DELETION AND LOGICAL FORM

by

IVAN ANDREW SAG

B.A., University of Rochester
1971

M.A., University of Pennsylvania
1973

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

at the

MASSACHUSETTS INSTITUTE OF
TECHNOLOGY

June, 1976

Signature of Author...
Department of Foreign Literatures and
Linguistics

Certified by...
Thesis Supervisor

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

Deletion and Logical Form

Ivan Andrew Sag

Submitted to the Department of Foreign Literatures and Linguistics on June 22, 1976 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

This thesis incorporates a detailed study of two deletion rules in English: Verb Phrase Deletion and Gapping. It establishes a fundamental connection between the familiar notion of the recoverability of deletion and matters of logical form.

Chapter One is an investigation of the syntax of Verb Phrase Deletion. Previous proposals for the constituent structure of VP and AUX are examined and rejected in favor of one that accommodates a proper formulation of that rule.

In the second chapter, it is argued that logical considerations play an important role in determining the applicability of such rules as Verb Phrase Deletion. A theory of logical form is sketched, as well as a theory of the recoverability of deletion. Those theories are justified on the basis of many new facts about the interaction of deletion and quantifier interpretation and certain well-known facts such as the problem of "sloppy identity".

Chapter Three concerns the rule of Gapping. A reformulation of that rule is offered in light of various inadequacies of previous proposals that are observed. A revision of the A-over-A Principle is also proposed, as well as a rather novel proposal for its interaction with the principle of the recoverability of deletion. The logical form of Gapped sentences is also examined.

The last chapter justifies the basic approach of the first three. A "mixed" theory of anaphora utilizing both deletion rules and interpretive rules is seen to be preferable to recent proposals which have attempted to treat all anaphoric processes interpretively.

Thesis Supervisor: Noam Chomsky
Title: Institute Professor of Linguistics

To My Parents

for both love

and friendship

Preface

This all began, strangely enough, as an investigation of English prosody and its relation to syntactic structure. Very early on, it became clear that a chapter would have to be set aside to study the relation between prosody and ellipsis. For some reason that still remains a mystery, I decided to write that chapter first. The most interesting thing about ellipsis, I soon discovered, was not its prosody, but rather its interaction with matters of syntax and matters of logical form. That, therefore, is what this essay is primarily about.

A have assumed the framework of transformational grammar as it has been developed by Noam Chomsky and others. I'd like to think, however, that most of the observations and even the conclusions will be of interest to those who do not embrace that framework. My primary concern in this thesis is to raise certain fundamental questions about the nature of ellipsis. I hope it will not seem that I am overly concerned with purely notational matters.

I am very fortunate to have had a thesis committee consisting of Noam Chomsky, Ken Hale, and Hans Kamp. Each of them has provided careful guidance, sound criticism, and constant encouragement. To my thesis supervisor, Noam Chomsky, I am particularly indebted for his extreme patience and his willingness to discuss things with me for hours on end. Few, if any, of the ideas contained herein would have ever been developed were it not for him.

I am also extremely grateful to Jorge Hankamer. Without the benefit of the many hours spent with him discussing the nature of deletion rules and other anaphoric processes, I surely would have been in no position to even begin this investigation. His detailed comments on Chapter Four were also of extreme value to me.

Others who have given freely of their time, and who I have profited greatly from discussions with, include Sylvain Bromberger, Larry Horn, Susumu Kuno, Phil LeSourd, Mark Liberman, Haj Ross, and Erich Woisetschlaeger. In addition, a week in Vermont, spent discussing the contents of my thesis with Geoff Nunberg proved to be extremely valuable.

After many of my ideas had crystallized, I benefited greatly from conversations with Edwin Williams, who had independently come up with many of the same ideas.

Without Barbara Partee's 1974 Linguistic Institute course in Montague Grammar, I don't think I could have undertaken the topic that I did. A better thesis would have no doubt resulted, had I discussed it with her as I went along.

Occasional conversations with the following people were also crucial in making this enterprise a reality: Barbara Abbott, Guy Carden, David Duncan, Bob Fiengo, Lee George, John Goldsmith, Morris Halle, Ray Jackendoff, George Lakoff, Terry Langendoen, Barbara Lust, David Nash, David Perlmutter, Tanya Reinhart, and Annie Zaenen.

There are still others who provided necessary encouragement before I came to MIT. These include George Cardona, Henry Hiż, Antanas Klimas, Bill Labov, and Arnold Zwicky. In addition, I am grateful to O.L. Chavarria-Aguilar, who is the one who got me interested in theoretical linguistics in the first place. Without his prodding, I would probably still be hustling rock gigs in upstate New York.

I am indebted for personal support and encouragement to Lois Betz. My debt to her for that far exceeds the one accrued for her technical assistance in preparing and typing the final draft of this thesis.

Mary-Louise Kean was kind enough to read the final draft and provide further editorial assistance. Thanks go to her for enjoyable conversations as well.

Perhaps my greatest debt, however, is the one I owe my friends who put up with me and my state of mind as I actually wrote this thing. If I had been in their position, I would have tolerated far less. They are: Annie, Ava, Barbara, Betsy, Carol, Fran, Jill, Judy, Jim, Margie, Norma, Peter, Ronnie and Wilson. Any similarity between these names and those that appear in the example sentences of this thesis is not surprising.

Table of Contents

Abbreviations

AE	Auxiliary Ellipsis	§ 1.2
A + W	Akmajian and Wasow (1975)	§ 1.2
BAC	The Backwards Anaphora Constraint	§ 4.2
CD	Comparative Deletion	§ 2.3
CE	Comparative Ellipsis (1 and 2)	§ 2.3
ESH	(Wasow's) Empty Structures Hypothesis	§ 4.0
IDP	The Immediate Domination Principle	§ 3.3
MDP	The Minimal Distance Principle	§ 3.2
NAC	(Hankamer's) No-Ambiguity Condition	§ 3.2
NCA	(Sag and Hankamer's) Null Complement Anaphora	§ 4.2
NLPNPC	(Langendoen's) Non-Left-Peripheral Noun Phrase Constraint	§ 3.2
N^2PHA	Null NP-Head Anaphora	§ 4.3
PRO → BV	(name of a rule which converts) Pronouns (into) Bound Variables	§ 2.2
Q	quantifier	§ 1.2
RAOAP	(Bresnan's) Relativized A-over-A Principle	§ 1.2
RSSR	(Kuno's) Requirement for Simplex-Sentential Relationship	§ 3.2

SAI	Subject-Auxiliary Inversion	§1.2
TSPI	(Kuno's) Tendency for Subject-Predicate Interpretation	§3.2
VPD	Verb Phrase Deletion	§1.1

CHAPTER ONE

The Syntax of Verb Phrase Deletion

1.1 Syntactic Overview

It is well known that English verb phrases are often left unexpressed. The standard examples are those like the following.

(1.1.1)(a) Gwendolyn made the team but Betsy didn't ∅.

[∅ = make the team]

(b) John loves Mary, and Peter does ∅, too.

[∅ = love Mary]

(c) Either Betsy wrote on the blackboard, or else Sandy did ∅.

[∅ = write on the blackboard]

From the earliest work on transformational grammar (Chomsky (1955), Harris (1957, 1964)), such sentences have been presumed to arise from the application of a transformation that deletes verb phrases, i.e. that deletes the bracketed material in the examples in (1.1.1). This is the familiar transformation of Verb Phrase Deletion (henceforth VPD).

Several recent attempts have been made to eliminate this transformation (Wasow (1972), Fiengo (1974), Williams (1976)), replacing it with an interpretive rule that operates to "interpret" sentences containing base-generated empty nodes. These proposals have actually been quite general, and concern anaphoric processes other than VPD as well. For this reason we will postpone our discussion of them until Chapter Four, where we will conclude that with respect to a coherent overall theory of anaphora, the deletion analysis is preferable to an interpretive one for the VPD phenomenon. Thus we will couch our discussion not only in the framework of standard transformational grammar, but also in terms of a specific version of that theory that employs deletion transformations.

The examples in (1.1.1) all involve Chomsky's (1957) rule of Do-Support. This rule inserts the morpheme do after "stranded" tense morphemes (base-generated in the auxiliary) which, for one reason or another, have no immediately following verb to attach to. VPD is thus one rule that creates inputs for the rule of Do-Support (and subsequent attachment of a preceding tense morpheme to the inserted do by "Affix-hopping" (Chomsky (1957)). Other elements of the auxiliary node may also precede a VP-Deletion site, as the following examples show:

(1.1.2)(a) John won't go to the store, but Bill will ∅.

[∅ = go to the store]

(b) Peter can hit a home run, and Betsy can ∅, too.

[∅ = hit a home run]

(c) Sandy should go to Boston, and Betsy should ∅, too.

[∅ = go to Boston]

When an element of the auxiliary is also deleted (thus in certain cases blocking Do-Support), as in the following examples, the deletion is demonstrably the result of a different rule, namely Gapping, which is the subject of Chapter Three.

(1.1.3)(a) Betsy will go home at four, and Sandy ∅ at five.

[∅ = will go home]

(b) Peter got on base in the first inning, and Betsy ø in the second inning.

[ø = got on base]

VPD can also apply in tenseless clauses, as in the following example.

(1.1.4) Betsy wanted to go home, but Peter didn't want to ø.

[ø = go home]

When it does, however, the infinitive marker <u>to</u> always precedes the deletion site. Since this morpheme otherwise occurs in auxiliary position in tenseless clauses:

(1.1.5) We wanted very much for Betsy to stay,

it seems reasonable to assume that it is in fact generated under AUX. It therefore also seems reasonable to conclude that a VP Deletion site must immediately follow AUX. We will have more to say about this a little later.

All the above examples have been instances of VPD in coordinate clauses. VPD also operates in subordinate clauses, and across speakers in discourse:

Subordinate Clauses:

(1.1.6)(a) Gwendolyn hit a single after Sandy did ø.

[ø = hit a single]

(b) The fact that Betsy said she didn't break the window made me wonder who did ø.

[ø = break the window]

In Discourse:

(1.1.7)(a) Speaker A: Who hit the home run?

Speaker B: Betsy did ∅. [∅ = hit the home run]

(b) Speaker A: Why did you go home?

Speaker B: Because John did ∅. [∅ = go home]

Additionally, VPD operates into relative clauses, flagrantly violating Ross's (1967a) Complex Noun Phrase Constraint:

(1.1.8) John didn't hit a home run, but I know a woman who did ∅.

[∅ = hit a home run]

Equally flagrantly, VPD violates Ross's Sentential Subject Constraint:

(1.1.9) That Betsy won the batting crown is not surprising, but that Peter didn't know she did ∅ is indeed surprising.

[∅ = win the batting crown]

With respect to Ross's Coordinate Structure Constraint, VPD is somewhat more respectful. Adapting the terminology of Grosu (1973), who distinguishes two subcases of this constraint, we can say that VPD freely violates the Element Constraint (that concerns affected elements within conjuncts):

(1.1.10) Peter never hit a home run, but Betsy did ∅ and she was very happy about it. [∅ = hit a home run],

but obeys the Conjunct Constraint (that concerns affected elements that are themselves conjuncts):

(1.1.11) *I couldn't lift this rock, but I know a boy who can ø and bend a crowbar, too. [ø = lift this rock]

(For more discussion of such matters, see Neubauer (1970), Neeld (1973), Grosu (1973, 1975, to appear).)

Another property of VPD is that it sometimes operates "backwards", i.e. right to left, as in the following examples.

(1.1.12)(a) Although Sandy said she didn't ø, Besty actually did go to the store. [ø = go to the store]

(b) Because Betsy didn't ø, Sandy went to the grocery store (in her place). [ø = go to the grocery store]

Since it is arguable that the adverbial clauses in these examples are fronted via transformation, one might be tempted to argue that VPD applies before such fronting rules, and hence, need never apply "backwards". Alas, such an argument is untenable, as is shown by the following examples of right-to-left VPD, which do not involve fronting rules.

(1.1.13)(a) The fact that Peter said she did ø doesn't mean that Betsy actually went to the movies.

[ø = go to the movies]

(b) The man who said she didn't ø, knew that Betsy really did go to the movies.

[ø = go to the movies]

(c) Anyone who can ø should go to see this movie.

[ø = go to see this movie]

A proper formulation of VPD must therefore allow bi-directional application.

"Backwards" VPD seems to obey well-motivated constraints, however. In particular, sentences like the following indicate that the Backwards Anaphora Constraint, whatever its proper formulation (Ross (1967a, 1969a), Langacker (1969), Reinhart (1974), Lasnik (1976), Sag and Hankamer (1976)), affects VPD.

(1.1.14)(a) *Betsy did ø after Peter went to the store.

[ø = go to the store]

(b) *Betsy didn't ø when Lois told her to get off the phone.

[ø = get off the phone]

This matter is actually somewhat more complex, and we will return to it in Chapter Four.

There are several other syntactic properties of VPD that are of relevance. For ease of exposition, however, we will now turn to the various formulations of VPD that have been given in the literature, and we will bring these various additional properties to the fore as they become relevant.

1.2 Theories, Objections, and a Proposal

The first explicit formulation of the VPD transformation that I am aware of is by Bouton (1970), though his formulation is pretty clearly what many people before him had in mind. He writes the following rule, which I will henceforth refer to as the "standard" formulation of VPD.

(1.2.1) Verb Phrase Deletion (optional)

	X	- VP	- Y	- VP	- Z	
S.D.:	1	2	3	4	5	
S.C.:	1	2	3	ø	5	condition: 2 = 4

The condition "2 = 4" is intended as an identity condition. In order for a VP to be deletable, there must be a syntactically identical VP present elsewhere (if we take this formulation literally, earlier) in the sentence. This condition invokes the standard notion of <u>recoverability of deletion</u>, which we will examine in more detail in the next chapter. For our purposes here, let it suffice to note that this notion, as it is normally thought of, amounts to requiring that a given element (we will follow Hankamer in referring to such entities as <u>deletion targets</u>) can be deleted only in the context of another element which it is identical to (which we will call the <u>deletion trigger</u>).

Various problems with this notion of identity have been pointed out in the literature. It is generally assumed, following Chomsky (1965), that a somewhat weaker notion of "non-distinctness" is in fact the relevant notion.

We will return to such matters at length in the next chapter. What is of immediate interest here, with respect to VPD, is that there is a further kind of non-distinctness that is relevant, namely identity of verbal affixes. This matter has been studied in considerable detail by Halliday and Hasan (1973) and by Quirk et al. (1972). The latter investigators cite the following paradigm as evidence that VPD does not in general require affixal identity. (Quirk et al. (1972, pp. 580-581).)

(1.2.2)(1) Present and Modal

John <u>understands</u> the situation and surely Peter should ∅.

[∅ = <u>understand</u> the situation]

His friends already <u>belong</u> to the club and he will ∅ soon.

[∅ = <u>belong</u> to the club]

(2) Past and Modal

Bob entered the competition and Paul may ∅.

[∅ = enter the competition]

(3) Progressive and Modal

Peter is complaining about the noise, but John won't ∅.

[∅ = complain about the noise]

(4) Perfect and Modal

John hasn't met my brother yet, but (he) will ∅ soon.

[∅ = meet my brother]

(5) Progressive and Perfect

John may be questioning our motives, but Peter hasn't ∅.

[∅ = questioned our motives]

(6) Past and Perfect

Peter saw your parents last week, but he hasn't ∅ since.

[∅ = seen your parents]

Paul apologized, but Bob won't have ∅.

[∅ = apologized]

The one exception they note[1] to this general disregard for affixal identity concerns active and passive VP's. If, say, the trigger VP is active, and the target VP is passive, VPD is impossible:[2]

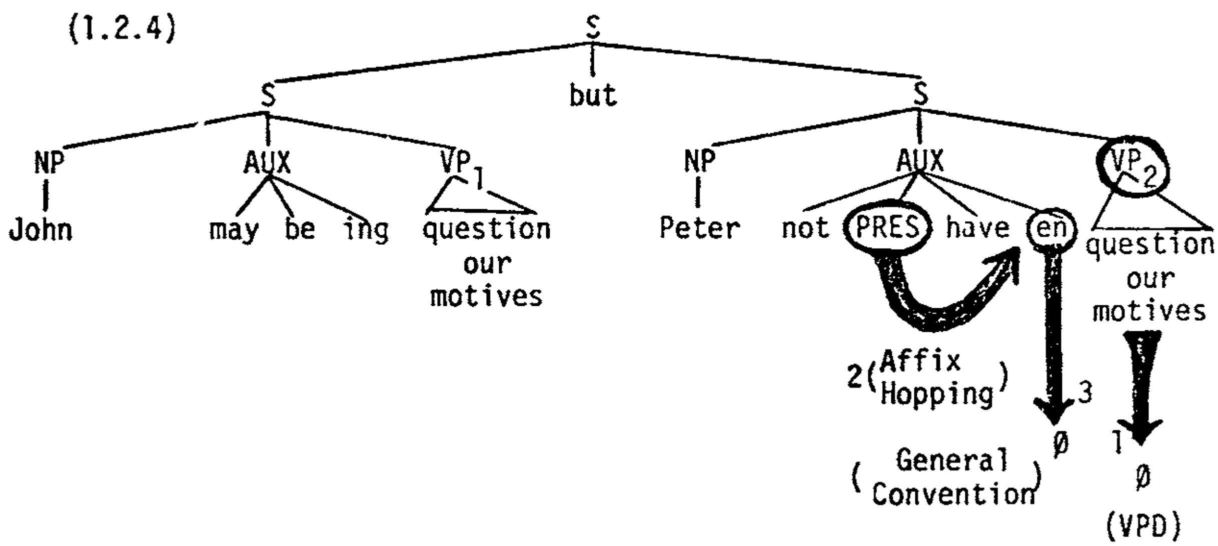
(1.2.3)(a) *Paul denied the charge, but the charge wasn't ∅ by his friends. [∅ = denied]

(b) *John had observed many of the enemy's soldiers, but hadn't been ∅ by them. [∅ = observed]

This is a curious state of affairs. If we are to subsume affixal non-distinctness within our general (metatheoretical) notion of recoverability of deletion, why are Quirk et al.'s examples in (1.2.3) ungrammatical?

Notice that one could account for the cases in (1.2.2) by ordering VPD before Affix Hopping, and deleting, by convention, stranded affixes (other than tense). If deletion takes place before Affix Hopping, exact identity of the trigger and target VP holds at the time VPD applies. This is illustrated in (1.2.4) for (1.2.2)(5).

(1.2.4)



This solution, by the way, would account for all the facts noticed by Akmajian and Wasow (1975) in their discussion of the "ordering paradox" concerning VPD, Affix Hopping and Do-Support. They do not give a discussion of such alternatives in their paper. Be that as it may, though it is clear that this kind of a solution comes close to accounting for the facts observed by Quirk et al., it still fails to explain the ungrammaticality of (1.2.3)(a) and (b).

In fact, I see no solution to that problem forthcoming from any of the current views of how VPD works. We will, however, return to this problem with a solution in Chapter Two.

Another general problem with the standard formulation of VPD concerns the left-hand context of the target VP. Bresnan (1976a) considers this problem, noting the following contrast.

