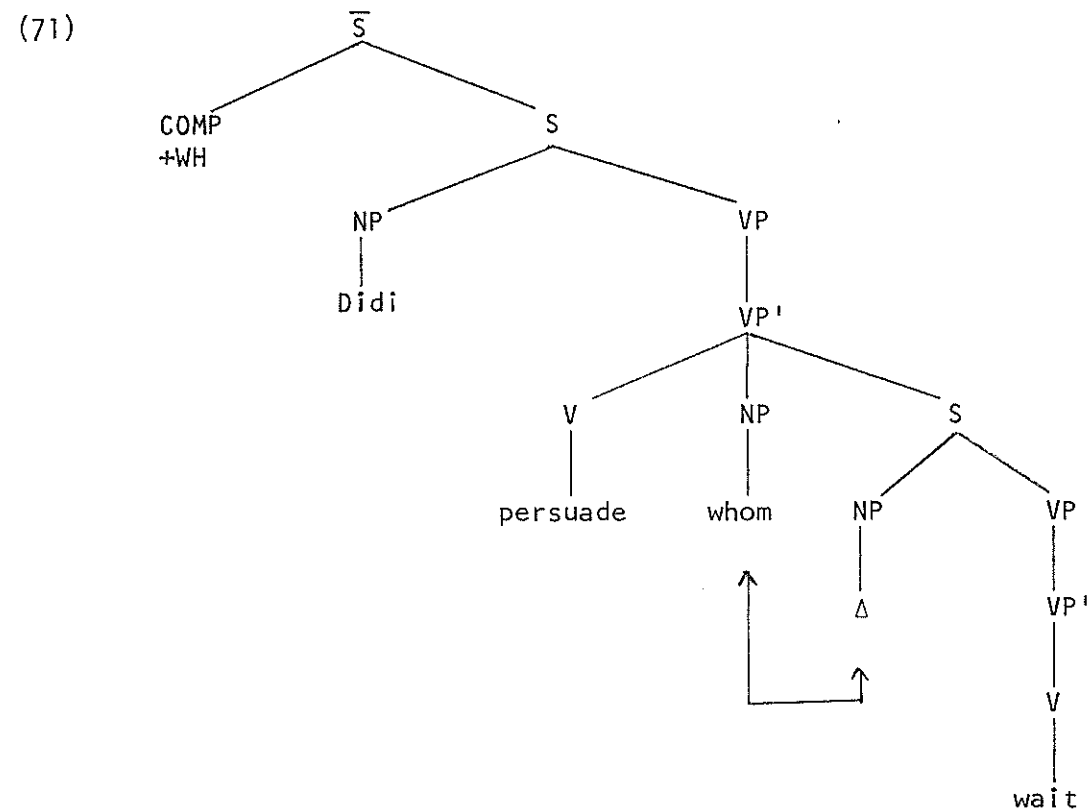
(70)



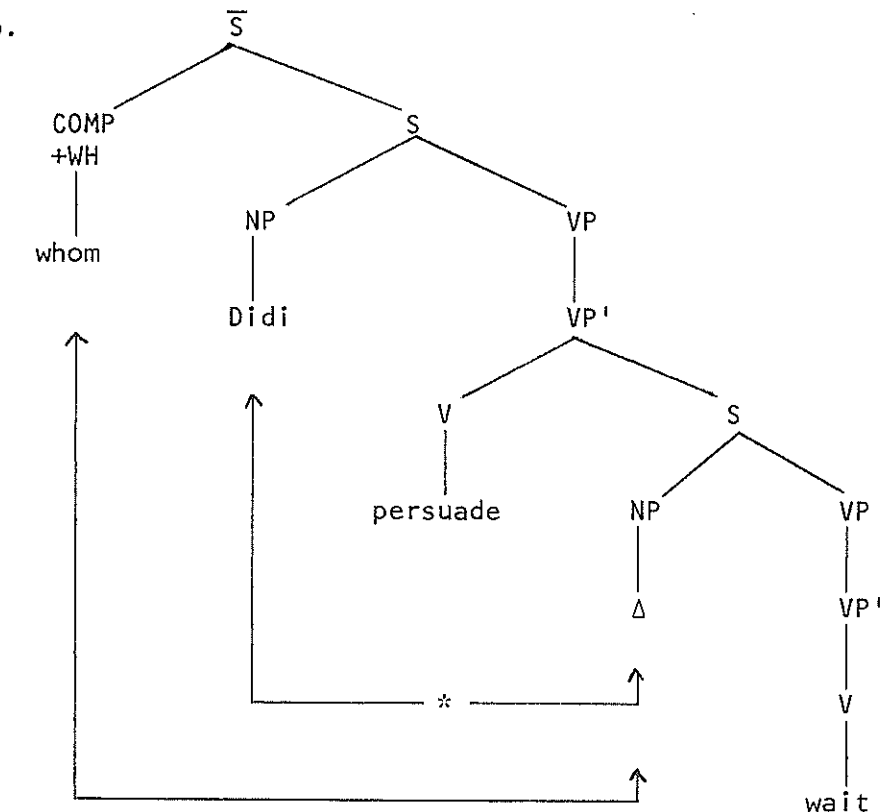
In (68b), (69), and (70) *beg* and *ask* exhibit subject control because there is no NP in the grossest analysis to block it. Where there is such an NP, (68a), these verbs exert object control. These verbs differ from *promise*, which requires subject control, in that their S complements may be either a sister to, or a daughter of, VP'. Where the NP object of *beg* or *ask* must be an accusative, as in (68a), then the VIC and rule (48) properly predict object control.

By use of the VIC it is thus possible to account in a general way for certain semantic details of particular lexical items.²² The fact that rule (48) functions from left to right, that is, from the leftmost lexically designated NP to the delta, might turn out to be true for all semantic rules of this type. It would then be predictable from the form of the rule that it functions differently from movement rules. (For further discussion of the predictable directionality of rule types see Chapter 6.) What remains unaccounted for here still is the fact that *promise* requires agent-subject control. This was illustrated in (55). *Promise* just may not undergo passivization where its S complement contains a delta subject to be interpreted.

We now turn, once more, to a *persuade* example. In (52) we saw that when, as a result of passive, there was no overt object for *persuade*, then there was subject control. Where there is no overt object as the result of WH-fronting however, the result is different.²³

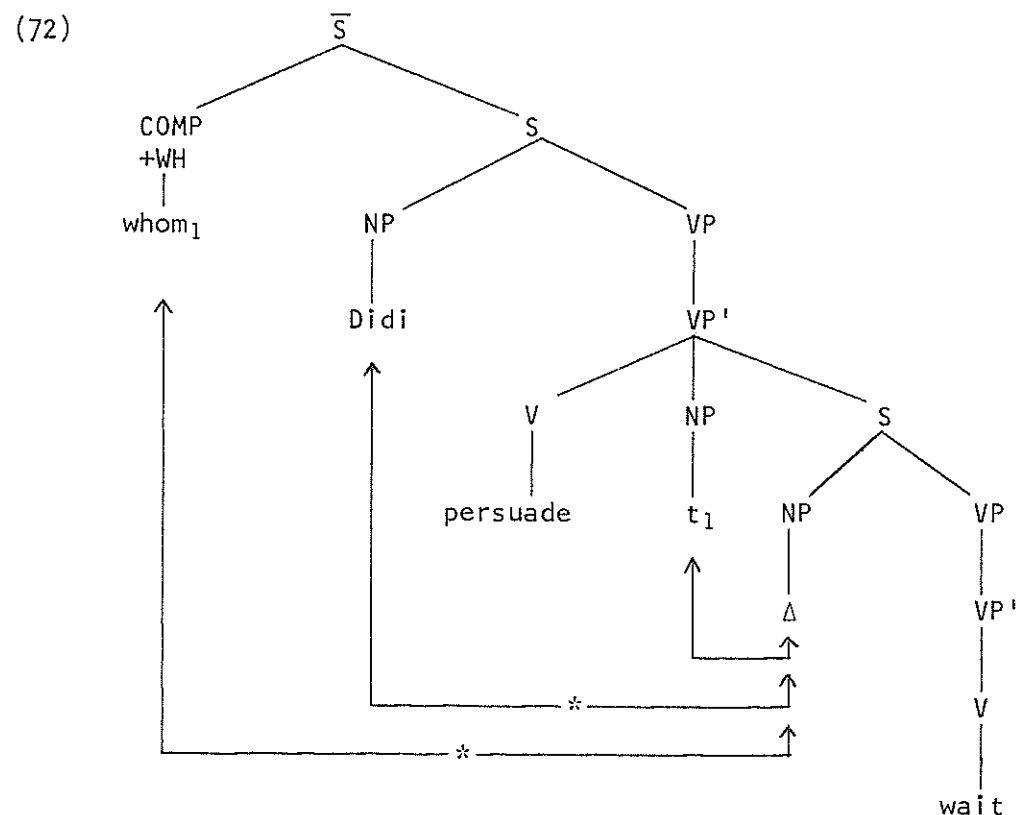


(71) b.



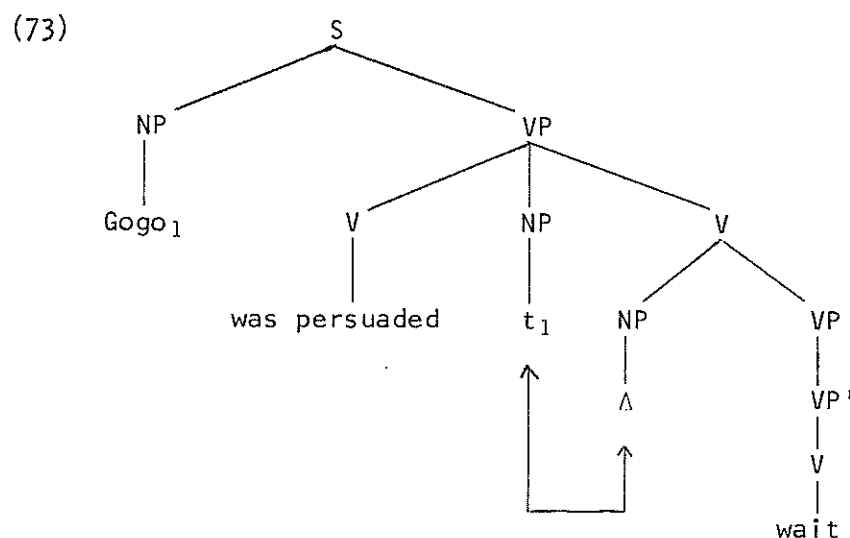
c. Whom did Didi persuade to wait?

In (71c) the controller of the subject of *wait* is *whom*, not *Didi* as would be predicted by the VIC. An account of this might well be that after WH-movement a bound trace is left in the position of the moved NP (Chomsky 1973, 1976 and references mentioned there). The result of the WH-fronting would then be (72) rather than (71b).



In (72) the NP dominating the trace would be the NP controlling the interpretation of the delta. Neither the subject NP nor the WH-term can be the controller.

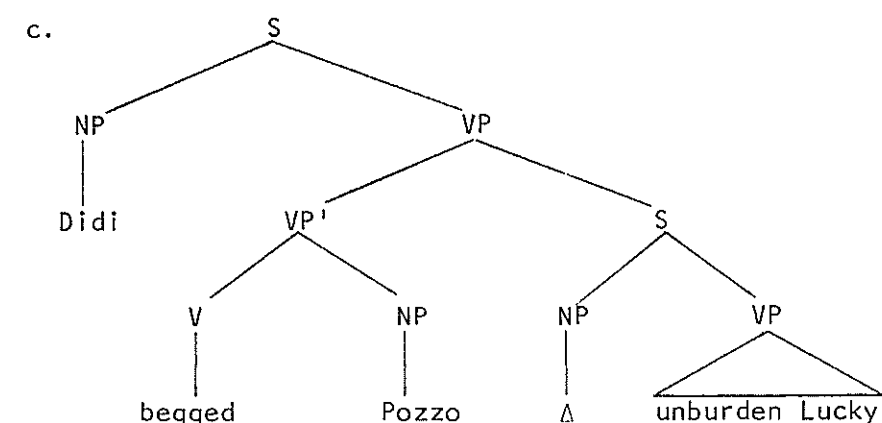
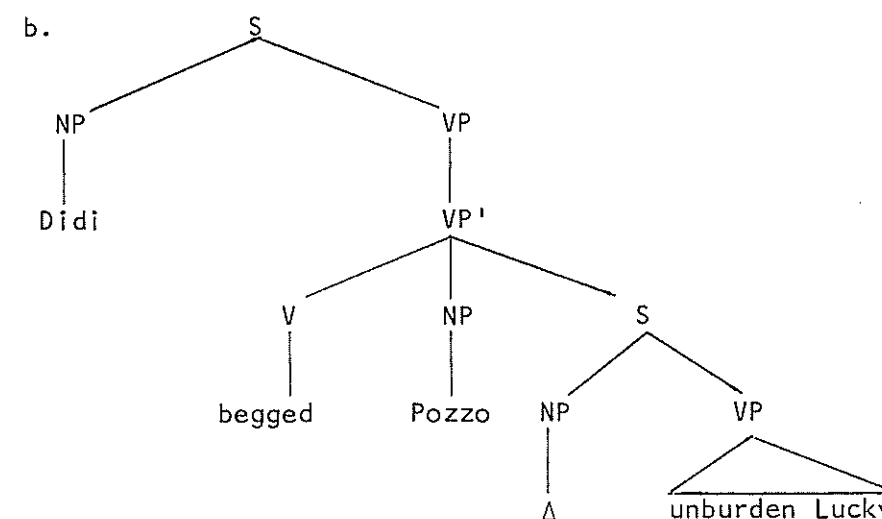
If this in fact is the correct account of (71c) then a similar account can be used for (52c). This is illustrated in (73) where the movement of the NP in passive leaves behind a bound trace.



Here *persuade* is still exhibiting object control. Although the surface result is the same, the claim illustrated in (73) is different from that illustrated by (52b) where after passive *persuade* would be allowing for subject control.

The facts become more complicated, and more interesting, if we now consider the *beg*-type examples. Consider again the ambiguity of (65) and the two deep structure sources (b) and (c), here repeated as (74).

(74) a. Didi begged Pozzo to unburden Lucky.



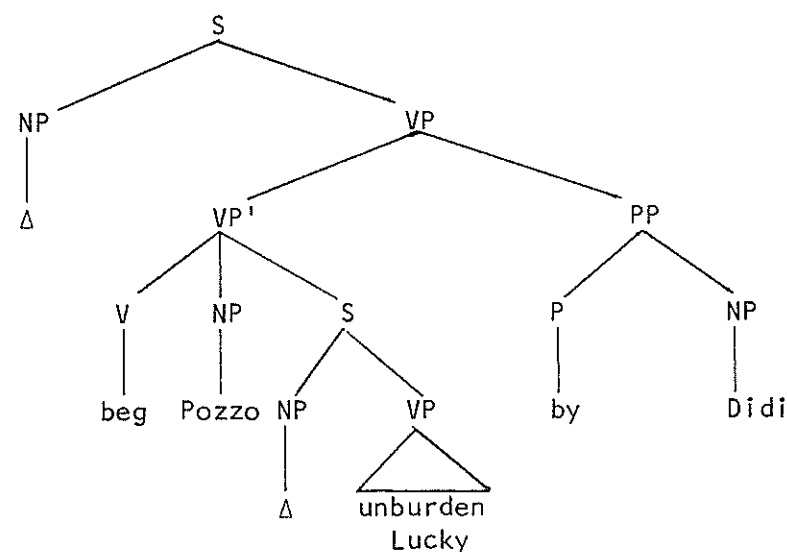
Reconsider also example (53b), here repeated as (75), which is not ambiguous.

(75) Pozzo was begged to unburden Lucky by Didi.

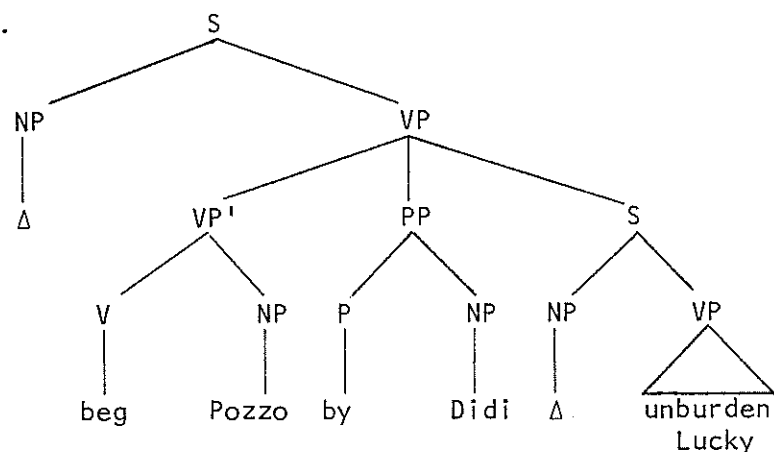
This lack of ambiguity is explained in an interesting way in this model if passive leaves behind a trace which can control interpretation. (76a) and

(76b) are the two possible base configurations for (75). (See footnote 19 with respect to the different placement of the PP's in the examples of (76). For the discussion here it is basically irrelevant.)

(76) a.