(1.2.5) *First people began pouring out of the building, and then smoke began ∅. [∅ = pouring out of the building]

(1.2.6) First people began to pour out of the building, and then smoke began to ∅. [∅ = pour out of the building]

(1.2.5) is ungrammatical, Bresnan argues, because the target VP is not immediately preceded by AUX. To account for this contrast (as well as for other reasons that invoke her extremely interesting notion of a "relativized" A-over-A Principle which we will turn to in a moment), she proposes the following modification of the standard formulation of VPD:

(1.2.7) Verb Phrase Deletion (Bresnan)

$$\begin{array}{ccccccccc} [_S W & - \text{AUX} & - \text{VP} & - X] & \text{-CONJ-} & [_S Y & - \text{AUX} & - \text{VP} & - Z] \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \Rightarrow 1 & 2 & 3 & 4 & 5 & 6 & 7 & \emptyset & 9 \end{array}$$

Although I am certainly in sympathy with Bresnan's idea, it is not without problems. First of all, although there is certainly a need to restrict the left-hand environment of the target VP, any such constraint on the left-hand environment of the trigger VP has undesirable consequences, as the following sentences show.

(1.2.8)(a) *Harry seems (to be) upset, but Bill doesn't seem ∅.

[∅ = upset]

(b) Harry seems upset, but Bill doesn't seem to be ∅.

[∅ = upset]

VPD cannot apply in (1.2.8)(a) because the VP is not preceded by AUX. VPD can apply in (1.2.8)(b), even though the trigger VP is not preceded by AUX.

Secondly, as we noted earlier, VPD is not restricted to coordinate structures, or even to subordinate clause structures _cum_ conjunction as Bresnan's rule would imply. Consider the following instances of VPD that are not analyzable by Bresnan's formulation.

(1.2.9)(a) Although John went to the store, Betsy didn't ∅.

[∅ = go to the store]

(b) Anyone who asks Sandy to hit a home run knows she will ∅.

[∅ = hit a home run]

These two difficulties are easily handled by the following modified version of VPD.

(1.2.10) X - VP - Y - AUX - VP - Z

	1	2	3	4	5	6
⇒	1	2	3	4	∅	6

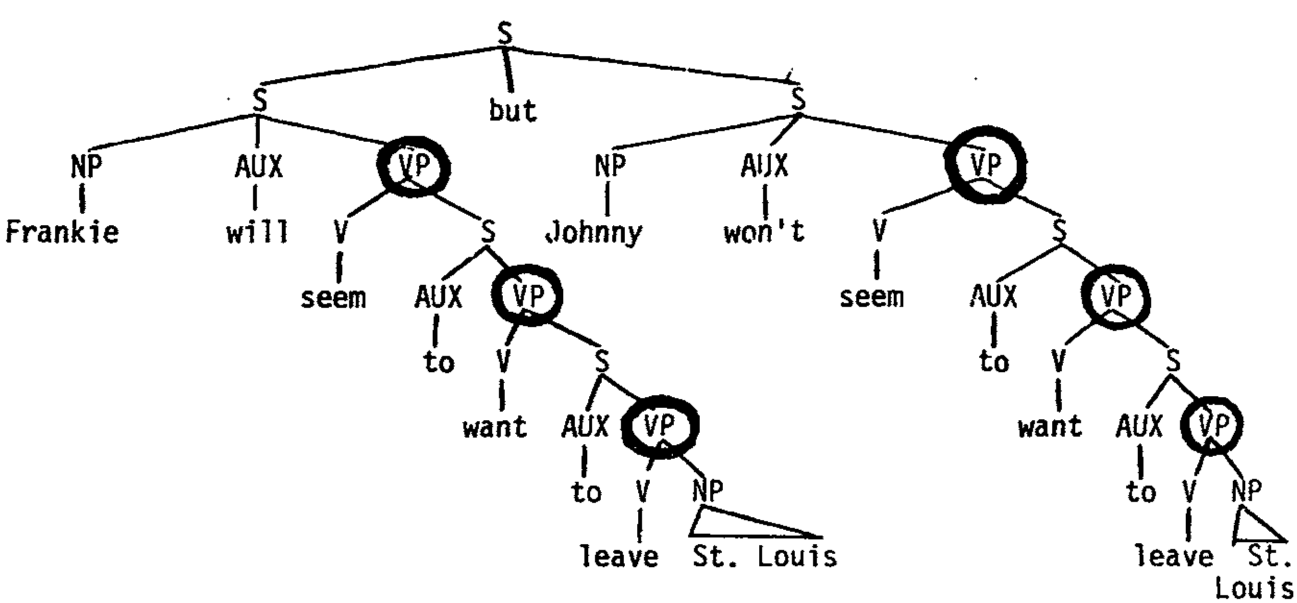
However this still does not solve the problem of non-distinctions of affixes noted above, nor does it provide for backwards deletion.

Stipulating AUX as a context term in the SD of this transformation sets the stage for a discussion of Bresnan's Relativized A-over-A Principle (henceforth RAOAP). The basic idea is as follows.

Certain empirical inadequacies of the earlier formulation of the A-over-A principle (Chomsky (1964, 1968, 1973)), can be remedied if one does not require absolute maximization of a proper analysis of a transformation, but rather, require that "a transformation T = (C,M) [Bresnan's entire discussion draws heavily from the formalization of transformational grammar found in Peters and Ritchie (1973) - I.A.S.] apply under a proper analysis Π which is maximal relative to all proper analyses which agree with Π on all context predicates in C." [3] In other words, the target predicate of a transformation must be maximal, but only with respect to particular fixed values of the context predicates.

With respect to VPD, this means that any auxiliary can be assigned to the context predicate AUX, and relative to that assignment, the following target predicate VP must have maximal value. Thus in the following phrase structure tree, each of the circled VP's is maximal relative to the auxiliary that immediately precedes it. RAOAP, therefore, correctly allows for three proper analyses, and, hence, for the three possible deletions in (1.2.12).

(1.2.11) [= Bresnan's (31)]



(1.2.12) [= Bresnan's (32)-(34)]

(a) Frankie will seem to want to leave St. Louis, but Johnny won't.

(b) Frankie will seem to want to leave St. Louis, but Johnny won't seem to.

(c) Frankie will seem to want to leave St. Louis, but Johnny won't seem to want to.

Our rule (1.2.10), together with Bresnan's RAOAP, seems to provide a reasonable account for the well-known fact that VPD has multiple output possibilities, as in the last three examples. This "well-known fact", however, has recently been challenged by Akmajian and Wasow (1975), who wish to defend the standard formulation of VPD. Their analysis is quite intricate, and hence requires a detailed exposition to which we now turn.

The essence of Akmajian and Wasow's (henceforth A + W) analysis is a rule they call BE-Shift, whose formulation they give (p.220, nt. 9) as the following:

(1.2.13) BE-Shift (Obligatory)

$$\begin{array}{lll} \text{SD:} & \underbrace{\text{Tense (Modal) (Have en)}}_{1} \text{ -} & \underset{2}{\text{be}} \\ \text{SC:} & \quad\quad\quad 1 + 2 & \emptyset \end{array}$$

This rule solves the "paradox" A + W observe with regard to passive be. As they put it:

> passive be acts as though it is part of the VP when the progressive be is present within AUX; but when passive be occurs alone (or with other nonprogressive auxiliaries) it acts as though it is part of the AUX (or at least outside of the VP). [A + W (1975, p. 220]

Notice that BE-Shift has precisely the effect A + W discuss because, as formulated, it cannot apply if there already is another be in AUX.

The facts under consideration in A + W's discussion are the following:

(1.2.14) John said there wouldn't be many people being examined by the doctor, but there were ∅.

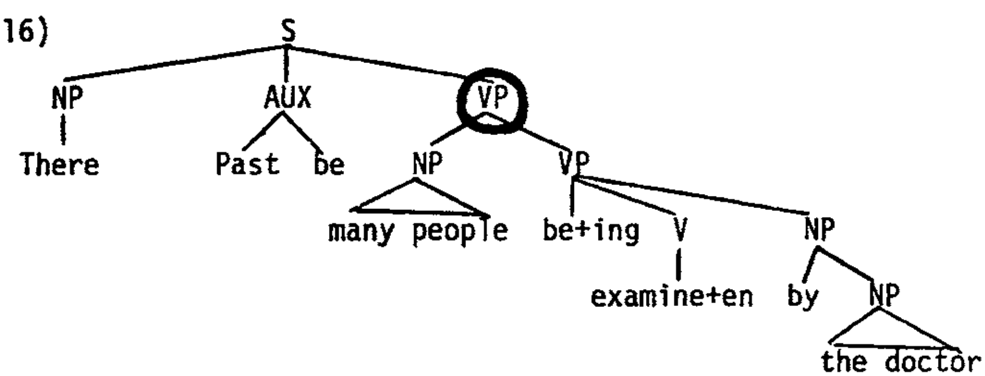
[∅ = many people being examined by the doctor]

(1.2.15) Ford was examined by the doctor, and Nixon was ∅, too.

[∅ = examined by the doctor]

They make two crucial assumptions: (1) Passive applies before BE-Shift and inserts be + en not under AUX, but under the VP, and (2) VP deletion deletes only VP's. Given these assumptions, (1.2.14) shows (1) that There-Insertion adjoins the subject NP to the VP, and (2) that passive be is in the VP at the time VPD applies. They give the following picture of the structure of the right conjunct of (1.2.14) prior to deletion.[4]

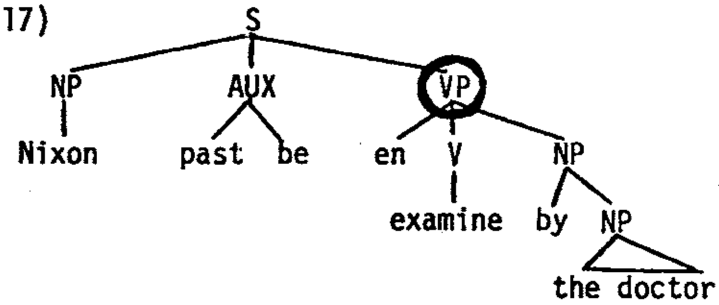
(1.2.16)



The circled VP is deletable.[5]

On the other hand, (1.2.15), given the same assumptions, shows that passive be has been shifted out of VP prior to VPD. The relevant pre-deletion structure is given:

(1.2.17)



The circled VP is again deletable. The rule of BE-Shift, crucially ordered after Passive, but before There-Insertion, accounts for these facts. In (1.2.16), progressive be has undergone BE-Shift (before There-Insertion), but passive be cannot (because both progressive be and the NP many people block shifting of passive be). No such blocking occurs in (1.2.17); passive be has shifted (obligatorily) into AUX.

As further evidence for the rule of BE-Shift, A + W cite the following sort of examples:

(1.2.18) Sam was being examined by a psychiatrist at that time, and

(a) Bill was ∅ too.

(b) *Bill was being ∅ too.

Their analysis predicts such facts, for once progressive be shifts into AUX, passive be cannot, and must remain in the VP. Thus if VPD applies, it must delete passive be (which has in this example received the progressive ing suffix by A + W's rule of EN/ING-Hopping, which they argue is distinct from Tense-Hopping) as well.

A + W's discussion proceeds to examples involving multiple deletion possibilities that they cannot regard as instances of VPD. They cite this pair of sentences, for instance.

(1.2.19)(a) John must have been using drugs, and Bill must have been ∅ too. [∅ = using drugs]

(b) John must have been using drugs, and Bill must have

ø too. [ø = been using drugs]

Since BE-Shift has obligatorily shifted <u>been</u> into AUX prior to deletion, VPD can be held responsible only for (1.2.19)(a). (1.2.19)(b), A + W argue, is the result of another rule, Auxiliary Ellipsis (AE), which "operates after VP-Deletion, and has the effect of deleting the auxiliary verbs which remain after VP-Deletion, up to, but not including, the left-most auxiliary verb." (p. 235). (1.2.19)(b) has thus undergone both VPD and AE, in that order.

This is a very curious rule. Why should there be two separate rules that do such similar things? One kind of justification for a separate rule of AE, of course, would be to show that it has an independent domain of application. As far as I can see, however, the fact that AE cannot be shown to have an independent domain of application remains a grave problem for any theory that claims such a rule exists. Notice that whereas both of the deletions in (1.2.19) are optional, AE must be constrained to apply only if VPD has applied (it is unclear whether A + W have a global constraint of this sort in mind). Otherwise the following ungrammatical string will be produced from the same structure underlying the two sentences in (1.2.19).

(1.2.20) *John must have been using drugs, and Bill must have ø

using drugs too. [ø = been]

A + W do try to motivate one difference between these two rules. They cite the following examples as evidence that AE, but not VPD, cannot apply in tenseless clauses.

(1.2.21)(a) Which bothers you more: John's having been arrested for drug dealing, or Bill's having been $\emptyset$?

(b) *Which bothers you more: John's having been arrested for drug dealing, or Bill's having $\emptyset$?

(1.2.22)(a) Which would bother you more: for John to have been arrested for drug dealing, or for Bill to have been $\emptyset$?

(b) *Which would bother you more: for John to have been arrested for drug dealing, or for Bill to have $\emptyset$?

[the judgement is A + W's]

(1.2.21)(b) is certainly not an acceptable sentence, but most informants are not so quick to reject (1.2.22)(b). Moreover, all informants accept the following sentence, which, in A + W's system would have to result from AE applying in a tenseless clause.

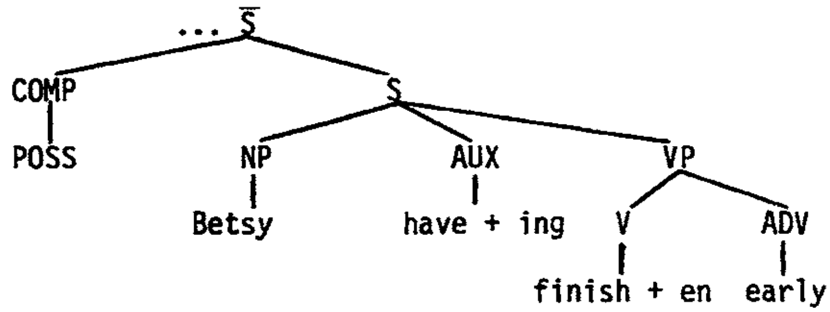
(1.2.23) Peter seems to have been careful, and Bill seems to have $\emptyset$, also.

The primary fact to be accounted for here, seems to by why (1.2.21)(b) is ungrammatical. Here is another example that bears on this.

(1.2.24) *John's having finished early was surprising, but Betsy's having $\emptyset$ came as no surprise.

This (non-)sentence is generated in A + W's system (as well as in any other system we have discussed) not by AE, but rather by VPD. Its assumed structure is sketched in (1.2.24)'.

(1.2.24)'



It seems to be a fact that VPD can never take place immediately after -ing.

Now there are several ways one might state a constraint to account for this fact, none of them particularly elegant (e.g. a surface constraint, an AUX-pruning convention. We will return to this matter in a moment). The crucial point to be made is that any theory of VPD, including A + W's, needs some such constraint.

With respect to A + W's analysis, this has far-reaching consequences. The necessity of this constraint undermines their earlier argument for BE-Shift concerning (1.2.18)(b) whose ungrammaticality would follow under any assumptions about the constituent structure of VP and AUX, given the constraint in question.

What's more, such a constraint also undermines A + W's argument that AE differs from VPD in not applying in subordinate clauses. Their example (1.2.21)(b), the only clear example they offer of ungrammatical deletion in a tenseless clause, also involves an -ing immediately preceding the deletion target. The facts of deletion in tenseless clauses, therefore, provide no evidence for A + W's claim that there are two separate deletion rules. Rather, they illustrate a property the two rules would have in common. The prudent conclusion, I would say, is that there is only one rule operating in these examples: VPD.

Consider now these examples, which, A + W claim, provide evidence for "some special restriction" that would have to be stated for AE, but not for VPD:

(1.2.25) If Bill had been using drugs, then his brother Sam

(a) must have been ∅

(b) ?must have ∅

(c) *must ∅

The ? on (1.2.25)(b) is unnecessary, I think (add <u>also</u> and it certainly disappears). The * on (1.2.25)(c) is quite real. The "special restriction" on AE, they claim, is that AE "may delete one auxiliary verb (i.e. the rightmost), but deletion of more than one usually leads to unacceptability" (p. 237). A + W did not consider the following sentences:

(1.2.26)(a) John must have eaten and Bill must $\left\{ \begin{matrix} \text{have } \emptyset \\ {*}\emptyset \end{matrix} \right\}$, too.

(b) Although Mike shouldn't have eaten, Betsy should $\left\{ \begin{matrix} \text{have } \emptyset \\ {*}\emptyset \end{matrix} \right\}$.[6]

These examples make clear that restricting AE so that it "may delete only one auxiliary verb" is the wrong explanation for the ungrammaticality of (1.2.25)(c). In fact, these facts present a grave problem for any formulation of AE, for in (1.2.25)(b) it must delete the rightmost element of AUX, but in (1.2.26), it cannot do so. Nor can AE be restricted to deleting only a third element of AUX because of the following example, where AE would have to delete <u>been</u> as a second element of AUX:

(1.2.27) John has been being hassled by the cops, and Betsy has ∅, too.
[∅ = been being hassled by the cops]

In short, A + W have presented no convincing arguments for a separate rule of AE, and their supposition of such a rule would seem to create more problems than it solves. Nevertheless, certain aspects of their system are quite appealing, and certain of their observations require explanation. We will not pursue this matter, arguing that all the facts noted in this section can be accounted for by assuming a single deletion rule: VPD.

The right account for the ungrammaticality of the sentences in (1.2.26) and (1.2.25)(c), it seems to me, results from the fact that the sequence Modal + have always forms an AUX constituent, presumably base-generated as such. Since our rule of VPD ((1.2.10) above) must analyze an AUX as a left-hand context of a VPD target, it automatically follows that auxiliary have, which is never inside VP, can never be deleted by VPD.[7]

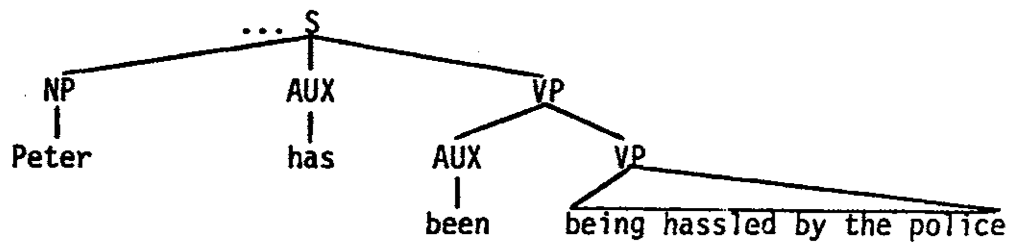
The following multiple deletion possibilities are very telling given our formulation of VPD, for each element that precedes a possible deletion site must be an AUX.

(1.2.28) Betsy has been being hassled by the police, and Peter $\left\{\begin{array}{l}\text{(a) has been } \emptyset \\ \text{(b) has } \emptyset\end{array}\right\}$ too.

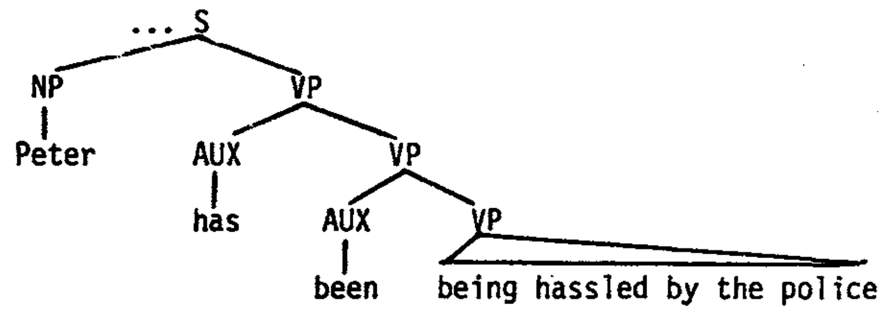
One approach we might take to the multiple output problem illustrated by examples like (1.2.28) is to assume that the optional deltion of been in (1.2.28) is the result of some optional rule that promotes be to AUX from an underlying position in VP. Actually I think it is possible to give a coherent account along these lines, but we will not pursue that matter here, because a much simpler account seems to be possible.

Let us assume that (1.2.28)(a) and (b) are the result of two successful proper analyses of the same terminal string with respect to our rule of VPD. This assumption would require positing one of the following two constituent structures, prior to deletion:

(1.2.29)



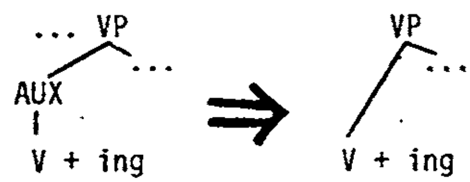
(1.2.30)



Both of these constituent structures (in their essentials) have been bandied about the literature (Ross (1967c), Emonds (1970), Jackendoff (1972, 1976), Kuno (1975a), Grosu (1975, to appear), Burt (1971), Akmajian and Heny (1975)); but no <u>conclusive</u> argument for one rather than the other has to my knowledge ever been given. We will adopt the latter structure, perhaps somewhat arbitrarily. If some argument could be constructed to show that the structure in (1.2.29) is to be preferred, that would not affect our subsequent conclusions. It is clear that the structure in (1.2.30) correctly accounts for the facts of (1.2.28) given our formulation of VPD (and RAOAP).

Let us further make explicit our account of the impossibility of VPD after -<u>ing</u>. Recall that A + W observed the impossibility of VPD after progressive -<u>ing</u>, but, as we observed earlier, the correct generalization seems to be that VPD is never possible after any -<u>ing</u>. One way of accounting for this is to posit a pruning convention that simply says: prune AUX when it dominates -<u>ing</u>. The effect would be this:

(1.2.31)



This automatically accounts for A + W's examples and for (1.2.24) above, for our formulation of VPD requires an AUX before the deletion target.

Another alternative would be to follow the intuitions of Chomsky (1955, §113.2) or Bolinger (1971a, 1971b) about progressive forms in -ing, and those of traditional grammar about gerund and gerundive forms in -ing. The kind of account I have in mind is one that analyzes forms in -ing as belonging to a particular lexical category, say noun, or perhaps (in the case of gerundives) adjective (see also Harris (1946, 1951)). Whether the transfer of grammatical category is done by a morphological rule (see Aronoff (1974), Siegel (1974)), some other lexical rule, or by transformation is not of direct relevance. Suffice it to say that there are many ways one might propose to bring it about that the sequence $\begin{bmatrix}+\text{Verb}\\+\text{AUX}\end{bmatrix}$ +ing is not dominated by AUX while at the same time capturing the intuition that such sequences belong to a more intuitive category. Notice that any treatment along these lines is quite consistent with a theory like that of Fiengo (1974), who eliminates EN/ING Hopping from the syntax entirely. With some such account in mind, let us turn to some additional data.

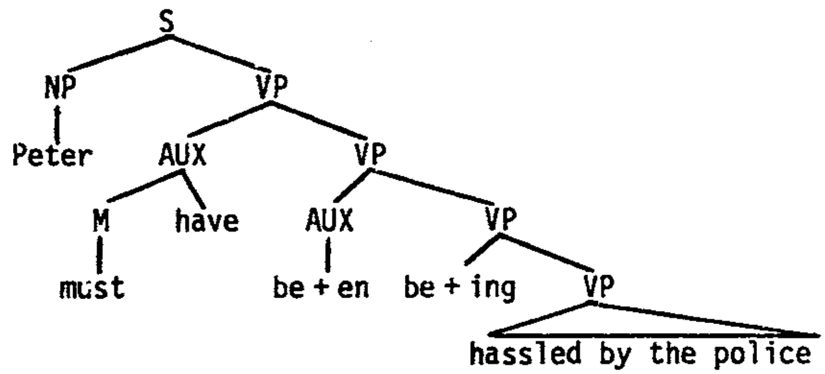
The following facts are also accounted for in our analysis:

(1.2.32) Betsy must have been being hassled by the police, and Peter...

$$\left\{\begin{array}{l}\text{(a) *must } \emptyset\\ \text{(b) must have } \emptyset\\ \text{(c) must have been } \emptyset\\ \text{(d) *must have been being } \emptyset\end{array}\right\}\text{ too.}$$

Prior to deletion (and after pruning of AUX over -ing), the right conjunct of (1.2.32) would have the following structure:

(1.2.33)



In (1.2.32), (a) is impossible because have is not dominated by VP and therefore is undeletable; (b) and (c) are legitimate instances of VPD because each deleted VP is immediately preceded by AUX; (d) is impossible because the AUX-node over be + ing has been pruned (or, alternatively, never existed).[8]

We are now in a position to sketch part of the grammar of English that is emerging from this discussion. We can formulate the following phrase structure rules, keeping in mind that the decision to generate -en and -ing with the auxiliaries that subcategorize them (hence assuming a rule of EN/ING Hopping) is in no way crucial. The decision to base-generate passive be-en isn't crucial either.

(1.2.34)(a) S → NP - VP

(b) VP → AUX - VP (...)

(c) AUX → tense - (M) - (have-en)

(d) AUX → be ($\left\{\begin{matrix}\text{-ing}\\ \text{-en}\end{matrix}\right\}$)

(e) VP → V - (NP) - (PP) (...)

The appropriate ordering of auxiliaries is handled by a subcategorization mechanism. Modals, for instance, might be marked with the feature [-[[be] __]]. Such a mechanism is unobjectionable in principle, though the proposed account is perhaps somewhat less elegant than that of Chomsky (1957). We generate copula and passive be both under AUX. Since AUX is generated under VP, this still guarantees that sequences like be happy, been eaten, be a man are constituents (VP). The fact that these be's are in AUX, however, correctly predicts the following instances of VPD:

(1.2.35) Peter was $\left\{\begin{array}{l}\text{a good baseball player}\\ \text{happy}\\ \text{hassled by the police}\end{array}\right\}$ and

Betsy $\left\{\begin{array}{l}\text{was}\\ \text{must be}\\ \text{must have been}\end{array}\right\}$ ∅, too.

Each of the deleted constituents is a VP immediately preceded by AUX.

If evidence could be provided that at some earlier stage of derivation, copular and passive be should not be in AUX, then we would be forced to assume some form of a "promotion to AUX" rule. I know of no such evidence, however, and will not discuss that matter further.[9]

A + W's formulation of There-Insertion fits into our system nicely:

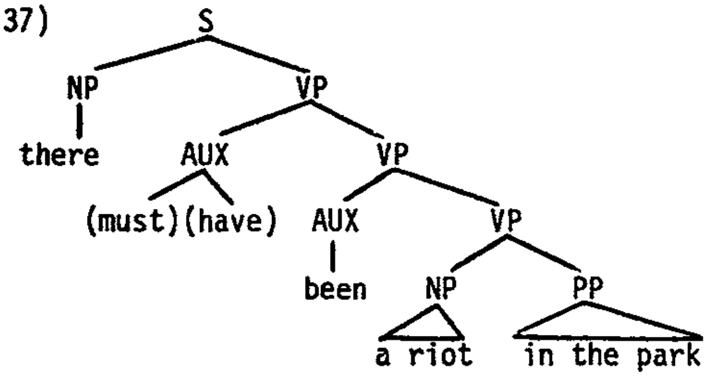
(1.2.36) There-Insertion (optional)[10]

$$\underset{[-\text{DEF}]}{\text{NP}} - \underbrace{\text{Tense (Modal) (have-en)}} - \text{be} - \text{VP}$$

S.D.	1	-	2	-	3 -	4
S.D.	There	-	2	-	3 -	1 + 4

Post-There-Insertion structures have the following general shape.

(1.2.37)



Our rule of VPD thus predicts the array of facts in (1.2.38).

(1.2.38) Peter said there couldn't have been a riot in the park,

but Betsy knew there $\left\{\begin{array}{l}\text{(a) must have been } \emptyset \\ \text{(b) must have } \emptyset \\ \text{(c) *must } \emptyset\end{array}\right\}$

This seems to be just right.

We must also accept A + W's version of Subject-Auxiliary Inversion (S A I), whose structural description contains the term: tense $([^{+V}_{+AUX}])$.[11] That is, because we analyze modals and auxiliary <u>have</u> as daughters of the same AUX when they co-occur, we cannot avail ourselves of the elegant simplification of this term of the rule to AUX, as Jackendoff (1972) proposed. We have not treated Jackendoff's analysis here in any detail because, as A + W show, his analysis provides no basis for handling the facts of VPD.[12]

Notice that since this formulation of SAI fronts elements dominated by AUX, not AUX itself, there is always an AUX node left to the right of the subject NP after SAI. We assume no "pruning" conventions of the sort frequently discussed in the literature. Therefore the grammaticality of such sentences as these, show nothing about the relative ordering of VPD and SAI:

(1.2.39)(a) I'm going home; are you ø?

(b) I've finished; have you ø?

The target VP of VPD would still immediately follow AUX, even after SAI has applied.

One fringe benefit of our analysis is that it provides a basis for a principled account of the possible positions of so-called "floated" quantifiers. Sentences like those in (1.2.40) are generally assumed (Dougherty (1970, 1971), Postal (1974a,1976), Fiengo and Lasnik (1976), Perlmutter and Postal (forthcoming)) to be derived from structures like those that underlie the corresponding sentences in (1.2.41).

(1.2.40)(a) My brothers (all) have (all) finished.

(b) My sisters (all) must (all) be talking to Professor Smith.

(1.2.41)(a) All (of) my brothers have finished.

(b) All (of) my sisters must be talking to Professor Smith.

Either one or two rules bearing various names (Quantifier Floating, Quantifier Movement, Q-Float, etc.) are generally assumed to be responsible for such derivations. Typically, a first rule is supposed to move quantifier words (henceforth Q's) to pre-AUX position, and a subsequent rule reorders Q and a following element of AUX.

The existence of the first rule is rather dubious, as Mark Baltin reminds me (see Baltin (forthcoming)), for sentences like the following would have no source (the non-source is provided).

(1.2.42)(a) the lion, the bear, and the monkey are all mammals. [13]

(b) *all the lion, the bear and the monkey are mammals.

One could, of course, rule out sentences like (1.2.42)(b) by an "output condition" of some kind (Perlmutter (1971), Ross (1972)), but I will not pursue that approach here.

A reasonable alternative approach would be to base-generate Q's in VP-initial position (when an appropriate subject NP precedes; numerous ways of filtering out bad co-occurences are possible), say, as sisters of AUX. This is always a possible position for such Q's.

It does seem reasonable to account for sequences of $\begin{bmatrix}+V \\ +AUX\end{bmatrix}$ - Q as a reordering of Q and the auxiliary verb. Let us take a look at the relevant facts such a reordering analysis is up against.

When there is only one auxiliary verb, it and Q freely reorder:

(1.2.43)(a) They have all arrived.

(b) They are all happy.

(c) They must all leave.

(d) They are all talking.

(e) They are all proctors.

When there are two auxiliary verbs, one reordering is always possible, but a second reordering seems to be possible only with <u>have</u>:

(1.2.44)(a) They may (all) have (all) arrived.

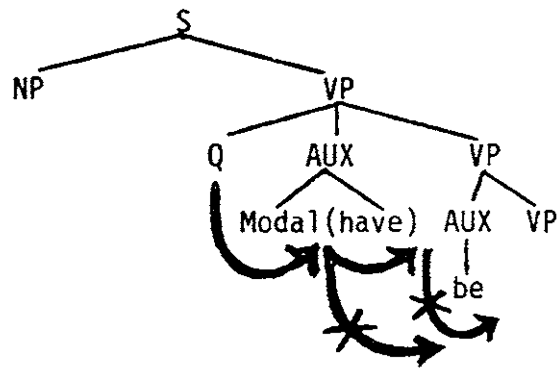
(b) They have (all) been (*all) $\left\{\begin{array}{l}\text{happy} \\ \text{loved} \\ \text{proctors} \\ \text{talking}\end{array}\right\}$

(c) They should (all) be (*all) $\left\{\begin{array}{l}\text{proctors} \\ \text{talking} \\ \text{talked to} \\ \text{happy}\end{array}\right\}$

(d) They are (all) being (*all) $\left\{\begin{array}{l}\text{considerate} \\ \text{talked to} \\ \text{patsies}\end{array}\right\}$

Notice that the constituent structure we have proposed provides exactly the right constituent structure distinctions to state this distribution. We can say that Q moves over one auxiliary verb (hence moving into an AUX) and then freely reorders with any element of that AUX. All the ungrammatical examples in (1.2.44) would in our analysis be instances of Q reordering with respect to an auxiliary in the "next AUX over". A sketch may be helpful:

(1.2.45)



Nor can Q's occur to the right of auxiliary verbs that are in more deeply embedded AUX's:

(1.2.46)(a) *They should have been being all more careful.

(b) *They have been being all very careful.

(c) *They must have been being all hassled by the police.

Interestingly, certain adverbs have the same distribution:

(1.2.47)(a) If they (ever) had (ever) been (*ever) able to get a visa,...

(b) They (simply) must (simply) have (simply) been (*simply) being (*simply) hassled by the police.

(c) If he (only) had (only) been (*only) being (*only) more careful,...

(d) He (merely) has (merely) been (?*merely) walking in the park.

These facts suggest that perhaps Q's should be treated as adverbs. In that event the same rule(s) could perform all these reorderings.

The negative morpheme *not* also partakes of this same distribution; except, of course, it cannot remain before a tensed auxiliary verb.

(1.2.48)(a) Peter (*not) must (not) have (not) been (*not) hassled by the police.

(b) Peter (*not) has (not) been (*not) very helpful.

The rule that obligatorily places *not* after the first tensed auxiliary (NEG-Placement: see Klima (1964), Lasnik (1972)) is presumably a separate rule from the optional rule that reorders Q (and adverbs) with respect to auxiliary verbs. But notice what happens once *not* and a Q or a Q and an adverb wind up inside the same AUX:

(1.2.49)(a) The guys down the street must $\left\{\begin{matrix}\text{not all}\\ \text{all not}\end{matrix}\right\}$ have been [14] hassled by the police.

(b) The guys down the street must have $\left\{\begin{matrix}\text{not all}\\ \text{all not}\end{matrix}\right\}$ been hassled by the police.

(c) *The guys down the street must have been $\left\{\begin{matrix}\text{not all}\\ \text{all not}\end{matrix}\right\}$ hassled by the police.

(1.2.50)(a) If they could $\left\{\begin{matrix}\text{ever all}\\ \text{all ever}\end{matrix}\right\}$ have been standing in the same place...

(b) If they could have $\left\{\begin{matrix}\text{ever all}\\ \text{all ever}\end{matrix}\right\}$ been standing in the same place...

(c) *If they could have been $\left\{\begin{matrix}\text{ever all}\\ \text{all ever}\end{matrix}\right\}$ standing in the same place...

It seems to be true in general that free order reigns vis à vis elements that are inserted into AUX. There are interpretational differences, to be sure, ranging from scope preferences to complete disambiguation. Moreover, certain co-occurrences must be barred (e.g. *If they could not have only listened). Nevertheless a wide variety of facts are accurately described by what we might call an AUX Reordering Principle:

(1.2.51) AUX Reordering Principle

Freely move non-verbs to the right within AUX.[15]

This principle, which is somewhat reminiscent of Keyser's (1968) "transportability convention", need be supplemented only by the standard rule of NEG-Placement, and the following formulation of Quantifier Shift (we leave open the question of whether this rule should be collapsed with the rule that shifts adverbs into AUX).

(1.2.52) Quantifier Shift (optional)

$$X - Q - (\text{tense}) \begin{bmatrix}+\text{Verb} \\ +\text{AUX}\end{bmatrix} - Y$$

S.D.:	1	2	3	4
S.C.:	1	∅	3 + 2	4

(tense) could be eliminated from the formulation of the rule if it could be shown that it need not apply until after Tense Hopping. That matter is inessential. Quantifier Shift might (but not necessarily) be classified as a Minor Movement Rule (in the sense of Emonds (1970)).

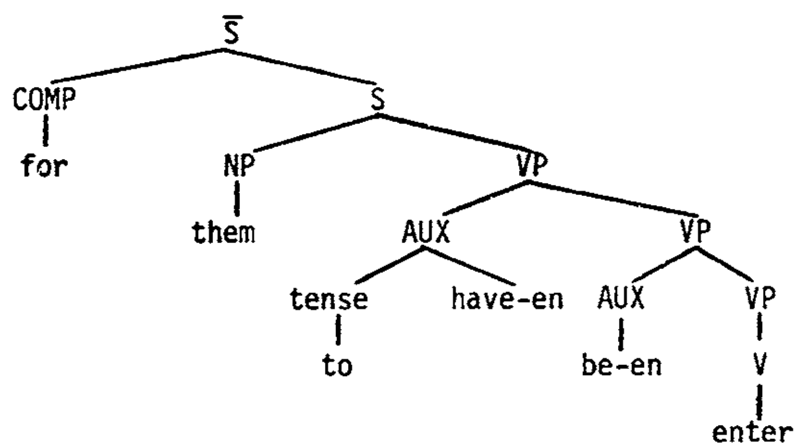
Notice that this entire array of facts poses quite a problem for the A ÷ W analysis. Since be will always be shifted into AUX, some ad hoc statement will be necessary to account for why Quantifiers, Adverbs, and not can occur after be only when it is the first auxiliary verb. The AUX

Reordering Principle would have to be given up. It's of course possible that a careful formulation of NEG-Placement and Quantifier Shift (with the possible addition of another rule that jumps <u>not</u> (and adverbs and Q's?) over <u>have</u>) could provide the A + W proposal with a proper description of the facts. Our proposal seems to account for this entire collection of facts, however, with little difficulty, if not elegantly.

Turning now to tenseless clauses, let us assume that the infinitive marker <u>to</u> is inserted under tense, as A + W suggest (p. 234, nt. 15). The non-co-occurrence of <u>to</u> and modals is again presumably a matter for subcategorization. (Alternatively, we could assume a base rule: $\text{AUX} \rightarrow \left\{\begin{matrix} \text{tense (M)} \\ \underline{\text{to}} \end{matrix}\right\}$ (<u>have</u>), as suggested in Chomsky (1965).)

We then generate tenseless clauses of the following general shape:

(1.2.53)



From structures like these, we predict the following pattern of VPD applicability.

(1.2.54) Peter seems to have been careful, and Bill seems

$$\left\{\begin{array}{l} \text{(a) to have been } \emptyset \\ \text{(b) to have } \emptyset \\ \text{(c) *to } \emptyset \end{array}\right\} \text{ also.}$$

(1.2.55) Peter was pleased that the drawings had been entered (in the contest), but Betsy would have preferred for them...