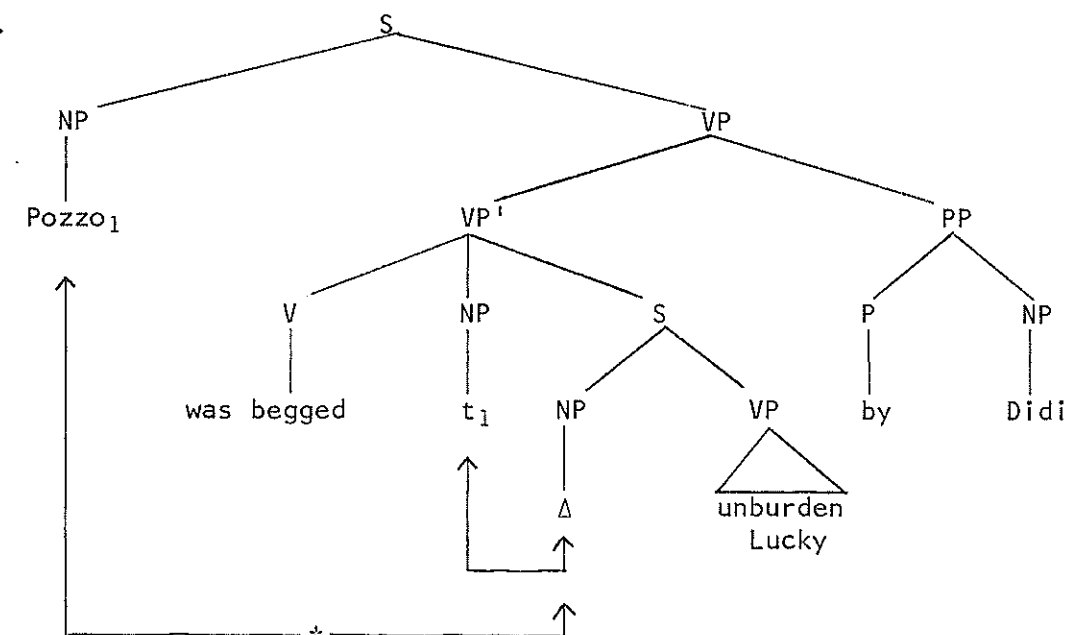


b.

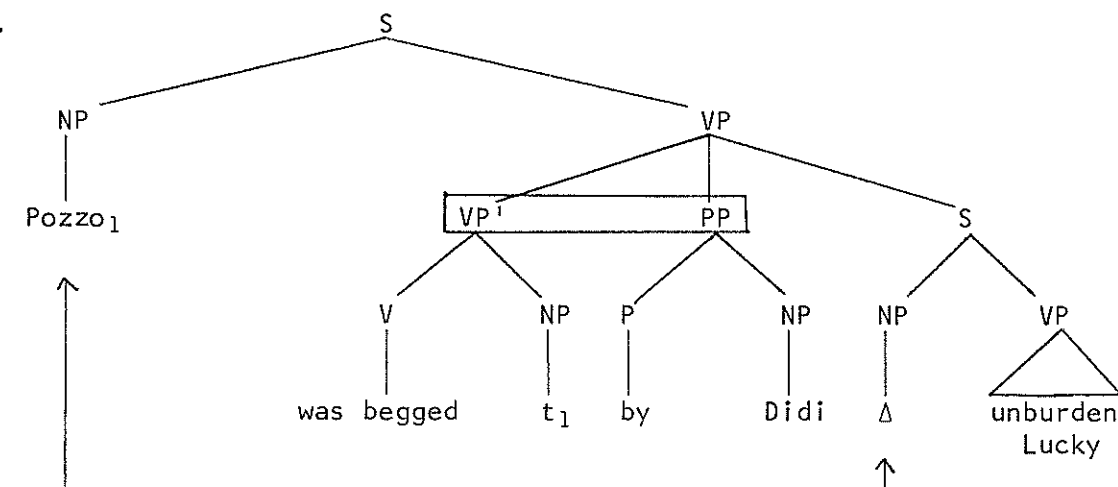


The result of passivization applying to (76a) and (76b) is (77a) and (77b), respectively.

(77) a.



b.



For (75) only the subject control reading is possible. In (77b) it is in fact the subject NP which, by (48) functioning in accord with the VIC, controls the interpretation of delta. In (77a) the VIC blocks interpretation by the subject. Control is exerted by the NP object of *was begged*. This object, however, is a trace bound by the subject so the interpretation of delta is properly coreferential to the subject. From either structure (77a) or (77b) the resulting meaning will be where *Pozzo* is the understood agent of *unburden*. Thus the unambiguous nature of (75) is accounted for.

This subsection of this chapter has successfully shown, I believe, that the VIC can provide for an interesting account of certain semantic properties. I hope to show in future work that it has broader implications for semantic theory as well.

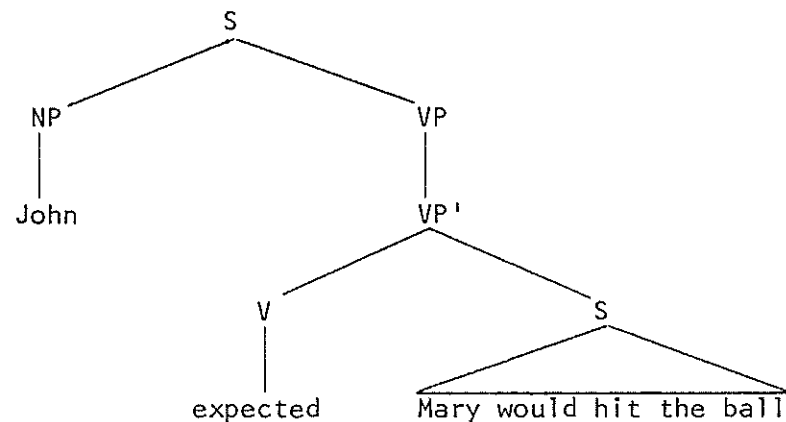
5.4. The Interaction of Passive and Raising

In any relatively detailed account of passivization it should not seem surprising that the question of the process of raising to object presents itself. An object-creating process could be used to account for sentences like (21f) (here repeated for ease of reference), as was shown in section 2.2 of Chapter 2.

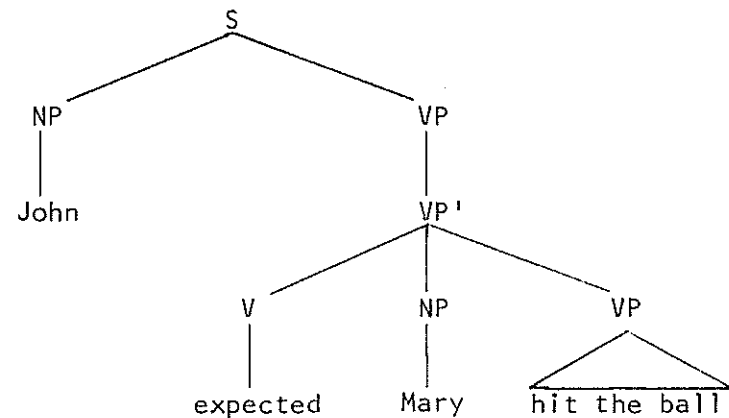
(21) f. John expected the ball to be hit by Mary.

Also, since passivization has been assumed (for instance by Postal 1974) to be a rule which is restricted to a single cyclic domain, raising to object could be considered to feed the passive rule. Above, in 5.3.1, an account of (21f) was given which did not involve a raising to object rule. It involved instead a deep structure for verbs like *expect* where the verb could occur either with a complement S or with an optional NP and a VP complement. This was shown in the examples of (29) which are repeated here.

(29) a.



b.



This type of account of the raising to object phenomenon is possible because in this model there is no reason to externally restrict the passive rule to any particular domain. Where the object position following *expect* is left empty after lexical insertion, it can be filled by the passive rule, to yield sentences like (21f). (This was discussed in detail in 5.3.1). In 5.3.2 it was shown that the passive rule predictably applies in lower domains before higher ones (where it should do so) and in this sense feeds itself (just as raising to object could be considered to feed passive).

What remains to be explained here is the exceptional nature of the objects of verbs like *expect* with respect to the raising process discussed in depth in Chapter 3. Whether the NP position following *expect* is filled by lexical insertion, or by the passive rule, that NP generally cannot be raised. As was pointed out in section 3.5. (Chapter 3) Berman (1973) suggests that the correct account of this fact is that objects created by the raising to object process just can never undergo *tough*-movement. For many obvious reasons this is not a possible account in the model presented here. There is, however, an alternative account.

First consider the following results of raising:

(78) a. *John is impossible to expect to understand that book.

b. *Smith was easy for Jones to expect to recover.

c. *Mary was easy for me to want to hit the ball.

d. *The ball was easy for me to want to be hit by Mary.

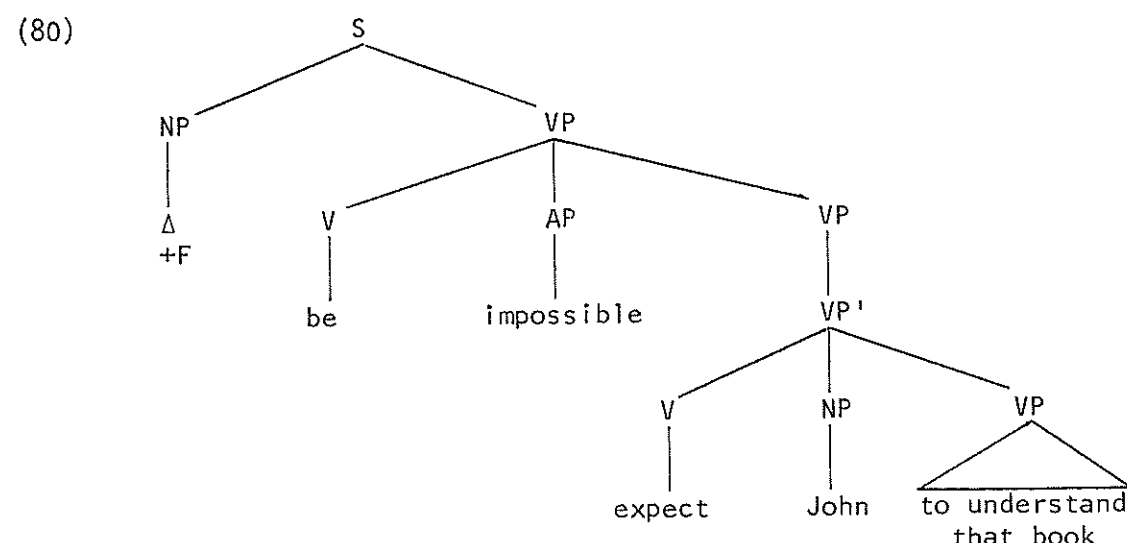
e. ?John is difficult to believe to have made such a mistake.

f. Smith was easy for Jones to believe to have made such a mistake.

g. Smith was easy for me to believe to be a fool.

Sentences (78f) and (78g) are clearly more acceptable than (78a) through (78e). Yet the structural input to raising for all of these examples is basically the same. The structural description is met in each case. The raising rule, repeated here as (79) should apply to a phrase marker such as (80).

$$(79) \Delta_{+F} - v_{\max}^{[NP]} = 2 - \emptyset$$



The prediction is that all the examples of (78) should be well-formed.

What I suggest here is that in fact raising does apply to produce all the examples of (78). The unacceptable status of some of the results of this rule is then due to the fact that there is often a basic semantic (of possibly pragmatic) conflict in the meanings of the adjectives which govern raising and the verbs which have been said to govern raising to object.²⁴ The acceptability of (78f) and (78g) is accounted for by the fact that the state of affairs of which *easy* or *hard* is predicated is in fact something that it is possible to find easy or hard to do. *Expect* and *want* just do not represent activities which it is either hard or easy to do, hence, the unacceptability of (a) through (d). By contrast, "believing" is something which might be hard to accomplish. There is nothing semantically odd about saying that someone is easier or harder to believe to be a fool, say, than someone else is.²⁵

This type of account of the facts of (78) would lead us to assume that even where object raising does not take place the same semantic conflict would arise where these adjectives and verbs co-occur in a string. But this is not true, as is shown in (81) (although to my intuitions (81a) through (81d) are in fact semantically rather odd).

- (81) a. It is impossible to expect John to understand that book.
 b. It was easy for Jones to expect Smith to recover.
 c. It was easy for me to want Mary to hit the ball.
 d. It was easy for me to want the ball to be hit by Mary.

- (81) e. It is difficult to believe John to have made such a mistake.
 f. It was easy for Jones to believe Smith to be the winner.
 g. It was easy for me to believe Smith to be a fool.

What is to be accounted for here is why all the sentences of (81) are well-formed for most English speakers if the unacceptability of some of the examples of (78) is due to a semantic conflict. There is in fact a plausible explanation for this difference between (78) and (81).

Noriko McCawley (McCawley 1977) argues that the speaker of any sentence is not just the agent or the performer of the speech act, but that the speaker is also an "experiencer". Additionally, McCawley claims, when a speaker uses a "tough-moved" sentence that speaker is saying that she/he has some sort of "evaluative judgement about [a] certain property of the subject NP" (p. 10). In other words, object raising (or *tough-movement* as McCawley refers to it) can only result in a well-formed string where the speaker has sufficient knowledge about, or experience with, the subject NP to make an evaluative judgement about that subject NP with respect to the action or state of affairs expressed by the rest of the sentence.

McCawley supports her claims in a number of ways. She points out, for instance, that there are two distinct meanings for the lexical item *impossible*: one which implies "extremely difficult" and the other which denotes the logical impossibility. Her examples to illustrate this are given here in (82) ((19a, b) in the original).

- (82) a. John is impossible₁ to live with.
 (John is extremely difficult to live with.)
 b. 17 is impossible to factor.
 (* > 17 is extremely difficult to factor.)

In general it is only the *impossible* which implies "extremely difficult" which can occur with a raised subject. McCawley illustrates this by the example in (83) ((4) in the original).

- (83) { a. It has been impossible to live with my husband.
 b. *My husband has been impossible to live with. }

He has been in prison for the last five years, you know.

Example (83b) is ill-formed because where the following sentence indicates that the *impossible* is the one denoting logically not-possible then there is no evaluative judgement to be made by the speaker about the subject NP.

With respect to example (82b), McCawley points out that it would only be a possible sentence for someone who knows how to factor and would not be a statement about the mathematical ability of the speaker. What McCawley leaves unclear, but what I believe to be the case, is that where the speaker's evaluative judgement coincides with (or confirms) the fact of logical impossibility, then raising can occur with this second meaning of *impossible*.

For further evidence that raising requires the experiencer (=speaker) to be making some judgement about the subject NP McCawley gives the following examples ((28) in the original).

- (84) John is impossible to talk to, for
- a. *he is in the bathroom right now.
 - b. *he is in a coma.
 - c. he is high on hashish.

She then says (p. 9):

Only (c) is acceptable. Notice that (a) and (b) are not the [sic] statements about [a] certain personal property or characterization of John. But (c) is. The speaker is saying that he finds John hardly communicable: you may talk to John; the problem is that he may not understand you or respond to you.

Returning now to the examples of (78), the claim is that (f) and (g) are well-formed because it is possible for the speaker to judge the subject NP, *Smith*, based on her/his experience of *Smith* with regard to the action or state of affairs of the rest of the sentence. It would be possible for it to be easy to judge *Smith* as "believably a fool". (78a) through (78d), on the other hand, present semantic contexts where an evaluative judgement on the part of the speaker is impossible. For instance, with respect to (78a), no amount of experience with *John* would make it possible for the speaker to judge whether an "expectation" is "impossible". "Expectations" are just not judged in this way.

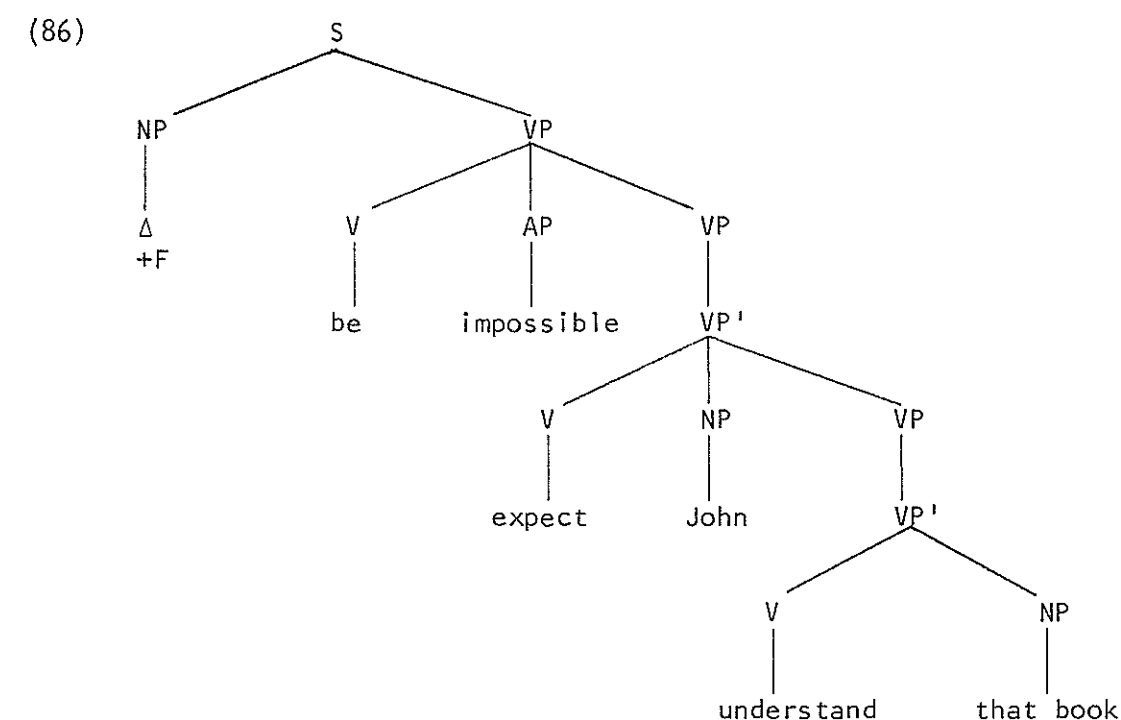
Now we return to the examples of (81). I suggest that these are all well-formed because since the subject is a semantically empty *it*, there is no requirement on how the speaker has experienced the subject NP. The sentences may represent true or false states of affairs (e.g. maybe it is false that the "expectation" is "impossible" in (81a) since any expectation is in fact possible). What is important is that there is no requirement of "evaluative judgement" on the part of the speaker with respect to the subject NP because the subject NP in these cases is just a grammatical formative. It is irrelevant, therefore, that an expectation, for instance, should not be judged as hard or easy. (This observation about examples like (81) is my own and is not meant to indicate that McCawley would necessarily agree.)

Also in section 3.5 of Chapter 3 examples like the following were pointed out.

- (85) a. *That book is impossible to expect John to understand.
b. *Such a mistake is difficult to believe John to have made.

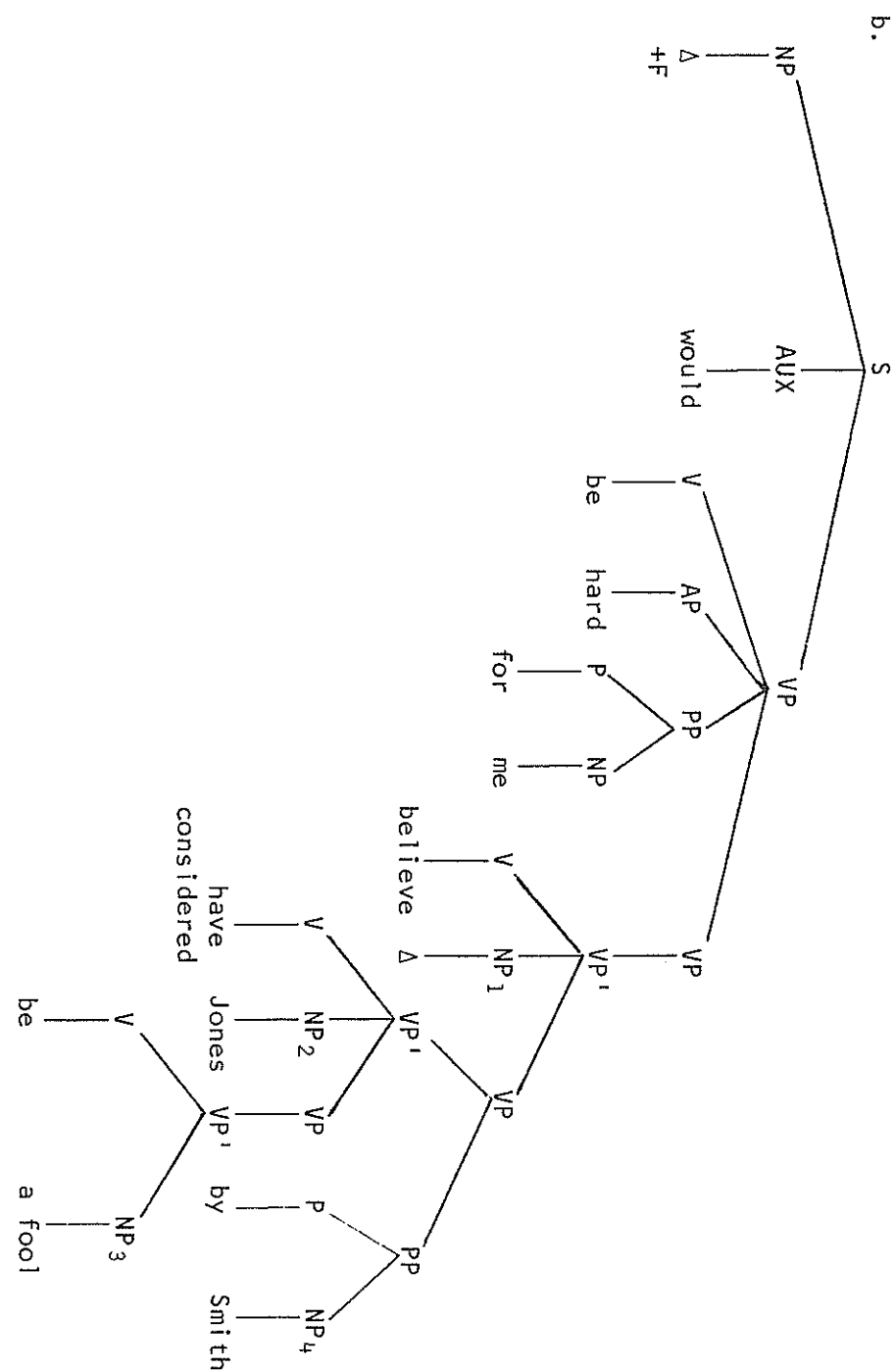
It was shown there that if the NP's immediately following verbs like *expect* were actually subjects of embedded clauses then the VIC would properly predict the ungrammaticality of (85a) and (85b). The objects of these embedded clauses would be prevented from raising because the subject NP would lie in the grossest analysis.

Now if examples like (85) were derived their unacceptability could be accounted for by the semantic considerations discussed above. However they will never be generated. Since the VP complement of verbs like *expect* occurs inside the VP' along with the direct object NP, this object NP blocks the application of raising. This is illustrated in (86).

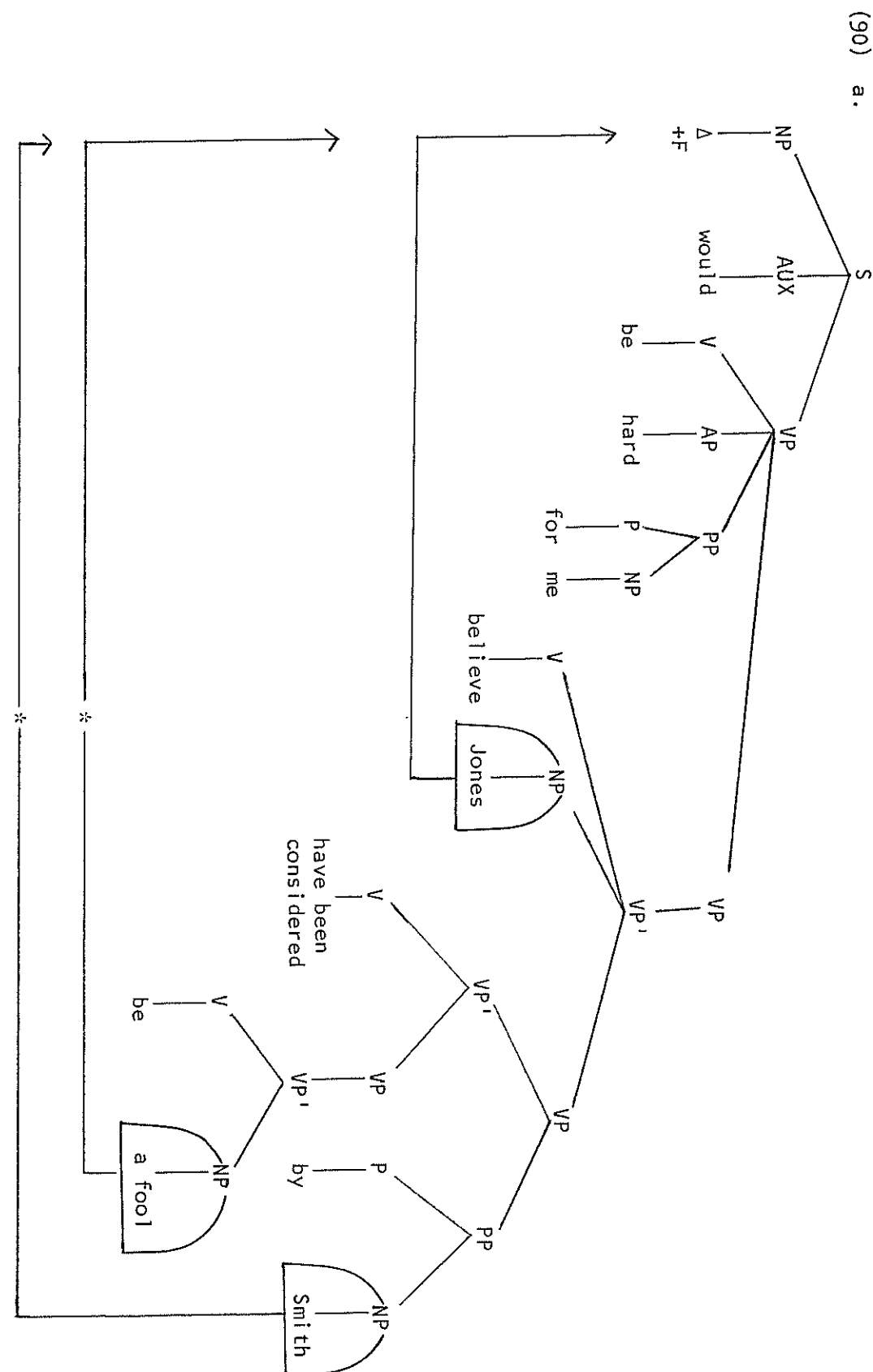


For a verb like *believe*, where no necessary semantic conflict is proposed above, a more detailed look at the interaction between object raising and passivization, rule (8), is warranted. Consider sentence (87a) and its deep structure (87b).

- (89) a. It would be hard for me to believe Jones to have been considered (to be) a fool
by Smith.



Where passive applies first to (89b) the result is (90a). Raising can then properly apply, as is illustrated, to derive (90b). (90c) and (90d) are properly blocked (the NP *Jones* is in the grossest analysis).



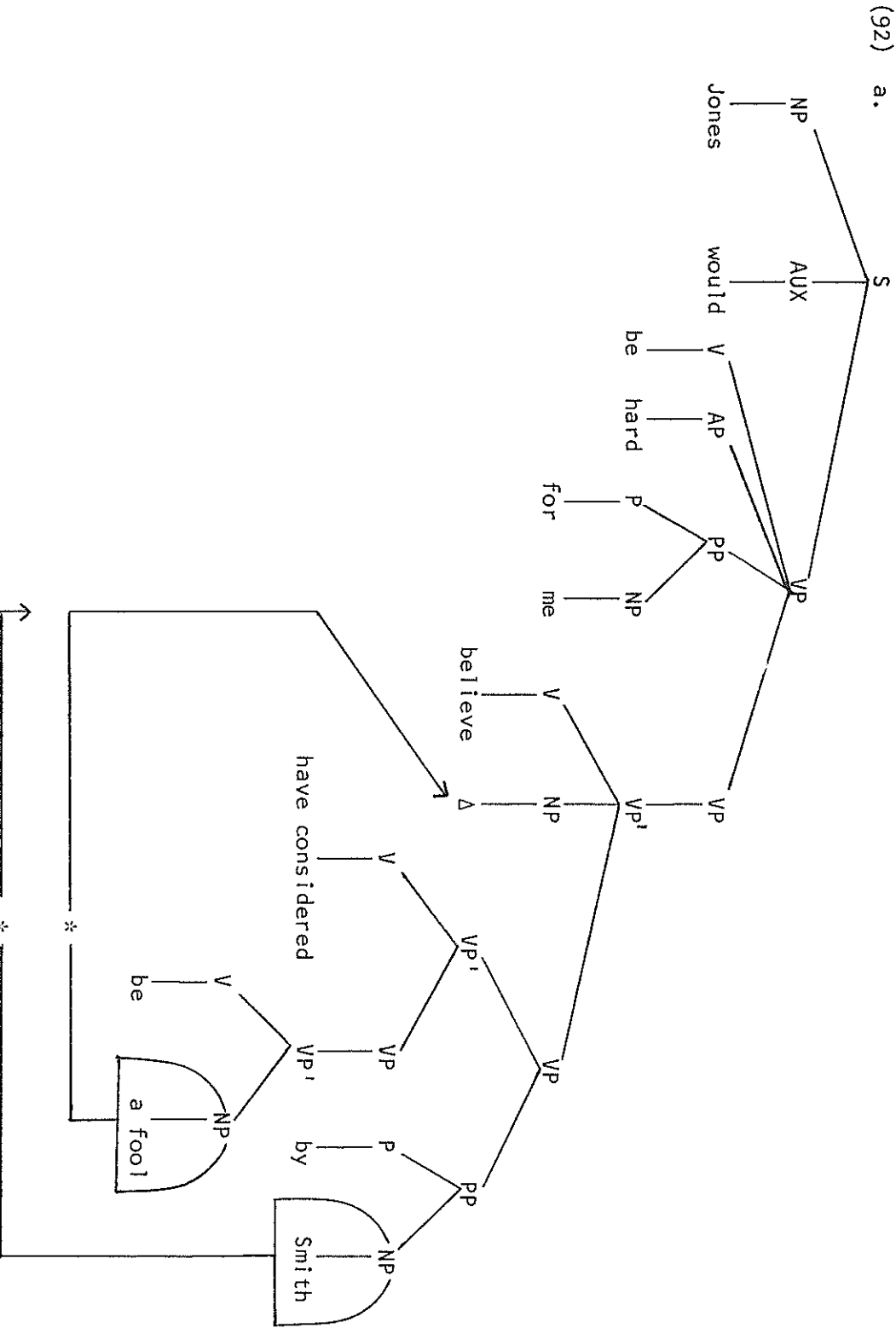
- (90) b. Jones would be hard for me to believe to have been considered (to be) a fool by Smith.
- c. *A fool would be hard for me to believe Jones to have been considered (to be) by Smith.
- d. *Smith would be hard for me to believe Jones to have been considered (to be) a fool by.

If, however, raising applies first in (89b) there are four NP's which would meet the structural description (they are numbers 1 through 4). NP₁ cannot be moved, by principle (49) of Chapter 3 which prohibits phonologically empty nodes from being moved by transformations. Nor may NP₂, NP₃, or NP₄ be moved as the ungrammaticality of (91a) through (91c) illustrates.

- (91) a. *Jones would be hard for me to believe to have considered (to be) a fool by Smith.
- b. *A fool would be hard for me to believe to have considered Jones (to be) by Smith.
- c. *Smith would be hard for me to believe to have considered Jones (to be) a fool by.

Since it is only an NP in the grossest analysis which can prevent raising from applying, we must examine the ungrammaticality of the examples of (91). If they are to be blocked by the VIC it might seem that it is NP₁ occurring in the grossest analysis which is the reason. NP₁, however, is empty. It has been assumed throughout this work that empty nodes "don't count" for the VIC.

It might be the case that the examples of (91) can be accounted for in some other way. For instance, after raising, if no other transformation applies then the occurrence of the delta in the surface string (NP₁ of (89b)) would cause it to be marked as ill-formed. This account will suffice for (91a) because after the application of raising, passive (the only other relevant transformation here) may not apply. This is illustrated in (92).



- (92) b. *Jones would be hard for me to believe a fool to have been considered (to be) by Smith.
- c. *Jones would be hard for me to believe Smith to have been considered (to be) a fool by.

The movements indicated in (92a) are blocked by the VIC. No NP can fill the delta, so the ungrammaticality of (91a) might be accounted for in this way.

The ungrammaticality of (91b) is easily explained because where NP₃ of (89b) is raised not only is NP₁ in the grossest analysis, but so is NP₂. The lexically designated NP₂ blocks the rule and the status of (91b) is explained.

For (91c), where NP₄ of (89b) has been raised, the explanation is not so simple. If no further rule applies, then the delta causes ill-formedness as with (91a). But in this case, after raising applies to NP₄ then passive may apply to NP₂. This is illustrated in (93).

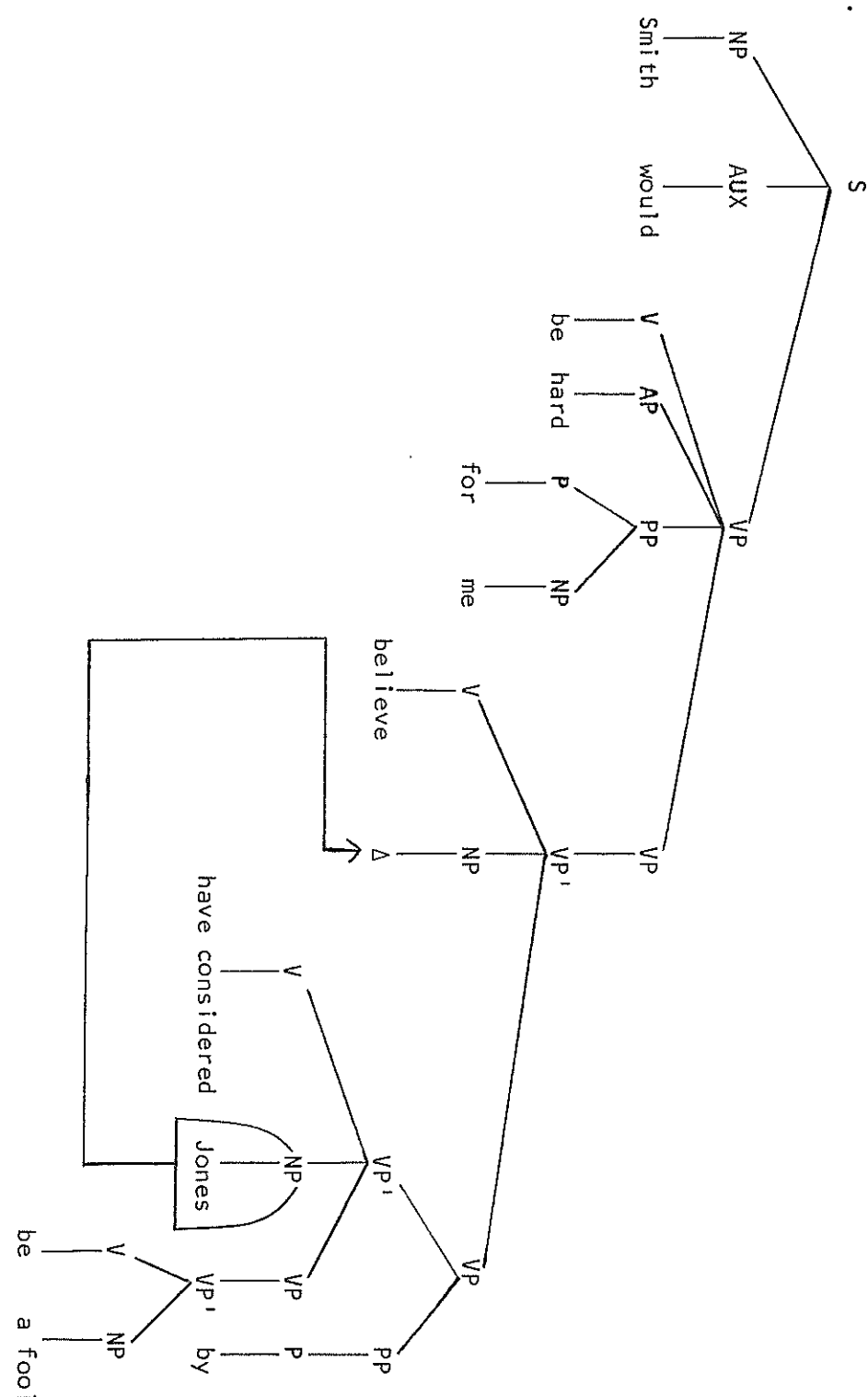
The ungrammaticality of example (93b) must be accounted for and this cannot be done, as yet, either by the VIC or by resorting to a delta in surface structure. There are three possible alternatives to explain the ungrammaticality of this example. Each of the three has broad theoretical implications.