(a) not to have been ø

(b) not to have ø

(c) *not to ø

Here the (c) examples are not as bad for all speakers as our analysis predicts. Nevertheless no speakers fail to see the appropriate contrast between the (b) and the (c) sentences.[16]

Note further the following double deletion possibility, which would be another case of A + W's rule of AE applying in a tenseless clause.

(1.2.56) I want Betsy to be careful, and I want Peter $\left\{\begin{matrix}\text{(a) to be } \emptyset \\ \text{(b) to } \emptyset\end{matrix}\right\}$ also.

In tenseless clauses, then, our formulation of VPD also gets the facts right.

Jackendoff (1972) has observed that in tenseless clauses, not does not immediately follow the first auxiliary verb as it does in tensed clauses:

(1.2.57)(a) John was not there in time...

(b) For John not to be there in time...

(c) *For John to be not there in time...

A + W cite these facts as a problem for their analysis. They suggest, somewhat half-heartedly, that "one might propose that not is generated as the first element of AUX and that it is moved after the first auxiliary verb just in case Tense is present". That would account for (1.2.57)(c), all right, where the absence of tense prevents not from moving after be, but it wouldn't account for these examples, which show that not can move to the right in tenseless clauses.

(1.2.58)(a) For John to not have been ready in time...

(b) For John to have not been ready in time...

(c) *For John to have been not ready in time...

Our view of the constituent structure of VP and AUX and the AUX Reordering Principle explain all these facts. (1.2.57)(c) is impossible because not cannot move outside of AUX; the same is true of (1.2.58)(c). (1.2.58)(a) and (b), however, are fine reorderings of not within AUX. (Some people object to sentences like (1.2.58)(a) because of a general distaste for "split infinitives". That problem is tangential to our concerns here.)

In all the above discussion, we have been assuming that shifted Q's, adverbs, and not, when they follow an auxiliary verb, are part of AUX. A potential counter-argument to this assumption arises from considering the following contrasts:

(1.2.59)(a) My brothers all have left, and my sisters all have ∅, too.

(b) *My brothers have all left, and my sisters have all ∅, too.

(1.2.60)(a) Betsy has written a novel, and Peter even has ∅.

(b) *Betsy has written a novel, and Peter has even ∅.[17]

(1.2.61)(a) Betsy must have called in sick, but Peter might not have ∅.

(b) Betsy must have called in sick, but Peter might have not ∅.

One might take the above contrasts to show that the Q in (1.2.59)(b) and the adverb in (1.2.60)(b) are outside of (i.e. sisters of) AUX, whereas not in (1.2.61)(b) is inside of AUX. That state of affairs would correctly account for the impossibility of VPD in the first two cases, since VPD can apply only when as AUX immediately precedes the deletion target.

Interestingly, no such argument is possible, because, for some reason, AUX-final Q's and adverbs exhibit a very general rule-inhibiting behavior which not does not. In particular, AUX-final Q's and adverbs inhibit Wh-Movement in questions:

(1.2.62)(a) Where are all your brothers?

(b) *Where are your brothers all?

(c) *How happy are they ever?

(1.2.63)(a) I don't know who they all are.

(b) *I don't know who they are all.

(c) I wondered how happy they ever were.

(d) *I wondered how happy they were ever.

Similarly, in the case of Wh-Movement in relatives:

(1.2.64)(a) Alan said they were bigots, which indeed they all were.

(b) *Alan said they were bigots, which indeed they were all.

(1.2.65)(a) Peter and Betsy wanted to be sympathetic to their children's problems, which their parents never were.

(b) *Peter and Betsy wanted to be sympathetic to their children's problems, which their parents were never.

AUX-final Q's and adverbs also inhibit Topicalization and Comparative Deletion:

(1.2.66)(a) Communists, they all were.

(b) *Communists, they were all.

(1.2.67)(a) Communists, they never were.

(b) *Communists, they were never.

(1.2.68)(a) Alan is in more trouble than the editors all are.

(b) *Alan is in more trouble than the editors are all.

(1.2.69)(a) Alan is in more trouble than the editors ever were.

(b) *Alan is in more trouble than the editors were ever.

Crucially, not has none of these rule-inhibiting effects:

(1.2.70) I know what you are, but...

(a) What are you not?

(b) I don't know what you're not.

(1.2.71)(a) Alan said they were bigots, which they were not.

(b) Communists, they were not.

(c) Betsy solved more problems than she did not.

I'm not at all sure how to state the appropriate constraint involving AUX-final Q's and adverbs, but it is clear that some such constraint exists. Whatever its proper formulation may be, it's[18] clear that this general constraint is responsible for the non-applicability of VPD in (1.2.59)(b) and (1.2.60)(b) above. Hence these instances of non-applicability of VPD provide no evidence whatsoever against the claims we have made about the constituent structure of "floated" quantifiers.

Another point should be mentioned. We bring it up here for lack of a better place. There is a class of sentences that has been taken by some (Shopen (1972), Jackendoff (1972)) to be instances of VPD (more precisely, of their interpretive analogues of VPD). The following sentences are typical.

(1.2.72) Someone has to take out the garbage,...

(a) so Alan volunteered ∅.

(b) but Sandy refused ∅.

As was shown in Sag and Hankamer (1976), however, these sentences do not involve VPD at all. Rather, they are instances of the process referred to there as Null Complement Anaphora, as is shown by their pragmatic controllability, their failure to exhibit the "missing antecedent" phenomenon (Grinder and Postal (1971)), and the range of (syntactic) antecedents they permit. The reader is referred to the discussion there, but the matter will be taken up again in Chapter Four.

This brings us to another interesting problem, that was first pointed out by Kuno (1975a). Kuno observed that VPD is not possible in examples like (1.2.73), which stand in marked contrast to sentences like (1.2.74), where VPD is perfectly acceptable.

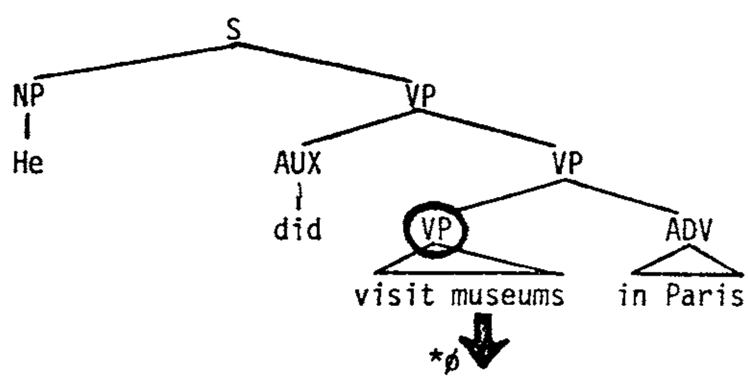
(1.2.73)(a) Speaker A: Where did John visit museums?

(b) Speaker B: He $\left\{\begin{array}{l}\text{visited museums}\\ \text{*did } \emptyset\end{array}\right\}$ in Paris.

(1.2.74) Mary did not visit museums in Paris, but she did ∅ in London.

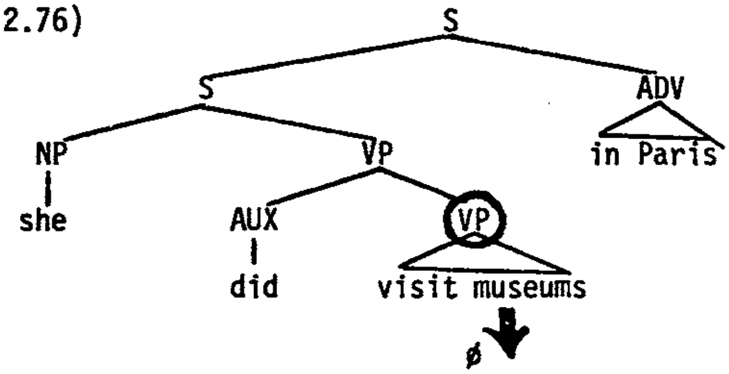
Kuno attributes the observed difference in grammaticality to a difference in constituent structure. (1.2.73)(b), he argues, has the following structure, where <u>in Paris</u> forms a constituent with the preceding VP.[19]

(1.2. 75)



(1.2.74), on the other hand, is argued to have the constituent structure shown in (1.2.76), where <u>in Paris</u> is not a modifier of VP.

(1.2.76)



The constraint Kuno formulates to account for these facts is this:

(1.2.77) Verb Phrase Deletion can apply only to the VP that is VP-final.

Grosu (1975, to appear) takes issue with Kuno's proposal. Although he accepts all of Kuno's arguments for the constituent structure difference between these two sentences, he argues that the correct constraint on VPD is a "Sisterhood Condition", which he states as follows:

(1.2.78) The Sisterhood Condition

VP-Deletion can apply only to a VP that is a sister of a VP-Specifier.[20]

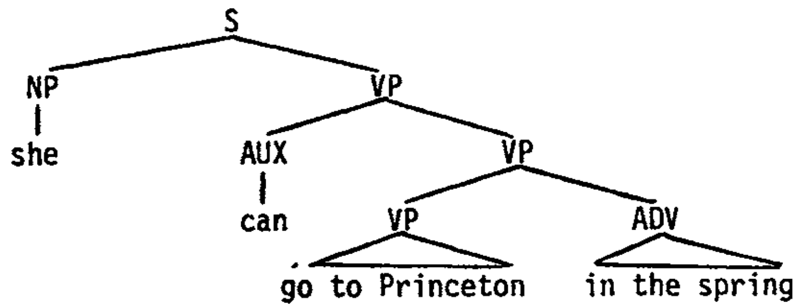
Notice, however, that neither Kuno's constraint, nor Grosu's Sisterhood Condition allows the following sentence to be generated.

(1.2.79) Mary can't go to Princeton in the fall, but she can $\emptyset_1$ in the spring, although if she does $\emptyset_2$, those who expect her in the fall will be very disappointed.

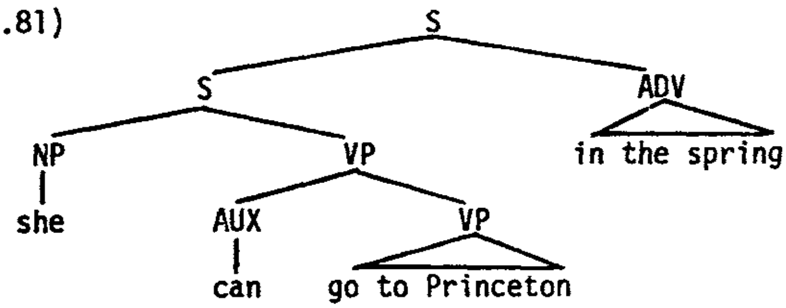
Here, it is clear that $\emptyset_1$ is interpreted as <u>go to Princeton</u>, and equally clearly, $\emptyset_2$ is interpreted as <u>go to Princeton in the spring</u>.[21] Both Kuno and

Grosu agree that in order for the first deletion to be possible, the structure prior to deletion must not be as in (1.2.80) where the target VP is neither a sister of AUX nor in VP-final position, but rather as in (1.2.81).

(1.2.80)



(1.2.81)



But the clause containing the second deletion site, $\emptyset_2$, must have the structure in (1.2.80) prior to deletion, under the assumption (shared by almost everyone, apparently) that VPD deletes only VP's. Furthermore, since $\emptyset_2$ presumably arises from deletion of a VP that is identical to one in the previous clause, it therefore follows, given normal assumptions about identity conditions, that the previous clause, i.e. the one containing $\emptyset_1$, also has the structure in (1.2.80) prior to deletion. The first deletion, therefore, must violate both Kuno's constraint and Grosu's Sisterhood Condition.

This is an interesting constituent structure paradox. If we accept Kuno's claims about the relevant constituent structures, then, it seems to me, there are two ways out of this dilemma. Either

(1.2.82) the identity condition on VPD must be weakened to allow deletion of a VP that exhaustively dominates a terminal string that is identical to another terminal (sub-) string whose constituent structure differs from that of the target VP.

or

(1.2.83) VPD must be allowed to delete non-constituents, i.e. sequences like VP, VP-ADV, VP-ADV-ADV, etc.

Under the first alternative we could account for our paradoxical sentence by assigning $\emptyset_2$ the structure in (1.2.80) prior to deletion, which would then count as identical to the string <u>go to Princeton in the spring</u> in the preceding clause, even though that string had the structure in (1.2.81) prior to deletion. If, instead, we accept the second alternative, but maintain the stronger identity condition, i.e. exact identity of constituent structure, then $\emptyset_1$ <u>in the spring</u> and $\emptyset_2$ could both have the structure in (1.2.81) prior to deletion, and (non-constituent) VPD could apply to produce $\emptyset_2$.

The question arises, then, whether there is any evidence that might decide between the two alternatives. I think there is. The first would have some unpleasant consequences. For instance, it would require finding some independent explanation for why sentences like the following one (where a VPD target and its (supposed) antecedent occur in different syntactic frames) are ungrammatical in spite of the superficial identity of terminal substrings.

(1.2.84) *Betsy likes blue cars, although her car isn't $\emptyset$.

[$\emptyset$ = blue]

Moreover, this sentence should be compared to (1.2.85), where minimal modification so as to create identity of constituent structure as well as identity of terminal sub-strings, restores grammaticality.

(1.2.85) Betsy likes cars that are blue, although her car isn't ∅.

This is just the kind of evidence one would need to decide against weakening the identity condition.

The second alternative, on the other hand, might conceivably be supported by the existence of sentences like these:

(1.2.86) Peter said it would turn out that he and Betsy would play on the same team, but Betsy knew it wouldn't ∅.
[∅ = turn out that he and Betsy would play on the same team]

(1.2.87) Although it didn't seem that Peter and Betsy would get married, Sandy claimed it did ∅.
[∅ = seem that Peter and Betsy would get married]

Such sentences would provide support for non-constituent deletion, of course, only if it could be shown that (<u>It-</u>) Extraposed clauses are not part of the VP, a position that is often assumed, but one that has never to my knowledge been justified (for the contrary view, see Emonds (1970)). If such a position could be justified, however, it is clear that these last two sentences would be instances of non-constituent VPD, and would hence provide direct support for (1.2.83).

It is therefore quite possible that the correct formulation of VPD is not as we gave it earlier (1.2.10), but rather something closer to (1.2.88).

(1.2.88) W_1 - X - W_2 - AUX - X - W_3

S.D.: 1 2 3 4 5 6

S.C.: 1 2 3 4 $\emptyset$ 6

condition: 2 = 5

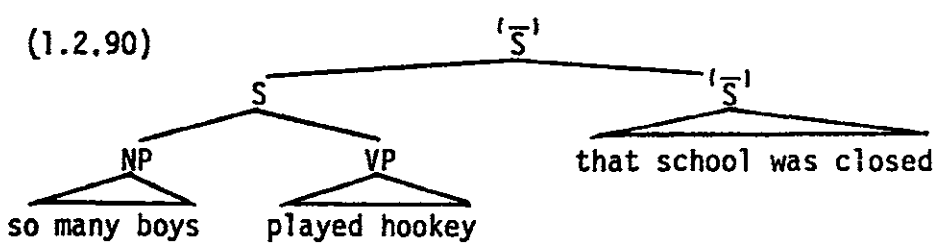
This formulation, <u>ohne weiteres</u>, overgenerates wildly, the problem being that the deletion target cannot be just any arbitrary stretch of tree, as the X-variable notation implies. In fact, it seems that VPD can delete only elements that are commanded by the AUX specified in term 4 of the rule. Put somewhat differently, term 5 of this rule cannot analyze $]_S$, i.e. the deletion target cannot be a sequence like: VP$]_S$ ADV or VP$]_S$ $\overline{S}$. This is shown by the following example:

(1.2.89)(a) On Wednesday, so many boys played hookey that school was closed, and on Thursday, so many girls played hookey that school was closed.

(b) *On Wednesday, so many boys played hookey that school was closed, and on Thursday, so many girls did.

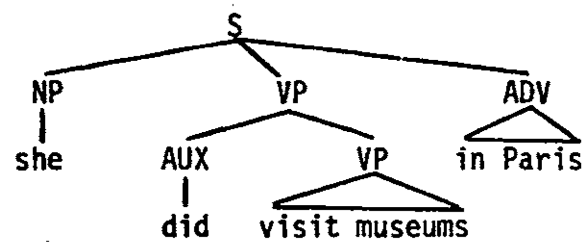
Result clauses like the one in (1.2.89)(a) have been shown (Williams (1974), Andrews (1975), Hankamer and Sag (1975)) to have the structure sketched in (1.2.90).

(1.2.90)



Therefore, a formulation of VPD like (1.2.88) would have to be suitably constrained. Moreover, in order to allow VPD to apply in our paradoxical sentence, but not in (1.2.89), it seems reasonable to modify Kuno's claims about the constituent structure difference between (1.2.73)(b) and (1.2.74). (1.2.74) (and the clause containing ϕ_2 in (1.2.79), mutatis mutandis) could be assigned the following structure, where the prepositional phrase in question is not in a "higher" S, as Kuno assumed, but rather is dominated by the same S, i.e. as a sister of VP.

(1.2.91)



A version of the rule in (1.2.88) that was suitably constrained to prevent the VPD target variable from analyzing S-boundaries could then apply to the sequence VP-ADV here, but would not delete result clauses like the one in (1.2.89).

This is of course only one possible solution to the constituent structure paradox we have observed. It is interesting to note that one way of imposing the restriction mentioned on the deletion target variable would be to introduce into syntactic theory the notion of restricted variable, analogous to the notion of "Q-variable" that has been proposed in recent work in phonology (Halle (1973), Vergnaud, Halle, and Prince (forthcoming)). Should such variables prove to be necessary in linguistic theory, then this would have rather far reaching consequences, for it would allow a rather straightforward means of incorporating the notion "clause-mate" into linguistic theory (though as even Postal notes (1974a, p. 76, nt. 24), much of the original evidence for the notion "clause-mate", i.e. reciprocals can be shown to have nothing to do with that notion).

Further evidence, by the way, for the necessity of formulating VPD as a variable-deleting rule might come from sentences like the following:

(1.2.92)(a) John could pull you out of a plane, like he did ø his brother.

(b) Mary hasn't dated Bill, but she has ø Harry.

If sentences like these, which most people accept, are the result of VPD, as the class of possible preceding elements would suggest, then we have cases where a subpart of a VP (just V) is a possible VPD target. I am not at all sure how the modification of the rule would have to proceed to allow for such examples. There are many possibilities that come to mind. Every one of them that I have given any thought to, however, creates more problems than I know how to solve at the moment.

As regards the status of such sentences, I observe that Quirk et al., who advocate no theory of this phenomenon in particular, cite sentences like these as fully grammatical, as do Halliday and Hasan (1973). The latter authors cite the following examples, for instance (Halliday and Hasan (1973, p. 49)):

(1.2.93)(a) Speaker A: Is she suing the hospital?

Speaker B: She is ø the doctor.

(b) Speaker A: Has he sold his collection yet?

Speaker B: He has ø some of the paintings; I'm not sure about the rest.

I even overheard the following discourse on the MIT campus (January, 1976):

(1.2.94) Speaker A: Gee, I've never seen you on campus before.

Speaker B: Yea! Neither have I ∅ you.

In short, although the data in question are not entirely clear, it is quite possible that there is empirical evidence that necessitates a formulation of VPD as a variable-deleting rule. I will not pursue the matter further and, for convenience, will continue to refer to the phenomenon as VPD (a better name might be Post-Auxiliary Ellipsis). Our subsequent discussions will not be affected by this possible inaccuracy.[22]

After all this, the careful reader will note that we have not chosen between Kuno's and Grosu's formulation of the appropriate constraint to block the generation of sentences like (1.2.73)(b) above. As Grosu points out, Kuno's constraint is incapable of accounting for certain facts of German which seem to be quite related. Thus in German, VPD targets in main clauses follow an AUX, but in subordinate clauses, they precede the clause-final AUX:

(1.2.95) Walter muss nicht gehen, aber Peter muss $\{{\text{gehen} \atop \emptyset}\}$.

Walter must not go, but Peter must go.

"Walter needn't go, but Peter must ∅"

(1.2.96) Walter glaubt nicht dass er gehen muss, aber Peter glaubt

Walter thinks not that he go must, but Peter thinks

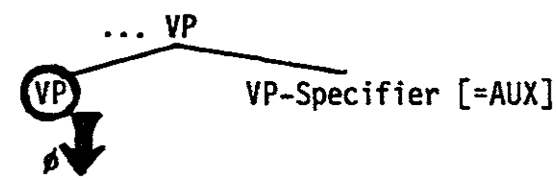
dass er $\{{\text{gehen} \atop \emptyset}\}$ muss.

that he go must.

"Walter doesn't think that he must go, but Peter thinks that he must ∅."

Examples like this last one, Grosu remarks, counterexemplify Kuno's constraint, for they exhibit the following constituent structure.

(1.2.97)



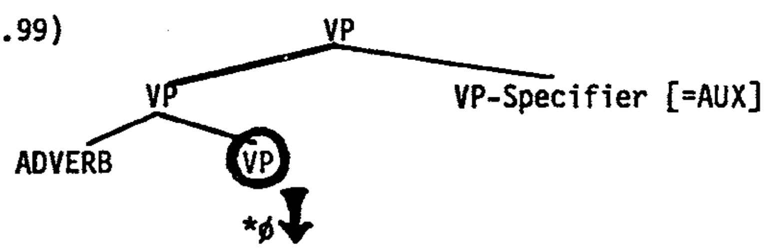
A non-VP-final VP has been deleted.

On the other hand, Grosu continues, there are cases where VPD cannot apply in German subordinate clauses:

(1.2.98) Walter glaubt dass er langsam gehen muss, aber Peter glaubt
Walter thinks that he slowly go must, but Peter thinks
dass er schnell {gehen / ø} muss.
that he quickly go must.
"Walter thinks that he must go slowly, but Peter thinks that he must go/*ø quickly."

Sentences like these are taken to have the following constituent structure:

(1.2.99)



The conclusion Grosu reaches is that if Kuno's constraint is correct, then German would need both that constraint (for main clauses) and another principle restricting VPD in subordinate clauses to VP-initial VP's. No such loss of generalization is necessary in Grosu's proposal, for in all these sentences, the only deletable VP's are those that are "sisters" of AUX.

Thus at this stage of the deliberation, Grosu's hypothesis would seem to be the leading contender.

I would like to argue, however, that neither Grosu's nor Kuno's constraint is necessary. Notice that if we accept all of Kuno's claims about the constituent structures in question (modulo the refinement mentioned earlier), then each case where VPD is impossible would involve a proper analysis wherein the deletion target was non-maximal. Could Bresnan's RAOAP, which we have tentatively adopted here, be the explanation for both Kuno's examples for English and Grosu's for German?

Grosu considers this possibility (he actually discusses only Chomsky's (1964, 1973) notion of a (non-relativized) A-over-A Principle). His conclusion is that, as regards an explanation of Kuno's examples, the A-over-A Principle is "irrelevant" because

> The highest phrase to which the rule can apply in this derivation [= the derivation of sentences like (1.2.73)(b) and (1.2.98) above - I.A.S.] is the actual deletee; the phrase which immediately dominates the deletee is not a suitable candidate for deletion, since it does not satisfy the identity requirement (i.e., there is no identical phrase in the first conjunct identical with it which could conceivably act as deletion controller. (Grosu, forthcoming, nt. 13)

Grosu also gives a second argument which concerns the general possibility of multiple outputs with VPD. Such multiple outputs are of course a problem for a non-relativized version of the A-over-A principle, but this is precisely the difficulty remedied by RAOAP and our formulation of VPD that "mentions" AUX as a context predicate. Therefore only Grosu's first objection is of interest.

His argument presupposes that it is clear in advance how the principle of recoverability of deletion and other general principles like RAOAP interact. This matter is not at all clear in advance. Grosu seems to be assuming that the class of proper analyses of a given sentence with respect to some transformation is delimited first by the principle of recoverability of deletion, and subsequently further delimited by other general principles like RAOAP. Only after this hidden premise is made explicit can Grosu's argument be properly assessed.

One might ask what empirical evidence exists for this presumption. I know of none. This is an interesting problem, however. Although, as far as I can tell, this matter has never been taken up in the literature,[23] one might conjecture that the following view of constraints on rules of grammar is a viable one.

Let us consider the set of proper analyses of some string S with respect to some transformation T. Let us call that set P. Without any general constraints on transformations, we would expect every member of P to be a legitimate proper analysis of S with respect to T.

Let us view each general constraint on transformations as determining a set of inadmissible proper analyses of S with respect to T. RAOAP thus might determine a certain set of inadmissible proper analyses, call that set I_{RAOAP}. I_{RAOAP} is necessarily a proper subset of P. The principle of recoverability of deletion might determine another set of inadmissible proper analyses of S with respect to T. Let us call this set I_{ROD}. I_{ROD} is a subset of P, but not necessarily a proper subset (that is, with respect to some sentences, no recoverable deletion is possible). If we view the matter in this way, with each general constraint on transformations determining a set I_i of inadmissible proper analyses of S with respect to T, then the set of well-formed proper analyses we ultimately arrive at is this:

(1.2.100) $P - (I_{RAOAP} \cup I_{ROD} \cup \ldots \cup I_n)$

This theory of constraints on transformations seems to be a perfectly plausible one (we will have more to say about this matter in Chapter Three). Furthermore, mirabile dictu, such a theory solves the problem raised by Kuno's and Grosu's examples. In the ungrammatical sentences they cite, non-maximal deletions have been attempted, i.e. the relevant attempted proper analyses are elements of I_{RAOAP}. Deletion of any larger VP in those examples would require proper analyses that, by the principle of recoverability of deletion, are elements of I_{ROD}. Thus with respect to all Kuno's and Grosu's sentences, the set P of possible proper analyses with respect to VPD is the same set as $(I_{RAOAP} \cup I_{ROD})$. In other words, $P - (I_{RAOAP} \cup I_{ROD}) = \emptyset$. The proposed theory of constraints on transformations therefore predicts that no deletion is possible in those examples. This is precisely the right result.

1.3 Semantic Problems for the Syntactic Theories

Having examined various theories of VPD, pondered various problems that beset each of them, and now having suggested a syntactic analysis of our own, we are in a position to examine certain additional data that have not entered into past discussions. The main thrust of this section will be to show that there are various semantic factors that affect the applicability of VPD.

Let us first examine sentences containing quantifier expressions, such as (1.3.1).

(1.3.1) Someone hit everyone.

As is well known, someone who uttered (1.3.1) would have told the truth in at least two possible situations. First, if there was some individual _a_ who hit every individual in the relevant domain of discourse (with the possible exception of himself), and secondly, if everyone (in the relevant domain of discourse) was hit by someone, but not necessarily by the same person.

The question of the proper semantic treatment of sentences like (1.3.1) is a vexed one, and it is not my intention here to become embroiled in a detailed discussion of the various proposals that have been made for representing the logical forms of such sentences. Let us simply note that there are many who would regard sentences like (1.3.1) as semantically unambiguous, i.e. as having no semantic reading that specifies that one person did all the hitting. For these people, the two situations just described would simply both be consistent with the truth of (1.3.1), or, more precisely, two different ways of satisfying its truth conditions.

A popular alternative view, which we will adopt here, is that (1.3.1) is indeed (semantically) ambiguous. The first situation described above corresponds to a reading which could be represented as in (1.3.2) (irrelevant details omitted), where the universal quantifier is within the scope of the existential quantifier.

(1.3.2) $(Ex)(\forall y)[x \text{ hit } y]$

The second situation described does not satisfy the truth conditions of (1.3.2) but does satisfy the truth conditions for the logical formula in (1.3.3), which represents the other reading of (1.3.1) where the existential quantifier is within the scope of the universal quantifier. The first situation described also happens to satisfy the truth conditions of (1.3.3).

(1.3.3) $(\forall y)(Ey)[x \text{ hit } y]$

We will not justify this approach any further, except to note that if one accepts quantification theory as a useful device for representing the logical forms of English sentences, and if one concedes that (1.3.1) can be truthfully uttered in either of the situations described above, then one is hard put to explain why (1.3.1) is not semantically ambiguous. That is, given that sentences like (1.3.4) would be represented roughly as in (1.3.5),

(1.3.4) Someone left

(1.3.5) $(Ex)[x \text{ left}]$

the proponent of non-ambiguity would have to explain why (1.3.1) can be represented only by (1.3.3), whose truth conditions are satisfied in both situations, and not by (1.3.2). In other words, some device would be needed to prevent (1.3.1) from being associated with a logical form like (1.3.2) by whatever mechanism associates sentences like (1.3.4) with representations like (1.3.5).

Furthermore, such a device seems unlikely, for it would be inconsistent with what is otherwise known about restrictions on quantifier scope assignment. In particular, the claim has frequently been made that the order of operators in logical structures tends to follow the order of the quantifier expressions in surface structure (Katz and Postal (1964), Lakoff (1969), Jackendoff (1968, 1972), Carden (1970), Dummett (1973, chap. 2)). Any device that requires (1.3.1) to be associated with a representation like

(1.3.3) forces a scope assignment that is exactly opposite to the surface structure order of quantifier expressions. Needless to say, this objection is perhaps answerable, but it is offered here as a plausibility argument for adopting the position that (1.3.1) is indeed semantically ambiguous.

We are now ready to consider the relevance of (1.3.1) to VPD. Observe that sentences (1.3.6) and (1.3.7) are ambiguous.

(1.3.6) Someone hit everyone, and then Bill hit everyone.

(1.3.7) Someone hit everyone, but Bill didn't hit everyone.

The right conjuncts in (1.3.6) and (1.3.7) are not ambiguous in the relevant respects, of course, because <u>Bill</u> is not a quantifier expression, and does not admit of scopal variation. The left conjuncts, however, admit of both readings, i.e. (1.3.2) and (1.3.3) above. There is the usual preference for the interpretation where the scopal assignment reflects the surface order of quantifier expressions, i.e. (1.3.3), but all speakers I have checked with agree the other reading ((1.3.2)) is still possible for the left conjuncts in (1.3.6) and (1.3.7).

Now as far as I can see, any purely syntactic theory of VPD makes the claim that the VP's in the right conjuncts of (1.3.6) and (1.3.7) are deletable no matter which interpretation is assigned to the left conjunct (presuming that that conjunct is <u>syntactically</u> unambiguous, and given the standard notion of deletion under identity, i.e. the standard notion of the recoverability of deletion, which we will discuss in more detail in Chapter Two).

This prediction, however, is clearly false. Sentences (1.3.8) and (1.3.9), which result from applying VPD to (1.3.6) and (1.3.7) respectively, are unambiguous.

(1.3.8) Someone hit everyone, and then Bill did.

(1.3.9) Someone hit everyone, but Bill didn't.

The left conjuncts in (1.3.8) and (1.3.9) can only be interpreted as in (1.3.2), where the existential quantifier has wide scope.

We might look at this as a parallelism requirement on VPD: Quantifier scope must be parallel if VPD is to apply. (1.3.10) represents the parallel interpretation for (1.3.9), and (1.3.11), the non-parallel interpretation of (1.3.7) that VPD has deprived (1.3.9) of.

(1.3.10) $(Ex)(\forall y)[x \text{ hit } y]$ but $\neg(\forall z)[\text{Bill hit } z]$

(1.3.11) $(\forall y)(Ex)[x \text{ hit } y]$ but $\neg(\forall z)[\text{Bill hit } z]$

The facts of (1.3.8) and (1.3.9) could conceivably be handled in an Aspects theory or within a Generative Semantics framework by means of a global condition on VPD that made reference to the level of logical form,[24] or, in an interpretive theory, by a condition on the interpretive rules for quantifier scope assignment that required the scope assignments to be parallel if VPD has applied.[25] Either of these approaches, however, would be entirely ad hoc as rule-specific mechanisms. In the next chapter we will see that they are inadequate as well. The question to ask, therefore, is whether this set of examples is an idiosyncratic case of VPD making reference to logical structures, or a particular instance of some more general relation between VPD and logical form. In what follows immediately, we will see that the latter is the case.

Consider now sentences like (1.3.12).

(1.3.12) Sandy greeted everyone when Betsy did.

The syntactic theory of VPD would derive (1.3.12) from (1.3.13).

(1.3.13) Sandy greeted everyone when Betsy greeted everyone.

Indeed (1.3.12) and (1.3.13) share a reading, the logical form of which we can represent as (1.3.14) (leaving aside the question of the appropriate representation for <u>when</u>).

(1.3.14) $(\forall x)$[Sandy greeted x] when $(\forall y)$[Betsy greeted y]

(1.3.14) would be true if, say, Betsy walked in and said "hello everybody" at the same time that Sandy did so.

Imagine the following situation, however: Betsy and Sandy walk into the room together. Mr. X, Ms. Y, and Ms. Z are the only people in the room. Sandy and Betsy together accost each person and say "Hello Mr. X", "Hello Ms. Y", and "Hello Ms. Z" in two-part harmony. In this situation, it seems to me, (1.3.13) could not be truthfully uttered, but (1.3.12) could be.

In other words, (1.3.12) has an additional reading that (1.3.13) does not have, one that seems adequately represented as (1.3.15), where there is only one universal quantifier binding a variable in each sentence.

(1.3.15) $(\forall x)$[Sandy greeted x when Betsy greeted x]

Note further that there is an undeleted English sentence that also has this interpretation, namely, (1.3.16).

(1.3.16) Sandy greeted everyone when Betsy greeted $\left\{\begin{matrix}\text{\%him} \\ \text{them}\end{matrix}\right\}$.

Should the syntactic identity condition on VPD be modified to allow (1.3.12) to be derived from (1.3.16) as well as from (1.3.13), or is this just another case of surface structure influencing quantifier scope interpretation? Let us postpone this question temporarily, while we unturn some further difficulties for the purely syntactic theory of VPD.

The next set of examples to be considered concerns the interaction of VPD and extraction rules. Take Question Movement as a first example. If Question Movement removes, say, the objects from two VP's, creating two identical VP fragments, then the syntactic identity theory predicts that deletion is possible. The relevant examples are sentences like (1.3.17) which might conceivably occur in contexts like (1.3.18).

(1.3.17) What did Bill?

(1.3.18) 1st Speaker: What did Harry take a picture of?
2nd Speaker: An elephant.
1st Speaker: What did Bill?
2nd Speaker: A tiger.

The (un)acceptability of such sentences was not at all obvious to me. I therefore constructed a questionnaire which was used to test twenty subjects. The questionnaire also contained sentence (1.3.19) presented in the context shown in (1.3.20).

(1.3.19) What was Harry able to?

(1.3.20) 1st Speaker: What was John able to take a picture of?
2nd Speaker: An elephant.
1st Speaker: What was Harry able to?
2nd Speaker: A tiger.

The results of this small experiment were rather clear. The sentences were evaluated on a three-point scale (1 = best, 3 = worst). (1.3.17) and (1.3.19), which I take to be typical cases of deletion of fragmented VP's, received average ratings of 2.5 and 2.6 respectively, indicating that they were consistently judged to be of low acceptability. Now the low acceptability of these sentences could, of course, be due to a performance factor, or to some other extra-grammatical phenomenon, as yet uncovered. I can find no motivated explanation along these lines, however, so we will take these results as showing that the sentences are ungrammatical. Why should this be? As noted above, a purely syntactic theory of VPD, ohne weiteres, predicts that these should be instances of legitimate deletion.

Leaving questions aside momentarily, let us consider the interaction of VPD and relativization. What is at issue is sentences like (1.3.21) and (1.3.22) (which were brought to my attention by Edwin Williams).

(1.3.21) We finally got in touch with John, who my brother Al tried to visit, but couldn't ∅.

(1.3.22) We finally got in touch with John, who my brother Al tried to visit, but who he couldn't ∅.

Again unsure of my own intuitions, I report the results of my questionnaires. (1.3.21) was given an average rating of 1.1, which attests to its full grammaticality. (1.3.22), on the other hand, received an average rating of 2.2, implying deviance, but somewhat less than the "green sleep ideas furiously colorless the" variety of deviance. Nevertheless, we will consider this contrast significant enough to warrant accounting for.

The obvious difference between (1.3.21) and (1.3.22), one versus two wh expressions, seems to be significant. The deviance of (1.3.22), where it is clear that a distinct wh-word has been extracted from the VP where deletion has been attempted, seems, furthermore, to be related to the deviance of (1.3.17) and (1.3.19) above, where the fragmented VP's also involved distinct wh-words. The plausibility of this being the relevant factor is further supported by the following facts involving pseudo-clefts.

(1.3.23) *What Sandy carried was the baseball bat, and what Betsy did ∅ was the catcher's mit.

(1.3.24) What Sandy wanted to buy but couldn't ∅, was the catcher's mit.

(1.3.24) where one wh-word "binds" both positions, i.e. where only one wh-word appears on the surface, is fully grammatical. (1.3.23), on the other hand, which contains two wh-words, each "binding" a position in the clause that contains it, is not. The syntactic identity requirement on VPD seems to be in serious trouble, because it is inherently incapable of distinguishing cases like these. That is, in both cases the target verb phrases are sequences of the form "verb + extraction site", as are the corresponding VPD triggers. No purely syntactic distinction between the two cases exists.

There is, however, a level where these cases are distinguished. If we assume, as most of the literature on the logic of questions has assumed (Hiż (1962), Belnap (1963), Åqvist (1965), Hull (1972), Chomsky (1973), for example) that wh-words in questions are to be treated on a par with quantifiers (in particular in that they both bind variables), and if we assume, following Chomsky (1973), that wh-words in relative clauses are to receive similar treatment, then the examples we have been considering are relatively easily distinguished at the level of logical form.

To be more precise, if we assume that each wh-word is to be regarded as a quantifier-like expression binding its own variables, then the undeleted sources for the two questions in (1.3.18) above would have logical representations roughly like these:

(1.3.25)(a) What did Harry take a picture of?

(b) (What x)[Harry took a picture of x]

(1.3.26)(a) What did Bill take a picture of?

(b) (What y)[Bill took a picture of y]

If, as seems plausible, we treat the logical forms of pseudo-clefts as involving set abstraction, we can represent the logical forms of the undeleted sources of (1.3.23) and (1.3.24) as follows.

(1.3.27)(a) What Sandy carried was the baseball bat, and what Betsy carried was the catcher's mit.