First, it might be that passive must be extrinsically ordered to precede raising. If this were done the examples of (91), (92), and (93) could never be generated. In the framework of this thesis this alternative is particularly undesirable because in no other case must transformational operations which move terms be extrinsically ordered with respect to one another. The only type of ordering which has been necessary, thus far, has been for classes of operations. For example, it seems that any rule which only induces a morphological change (e.g. *that*-insertion, or subject-verb agreement) must follow all the movement transformations. Ideally there will be no need for extrinsic ordering of movement rules.

The second alternative would be where the phonologically empty NP in (89b), NP₁ would "count" for the VIC. This would mean that raising could not apply in any way to (89b). NP₁ itself cannot move because it is phonologically empty, but it could serve to block movement of NP₂, NP₃, or NP₄ because in each case it lies in the grossest analysis. With this second alternative passive would in fact have to apply before raising (as shown in (89) and (90)) but it would do so automatically. There would be no need for an ordering statement. This account of the interaction between raising and passive would cast doubt on the account of WH-fronting presented in Chapter 4. There it was important that empty nodes, in particular empty COMP nodes, NOT prevent the application of the WH-fronting rules.

The third alternative which I now suggest is one which seems to have the most interesting implications. The ungrammaticality of (93b) is to be accounted for by trace binding, like both movement rules and the rule of control already discussed (5.3.3), being subject to the VIC. If we again assume that NP's which move leave behind traces, then (93b) is generated but it can never be given an interpretation because there is an NP in the grossest analysis blocking the binding of the trace. This is shown in (94) which is the structural result of both raising and passive having applied to the relevant NP's of (89b).

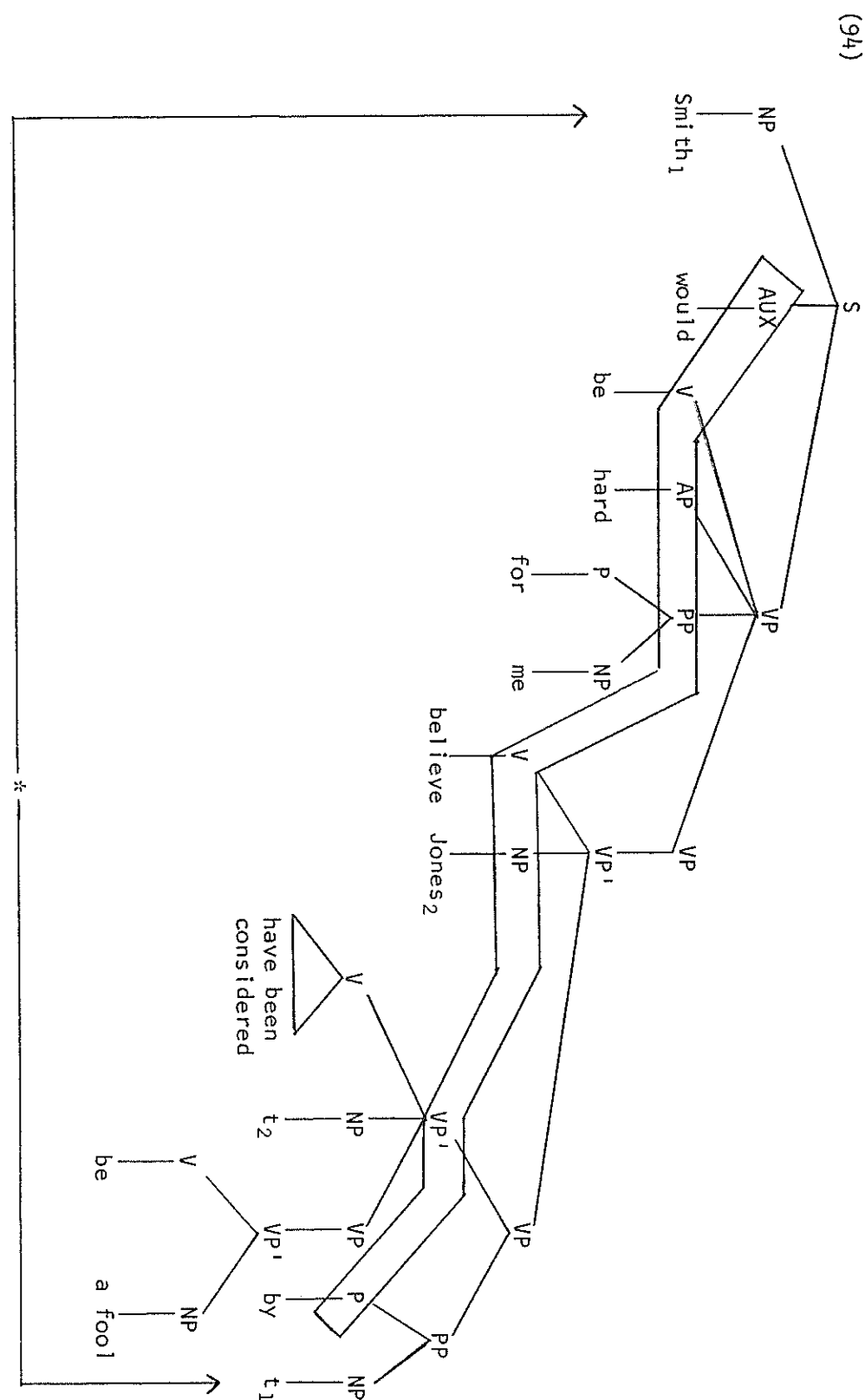
(93) a.



b. *Smith would be hard for me to believe Jones to have been considered (to be) a fool by.

The trace t_1 must be bound by NP_1 . This binding is blocked in (94) by the occurrence of the NP *Jones* in the grossest analysis. Where a trace cannot be properly bound the result is ungrammatical.

This last alternative for an account of the ungrammaticality of (93b) must, at this stage, be considered a tentative suggestion. I think, however, that future research into the use of the VIC will show this to be the correct account.



FOOTNOTES TO CHAPTER 5

¹This is not meant to indicate that there are no instances of "complex verbs". See, for instance, examples (10) and (11) of Chapter 2. Here in (1) and (3) the constituency of the PP must be maintained, however. Compare cases of P NP which are NOT constituents, e.g. as would arise after particle movement.

- (i) They searched John out (in order to congratulate him).
- (ii) They searched out John.
- (iii) *Out whom did they search?
- (iv) *It was out John that they searched.

I was reminded of the ungrammaticality of examples like (iii) in this regard by Joel Rotenberg.

²For discussion of restructuring processes see Wexler and Culicover (in preparation).

³This notion of reanalysis was suggested by Chomsky, class lectures, 1976.

⁴Peter Culicover (personal communication) suggests that restructuring applies only locally (may not apply over a variable), may move only terms at constituent boundaries, and may attach terms only at boundaries.

⁵See Chomsky (1957).

⁶This analysis obviously depends on a restructuring taking place before passivization. If (6) applied to a structure like (5) the VIC would predict that both (1b) and (1c) should be grammatical. Restructuring, however, would have to be optional or the facts of (3) could not be accounted for. Perhaps a "reanalysis" account, although formally less precise, is to be preferred here. This would mean that the structure underlying (1a) would be only that given in (5) but that for certain combinations of terminal elements (here particular verbs followed by particular prepositions) a transformation could analyze them as a single unit, without actually altering the phrase marker. If neither the restructuring nor the reanalysis account can be properly worked out, the facts of (1) and (3) will be handled as they were in earlier versions of this work. In Wilkins 1976 there is a suggested addition to the VIC which states that "where A or B is a lexical category then X does not contain any phrasal category."

⁷The *by* mentioned in this rule can of course be only the agentive preposition. Perhaps the rule could more accurately be written with feature specifications such as

[PREP
+agentive
.....]

rather than the actual morpheme *by*. The rule will not be amended here because it is discarded below.

⁸*Untouched* in (12) is probably more accurately considered an AP. Compare: *Human hands* (*un)touched this product ~ This product was (un-) touched by human hands. For discussion, see Wasow (1976).

⁹*The king of England* is a full NP, in the sense of Nⁿ, since it can occur here with the full range of specification, modification, and complementation. *The king* in (15a) and (15b) is properly not indicated as an NP. It is part of the NP which contains it and the PP complement.

¹⁰For a discussion of a "relativized" version of the A-over-A condition, particularly with respect to deletion rules which are not discussed in this thesis, see Bresnan (1976a).

¹¹The modal *would* is used in the embedded clauses here for semantic well-formedness. The same is true for (27).

¹²For an interesting recent discussion of the structure of the verb phrase and deep structures similar to that presented in (29b), see Brame (1976).

¹³We are ignoring for the moment, but will return in 5.3.3, to the issue of the delta subject of *love* in an example like (34).

¹⁴Since passivization is a structure preserving rule the delta which is replaced by the NP is also predictably an NP. Therefore the NP's in the grossest analysis in (31) would also be sufficient to block application of the rule, if this aspect of the formalism were more fully worked out. We will return to some speculations on this issue in Chapter 6.

¹⁵Evidence that *wanted* does not passivize is provided by the ungrammaticality of (i) and the lack of synonymy between (iia) and (iib).

- (i) *The job was wanted by the student.
- (ii) a. That man is wanted by the police.
- b. The police want that man.

¹⁶Notice that for (31) it doesn't matter in which domain the rule applies first.

¹⁷Although this sentence sounds a little odd without a context I think it is fully grammatical.

¹⁸With respect to the discussion of control properties, which follows, I would like to thank Bob Freidin both for the suggestion to pursue this line of investigation and for helpful comments and criticism.

¹⁹It has already been argued that the S complement of *persuade* is in VP', see Chapter 3, section 3.4. Agent *by*-phrases are properly attached outside VP', see examples (10), (14), (15), and (17) through (19) of this chapter, for instance. The fact that on the surface the agentive prepositional phrase usually precedes an S or VP complement is evidence that there is a late rule which reorders constituents of the VP so that the usual order is that shorter items precede longer ones.

²⁰In dialects where (56a) is not acceptable (and, reportedly, there are such dialects) then it of course provides no evidence.

²¹In dialects where (65) cannot be ambiguous there is nothing here to be accounted for.

²²Bob Freidin has pointed out to me that there is a great deal of overlap in the predictions made by the VIC and those made by the use of the notion C-commands (see, in particular, Reinhart (1976)). For movement rules, the relevant node in the grossest analysis which blocks the movement is often a C-commanding node of the same category as the constituent to be moved. For the rule of construal discussed here, if all the underlying structures presented are presumed to be accurate, then it seems that control of the interpretation of delta can only be exerted by an NP which C-commands delta. In (63a) and (63b), for instance, the NP *Holmes* does not C-command delta and no control from it is possible. The control fact in (64) can be accounted for by saying that it must be THE CLOSEST C-COMMANDING NP WHICH CONTROLS THE INTERPRETATION OF DELTA. If the underlying structures for *persuade* and *promise* (also for the two cases of *beg*) are not recognized to be different then this elegant generalization cannot be made.

²³This fact was called to my attention by Paul Schachter.

²⁴This approach to accounting for facts like those of (78) was suggested to me by Joan Bresnan.

²⁵The difference in grammaticality between (78e) and (78f) should be accounted for, I think, by the fact that a "state of affairs" is something which is more readily hard or easy to believe in than is an action.

²⁶On the optionality of *to be* see footnote 11 of Chapter 3.

CHAPTER 6

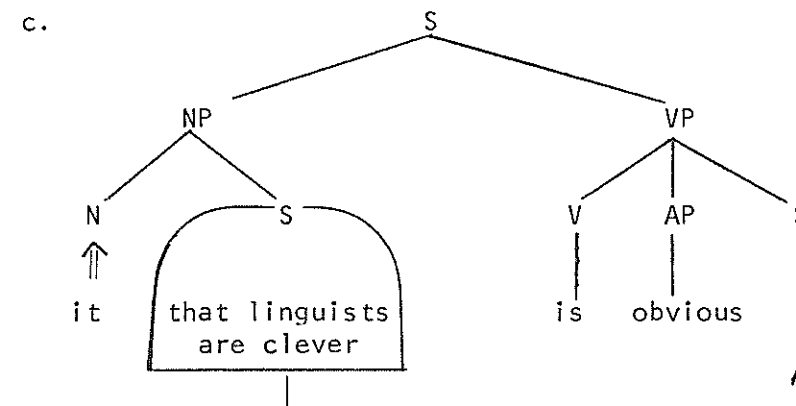
RIGHTWARD VERSUS LEFTWARD AND LOCAL VERSUS NON-LOCAL MOVEMENTS: SOME SPECULATIONS

6.1. Extraposition

In the preceding chapters we have discussed only leftward movements, aside from the brief consideration of passivization in Chapter 2 which included an agent postposing process. In this chapter we will look at two rightward movements, extraposition of S and complex NP shift, and consider a proposal for a basic dichotomy between rightward and leftward processes.

As was pointed out briefly in Chapter 3, extraposition is the rule which relates the (a) and (b) examples of (1) via the process illustrated in (1c).

- (1) a. That linguists are clever is obvious.
b. It is obvious that linguists are clever.



The rule of extraposition, as formulated by Emonds (1976:122), is:

$$(2) X - \text{NP}[\Delta - S] - Y - \text{S}[\Delta] - Z =$$

$$1 - it - \emptyset - 4 - 3 - 6$$

What we will do here, as we have done with other rules, is consider what can occur in the internal variable, Y of (2). Before we do that, however, we must consider Emonds' use of brackets in this rule. As was pointed out in Chapter 3, in a model where variable material is predicted, and therefore never written into any rule, it is not possible to stipulate, in a structural description, the notion of exhaustive domination (see section

3.1.2). Therefore, the intended use of the brackets in (2), so that they would be relevant for the structure presented in (1c), is not possible in the formalism used in this thesis.

For now the rule will be stated as:

$$(3) \quad N[S] - \Delta \Rightarrow \emptyset - 1$$

where certain details remain to be worked out. This rule is interpreted to mean that an S is extracted from NP, and moved to the right. (See discussion of the use of brackets in structural descriptions in Chapter 3, section 3.1). In general, this rule will function where the S is contained in an NP and is a sister to Δ . However, if rule (3) can also be shown to be an adequate statement of the process known as extraposition from NP (see examples (13) through (15) below) then this rule properly excludes the NECESSARY construal of $_{NP}[S]$ as "an S which is also an NP".

Not just any S contained in an NP can be extraposed, however. For instance, if there are conjoined S's as a sentential subject, then the extraposition of just one of them must be blocked. (4b) must be prevented as the result of (3) applying to the structure underlying (4a).

- (4) a. That linguists are clever and that philosophers are cunning is obvious.
- b. *That linguists are clever and (it) is obvious that philosophers are cunning.
- c. It is obvious that linguists are clever and that philosophers are cunning.

For (3), as compared to Emonds' rule (2), it must be that the insertion of *it* is a separate process. Ideally this *it*-insertion would be part of the same rule which inserts an *it* where the rule of raising does not apply (see discussion in Chapter 3, section 3.4). The conditioning for this insertion would then have to be more complicated, however, since it could not be written as just a replacement of the particular lexical item Δ .

+F

Finally, for the statement of rule (3), there must be a way to assure that the delta which is replaced is a delta dominated by S. We will be returning to this issue below (section 6.4) where we make use of the fact that extraposition is a structure preserving movement.

Assuming that the problems in the formal statement of (3) can be resolved we can go on to consider what can occur in the variable material over which an S is extraposed. As is obvious from (1) the process can readily apply where the variable contains a V and an AP. That it can also contain a PP is shown in (5).

- (5) a. It reminds me of my youth $\left\{ \begin{array}{l} \text{that Kira now takes} \\ \text{for Kira to take} \end{array} \right\}$
ballet lessons.
- b. It appeals to balletomanes $\left\{ \begin{array}{l} \text{that Twyla choreographs} \\ \text{for Twyla to choreograph} \end{array} \right\}$
for Mikhail.