(b) {the baseball bat} = $\hat{x}$(Sandy carried x) & {the catcher's mit} = $\hat{y}$(Betsy carried y)

(1.3.28)(a) What Sandy wanted to buy, but couldn't buy, was the catcher's mit.

(b) {the catcher's mit} = $\hat{x}$(Sandy wanted to buy x but Sandy couldn't buy x)

As we have seen, VPD is possible in (1.3.28), but not in (1.3.25-26) or in (1.3.27). Could the reason for this be that the two VP's in (1.3.28)(a) correspond to two parts of the logical form that contain the same bound variable, i.e. two variables bound by the same operator, whereas the logical forms for the other sentences contain distinct variables, i.e. variables bound by different operators? Suppose that this aspect of logical form is indeed relevant in determining the applicability of VPD. We might then ask how much of the identity requirement on VPD is identity of surface syntactic form, and how much is identity of logical form. The following discussion may shed some light on this question. [26]

Consider sentences like these next few, which are similar to examples first observed by Bouton (1970).

(1.3.29)(a) Alan will eat anything you want him to (eat).

(b) Sandy hit everyone that Bill $\left\{\begin{matrix}\text{hit}\\\text{did}\end{matrix}\right\}$.

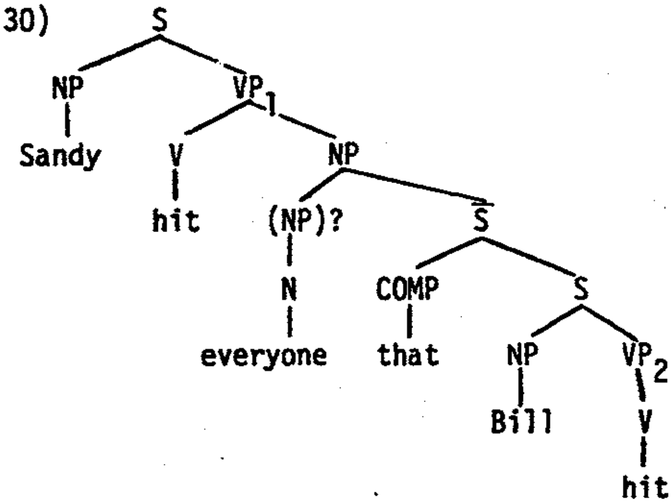
(c) Sandy ate whatever Tom $\left\{\begin{matrix}\text{ate}\\\text{did}\end{matrix}\right\}$.

(d) Betsy grabbed whatever she could (grab).

Bouton argues, correctly I think, that examples like these are also instances of VPD. Their distribution is certainly quite analogous to that of other undisputable instances of VPD, for they involve stranded *to*, tense (with *Do*-Support), or modal auxiliaries.

The syntactic structure of these examples, however, is of considerable interest. Let us take (1.3.29)(b), whose structure is typical of this class of sentences, as a first illustration. After Relativization but before deletion, (1.3.29)(b) has the intermediate structure sketched:

(1.3.30)



Such a structure, as Bouton remarked, does not meet any version of the structural description of VPD that has ever been proposed. That is, (1.3.30) cannot be analyzed by any structural description like X - VP_1- Y - VP_2- Z, because VP_2 in (1.3.30) is contained within, i.e. dominated by, VP_1. Transformational grammar, as it was originally conceived, simply does not allow transformations to refer to dominance relations of this kind. Note additionally that even within the less-constrained theory of transformational grammar formalized in Peters and Ritchie (1973), dominance relations of the type that are necessary are not countenanced.

Bouton, in fact, proposes two separate rules of VPD. One is the conventional rule that is usually assumed (S.D. = X - VP - Y - VP - Z), and the other rule, which is supposed to account for examples like those in (1.3.29), he formulates as in (1.3.31).

(1.3.31)	W	$[_{VP}$ X	VP	Y]	Z	
S.D.:	1	2 3	4	5	6	
S.C.:	1	2 3	∅	5	6	

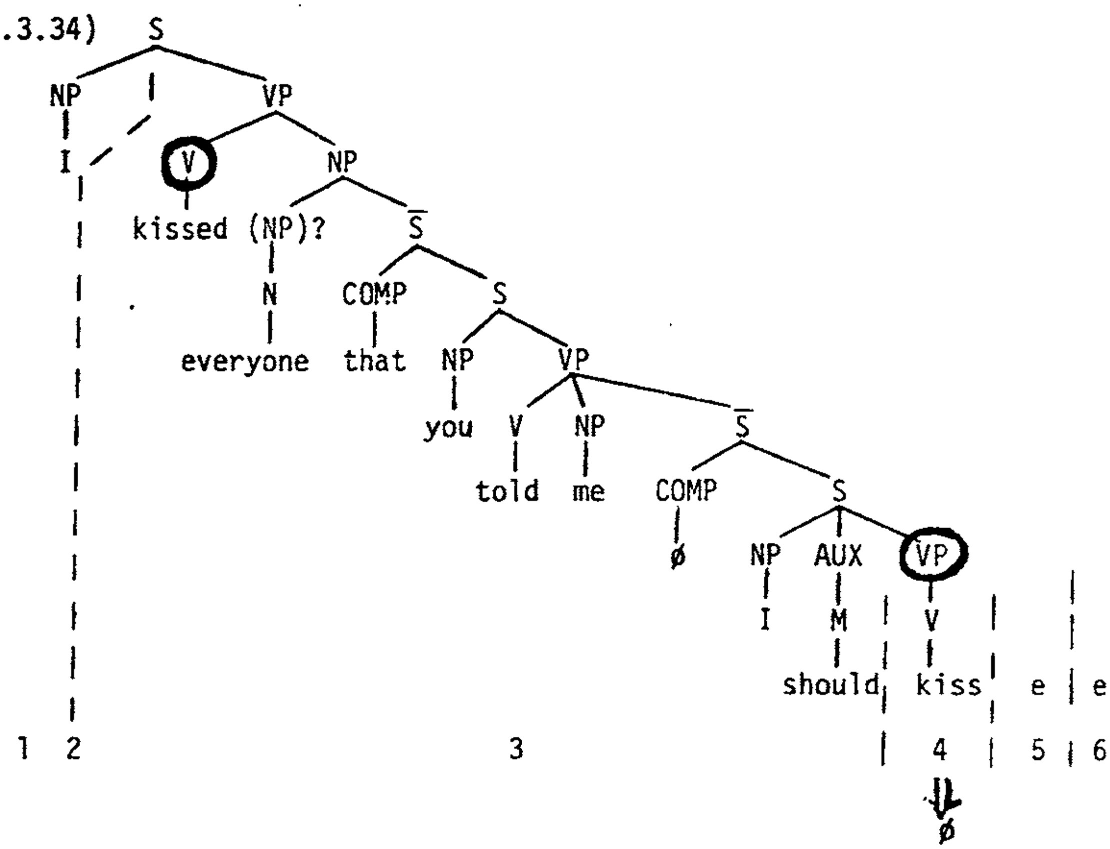
This proposal seems to me to be rather confused. Specifically, even if one assigns an interpretation to rules written in this way, there seems to be no way of ensuring recoverability of deletion. Surely some identity must obtain between the deleted VP and other elements of the containing VP. Otherwise Bouton's rule will generate such sentences as (1.3.32) from sources like (1.3.33).

(1.3.32) *I saw everyone who John was ø.

(1.3.33) I saw everyone who John was $\left\{\begin{array}{l}\text{hit by}\\ \text{happy with}\\ \text{etc.}\end{array}\right\}$

The appropriate identity furthermore would have to obtain between the target VP and a sub-part of the string analyzed by term 3 of (1.3.31), as shown rather dramatically by examples like the next one, where the necessarily identical constituents are circled.

(1.3.34)



In short, I can see no way to make sense of Bouton's proposal. Nevertheless, I am in sympather with his consluding remark that we must "somehow provide a means of collapsing structural descriptions like those in (30) [=(1.3.31) above - I.A.S.] and (32) [=X - VP - Y - VP - Z - I.A.S. (Bouton (1970, p. 164)). By "in sympathy with" I mean that I share Bouton's intuition that these "antecedent-contained" deletions are instances of the same phenomenon as ordinary VPD.

The only other account of sentences like these that I am aware of is that of Hankamer (1972b). Hankamer follows Morgan (1972) in assuming that one strategy for relative clause formation in English is deletion in situ. The relative deletion rule, Hankamer argues, can sometimes delete more than just the target, because it is subject to a principle which he calls "Pied-Wiping", on the analogy of Ross's (1967a) well-known "Pied-Piping" convention. Hankamer states this principle as the following:

(1.3.35) The Pied Wiping Effect:

> A deletion rule which is formulated to delete a specific target constituent can, under not very well understood conditions, also delete a larger constituent containing the target, if that larger constituent is identical to a constituent containing the controller of the target deletion.

Hankamer's claim, then, is that VPD is not involved in the derivations of the sentences of (1.3.29), which, under his account, arise from an over-zealous application of Relative Deletion.

This proposal encounters difficulties. First, it predicts that all cases of the deletion in question are inextricably linked with Relative

Deletion. Thus it predicts that no such deletions will be found if a wh-word is actually present, for all transformational analyses of relative clauses agree that relative clauses containing overt wh-words (instead of that or ø) are derived by movement, not by deletion. Thus the "Pied-Wiping" theory predicts that sentences like the next two, which most people find acceptable, are ungrammatical.

(1.3.36)(a) I spoke with everyone who Tom did ø.

(b) I did the very things which Tom warned me not to ø.

Secondly, it is not the case that all of these "antecedent-contained" cases of VPD involve deletion of "a larger constituent containing the target". That is, there are cases where a VP to the right of the relativization site is deleted, such as (1.3.37)(a)-(e), which were observed by Bouton (1970, p. 156).

(1.3.37)(a) Bill visited only those girls who invited him to ø.

(b) Sam apologized to those who expected him to ø,--but to no one else.

(c) Mark is going to try to kiss a girl tonight who slapped him for trying to ø just yesterday.

(d) Hiram seldom raves about people who don't want him to ø.

(e) Andrew sent a program to everyone who asked him to ø.

The missing VP's are clearly reconstructable in these examples (in (1.3.37)(a), ø = visit them, in (1.3.37)(b), ø = apologize to them, etc.), yet they cannot be accounted for by the "Pied-Wiping Effect", for none of them is even contiguous to the relativization site, let alone "a longer constituent containing" it.

A third objection to the "Pied-Wiping" theory arises when we observe that there are semantic restrictions on the deletion of antecedent-contained VP's. To illustrate this, let us consider (1.3.38).

(1.3.38) Betsy's father wants her to read everything her boss wants her to read.

This sentence, in which no deletion has taken place, is ambiguous. It is true, on one reading, if, say, the set of things Betsy's boss wants Betsy to read is a subset (possibly improper) of the set of things her father wants her to read. This is roughly paraphrasable as: "everything that is such that Betsy's boss wants her to read it is also such that her father wants her to read it". We might represent this reading more formally as in (1.3.39).

(1.3.39) $(\forall x)$[Betsy's boss wants (Betsy read x) $\rightarrow$ Betsy's father wants (Betsy read x)]

The other reading of (1.3.38) is one that might be true, say, in a situation where Betsy's father says to her: "Betsy, read everything your boss wants you to read!" Omitting details of tense and modality, as well as many other details, this reading seems adequately represented by

(1.3.40) Betsy's father wants [$(\forall x)$[Betsy's boss wants (Betsy read x) $\rightarrow$ Betsy read x]].

Now the "Pied-Wiping" theory predicts, since it is a purely syntactic theory, that if the VP: wants her to read (in the relative clause of (1.3.38)) is deleted, then both readings will still be possible. Indeed, any attempt to blame this deletion on a syntactic rule like Relative Deletion makes this same claim, at least tacitly.

The claim is strikingly false. Consider the result of the deletion in question, namely,

(1.3.41) Betsy's father wants her to read everything her boss does.

This sentence has another reading, which might be true if Betsy's father says to her: "Betsy, read everything your boss reads". This reading can be represented as in (1.3.42), and of course, it is also a reading of (1.3.43), which, presumably is the source for (1.3.41) on this reading.

(1.3.42) Betsy's father wants [$(\forall x)$[Betsy's boss read x $\rightarrow$ Betsy read x]]

(1.3.43) Betsy's father wants her to read everything her boss reads.

It is clear, however, that the elliptical (1.3.41) also has the reading sketched in (1.3.39). That is, it shares one reading with its presumed source (1.3.38). Crucially, the elliptical (1.3.41) cannot have the other reading of (1.3.38), the one sketched in (1.3.40). In other words, in a situation where it is clear that what Betsy's father wanted was for her to read whatever her boss tells her to read, there is no reading of (1.3.41) whose truth conditions would be met. The matter may require some reflection on the part of the reader, but these facts, I believe, are quite verifiable.

The "Pied-Wiping" theory now has three strikes against it. The antecedent-contained deletions we have been considering are not confined to cases of relativization analyzable as Relative Deletion, they are not always contiguous to the relativization site, and they restrict the possible interpretations of their sources in a way that the "Pied-Wiping" analysis is inherently incapable of accounting for.

In this section, then, we have raised various problems for any purely syntactic formulation of VPD. On the basis of the facts we have examined, we must, at the very least, conclude that overt syntactic identity between two VP's is neither a necessary nor a sufficient condition for VP-deletability. In the next chapter, we will give a unified account of all the observations made in this section. In the process of doing that, we will see even more sensitivity on the part of VPD with regard to matters of logical form.

1.4 Summary of Chapter One

In this chapter we have concerned ourselves with previous transformational analyses of VPD. Critical examination of these analyses has led us to a new proposal for the constituent structure of VP and AUX. Additionally, we have offered a new formulation of VPD, which interacts with Bresnan's notion of the Relativized A-over-A Principle in such a way as to account for the various problematic facts we have observed. This has led us to offer a hypothesis as to the nature of the interaction of constraints on rules. We have shown that that hypothesis, if correct, renders unnecessary two constraints on VPD that have been proposed by Kuno and Grosu. Although the syntax of VPD is accounted for under our proposal, the various facts examined in the last section of this chapter suggest that there is something more to be said about the interaction of VPD and matters of logical form. That, in fact, is what Chapter Two is all about.

Footnotes to Chapter One

1. Quirk et al. also cite as dubious examples of modal followed by progressive:

(*i*) ?*John won't <u>enter</u> the competition, but Peter is ∅.

[∅ = <u>entering</u> the competition]

It seems, however, that it is possible to construct acceptable examples of this sort, such as (*ii*).

(*ii*) John said he would never <u>take</u> money on the side, but I knew he was ∅. [∅ = <u>taking</u> money on the side]

It is clear that there are many subtle factors at work here which interact to induce a complex distribution of acceptability. For an attempt to navigate these treacherous waters (with a theory of markedness as their compass), see Halliday and Hasan (1973).

2. Although this observation is surely in general correct, I have nevertheless noted the following peculiar examples of VPD ignoring the difference between active and passive.

(*i*) Botanist: That can all be <u>explained</u>.

Mr. Spock: Please do ∅. [∅ = <u>explain</u>] (Star Trek rerun)

(*ii*) It should be <u>noted</u>, as Dummett does ∅, that... [∅ = <u>note</u>] (Lust (ms.))

Also note the following general type of discourse:

(*iii*) Speaker A: Someone mugged Tom yesterday.

Speaker B: Oh yeah?

Speaker C: You know, the same thing happened to Mary.

Speaker B: Wow!

Speaker A: You know, now that I think of it, Sandy was ∅, too.

[∅ = mugged]

This last kind of discourse, which I suspect is rather common, probably shows more about memory (or processing) than it does about grammar. It's clear that there is much more going on here than can be explained at the moment.

3. Bresnan's RAOAP is thus a "disambiguating" version of the A-over-A Principle, rather than an "absolute" one (this terminology is due to Kayne (1975, 115 ff.)). That is, it does not dictate that elements in certain syntactic configurations are absolutely "frozen" (i.e. unable to be extracted), but rather, when a transformation applies ambiguously to a given string, it rules out certain proper analyses.

4. I note parenthetically that A + W's assumption that by-phrases are NP's and not PP's leaves unexplained why those phrases can be conjoined with PP's:

(*i*) This was given by the people, to the people, and for the people.

General considerations of parallelism in conjoined structures disallow conjunction of elements of different syntactic categories:

(*ii*) *The book was given Harry and to Sam.

By-phrases never conjoin with NP's:

(*iii*) *The book was given Harry and by Sam.

All this suggests that by-phrases are PP's, not NP's.

5. A + W argue that the indefinite NP is Chomsky-adjoined to the VP by There-Insertion. The deleted meterial in sentences like (*i*) will then still be a VP, and hence able to undergo VPD.

(*i*) They said there was a boy on the dock, but there was a girl ∅.

[∅ = on the dock]

The claim that sentences like (*i*) (which incidentally many informants reject), are to be derived by VPD is dubious. Notice that this variety of deletion is rather restricted (unlike VPD). It does not apply backwards for instance:

(*ii*) *Those men who claimed there was a boy ∅ actually knew that there was a girl on the dock.

(*iii*) Those men who claimed there wasn't ∅, actually knew that there was a boy on the dock.

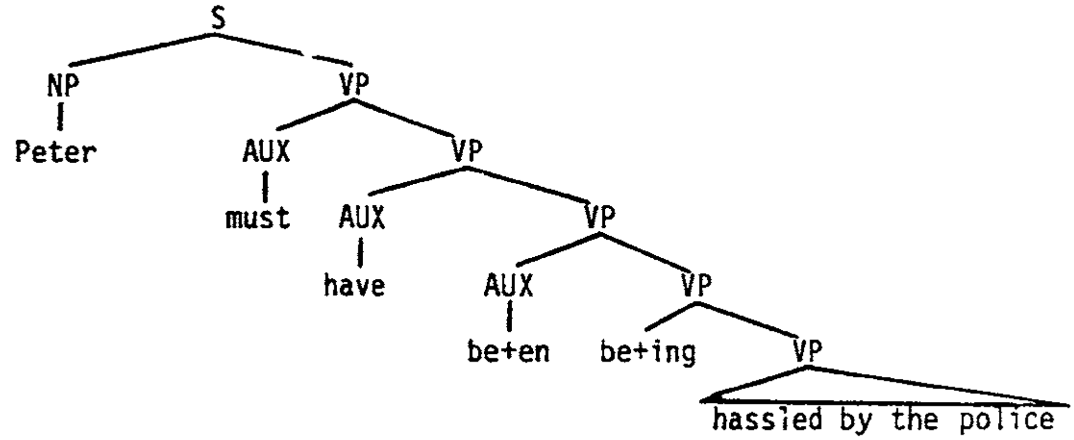
I do not pretend to understand all the idiosyncracies of the deletion rule in question (if indeed it is a deletion rule).

6. There is a certain amount of dialect variation with respect to these sentences. See note 8.

7. We will consistently speak of a rule (or a term of the structural description of a rule) "analyzing" either a terminal sub-string or a category (e.g. AUX). This somewhat imprecise usage should be understood as a short-hand for speaking of a terminal sub-string (analyzable as some category (e.g. AUX)) being a value for some predicate in the structural condition of a rule. For a more precise discussion of such matters, see Bresnan (1976a).

8. As mentioned previously, some people accept examples like (1.2.32)(a) (and (1.2.26)). These examples would be generated of course if we assumed the minimally different structure for such sentences sketched in (*i*).