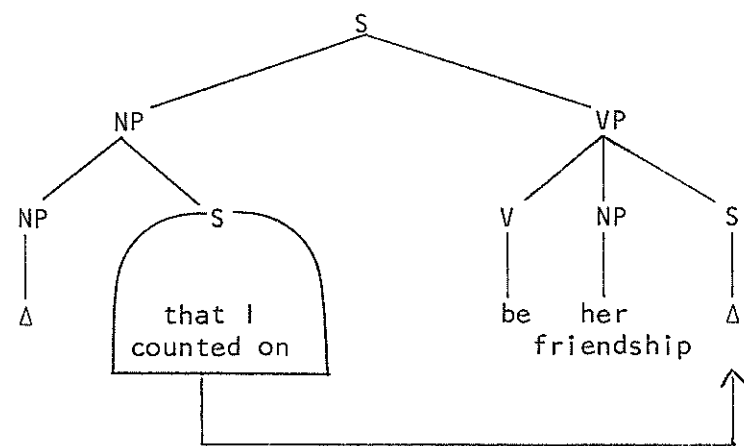
An S cannot extrapose over an S or a VP. The examples of (6) are from Emonds (1976:123).

- (6) a. That John has blood on his hands proves (that) Mary is innocent.
- b. *It proves (that) Mary is innocent that John has blood on his hands.
- c. To see this movie is to relive the past.
- d. *It is to relive the past to see this movie.
- e. That John is late persuades me that the train was delayed.
- f. *It persuades me that the train was delayed that John is late.

Emonds uses these examples as confirmation of the fact that extraposition is a structure preserving rule. Since there is only one VP-final S position, where it is already filled extraposition is not possible.

Notice additionally that even where there is another VP and hence another VP-final S position extraposition still may not apply.

(9) a.



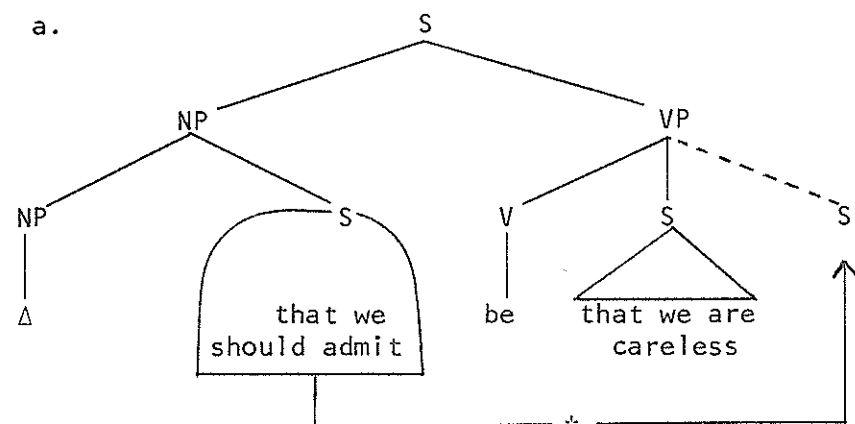
b. It was her friendship that I counted on.

The fact that examples like (10), which follows, are ungrammatical is evidence that cleft extraposition is structure preserving. There is no post-VP or post-S position in a VP for an S to move into in structure preserving fashion. The structures in (11) illustrate this. (The examples in (10) are from Emonds (1976:140).)

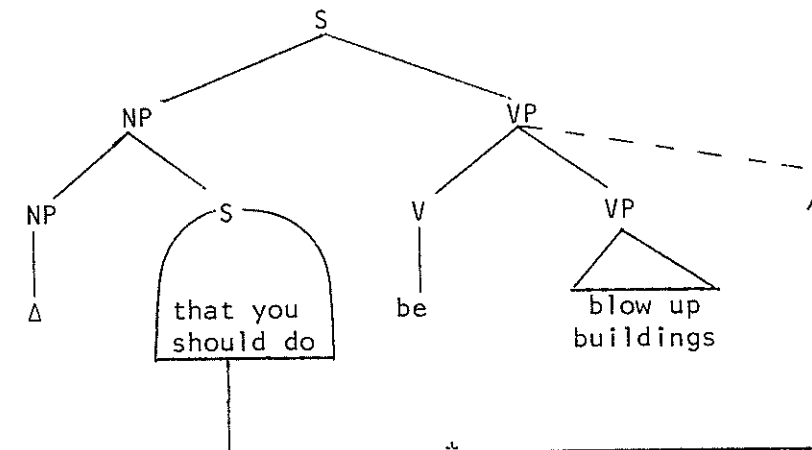
(10) a. *It is that we are careless that we should admit.

b. *It is blow up buildings that you should do.

(11) a.

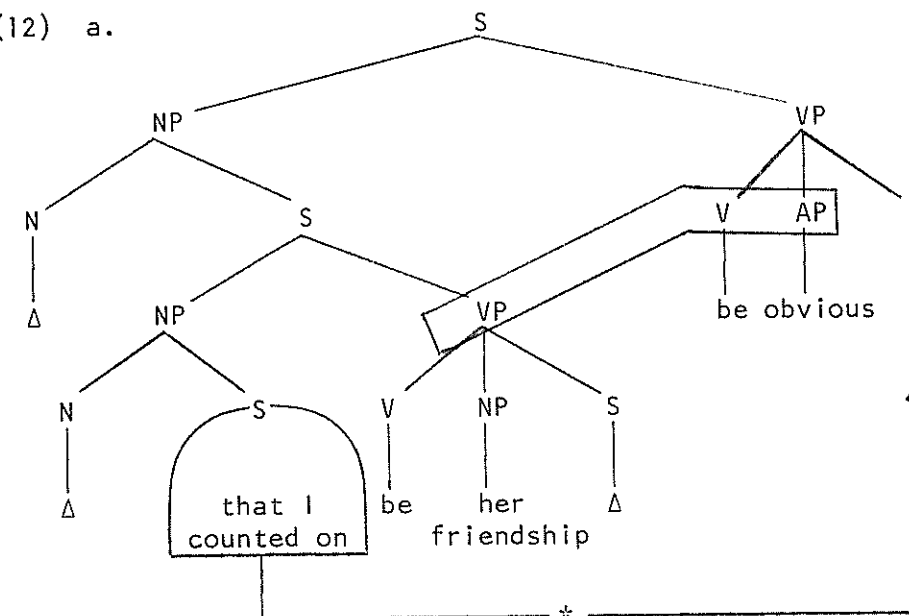


(11) b.



The examples of (12) show that even where there is an available VP-final S position, because the variable may not contain a VP, extraposition may not take place.

(12) a.



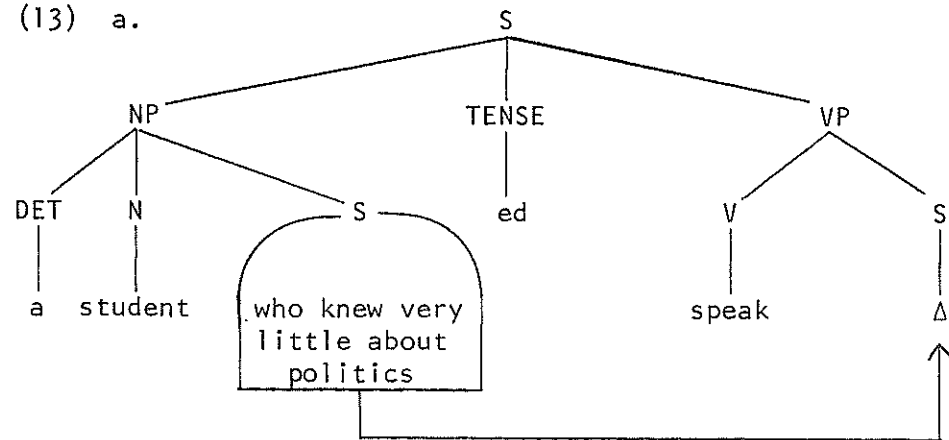
b. *That it was her friendship is obvious that I counted on.

c. That it was her friendship that I counted on was obvious.

d. It was obvious that it was her friendship that I counted on.

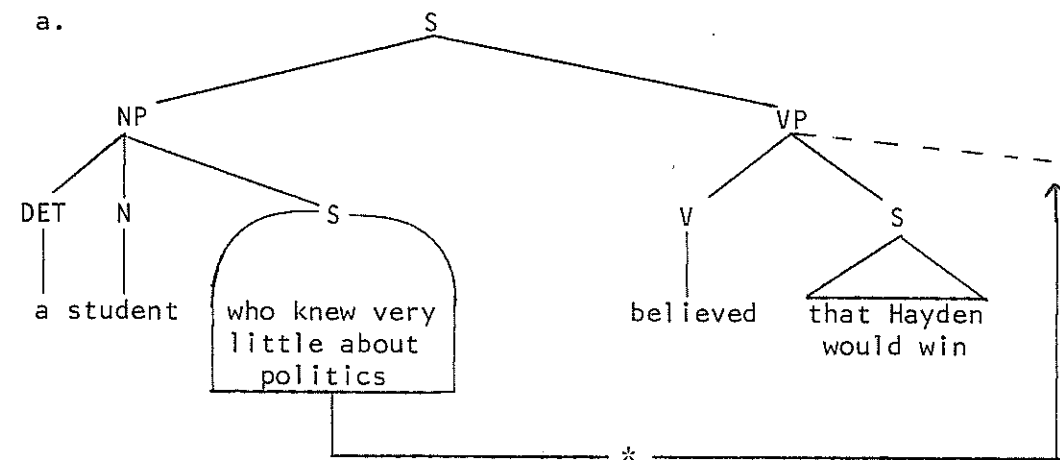
Extrapolation from NP, just like the two previous extrapolation processes, is both structure preserving and upward bounded.² Examples (13) through (15) illustrate this. (Example (13) is from Emonds (1976:145).)

(13) a.



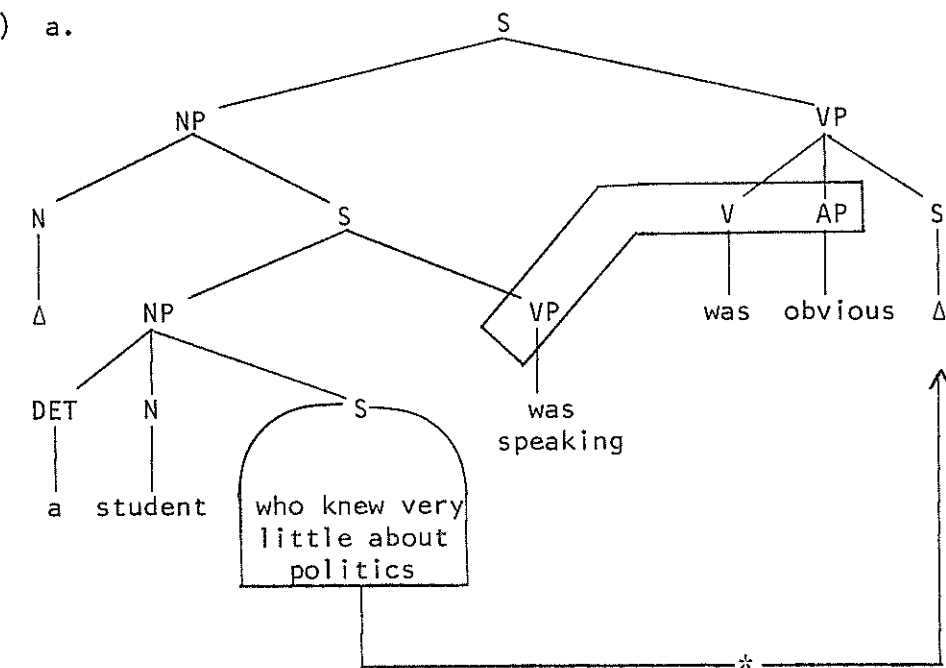
b. A student spoke who knew very little about politics.

(14) a.



b. *A student believed that Hayden would win who knew very little about politics.

(15) a.



b. *That a student was speaking was obvious who knew very little about politics.

What has been illustrated here so far is that the rightward movement of S is structure preserving and also that an alternative to saying that it is upward bounded is to say that it may not operate where the internal variable contains a VP.³ By adhering to the assumption, used in Chapter 3, where C^{\max} was defined, that VP is the head of the S, this upward bounding phenomenon is readily predicted by the Variable Interpretation Convention. An S may not be moved where the grossest constituent analysis would contain either an S or the head of an S, that is, a VP. It seems because of the possible phrase structure configurations of English the first alternative, where the grossest analysis would contain an S, never arises, except of course where an S may not extrapose within a VP because the final position is already filled. This would be prevented by the VIC but need not be since extrapolation is a structure preserving movement.

Notice, now, that while extrapolation is properly prevented where the variable contains a VP, it must not be blocked where the variable contains just a V: see example (1c). What this illustrates is that V is not the head of S. S is not to be considered the maximal projection of the category V (cf. Jackendoff 1971). V is the head of VP, VP is the head of S, but V is not the head of S. An S can occur as V^{\max} (as in (25b) of Chapter 3) in a given phrase marker where it is not the head of any bigger

constituent, but S is NOT V^n where n = the maximal number of bars. For discussion of the claim that C^{\max} should not be considered to be synonymous with C^n see footnote 14 of Chapter 3. Whether or not a constituent is a C^{\max} can be determined only with respect to a particular phrase marker and by examination of the configuration in which C occurs. C^n , on the other hand, can be defined without regard to any particular phrase marker.

Before we go on to discuss a second rightward movement, complex NP shift, in section 6.3, we will look at the consequences of incorporating into this model two constraints on transformations which have been independently suggested. What follows in this chapter is meant to be suggestive, indicative of directions for future research, rather than theoretical claims whose ramifications have been fully considered.

6.2. The Boundary Attachment Condition and the Fixed Head Principle

In this thesis, thus far, we have considered several movement processes and one thing that has been true about each one is that it has moved terms to constituent boundaries. WH-fronting moves a term to the left boundary of \bar{S} , raising moves an NP to the left boundary of S , passive moves an NP to the left boundary of S (generally), and extraposition moves an S to the right boundary of VP. This is exactly as is predicted by Schwartz (1972)⁴ and Baltin (1976)⁵ and will be referred to from now on as the Boundary Attachment Condition. The only counterexamples to this condition, for the rules discussed in detail in this work, are where passive applies to move an NP into object position of verbs like *expect* (see section 5.3.1 of Chapter 5). Where passive applies in these cases it is affecting strictly adjacent terms. Perhaps it is only where rules function over a variable that they NECESSARILY obey the Boundary Attachment Condition. Without examining all cases of local versus non-local rules (see Emonds 1976) we will assume, for the sake of argumentation, that this is in fact the case.

By assuming also that we know what the phrase structure rules for English are, then conceivably every movement which may take place over a variable can be written as a replacement of some category where what that category is can be predicted from what can occur at constituent boundaries. Continuing to generalize from what we have seen so far, it is conceivable that leftward movements replace only categories at the left boundaries of constituents and that rightward movements replace only a category which is rightmost in a constituent. Where the movement indicated by a rule is strictly structure preserving, like raising, passive, and extraposition, then the category of the target for the movement is also, obviously, predictable from the category of the term which moves. We will see below, in section 6.3, that for rules which are not structure preserving, the Boundary Attachment Condition turns out to be very useful. Before we continue on with that, we will consider a second constraint proposed by Schwartz.

In the same paper mentioned above, Schwartz (1972:37) proposes another constraint on movement rules which has to do with the reordering of heads of phrases. His constraint is:

(16) The Fixed Nucleus:

The nucleus (= head) of a phrase cannot be moved within its phrase.

If we extend this constraint to include extraction of a nucleus, and then state it as a positive requirement, we will have the following condition:

(17) The Fixed Head Principle:

If a transformation T indicates the reordering of C where C is a lexical category and the head of C' , then T is to be construed so as to reorder C' .⁶

The Fixed Head Principle (FHP) entails basically two things. First of all, any movement which is meant to indicate the reordering of a whole phrase can be written as just the movement of the head of that phrase. Secondly, from the FHP can be derived Schwartz's Fixed Nucleus Constraint. No head could be moved within its phrase because the movement of a head would require the whole phrase to move.

By the FHP, now the movements of noun phrases which were discussed in Chapters 3 and 5 can be written as movements of N . A movement of N will always be construed as a movement of the NP of which that N is the head. In Chapter 4 it was shown that WH-fronting could be written as a movement of a specifier where, by the Revised Left Branch Condition, the whole phrase containing that specifier would be reordered. Thus far, then, leftward movements can be written as the movement of either a lexical category or a grammatical category, or in other words, a morpheme category.

The one rightward movement we have looked at so far, extraposition, can now be written as the movement of VP. Since VP is the head of S , the FHP would require that a movement of VP reorder the whole S . This change in the formal statement of extraposition and how it interacts with the VIC will be discussed in section 6.4. Notice now that in this model a rule such as VP-preposing would have to be written as a movement of V rather than of VP. Moving a V , the head of VP, would reorder the whole VP. If the rule were written to prepose VP then the whole S would move, as in extraposition. For leftward movements of VP this, of course, would be the wrong result.⁷

In summary, thus far, we can make the following generalizations. Leftward movements which may take place over a variable can be written as movements of a morpheme category. The fact that whole phrases are re-ordered is accounted for by the Fixed Head Principle and by the Revised

Left Branch Condition. Rightward movements, or at least the one rightward movement rule we have considered (which has three subcases), are stated as the movement of a phrasal category. Where the structural description of a rule indicates the reordering of C and C is a morpheme category the rule is predictably a leftward movement. Where the C to be reordered is a phrasal category the movement is to the right. Additionally, terms which are moved to the left are attached at left constituent boundaries and terms which are moved to the right are attached at right boundaries. The only movement we have discussed which moves a constituent away from the head of the phrase in which it was generated is the rightward rule of extraposition. We will again consider rightward versus leftward movements in 6.4, but now we will look at complex NP shift.

6.3. Complex NP Shift

There is a process in English which evidently can move object noun phrases to the end of the VP. Since by the phrase structure rules for VP there is no post-PP position for NP's, the (b) examples of (18) and (19) must be the result of a transformational operation.

(18) a. Alice brought the rabbit that ran down the hole to the Queen.

b. Alice brought to the Queen the rabbit that ran down the hole.

(19) a. Alice accompanied the cat with the big grin down the yellow brick road.

b. Alice accompanied down the yellow brick road the cat with the big grin.

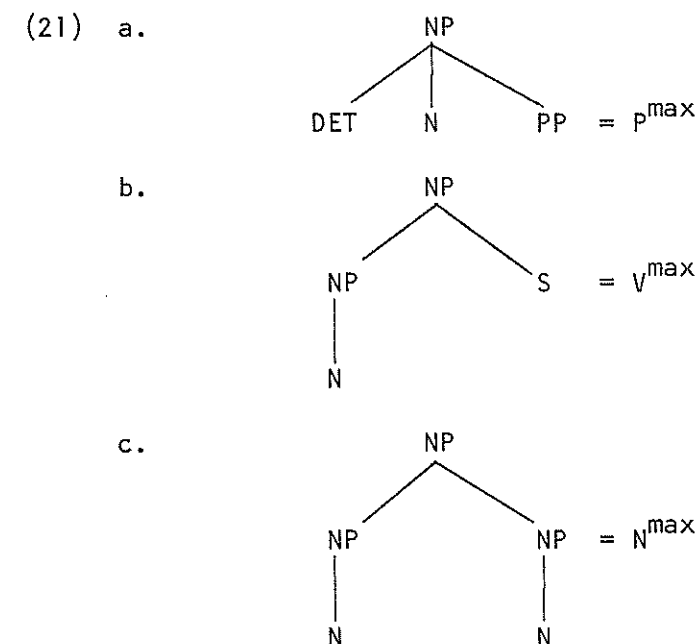
The transformation illustrated here is known as complex NP shift (Ross 1967). Object NP's may be moved to the end of VP as long as they are sufficiently "complex".

Since it is only NP's with certain internal structure which may undergo this rule the structural description for complex NP shift has often appeared in the literature with certain conditions pertaining to the NP term, for instance, that it must dominate S or PP (Ross 1967:32). This rule can also apply to conjoined NP's:

(20) a. Alice brought the Cheshire Cat and the Cowardly Lion to the Queen.

b. Alice brought to the Queen the Cheshire Cat and the Cowardly Lion.

It seems that an accurate description of the notion "complex NP", in the framework of this dissertation, is any NP which contains a C^{max} as a right sister to N or NP. This is shown in (21):⁸



The rule of complex NP shift can now be stated as (22).

$$(22) N + C^{max} - \Delta = \emptyset - 1^9$$

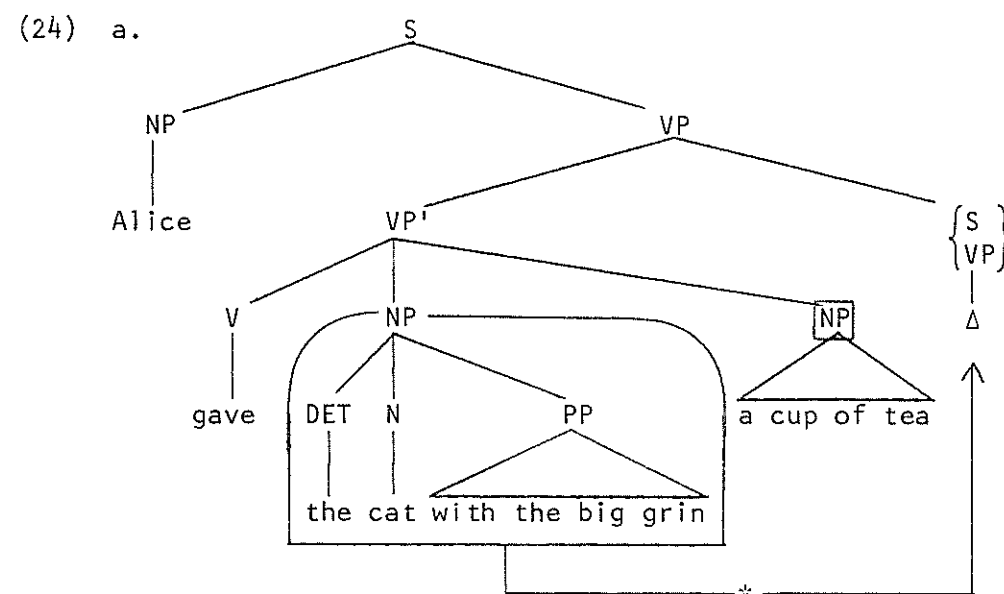
Assuming that there is a general condition that allows transformations to move only constituents (to prevent, for example **John gave yesterday Bill a book* as the result of (22) applying to *John gave Bill a book yesterday*), rule (22) can account for the postposing of any NP like those illustrated in (21). The fact that (22) indicates the movement of a whole NP, that is, including what precedes the N (DET, Adj) is accounted for by the FHP.

Now, by use of the Boundary Attachment Condition, we can consider whether it is possible to predict what the category of term 2 of the structural description of (22) must be. In examples (18), (19), and (20) the complex NP is in VP-final position. The final constituent in VP's is either an S or a VP. We can hypothesize, then, that complex NP shift moves an NP to replace a delta which is an instance of either category S or VP. By virtue of the VIC there is now a way to test this prediction.¹⁰ First of all, since (22) indicates the movement of an NP, the VIC would say that the variable could contain no NP or head of an NP. Additionally, if term 2 were an S, the VIC would say that the variable could contain no S or head of an S, VP. If term 2 were a VP then the variable material could contain neither a VP nor a V which is the head of VP.

To show first that the variable for (22) may not contain an NP we have to construct a situation where some NP could occur in the grossest analysis between the NP to move and the delta to be replaced. This might happen after indirect object movement. Example (23) is a case where by indirect object movement a complex NP precedes another NP.

(23) Alice gave the cat with the big grin a cup of tea.

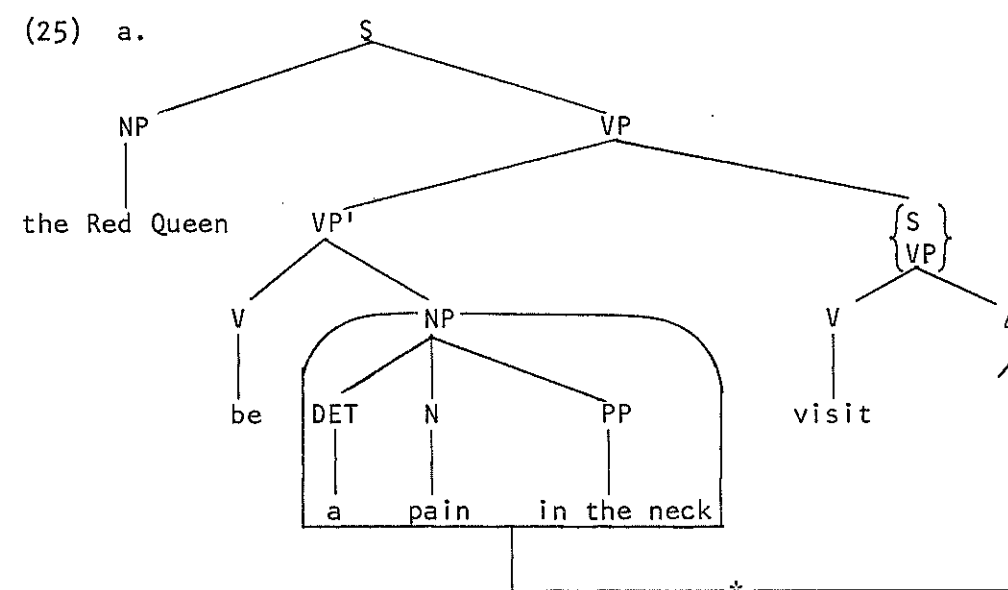
The phrase marker underlying (23) would be (24a).



b. *Alice gave a cup of tea the cat with the big grin.

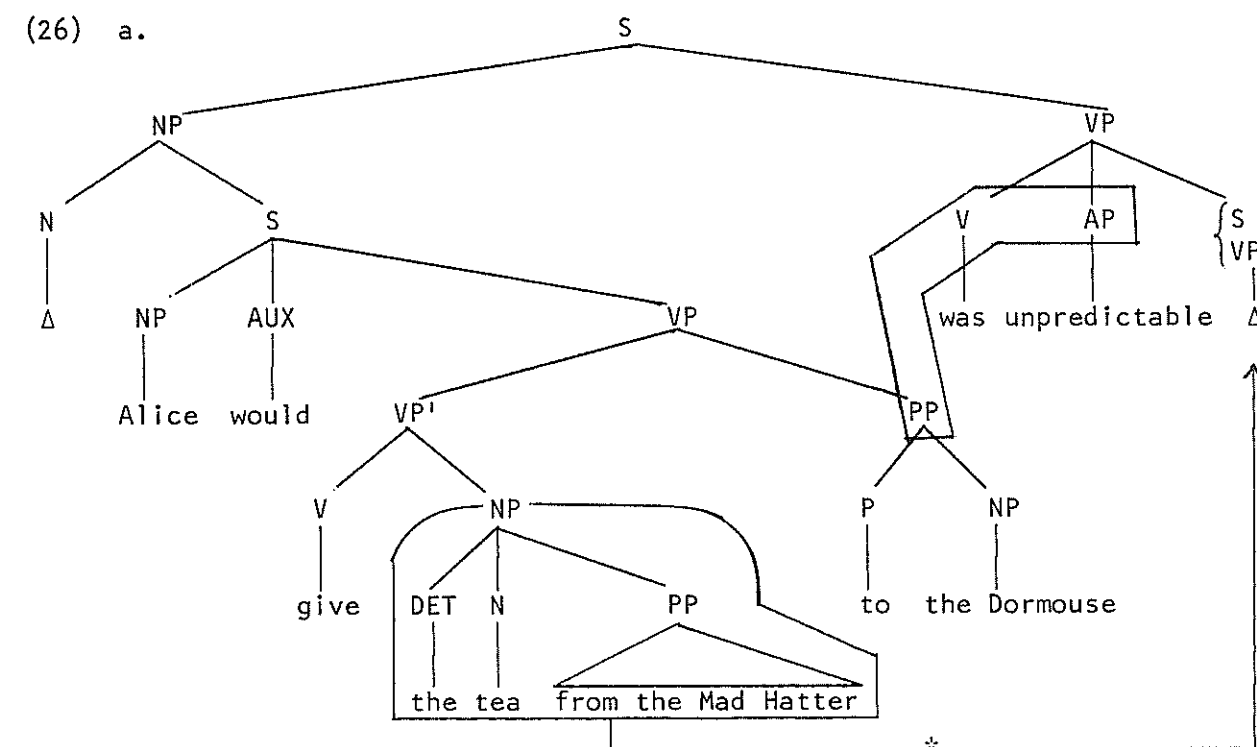
The movement indicated in (24a) is properly blocked by the VIC as is indicated by the ungrammaticality of (24b).¹¹

Phrase marker (25a) shows a structure where a V would occur in the grossest analysis for complex NP shift.¹²



b. *The Red Queen is to visit a pain in the neck.

Another type of structure where complex NP shift must be blocked is given in (26).



- (26) b. *That Alice would give to the Dormouse was unpredictable
the tea from the Mad Hatter.

What (25) and (26) show is that the variable for rule (22) may not contain a V. This would be naturally accounted for where term 2 of complex NP shift was a VP. Since V is the head of VP, but not the head of S, a V in the grossest analysis blocks a rule only where a VP (or V) is part of the structural description. If the VIC is relevant for rule (22), as of course it should be in this model, then term 2 is a VP and not an S.

Example (26) shows that complex NP shift, like extraposition, is an upward bounded rule. Unlike extraposition, however, complex NP shift is not structure preserving. (For discussion of this compare Culicover and Wexler 1976 and Emonds 1976, particularly Chapter III.7 of Emonds and the statement of the Sentence Boundary Condition.) In the framework of this dissertation this rule is in fact the only rule we have discussed which is not structure preserving. For Emonds (1976) WH-fronting, like complex NP shift, is not STRICTLY structure preserving. Here, where WH-fronting is written as a movement of a WH determiner, rather than a movement of a phrasal category, it can be considered structure preserving.¹³

What I now suggest as a viable formalism for transformations is as follows. Structure preserving movements are substitutions where the category of the target site is predictable from the category of the term which moves. Rules which may function over a variable but are not structure preserving are written as replacements of delta where the category of the delta must be indicated in the structural description. Rule (22) for complex NP shift is therefore revised to:

$$(27) N + C^{\max} -_{VP}[\Delta] \Rightarrow \emptyset - 1^{14}$$

where a labelled bracket containing only a delta is defined as a node which is empty of terminals. In other words, $_{VP}[\Delta]$ indicates a case where no rule applied to further expand VP; $_{VP}[\Delta]$ equals $_{VP}$. All movements, structure preserving or not, are subject to the Variable Interpretation Convention.

6.4. Two Hypotheses

The speculations of this chapter, taken together with the overall model worked out in depth in the previous chapters, allows the following four-part hypothesis.

1. Any reordering transformation either, (a) involves the movement of just a morpheme category, that is, a lexical category or a grammatical category; or, (b) involves the movement of a phrasal category, where C^{\max} is considered a phrasal category since its instantiation can be a phrasal category.

2. Any rule which indicates the reordering of JUST a morpheme category will be a leftward movement. This is true in English (by virtue of the Revised Left Branch Condition and the rules which happen to invoke the FHP) and it seems plausible that it will also be true at least in other languages with specifiers and complementizers which are leftmost in their constituents.

3. Any rule which can be written as involving the movement of a phrasal category will be a rightward movement. This is true thus far for English (extraposition and complex NP shift) and plausibly also true at least for other languages where complements occur to the right of their phrasal heads.

4. Additionally, by the Variable Interpretation Convention, leftward movements are predictably cyclic (possibly successive cyclic) or unbounded (at least for the processes investigated here in Chapters 3, 4, and 5) and rightward movements are predictably upward bounded (at least where they are replacements of VP-final S or VP). Again the speculation is that this will be true at least for languages which, like English, have pre-head specifiers and post-head complements.

This hypothesis, for this predictable dichotomy between leftward and rightward movements is, like the other suggestions of this work, subject to empirical verification. If this hypothesis can be verified, transformations can then be written in even further simplified form than previously suggested in this dissertation.¹⁵ For instance, a rule like raising might be written as:

$$(28) \text{ Move } N \text{ out of } S \text{ to } \Delta +F$$

Because for structure preserving rules it is not necessary to indicate the category of the target term, the form of (28) indicates that it is a structure preserving rule. It is predictably a leftward reordering of a full NP because it is written as the movement of a morpheme category. That it applies cyclically and only where no instance of N occurs in the grossest analysis is predictable from the VIC.

The rule for complex NP shift, given in (27) and restated here as (29):

$$(29) \text{ Move } N + C^{\max} \text{ to } _{VP}[\Delta]$$

is predictably a rightward movement since it is written as involving the reordering of a phrasal category. Since in the statement of (29) the category of the target is included the rule is not structure preserving. The rule moves a full NP, by the Fixed Head Principle, and is predictably upward bounded by the VIC.¹⁶

Before we continue on to discuss the reformulation of the rule of extraposition and a new rule not previously considered above, it should be reiterated that all the predictions which can be made about rule applicability here are based on the FORM OF THE RULE. How that rule interacts with the VIC is determined not only by the form of the rule itself, but also by the FORM OF THE PARTICULAR PHRASE MARKER to which the rule applies. The grossest constituent analysis for the variable material can only be determined with respect to a particular structural description and a given phrase marker. We will see next that it is also important to take into account the FORM OF PARTICULAR INSTANCES OF NON-VARIABLE TERMS WITH RESPECT TO GIVEN PHRASE MARKERS.

Let us consider the reformulation of extraposition.

(30) Move VP out of NP to Δ

This rule is written as a reordering of VP because, as was pointed out in section 6.2, by the FHP and the fact that VP is the head of S, a movement of VP in fact reorders the whole S. Since the category of the target site is not indicated in (30) this is meant to be a structure preserving rule. It was shown in section 6.1 that extraposition was a replacement of S over Δ . Therefore, in (30), even though the term to be moved is VP it must be that the delta is an S.