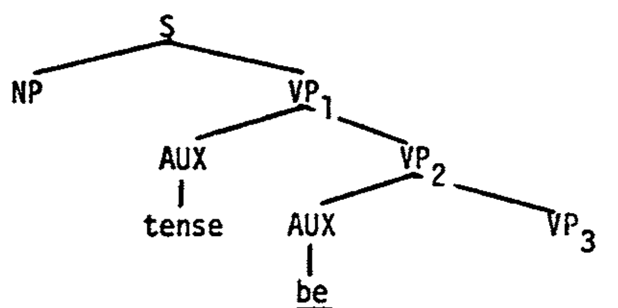
(*i*)



Perhaps dialects differ in this regard. Interestingly, this proposal makes predictions with respect to sentences involving "floated" quantifiers. See note 15.

9. The syntax of VPD, as it has been developed in this chapter has at least one apparent shortcoming. Since tense is always introduced in the "highest" AUX, and auxiliary be's occur in subordinate AUX's, structures like the following are generated:

(*i*)



Nothing has yet been said that would prevent the deletion of VP_2 in such structures, leaving stranded tense morphemes:

(*ii*) Betsy-past-be-talking when Peter past-be-talking

(*iii*) Betsy-past-be-talking when Peter-past.

To this, Do-Support must then apply, yielding...

(*iv*)(?)Betsy was talking when Sam did ∅.

In a theory of Deep Structure Interpretation, this problem is quite formidable (the second clause would be assigned a progressive interpretation). In a theory of Surface (or Shallow) Structure Interpretation of the sort we will propose in Chapter Two, however, it is not at all clear how much of a problem this is.

In any case, it seems that this entire matter can be satisfactorily accounted for if, following a suggestion of Howard Lasnik's (personal communication), we allow Δ as one possible expansion of M (Modal) in the base. Do-Support then replaces Δ by do. We have two possible options under this proposal with regard to the non-co-occurrence of do with have and be. Either, as Klima, Emonds, Jackendoff and Lasnik suggest, there is a rule that moves have and be leftward (to replace Δ) or else the matter is simply one of selectional restrictions. In either event, the presence of $[_M \Delta]$ in underlying structure will allow us to order Affix Hopping before Subject-Auxiliary Inversion and before VPD. Subject-Auxiliary Inversion will move $\begin{bmatrix}\text{+verb}\\ \text{+AUX}\end{bmatrix}$ ([Δ + Tense] will be eligible) and Do-Support can then freely apply to replace Δ by do. Since Subject-Auxiliary Inversion, Affix-Hopping, and Do-Support all apply before VPD, as is independently suggested by the general ordering hypothesis we will offer in Chapter Two, structures like (*i*) (with dangling tense) will have all been eliminated by the time VPD applies. Therefore the problem of (*ii*), (*iii*), and (*iv*) will not arise. At this writing, I am inclined to adopt Lasnik's proposal.

10. The proper formulation of this rule under A + W's proposal is actually more complicated than they indicate. Term 2 of the rule cannot be as indicated in (1.2.36) (in their proposal), because their rule of EN/ING Hopping is ordered before There-Insertion. Consequently, they would have to reformulate

the structural description of the rule to look something like (*i*).

(*i*) $\underset{[-\text{Def}]}{\text{NP}}$ - Tense (M) $\langle \text{have} \rangle_{\alpha}$ - be $\langle \text{en} \rangle_{\beta}$ - VP

α iff. β

I strongly suspect, however, that EN/ING Hopping is not a (syntactic) rule at all, in which case a simplification is possible.

In our analysis, however, an even further simplification is possible, for the sequence: tense - (M) - (have) is always an AUX. Thus, whereas A + W must specify an ad hoc set of conditions on Term 2 of their There-Insertion rule, we can simply specify that Term 2 is AUX.

11. We are not forced to this if we accept Lasnik's suggestion discussed in note 9.

12. We have intentionally omitted a discussion of VP-Preposing, which A + W appeal to for further support of their analysis. The data in this area seem to me extremely murky, for VP-Preposing is a marked stylistic process. Most people's intuitions about such sentences, or so they tell me, are intuitions about whether or not somebody else might say them, or about whether or not they have ever read a sentence like them.

Amidst this general unclarity of the data, it is nevertheless the case that A + W's analysis and ours both generate ungrammatical examples if VP-Proposing freely fronts any VP, as A + W suggest. Their analysis, for instance, incorrectly generates (*i*).

(*i*) *John claimed he had been being hassled by the police, and being hassled by the police he had been.

Our analysis also generates (*i*), but if VP-Preposing is unconstrained, we incorrectly generate non-sentences like (*ii*) and (*iii*) as well.

(*ii*) *They said he had been being careful, and been being careful he had.

(*iii*) ?*They said he should be careful, and be careful he should.

There is independent evidence that VP-Preposing must be constrained. Consider the ungrammaticality of the following example, which involves fronting of a verb-that clause sequence that is almost uncontroversially a VP.

(*iv*) *They said he would think that she liked him, and think that she liked him he did.

A possible solution would be to restrict VP-Preposing so that it can front only $[_{VP}V - (\left\{\begin{matrix}NP\\PP\end{matrix}\right\})]$ or $[_{VP}Adj - (PP)]$ (further collapsing is possible). This would account for the ungrammaticality of (*i*)-(*iv*), but further refinements may be necessary.

13. Curiously, Dougherty (1970, 1971) discusses sentences like (1.2.42)(a), but does not mention that they have no source in his analysis.

14. Some informants accept these sentences only with accent on not or all. I don't know why this should be the case. Extra-grammatical factors may well be involved.

15. Baker (1971, p. 168 nt. 3) appears to speak a different dialect. He rejects sentences where adverbs (and presumably quantifiers) are to the right of a sequence Modal + have (e.g. (1.2.44)(a)). Indeed one or two of my informants also reject such sentences.

Notice that if we assume the minimally different constituent structure mentioned in note 8, where modals and perfect have are each in a separate AUX, then the AUX Reordering Principle predicts precisely the facts of Baker's dialect.

This amounts to an empirical hypothesis that speakers who allow deletion of have after a Modal (by VPD) are the same people who reject sentences like (1.2.44)(a), where "floated" Q's and Adverbs appear to the right of post-modal have. On the basis of a limited investigation, this prediction seems to be correct.

16. We might again hypothesize that the variation in informants' judgements corresponds to variations in constituent structure.

17. Sentences like this one are, of course, acceptable if the adverb is preceded by a pause. In that case, the adverb in question follows the deletion site. Some such cases may be best regarded as parenthetical.

18. The proper formulation, as Ken Hale points out to me, may be intimately involved with prosodic matters. Consider (*i*):

(*i*) Where did you say they were all át?

Is it the case that a "floated" Q must simply have something stressed after it to "lean on"? Perhaps.

19. Kuno's arguments, which primarily involve the "do so test" (Lakoff and Ross (1966)), duplicate in many respects those offered by Emonds (1970).

20. Grosu's conception of VP-Specifier is extensionally equivalent to the standard conception of AUX. The terminological quibble is irrelevant to the present discussion.

21. Lest one object that underlying $\emptyset_2$ was only go to Princeton, consider

the different sense, and strangeness, of (*i*).

(*i*) Mary can't go to Princeton in the fall, but she can go to Princeton in the spring, although if she does go to Princeton, those who expect her in the fall will be very disappointed.

22. Actually it seems that whatever process deletes just verbs (leaving AUX's behind) is considerably more constrained than VPD. In general, such deletion cannot apply backwards, for instance:

(*i*) *The man who had ø the paintings, hadn't sold the drawings.

(*ii*) It doesn't bother Harry that Bill left, but it does ø me ø.

(*iii*) *The man who said it didn't ø him ø knew it bothered Harry that John left.

I have nothing enlightening to say about these cases.

23. Bresnan (1975, 1976a) makes the same assumption as Grosu, but supports it with an in-depth analysis of Comparative Deletion. We examine Bresnan's proposal in detail in Section 3.3 of Chapter Three.

24. One might construe certain remarks of Lees (1960, p. 76) (which are directed toward formulating the notion of recoverability of deletion in terms of identity of T-markers) as implying such a proposal. The same is true for the deep structure identity theories of Ross (1967a) and Lakoff (1968) (presuming the Lakoff-Carden theory of representing Quantifiers as "higher" predicates in underlying structure). We return to this matter in the next chapter.

25. One might be tempted to chalk this up to general conditions of parallelism required for deletion in coordinate structures to take place (see Hankamer

(1971)), Recall, however, that VPD is not restricted to coordinate structures (see Section 1.1 of this chapter).

26. However, as Larry Horn points out (personal communication), sentences like (*i*) seem much more acceptable than our previous examples with relative clauses containing more than one wh-word.

(*i*) My brother Al, who John liked, but who Harry didn't ø, was a nice guy.

If sentences like this are indeed fully grammatical, then that suggests that relative wh-words should not be treated as variable binding operators, and that some independent explanation must be found for the deviance of examples like (1.3.22) in the text.

Alternatively, we might choose to regard examples like (*i*), which crucially involve double contrastive foci , as ungrammatical, but derivativel generated. Multiple contrast environments like this are a common place for rules of grammar to relax. Notice, for example, that even the paradigm case of an ungrammatical passive, namely (*ii*) (cf. Chomsky (1965)) can become perfectly acceptable in a multiple contrast situation like (*iii*).

(*ii*) *England was died in by John.

(*iii*) Our beloved country, that has been líved in and díed in by so mány for so lóng, must not be forsaken.

The same is true for other standard cases of ungrammatical passives:

(*iv*) *John was resembled by his brother.

(*v*) John was much more closely resembled by his older brother than by his younger brother.

Given this, it is difficult to know how best to deal with Horn's example and the one given in the text.

CHAPTER TWO

Verb Phrase Deletion and Logical Form

2.0 Recoverability of Deletion

In the previous chapter, we touched briefly on the notion of recoverability of deletion.[1] Most of the problems noted in that chapter, it would seem, especially those of the final section, go to the very heart of that matter. Thus it will be of some value, before proceeding, to examine the various conceptions of recoverability that are to be found in the literature.

The standard formulation is due to Chomsky (1965) who writes:

> A deletion operation can eliminate only a dummy element, or a formative explicitly mentioned in the structure index (for example _you_ in imperatives), or the designated representative of a category (for example, the _wh_-question transformations that delete Noun Phrases are in fact limited to indefinite Pronouns...), or _an element that is otherwise represented in the sentence in a fixed position_. [emphasis added - I.A.S.] (Chomsky (1965, p. 144-145))

It is the underlined clause in this passage that concerns all the formulations of VPD discussed in the last chapter. Conditions like "2 = 4" specify that the VP analyzed by term 2 of the S.D. of the transformation is the "element represented in a fixed position." Put somewhat differently, the erasure operation performed by the rule _uses term 2 to delete term 4_ (cf. Chomsky (1965, p.145)), or _deletes term 4 under identity with term 2_.

However, as Chomsky also observed (1965, 177 ff, p. 234, nt. 38), exact identity of target and trigger seems to be too strong a requirement with respect to the deletion operation involved in such sentences as those in (2.0.1) (details are omitted).

(2.0.1)(a) I know several more successful lawyers than Bill.

[Bill is a lawyer]

(b) I have a friend from England.

[the friend is from England]

Chomsky proposes (1965, p. 181) that the notion of "identity", with respect to deletion rules must be weakened to a notion of "non-distinctness", where a target and a trigger which differ only with respect to transformationally-introduced features such as definiteness and plurality, are considered to be non-distinct.

Lees (1960) was, it seems, the first to point out that mere identity of terminal strings is insufficient as a recoverability condition. Lees cites the following example, which would be derived by deletion of one

(2.0.2) *Drowning cats, which is against the law, are hard to rescue.

of the two (grammatically different) constituents of the same shape in the following two sentences, if mere terminal string identity were sufficient.

(2.0.3) {Drowning cats are hard to rescue.
Drowning cats is against the law.

Lees concludes that "it is necessary to specify that the two constituents in question have the same phrase structure" (Lees (1960, p. 75). He further speculates that "identity of phrase structure must then mean something like "same internal constituent structure," i.e., the two constituents under consideration must be traceable back to the same node of identical derivation trees" (p. 76). Put somewhat differently, Lees' speculation is that in order for two constituents to be "identical" they must have the same underlying structure, and if they are transformationally-created constituents (Lees has in mind nominalization transformations), then they must be the result of applying the same transformations (in the same order).

One might ask whether the example Lees cites warrants such a conclusion. In particular, one might suggest that the two senses of NP's like drowning cats are reflected as a node-labelling difference in surface structure. One understanding might be, say, Adjective-Noun; the other, Noun-Noun (see also the discussion in Chapter One). If such a distinction, which is commonly assumed without argument, can be justified (see Chomsky (1955) for some initial justification), then one need not interpret Lees' example as showing any more than that identity of surface constituent structure (including identity of node labels) rather than mere identity of terminal string constitues an adequate identity condition.

Chomsky (1968) presents another example of this type (due to Ross), namely, the following:

(2.0.4) I know a taller man than Bill, and so does John.

This sentence, Chomsky notes, can be interpreted either as in (2.0.5)(a) or (b), but not as in (2.0.6)(a) or (b).

(2.0.5)(a) I know a taller man than Bill does, and John knows a taller man than Bill does.

(b) I know a taller man than Bill is, and John knows a taller man than Bill is.

(2.0.6)(a) I know a taller man than Bill is, and John knows a taller man than Bill does.

(b) I know a taller man than Bill does, and John knows a taller man than Bill is.

This is surprising, because independent applications of the rule(s) that effect ellipsis in comparative clauses should produce the following intermediate stage of derivation from any of the four sources indicated above.

(2.0.7) I know a taller man than Bill, and John knows a taller man than Bill.

But this intermediate stage should then give rise to (2.0.4) irrespective of how that intermediate stage was reached. The fact that (2.0.4) has only the two interpretations in (2.0.5) again shows that more than mere identity of terminal string is necessary. Chomsky maintains that in order to account for restrictions on deletions such as the case in question, we must know "that the two conjuncts of [(2.0.7)] derive from underlying structures in which the same element was deleted."

Again, it seems to me that this conclusion is not forced by the facts. One thorn in the side of the argument is the fact that the intermediate source for the deletion, i.e. (2.0.7) is itself only two, not four-ways ambiguous for most, if not all speakers (Chomsky himself points out the possibility of this (1968, p. 35, nt. 12)). This is especially true for a rendition of this utterance (2.0.7) where the second verb phrase is destressed.

(2.0.7)' I know a taller man than Bill, and
John knows a taller man than Bill.

At the end of Chapter Three we will discuss briefly what appears to be a prosodic principle of deletions, namely that all deleted material must be destressed. It is particularly striking that (2.0.7)', which bears precisely the prosodic contour that, by that prosodic principle, is necessary in order for deletion to take place, is precisely the rendition of (2.0.7) where the four-way ambiguity is least possible. Thus, it may be the prosodic system which induces the narrowing of the class of possible readings, and the non-four-way ambiguity of the deleted sentence (2.0.4) may simply result from the interaction of the deletion system with the prosodic system, i.e. from the prosodic principle just mentioned.

Alternatively, if Hankamer's (1973b) contention that there are two than's in English is correct, then there is another solution to this dilemma. Hankamer suggests that comparative sentences which contain simply than-NP which are interpreted as in (2.0.5)(b) might not be derived by deletion at all. Following this suggestion, these sentences would be analyzed simply as preposition than with prepositional object.

Each conjunct of (2.0.7) then would involve the preposition than when interpreted as ...than Bill is, but would involve the complementizer than when interpreted as ...than Bill does. Only when two complementizers or two prepositions are present would deletion be allowed. In this way identity of surface constituent structure, taking into account differences of node labelling could account for the observed facts without recourse to identity of deep structure.

Ross (1967a, p. 63, nt. 19) draws the same conclusion as Chomsky on the basis of these same sentences. He adds one more example, namely

(2.0.8) I divulged when Bill promised to call me, but I did so reluctantly.

Here the left conjunct is ambiguous (when can modify either promise or call), yet the example is only two, not four-ways ambiguous. Ross concludes that "it is clear that reference has been made to the deep structure...".

This would appear to be a more convincing example, but it should be noted that if one adopts either the indexing theory of Baker (1968) for indirect questions, or trace theory à la Chomsky, then again the necessary distinctions are represented in the surface (more precisely, the pre-deletion) phrase marker.

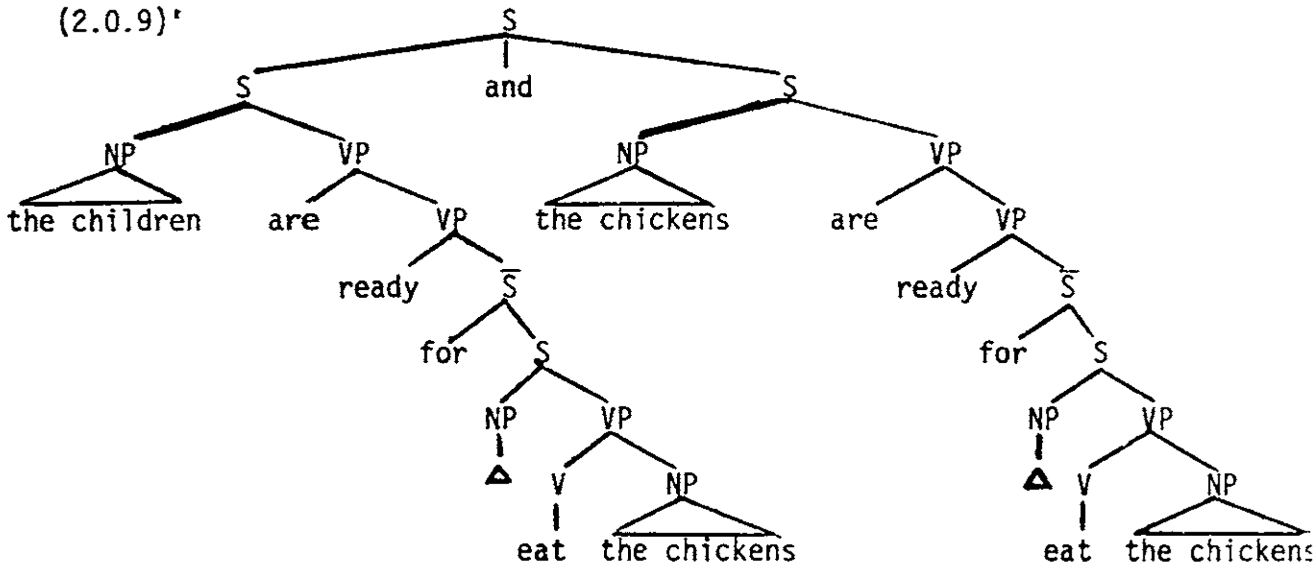
Lakoff (1968) also discusses certain difficulties of identity and recoverability of deletion.[2] He notes similar problems with regard to sentences like these, which, as one might expect, are also only two, not

(2.0.9) The children are ready to eat and so are the chickens.

four-ways ambiguous. (cf. the ambiguity of the <u>chickens are ready to eat</u> in isolation.)