Returning momentarily to (28) and (29), we now make explicit an implicit assumption. Even though (28) is written as the movement of N, since by the FHP it must involve the movement of NP, for both the VIC and the notion of structure preservation it is treated as the movement of NP. For (28) the target site is an NP, not an N, because if the result of the movement were $N[NP]$ this would not be a possible base configuration and hence, the rule would not be structure preserving. The VIC interprets "move N" as a movement of NP because it is still an NP or head of an NP which is relevant in the grossest analysis, not just an N. Rule (29), written as "move N + C^{max} ", is treated by the VIC as a movement of NP. The relevant constituents in the grossest analysis are NP's or heads of NP's (see example (24), section 6.3) or, based on the target site, VP's or heads of VP's (see examples (25) and (26) of section 6.3). A C^{max} or head of a C^{max} is not relevant for the VIC even though this is a non-optional term in the structural description. Just some C^{max} in the grossest analysis does not prevent the rule from applying. Where, by the FHP, a movement of the head of a phrase causes the reordering of the whole phrase, it is that whole phrase which is relevant for the Structure Preserving Hypothesis and for the VIC.¹⁷

Going back then to rule (30) for extraposition, we again explicitly state the same assumptions. Since by the FHP a movement of VP reorders the whole S the category of the target in (30) is understood as an S. For the VIC the relevant constituents in the grossest analysis are S and

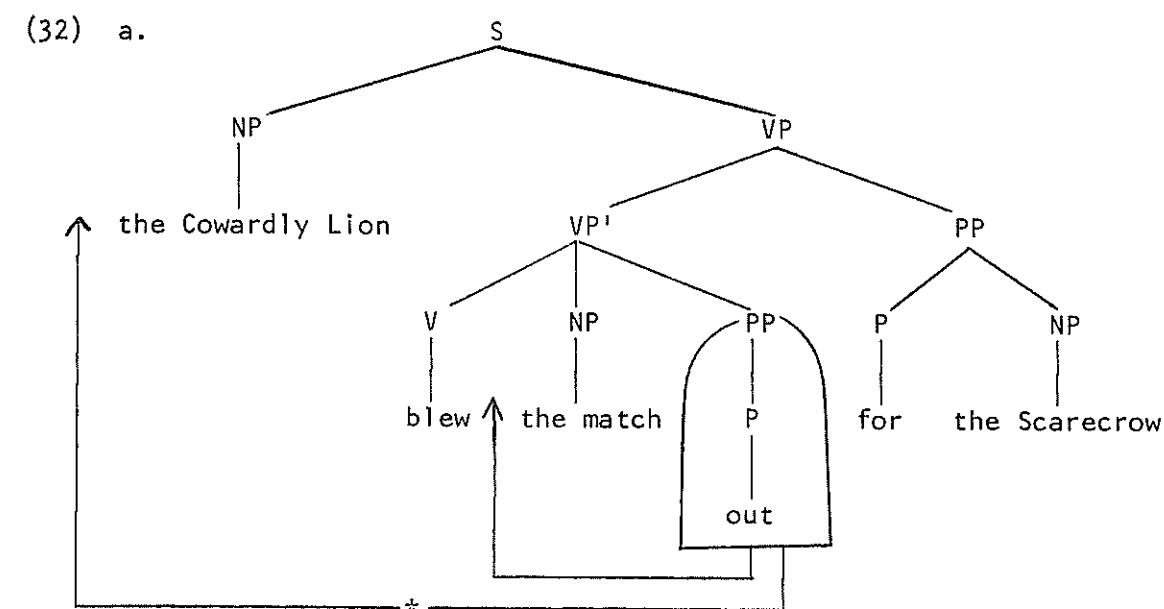
the head of an S: namely VP. This was illustrated in examples (7), (12) and (15). By the hypothesis at the beginning of this section, rule (30), written as the movement of a phrasal category, is predictably a rightward movement. By the VIC, as in 6.1, it is predictably upward bounded.

In this last discussion of (28), (29) and (30), we have used both the form of the rule and its actual affect on the phrase markers to which it applies. In other words we have used the nonvariable terms of the structural description and the actual instantiation of those terms where they are applied to a structure. Rule (30) is written as the movement of the phrasal category VP, but the instantiation of VP is S, in that what the rule actually does is extract (and then reattach) an S in a given phrase marker. With this in mind, let us now consider a rule which might be written as (31).

(31) Move $\left[\begin{smallmatrix} P \\ +directional \end{smallmatrix} \right]$ to immediately precede NP¹⁸

By the Fixed Head Principle and the hypothesis at the beginning of this section, (31) is a leftward movement of a directional prepositional phrase. Since by the base rules for English there is no pre-NP position for PP's, rule (31) cannot be a structure preserving rule. I think it can be shown to be predictably either strictly local (that is, the rule usually known as particle movement) or root (the rule usually known as directional adverb preposing).¹⁹

Consider the structure in (32a).



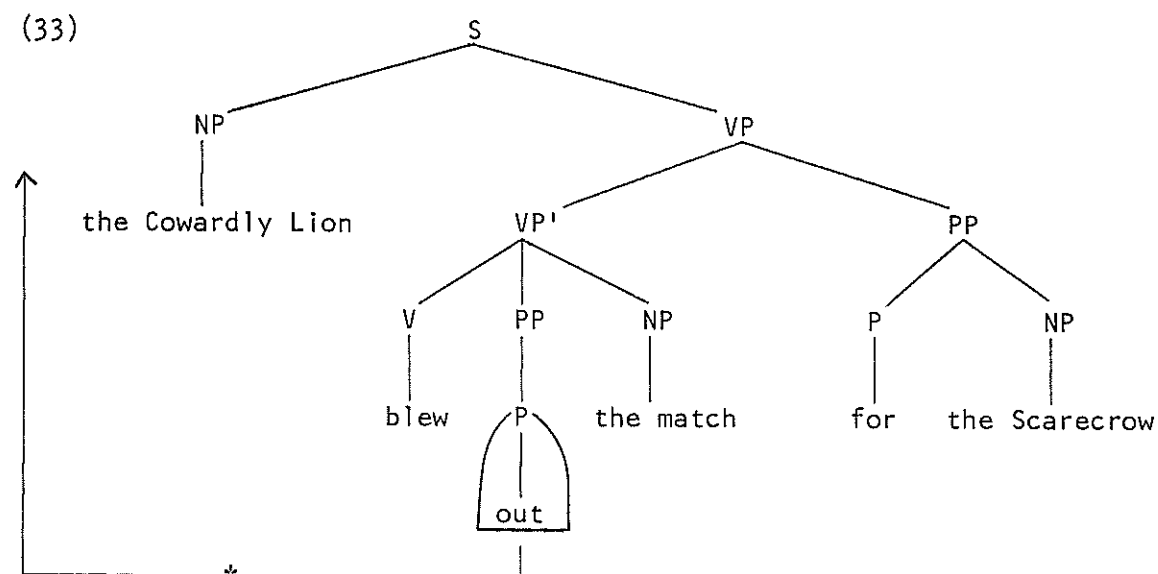
b. The Cowardly Lion blew the match out for the Scarecrow.

(32) c. The Cowardly Lion blew out the match for the Scarecrow.

d. *Out the Cowardly Lion blew the match for the Scarecrow.

Since the intransitive prepositions (or most of them) are a subset of those prepositions which occur in directional adverbial PP's we are considering the particles (such as *out* in (32)) as [+directional]. Since they are not always "directional" in any obvious sense there might be a better feature to use.²⁰ For now, however, we are considering the process illustrated in (32a) to be a case of the application of rule (31). The grammaticality facts illustrated in (32) are accounted for by the VIC requirement that transformations apply first where the terms of the structural description are strictly adjacent.²¹

Consider next, however, phrase marker (33) which is the result of an application of (31).



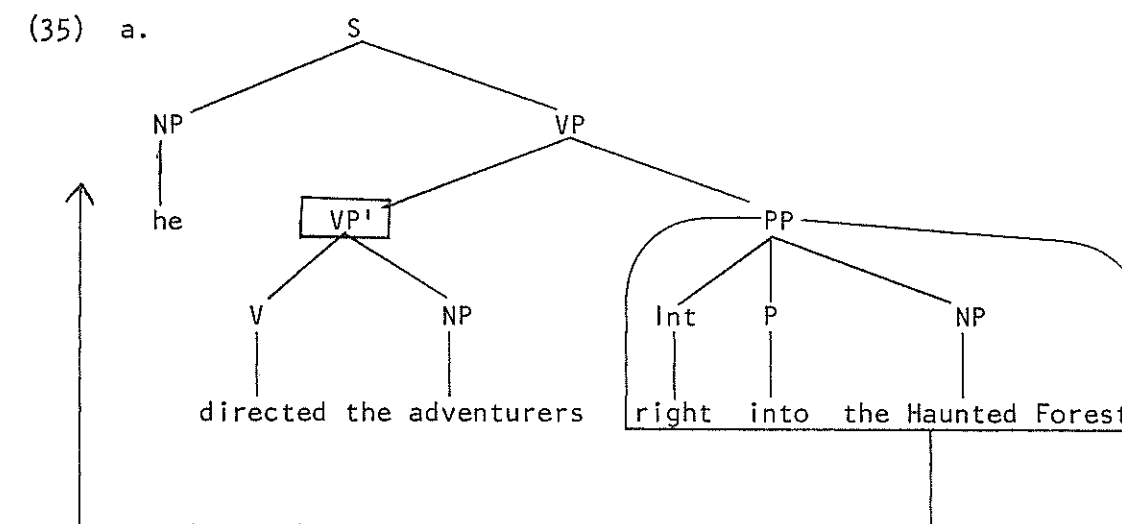
If the process illustrated in (33) were allowed the result would be the same as the ungrammatical (32d). So far there is no way to block the movement illustrated in (33). We might, perhaps, adopt the following hypothesis:

(34) If a transformation causes the reordering of a phrase where the particular instance of that phrase contains just the morpheme category head of that phrase, then the transformation may operate ONLY where the terms of the structural description are strictly adjacent.

(Cf. the definition of Local Transformation (Emonds 1964:4).) With this principle, (32d) cannot be derived from either (32a) or (33). Where the

instance of PP which is moved is just a P the movement may not take place over a variable, by (34). Sentence (32c) is properly derived by an application of (31) to (32a). The attachment of the moved PP in the derivation of (32c) properly need not be at a constituent boundary since the Boundary Attachment Condition is relevant only for movements over a variable.

Now compare phrase marker (35a) with (32a).



b. Right into the Haunted Forest he directed the adventurers!

Rule (31) applies to (35a) to derive (35b). Since the instance of PP being moved contains more than just P, (31) properly applies where the terms of the structural description are weakly adjacent. Rule (31) is functioning here over a variable where the grossest constituent analysis contains VP'. The PP in (35a) is correctly attached outside VP' since it is not obligatorily strictly subcategorized by the verb.

There remains still some detail to be accounted for. For instance, example (36) must not be generated.

(36) *He directed right into the Haunted Forest the adventurers!

Where (31) moves a PP which contains more than just a P, it, unlike other rules, may NOT apply to strictly adjacent terms. Interestingly enough, however, if (31) applies to derive (36) it would be applying where, for the strictly adjacent terms A and B of the structural description, A is more deeply embedded in the phrase marker than is B. In no other example of strict adjacency has this been the case. Perhaps then strict adjacency should be defined for A and B where either A and B are sisters, or A is higher in the phrase marker (= dominated by fewer nodes) than B. Example (36) is then prevented by the Boundary Attachment Condition which must be obeyed except where a transformation affects strictly adjacent terms.

Another type of case which must be accounted for is

(37)  Out John ran!

Here a PP containing just a P was moved over a variable and the result is well-formed. Thus far I cannot suggest a way to account for (37), although it will probably turn out to be useful that this movement of P, unlike the strictly local movement of P, can only occur in root sentences.

In this final section of this chapter two hypotheses have been proposed. The first is for a predictable dichotomy between rightward and leftward movements based on the form of structural descriptions. The second hypothesis is that there is a predictable distinction between rules which must function only where terms are strictly adjacent and those which may function over a variable. The second distinction can be made here only by utilizing both the form of the rule and the form (instantiation) of a nonvariable term with respect to a given phrase marker. Since this dissertation has already demonstrated the explanatory power of the VIC framework, these two hypotheses represent two promising directions for additional research using the Variable Interpretation Convention.

FOOTNOTES TO CHAPTER 6

¹In this discussion we will not be considering extraposition from object position since it operates over an empty variable. It will always properly move an S to final position in the SAME VP because by the VIC rules it will always function first in strictly local fashion.

In fact, as Joe Emonds points out to me, it seems that "local" extraposition applies even before gerund formation (for discussion of this rule see Emonds (1969:124-26)). This is illustrated in (i) and (ii).

- (i) a. That rules don't move variables suggests that we limit structural descriptions to constant terms.
- b. That rules don't move variables suggests limiting structural descriptions to constant terms.
- (ii) a. *It suggests that we limit structural descriptions to constant terms that rules don't move variables.
- b. *It suggests limiting structural descriptions to constant terms that rules don't move variables.

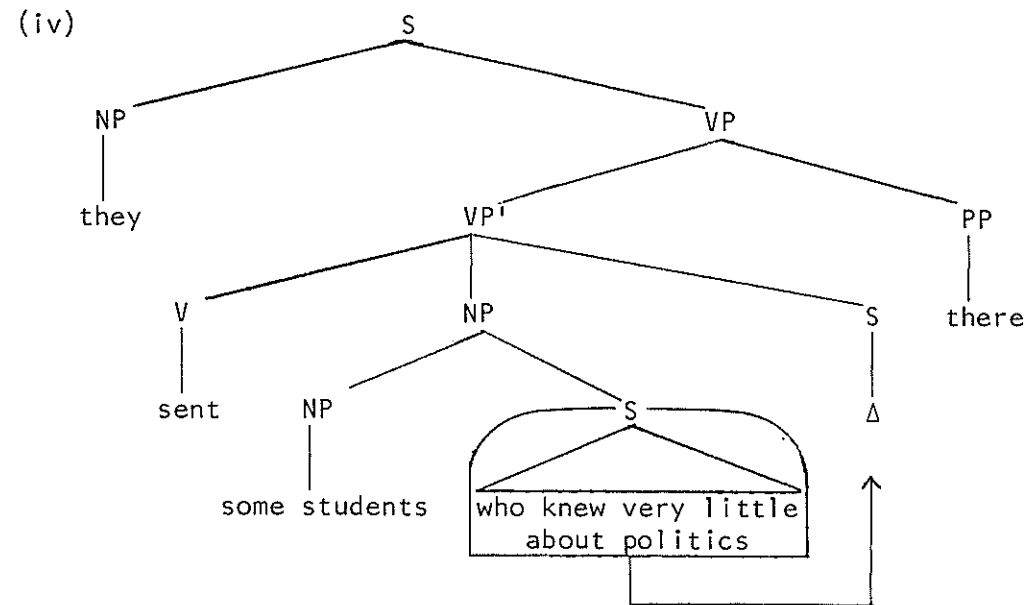
²Ross (1967:section 5.1.1.3) presents a rule ordering argument that extraposition and extraposition from NP must be separate rules because the rule for question formation must intervene between them. I think both extrapositions can be considered subcases of rule (3). This is possible here because of the formulation of WH-fronting and the Revised Left Branch Condition. WH-fronting in the model of this thesis moves a DET rather than an NP, and the Revised Left Branch Condition moves a left branch and the immediately dominating NP or AP and no larger phrasal category. Additionally, WH-fronting out of sentential subject *that*-clauses is blocked by the COMP in the grossest analysis (cf. Ross (1967:155-61)).

³Paul Schachter has called my attention to the following examples, with respect to extraposition from NP, where the relative clause to be extraposed is on an object NP.

- (i) They sent some students who knew very little about politics there.
- (ii) They sent some students there who knew very little about politics.
- (iii) *That they sent some students there was obvious who knew very little about politics.

Sentence (iii) cannot be accounted for by a VP in the grossest analysis. It is, however, prevented by the VIC requirement that rules function first

where the involved terms are strictly adjacent. See footnote 1 above. The fact that extraposition cannot then reapply to derive (iii) is discussed in the text of this chapter following example (8). What does have to be explained here is the derivation of (ii). In this model it would have to be the result of rule (3) applying first in strictly local fashion to (iv), as illustrated.



Then the fact that in (ii) the PP precedes the S is accounted for by the late (scrambling, or perhaps stylistic) rule, already alluded to in footnote 19 of Chapter 5. This rule optionally reorders constituents contained in VP so that the usual order is that shorter constituents precede longer ones. This account obviously depends on the fact that a locative PP which is not obligatorily strictly subcategorized by the verb occurs outside VP'.

⁴Schwartz's constraint would not, however, allow for any unbounded movement. See footnote 17 of Chapter 3 for the statement of Schwartz's boundary attachment condition.

⁵This was pointed out in section 3.4 and footnote 17 of Chapter 3.

⁶A similar condition is proposed in Wexler and Culicover (in preparation), where they suggest that transformations may not move heads away from their complements.

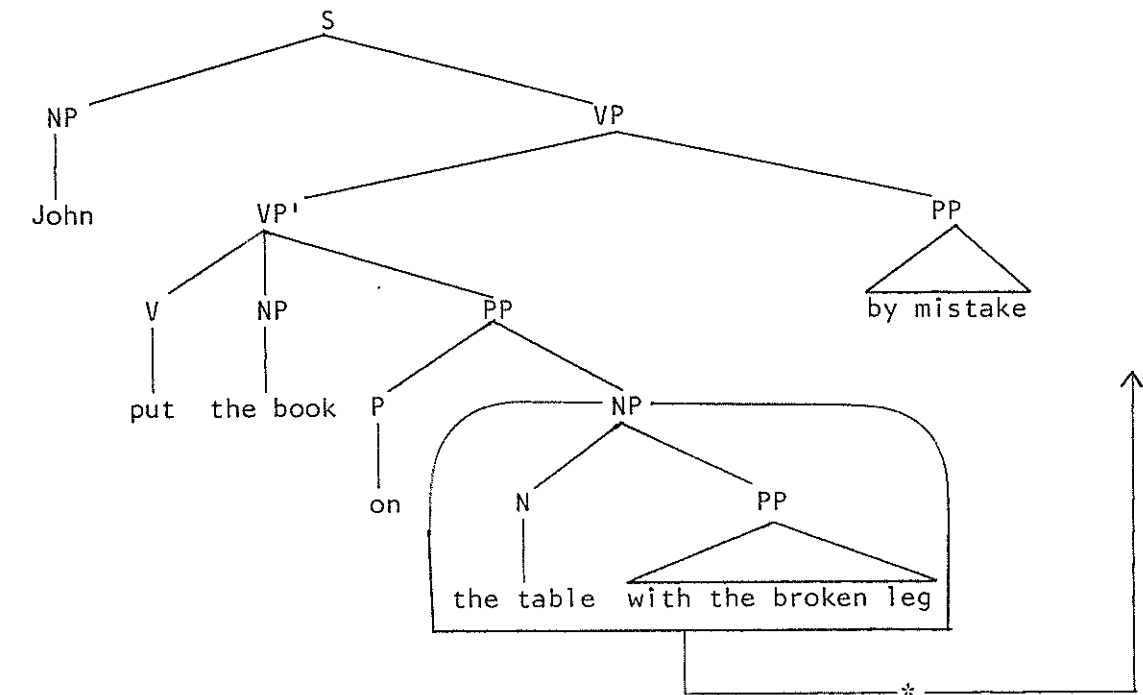
⁷Joe Emonds (personal communication) suggests that actually preposings of VP should be written as movements of TENSE.

⁸Paul Schachter has called to my attention the fact that some sufficiently complex compounds can also undergo complex NP shift:

Alice brought to the Queen The Cheshire-Cat-Look-Alike-Contest winner.

I have no suggestion here as to how to characterize the internal structure of such compounds.

⁹Actually there must be some way to restrict this rule to apply only to object NP's. It must not apply either to matrix or embedded subjects. Writing the rule as a movement of $VP' [N + C^{max}]$ would prevent matrix subjects from ever meeting the structural description but subjects of embedded clauses contained in VP' would meet the description. In all cases embedded subjects would be blocked from moving to VP-final position because of what would be contained in the variable (namely VP or V) as will be clear from the discussion below. The use of VP' is suggested here only tentatively because of the problem of prepositional objects which can occur in VP'. These object NP's may not be moved away from their prepositions.



If, however, extraposition of PP and complex NP shift can be collapsed into one rule, a not implausible possibility, then the whole PP in this example would be properly reordered.

¹⁰This is used here merely as discovery procedure because, in fact, the terms of the rule predict what can occur in the variable, not vice versa. In other words, just as with all other movements, when the structural description of complex NP shift is accurately stated the non-variable terms will determine what can occur in the variable material. It will not be the variable which determines what can occur as term 2.

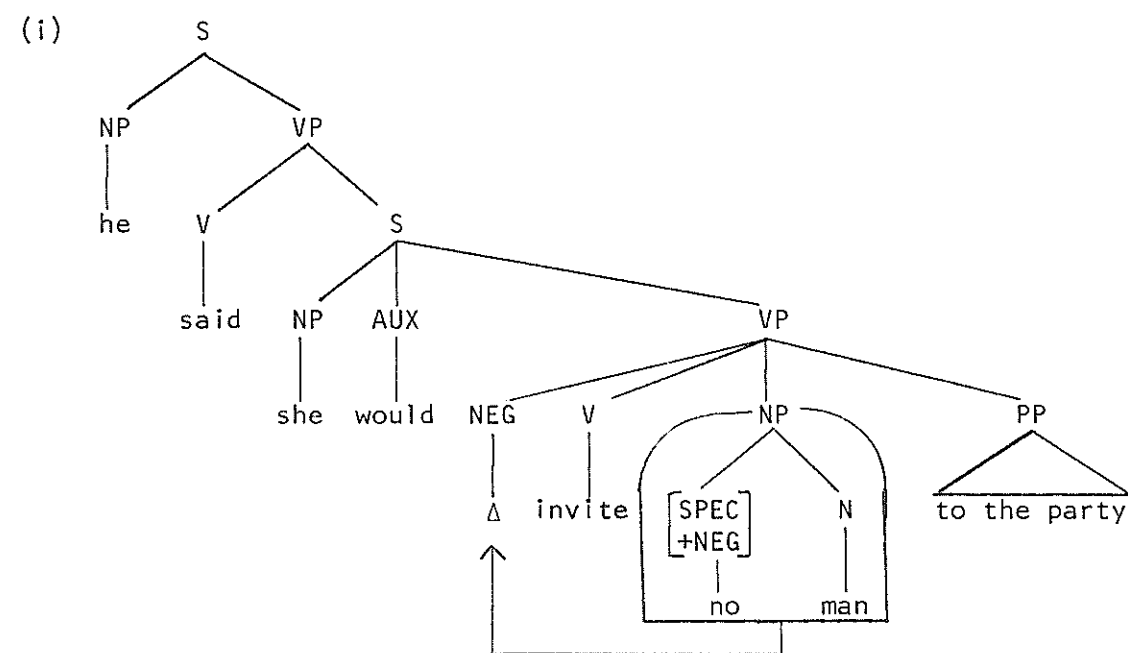
¹¹The VIC would predict that complex NP shift could never apply over an NP or head of an NP in the grossest analysis. The well-formedness of example (ii), pointed out to me by Paul Schachter, is therefore unaccounted for.

(i) They elected a man with a big grin president.

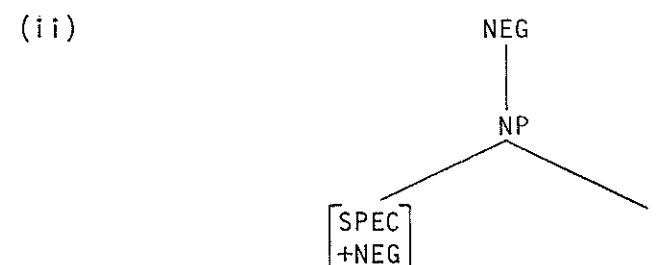
(ii) They elected president a man with a big grin.

¹²The structure shown in (25a) is the result of object raising. The fact that [*be a pain in the neck*] is considered a VP' is discussed in footnote 23 of Chapter 3. See also footnote 12 of Chapter 3.

¹³With the statement of the Sentence Boundary Condition, Emonds (1976: 112) allows for rules which are weakly structure preserving if they are substitutions for some constituent which is the rightmost or leftmost constituent of S. He has pointed out to me that perhaps in the model of this thesis the Sentence Boundary Condition would not be necessary. He points out also, however, that without it certain wrong predictions would have to be accounted for in some other way. For instance, assuming for English that NEG (= *not*) is VP-initial, then the process illustrated in (i) would be a POSSIBLE rule.



The movement of $\begin{bmatrix} \text{SPEC} \\ +\text{NEG} \end{bmatrix}$ would reorder the whole NP by the Revised Left Branch Condition. The result,



would be weakly structure preserving (NP replaced $\text{NEG}[\Delta]$ where NP dominated $[\text{+NEG}]$), but (iii) is ungrammatical:

(iii) *He said she would no man invite to the party.

With the Structure Preserving Hypothesis including the Sentence Boundary Condition this would not be a possible rule because it is a substitution but would not qualify as structure preserving, root, or local.

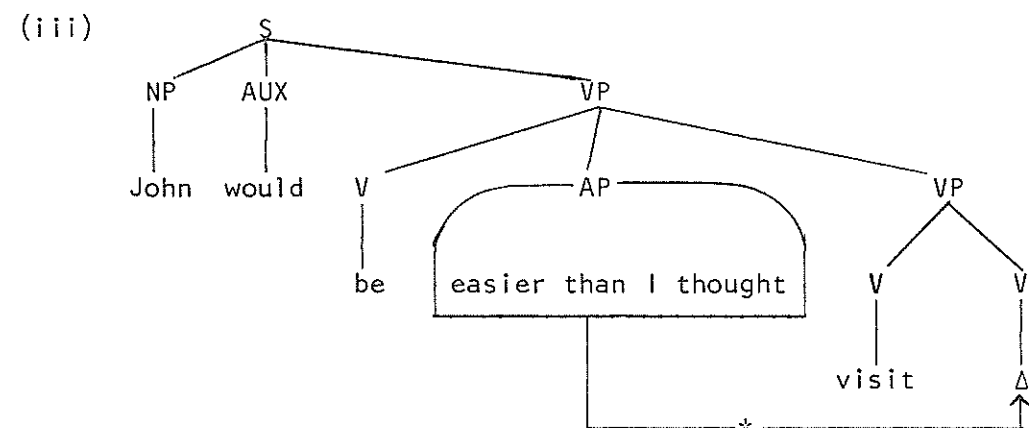
¹⁴As Joe Emonds has suggested to me, it seems that this rule can be conflated with the rule for postposing AP's. Stated as:

$$(i) \begin{Bmatrix} N \\ A \end{Bmatrix} + C^{\max} - \text{VP}[\Delta] = \emptyset - 1$$

it would also account for

(ii) The chicken smelled $\boxed{\text{to me } \text{AP}[\text{worse}]}$ than the fish did].

Properly blocked, by the V in the grossest analysis, would be:



¹⁵For the proposals which follow I am indebted to the influence of Noam Chomsky via recent publications, class lectures, and personal communication.

¹⁶For rules which insert (or delete) morphological material (e.g. passive) there will have to be some additional consideration, as there will have to be for rules like WH-fronting which might have to be considered weakly structure preserving. See footnote 13 above.

¹⁷Notice that where a whole phrase is reordered by use of the Revised Left Branch Condition it is still only the term actually mentioned in the structural description which is relevant for the VIC. WH-fronting is not blocked when an NP, for instance, occurs in the grossest analysis even though the movement of $\begin{bmatrix} \text{DET} \\ +\text{WH} \end{bmatrix}$ actually involves the reordering of an NP. This distinction between movements of left-branch grammatical formatives and heads of phrases does not, intuitively, seem implausible.

¹⁸This mentioning of the context term NP represents a weakening of the condition on nonvariables suggested in Chapter 2, section 2.3. For relevant discussion see, for example, Lasnik and Kupin (1976).

¹⁹The suggestion to attempt to collapse particle movement and directional adverb preposing into one rule was originally made to me by Joe Emonds. For discussion of the rule of particle movement see Emonds (1972) and for discussion of directional adverb preposing see Jackendoff (1973) and Hendrick (1976).

²⁰Perhaps there must be a better feature also because certain particles, for instance *on* or *about*, don't occur in directional adverbs.

²¹The PP in (32a) is properly contained in VP' since particles are very closely associated with verbs. They are obligatorily strictly subcategorized. Where the verbs they can co-occur with appear alone they have a different meaning.

BIBLIOGRAPHY

- Baltin, Mark. 1976. On the notation for movement rules. Manuscript, MIT.
- Berman, Arlene. 1973. A constraint on *tough*-movement. *CLS* 9.34-43.
- Brame, Michael K. 1976. *Conjectures and Refutations in Syntax and Semantics*. New York: Elsevier North-Holland.
- Bresnan, Joan. 1971. Sentence stress and syntactic transformations. *Language* 47.257-81.
- _____. 1972. *Theory of Complementation in English Syntax*. Doctoral dissertation, MIT.
- _____. 1976a. On the form and functioning of transformations. *Linguistic Inquiry* 7.3-40.
- _____. 1976b. Toward a realistic model of transformational grammar. Paper presented at Convocation on Communications, MIT.
- _____. 1976c. Nonarguments for raising. *Linguistic Inquiry* 7.485-501.
- _____. 1976d. Variables in the theory of transformations. Paper presented at MSSB Irvine Conference on Formal Syntax, revised version. To appear in *Formal Syntax*, ed. by P. W. Culicover, T. Wasow, and A. Akmajian. New York: Academic Press.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- _____. 1962. The logical basis of linguistic theory. *Proceedings of the Ninth International Congress of Linguistics*. The Hague: Mouton, pp. 914-1008.
- _____. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.
- _____. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- _____. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, ed. by S. Anderson and P. Kiparsky. New York: Holt, Rinehart & Winston, pp. 232-86.
- _____. 1976. On WH-movement. Paper presented at MSSB Irvine Conference on Formal Syntax. To appear in *Formal Syntax*, ed. by P. W. Culicover, T. Wasow, and A. Akmajian. New York: Academic Press.
- _____ and Howard Lasnik. 1977. Filters and control. Manuscript, MIT.

- Culicover, Peter and Kenneth Wexler. 1973a. An application of the Freezing Principle to the dative in English. Social Sciences Working Paper, no. 39, University of California, Irvine.
- _____. 1973b. Three further applications of the Freezing Principle in English. Social Sciences Working Paper, no. 48. University of California, Irvine.
- _____. 1976. Some syntactic implications of a theory of language learnability. Paper presented at MSSB Irvine Conference on Formal Syntax. To appear in *Formal Syntax*, ed. by P. W. Culicover, T. Wasow, and A. Akmajian. New York: Academic Press.
- Dean, Janet. 1967. Noun phrase complementation in English and German. Manuscript, MIT.
- Dougherty, Ray. 1970. Recent studies on language universals. *Foundations of Language* 6.505-61.
- Emonds, Joseph. 1970. *Root and Structure-Preserving Transformations*. Doctoral dissertation, MIT.
- _____. 1972. Evidence that indirect object movement is a structure-preserving rule. *Foundations of Language* 8.546-61.
- _____. 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.
- Erteschik, Nomi. 1973. *On the Nature of Island Constraints*. Doctoral dissertation, MIT. (Formerly distributed by Indiana University Linguistics Club.)
- Fiengo, Robert and Martin R. Gitterman. N.d. Remarks on French clitic order. Manuscript, City University of New York.
- Fillmore, Charles. 1965. *Indirect Object Constructions in English and the Ordering of Transformations*. The Hague: Mouton.
- Halle, Morris, Alan Prince, and Jean-Roger Vergnaud. 1976. Formal Phonology (FP) Colloquium handout, colloquium presented by Jean-Roger Vergnaud, MIT.
- Hendrick, Randall. 1976. Prepositions and the X' theory. Proposals for Semantic and Syntactic Theory. *UCLA Papers in Syntax* 7.95-122.
- Higgins, F. R. 1973. On J. Emonds' analysis of extraposition. In *Syntax and Semantics*, vol. 2, ed. by J. Kimball. New York: Seminar Press.

- Jackendoff, Ray. 1971. Gapping and related rules. *Linguistic Inquiry* 2.21-36.
- _____. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.
- _____. 1973. The base rules for prepositional phrases. In *A Festschrift for Morris Halle*, ed. by S. Anderson and P. Kiparsky. New York: Holt, Rinehart & Winston, pp. 345-56.
- _____. 1974. Introduction to the \bar{X} Convention. Bloomington: Indiana University Linguistics Club.
- _____. 1976. Lexicalist Syntax. Manuscript.
- Kayne, Richard. 1969. On the inappropriateness of rule features. *Quarterly Progress Report* 95. MIT Research Laboratory of Electronics.
- Koster, Jan. 1976. Subject sentences don't exist. Manuscript, to appear in *Linguistic Inquiry* Monograph no. 3.
- Kuroda, Sige-Yuki. 1965. *Generative Grammatical Studies in the Japanese Language*. Doctoral dissertation, MIT.
- _____. 1969. English relativization and certain related problems. In *Modern Studies in English: Readings in Transformational Grammar*, ed. by D. Reibel and S. Schane. Englewood Cliffs, N.J.: Prentice-Hall, pp. 264-87.
- Lasnik, Howard. 1976. Remarks on coreference. *Linguistic Analysis* 2.1-22.
- _____. and Joseph H. Kupin. 1976. A restrictive theory of transformational grammar. Manuscript, University of Connecticut.
- Lightfoot, David. Forthcoming. The theoretical importance of subject raising. Manuscript. To appear in *Foundations of Language*.
- McCawley, Noriko. 1977. Why Tough-Movement is impossible with *possible*. Manuscript.
- Oehrle, Richard. 1976. *The Grammatical Status of the English Dative Alternation*. Doctoral dissertation, MIT.
- Pope, Emily. 1972. *Questions and Answers in English*. Doctoral dissertation, MIT. (Formerly distributed by Indiana University Linguistics Club.)
- Postal, Paul. 1974. *On Raising*. Cambridge: MIT Press.



- Reinhart, Tanya. 1976. *The Syntactic Domain of Anaphora*. Doctoral dissertation, MIT.
- Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge: MIT Press.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Bloomington: Indiana University Linguistics Club.
- Schachter, Paul. 1973. Focus and relativization. *Language* 49.19-46.
- _____. 1976. A non-transformational account of gerundive nominals in English. *Linguistic Inquiry* 7.205-41.
- Schwartz, Arthur. 1972. Constraints on transformations. *Journal of Linguistics* 8.35-86.
- Selkirk, Lisa. 1976. Some remarks on noun phrase structure. Paper presented at MSSB Irvine Conference on Formal Syntax. To appear in *Formal Syntax*, ed. by P. W. Culicover, T. Wasow, and A. Akmajian. New York: Academic Press.
- Stockwell, Robert, Paul Schachter and Barbara Partee. 1973. *Major Syntactic Structures of English*. New York: Holt, Rinehart & Winston.
- Vergnaud, Jean-Roger. 1976. Formal properties of phonological rules. Manuscript University of Paris VII.
- Wasow, Tom. 1976. Transformations and the lexicon. Paper presented at MSSB Irvine Conference on Formal Syntax. To appear in *Formal Syntax*, ed. by P. W. Culicover, T. Wasow, and A. Akmajian. New York: Academic Press.
- Wexler, Kenneth, and Peter Culicover. 1973. Two applications of the Freezing Principle in English. Social Sciences Working Paper, no. 38. University of California, Irvine.
- _____. In preparation. *Formal Principles of Language Acquisition*.
- Wilkins, Wendy. 1976. On the proper interpretation of variables in syntactic transformations. *Proposals for Semantic and Syntactic Theory. UCLA Papers in Syntax* 7.123-45.

The articles in this volume are meant to be tongue-in-cheek and therefore should not be taken as attacks on any individual or group of individuals in linguistics or any other line of work.

The Editors