Lakoff proposes an account of the notion of deep structure identity in terms of a rather complex indexing mechanism which assigns sets of indices throughout transformational derivations. The mechanism Lakoff comes up with (collaborating with Ross), however, fails to account for the possibility of deletion in (2.0.9) on the <u>ready to be eaten</u> reading, as Lakoff notes (p. 73). The problem is that on this reading, the deep structure for (2.0.9) is roughly as follows.

(2.0.9)'



The theory of deep structure identity therefore incorrectly predicts that no deletion is possible (on this reading) since the circled deep structure VP's in (2.0.9)' are not identical.

It is not difficult to add to the list of examples that are unaccounted for by the deep structure identity requirement. In most syntactic theories, all the following examples have deep structure sources where the relevant identity does not obtain.

(2.0.10)(a) John was hassled by the police, and Sam was ∅, too.

(b) Betsy seems to me to be unhappy, and Sandy does ∅, too.

(c) Betsy is easy to talk to, and Peter is ∅, too.

We will examine these cases in more detail at the end of the next section. Here, we only flag them as quite problematic for the deep structure identity hypothesis.

Lakoff concludes his discussion without a precise resolution of the problem at hand, but with a suggestion that "items that do not appear in the derived structure are completely irrelevant to the question of linguistically significant identity" (1968, p. 74). Such a formulation, though perhaps adequate for the facts just considered, is hardly enlightening.

Although this problem with the idea of deep structure identity as a necessary condition for deletion has never been satisfactorily cleared up, there is nevertheless a certain body of literature that presupposes its correctness and goes on from there. The matter is further complicated by the fact that the literature I am referring to equates deep structure with "meaning". Thus Lakoff (1970) and Catlin and Catlin (1972) both presuppose that VPD requires identity of meaning, and enter into a debate about whether volitional vs. non-volitional interpretations of verbs like <u>hit</u> (<u>the wall</u>) should be represented as a (covert) semantic distinction in underlying structure. The sentences in question are those like the following, where, it is claimed, both conjuncts must be interpreted either volitionally or non-volitionally.

(2.0.11) John hit the wall, and so did Pete.

(A rather precise summary and evaluation of such arguments can be found in Sadock and Zwicky (1973).)

Now surely in some sense, "sameness of meaning" is what all the previous discussions of deletion identity conditions have been trying to get at. It seems to me, however, that none of the existing discussions has made precise just what is meant by sameness of meaning. Presumably, Lakoff has in mind that identity at the level of "underlying structure", as he conceives of that notion, is a necessary condition for deletion. Thus (as was pointed out in footnote 24 of the last chapter), some (but not all) of the observations made in the last chapter might be as expected, given a general dependence of deletion on underlying structure as Lakoff, McCawley, Postal and others conceive of that notion. But to assume that grammar is a mapping from semantic representations of the generative semantics sort directly into surface structure (i.e. to presume underlying structure where quantifiers are represented as "higher predicates" and to presume such transformations as "quantifier lowering", commits one not only to an unintelligible notion of quantifier, but also to such devices as arbitrary derivational constraints (Lakoff (1969, 1972), Postal (1972)). This unquestionably results in a methodologically worse theory, contrary to the remarks made by Postal (see the remarks by Chomsky in Parret (1974)). We will therefore not pursue such a theory here.

Notice, however, that underlying structures, in the generative semantics sense, are generally taken to be equivalent to semantic representations. In other words, underlying structures on this view are taken to be such that they determine such semantic matters as logical consequence, synonymy, and the like. Therefore the claim that deletion rules require identity of representation at this level is a very strong claim. It is in fact tantamount to the claim that such rules as VPD require identity at the level where logical consequence is determinable.

Such a claim, interestingly, seems to be incorrect. Consider, for example, the ambiguity of (2.0.12).

(2.0.12) They caned a child severely when I was a child.

This sentence may be understood _generically_, i.e. as conveying what normally was the case in my childhood if, say, a child behaved improperly. On another understanding, (2.0.12) is a statement about some _specific_ child who was beaten in my childhood. No one would disagree, I don't think, that determining which understanding of a sentence like this is being intended is relevant to determining what inferences can be drawn. On the generic reading, for instance, it does not necessarily follow that any child had ever actually been caned, as it does on the specific reading.

The following sentence, due to Geoff Nunberg (personal communication), is therefore of considerable interest.

(2.0.13) They caned a student severely when I was a child, but not like Miss Grundy did ø yesterday.

[ø = caned a student (severely)]

Here, it seems, the first clause, which contains the trigger VP, can be interpreted generically, while the clause containing the target VP is interpreted specifically. Such "mixed" interpretations occur in either order:

(2.0.14) Miss Grundy caned a student yesterday, just like they did when I was a boy.

Here the first clause may be given a specific interpretation, and the second clause, a generic interpretation.

Kuno (1974) also cites some examples of this kind, like this next one, where it is also possible to interpret the deleted NP non-specifically, and the NP in the trigger VP, specifically.

(2.0.15) Jane ended up marrying a doctor, although she didn't want to ø.

[ø = marry a doctor]

These examples, it seems to me, lead to the conclusion that <u>there is no</u>

one level of representation which determines both the identity relevant for deletion rules and such matters as logical consequence. This is rather surprising in view of the many facts examined in the previous chapter, where certain logical matters such as assignment of quantifier scope seemed to play an important role in determining deletability. Indeed the goal of giving a precise account of such matters as logical consequence lies at the very heart of quantification theory. What we are in need of then, is some principled account of the kind of semantic information that is relevant to determining the applicability of deletion rules. In the next section, we will offer an hypothesis that addresses this matter directly.

To summarize, we have examined various conceptions of the notion of recoverability of deletion as they have appeared in the literature. A common thread in the discussions of this notion has been that a proper view of identity for deletion must make reference to underlying structure. Some of the arguments for this have been seen to be uncompelling; moreover the requirement of identity of underlying structure makes certain incorrect predictions about deletability (most strikingly with respect to the ready to eat sentences observed by Lakoff). We have also seen that to require identity of semantic representation, where semantic representations are taken to be such that they determine logical consequence, is too strong a requirement, since at least one deletion rule, VPD, is not mindful of certain distinctions that must be represented at that level, i.e. the difference between specific and non-specific interpretations of indefinite NP's. Thus in this section we have left unanswered the question of at which level deletion is recoverable.

2.1 Logical Form[3]

As we have seen in the last section, neither identity of deep structure (in the sense of Chomsky (1965)), nor identity of underlying structure (= "logical structure") in the generative semantics sense (if we take those structures seriously, i.e. as being sufficient to determine matters like logical consequence) seems to be a workable criterion for ascertaining whether or not a particular deletion is recoverable. Moreover, we have still to contemplate a proper account of the observations made at the end of Chapter One, where certain matters of scope of logical operator also seemed to play a role in determining deletability. What appears to be necessary then is a level of representation between the syntax proper and whatever level is to provide a full representation of meaning.

Chomsky (1975a) proposes a theory which provides just such a level. He gives the following picture of grammar and its relation to "meaning" (p. 105).

(2.1.1)(a) Sentence Grammar:

$$\xrightarrow[B]{\text{(Base Rules)}} \begin{array}{c}\text{Initial}\\ \text{Phrase}\\ \text{Markers}\end{array} \xrightarrow[T]{\text{(Transformations)}} \begin{array}{c}\text{Surface}\\ \text{Structures}\end{array} \xrightarrow[\text{SR-1}]{\text{(Rules of Semantic Interpretation)}} \begin{array}{c}\text{Logical}\\ \text{Forms}\end{array}$$

$$\text{(b)} \left\{\begin{array}{c}\text{Other Semantic Rules}\\ \text{(SR-2)}\\ \text{other Systems}\end{array}\right\} : \begin{array}{c}\text{Logical}\\ \text{Forms}\end{array} \longrightarrow \text{"meaning"}$$

In Chomsky's system, "logical form" is one endpoint of sentence grammar. Each of the mappings indicated in (a) is performed by a component of sentence grammar. This output of the sentence grammar is mapped by various other rules, and possibly other cognitive systems into a fuller representation of "meaning". The kind of semantic rules that Chomsky sees as part of sentence grammar, i.e. the rules of SI-1, are the rules involving "bound anaphora, scope, thematic relations, etc."

Within such a theory of grammar, we might offer the following hypothesis as to the nature of the recoverability of deletion.

(2.1.2) Deletion is recoverable at the level of Logical Form. (Where Logical Form is viewed as the output of Sentence Grammar, and the input to other devices which then map Logical Form into "meaning", as in (2.1.1).)

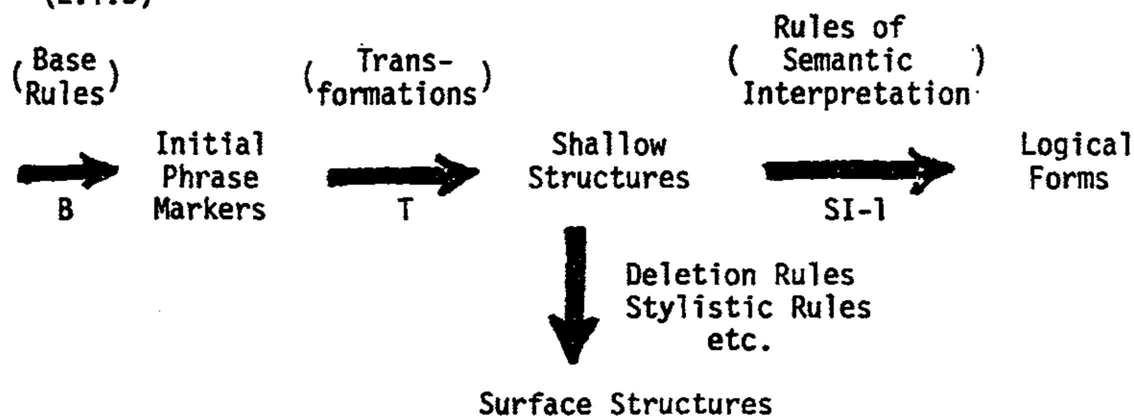
This hypothesis will require certain clarifications vis à vis the nature of the rule systems sketched in (2.1.1). For instance, in Chomsky (1975b, p. 36), it is suggested that "we might then try to devise general rules of inferences, truth conditions, etc. for such representations...[= logical forms - I.A.S.]." If we assume rules of inference are to give an account of specific vs. non-specific NP's, however, then, for the reasons discussed in the previous section, such rules of inference should not be defined on logical form, but rather on some subsequent level, i.e. at some point where perhaps more than mere sentence grammar is taken into account. Alternatively, one might assume that a certain class of inferences is definable on logical forms, and that whatever account we give of specific vs. non-specific NP's interacts in some principled way with an otherwise adequate theory of inference. We will not take a stand here on what other cognitive systems may be relevant in actually giving an account of inference or whether any further levels are definable between logical form and "meaning (say semantic representation in the sense of Katz and Postal (1964) or

Katz (1972)). Rather, we will address ourselves to giving an account of the nature of logical forms if those structures are to specify certain matters of scope, as Chomsky intends, and, at the same time, provide an adequate account of when deletion is possible.

The notion surface structure in (2.1.1) deserves some comment. As Chomsky (1975b) notes, surface structure no longer correspond exactly to the use of that term in the standard theory (Chomsky (1965)) or the Extended Standard Theory (Chomsky (1972a,b)) because it abstracts away from certain operations, such as stylistic reorderings. In languages that exhibit rather free word order ("scrambling" in the popular sense of that term), for instance, semantic interpretation must certainly take place at a level prior to the operation of such rearrangement processes.

Given the theory we have been developing, i.e. a theory that countenances deletion rules, it is certainly the case that semantic interpretation must take place at a level prior to deletion. The picture of linguistic systems that emerges, then, is one that countenances a level which we might call "shallow structure" (the term is due to Postal, though it is used somewhat differently by him), which serves as the input to rules of semantic interpretation (SI-1).

(2.1.3)



We may refer to such a view of the relationship of deletion rules, rules of semantic interpretation and logical form as the "theory of shallow structure interpretation".[4]

Given such a theory, we may then pose the question of what representations at the level of logical structure look like. Most linguists (McCawley (1970, 1973), Lakoff (1972), Postal (1971), Jackendoff (1972), to name just a few) and many philosophers are fond of representing simple sentences like "Betsy loves Peter" as something like (2.1.4) at the level of logical structure as they conceive of that notion (which, in general, is somewhat different from our conception of that notion), where love is treated as a two-place predicate.

(2.1.4) LOVE (BETSY, PETER)

Others, influenced primarily by the work of Montague (see Thomason (1974), Montague (1975), Partee (1972, 1975), Bresnan (1976b)), have maintained, or else simply assumed, that such relations as the grammatical relation of subject-predicate should be reflected in a logical representation, at least at some level. It seems to me that some of this latter research has made considerable progress toward dealing with the relationship of natural language sentences and logical formulae in a rigorous manner, though one might very well take issue with certain specific conclusions that have been reached (say, Montague's "proper treatment" of quantifiers).

Nevertheless, we will employ one useful device that has been utilized rather extensively in this literature, namely, Church's lambda calculus (Church (1941); see also Curry, Seys, and Craig (1958), Hindley, Lercher and Seldin (1972), and Curry, Seldin, and Hindley (1971)). We will assume that one rule of semantic interpretation assigns λ-representations to shallow structure VP's. This rule may be thought of as analogous to Partee's (1975) Derived Verb Phrase Rule. Thus at our level of logical form, a sentence like "Betsy loves Peter" will not be represented as in (2.1.4), but rather as something closer to (2.1.5).

(2.1.5) $\lambda x(x \text{ love Peter})(\text{Betsy})$

We might assume further that certain other rules operate prior to logical form to alter the structure inside our λ-expressions, say to indicate predicate-argument relations in some way, or to indicate thematic relations. The details of such rules will not concern us very much here, though we might write (2.1.5) in a somewhat more articulated manner as (2.1.6).

(2.1.6) $\lambda x(\text{love}(x, \text{Peter}))(\text{Betsy})$

Intuitively, a sentence like "Betsy loves Peter" says that Betsy has a certain property, namely, the property of loving Peter. λ-representations of the sort under discussion capture this intuition nicely. Our λ-representations will not always be so intuitive, for we will need to assign all surface verb phrases a λ-representation. Thus a sentence like "It's raining" will have a λ-representation at the level of logical form which predicates a property of "it". We might propose to treat transformationally-inserted entities like "it" as dummy elements at the level of logical form, and perhaps allow rules of SI-2 to eliminate such dummy elements, or perhaps to map logical forms of sentences like "it is raining" into propositional representations of another sort entirely. We will leave open the question of whether λ-expressions are in general eliminated, say, by λ-conversion

(= functional application), somewhere along the way from logical forms to fuller expressions of meanings. For a more detailed sketch of how complicated cyclic derivations, like, say, that of the following sentence might be treated within a λ-calculus framework that operates on very superficial

(2.1.7) There was believed to be a woman in the garden.

syntactic structures, the reader is referred to Bresnan (1976b).

A central motivation for employing λ-representations of the sort we are developing, is essentially to capture the intuition of McCawley (1967) about the recoverability of rules like VPD:

> The only way I know of stating this transformation [= VPD - I.A.S.] is to say that the deletion may take place only in a structure whose semantic representation is of the form $f(x_1) \wedge f(x_2)$.

Similarly, the program offered by Keenan (1971a) for sentences involving <u>so be</u>, which stresses the notion "identity of predication", has much in common with the theory we will develop here. McCawley and Keenan both are led to their position by a desire to solve the problem of "sloppy identity" first observed by Ross (1967a, 1969b), a problem which we will return to in Section 2.2 of this chapter. For the moment, however, we will motivate this general approach on independent grounds.

The notion "saying the same thing about" or "predicating the same property of" corresponds to the notion "equivalent λ-expression" in our theory of logical form. Thus the following sentence, which, following the intuitions of McCawley and Keenan, "says the same thing about" Sandy and Peter, will have a logical representation as in (2.1.9)

(2.1.8) Peter loves Betsy, and Sandy loves Betsy, too.

(2.1.9)(a) λx(x loves Betsy)(Peter) & λy(y loves Betsy)(Sandy)

or (b) λx(love (x, Betsy))(Peter) & λy(love (y, Betsy))(Sandy)

In each of (2.1.9)(a) and (b), λx(...) and λy(...) are equivalent. Given that it is identity at the level of logical form that is relevant for deletion, then assuming logical forms to have roughly the shape sketched in (2.1.9) allows us to correctly predict that (2.1.8) can undergo VPD to become (2.1.10) (assuming the syntactic rules developed in Chapter One, including automatic application of Do-Support (and tense Hopping) after VPD has applied).[5]

(2.1.10) Peter loves Betsy, and Sandy does ∅, too.

We will take a rather narrow view of the nature of the semantic rules of SI-1. That is, we will presume that such rules are primarily scope assignment rules that "extract" quantifiers creating bound variables (precisely as is assumed in Chomsky (1975b)). This is necessary because of what we might call "the correspondence problem". That is, we will assume that deletion rules are defined on syntactic objects, e.g. VP's, but are subject to a general recoverability condition that concerns the representations of those syntactic objects at the level of logical form. In order for this view to be coherent, we must be able to give a precise account of the nature of the correspondence between objects at two levels (here, say, between shallow structure VP's and λ-predicates at the level of logical form). Without very narrow constraints on what is a possible rule of SI-1, defining such a correspondence may become rather difficult.

Given this, logical representations really are creatures of two worlds. Their λ-predicate structure will reflect very closely the surface (more precisely, shallow) syntax, yet they will indicate scope, bound variables, etc. much in the way (it's reasonable to assume) that those matters will be indicated in fuller representations of meaning. It may well be the case, though, that certain differences in understanding are represented scopally at some level of meaning, but not at logical form. Specific vs. Non-specific understandings of the indefinite article, for instance, may be just such a case (see the discussion in Section 2.0).

There is one modification we will assume with regard to λ-calculus representations of logical forms. Finding very appealing Chomsky's (1975b) suggestion that there are certain generalizations that make reference to both logical form and surface order (in particular his proposal (following Wasow) that certain of Postal's (1971) cross-over violations and certain illegitimate instances of "backwards" variable binding by quantifiers can be given a unified account in terms of a constraint prohibiting logical structures containing a bound variable and a preceding anaphoric proform),[6] we will write logical forms such that λ-predicates follow their arguments. Our logical formulae will thus reflect surface word order extremely closely.

We give, for example, the logical form of a sentence like "Betsy loves Peter" (roughly) as in (2.1.11).

(2.1.11) Betsy, λx(x loves Peter)

We will discuss one justification for this notational quirk (having to do with the representation of stress in logical forms) in Section 2.2 and in Section 3.5 of Chapter Three.

In the following discussion, we will not be overly concerned with the internal structure of λ-representations. Since we will be primarily concerned with such notions as "identity of logical form", whatever further structure can be shown to be correct for such expressions will not be of immediate concern with respect to our arguments.

One more notion must now be brought to the fore. That is the standard notion of "alphabetic variance". Intuitively, two λ-expressions are alphabetic variants, if they differ only with regard to variable letters. The notion is not quite this simple, however. For two λ-expressions, $\lambda x(A)$ and $\lambda y(B)$, to be alphabetic variants, every occurrence of x in A must have a corresponding instance of y in B, and vice versa. Also, any Quantifier in A that binds variables (in A) must have a corresponding (identical) Quantifier in B that binds variables in all the corresponding positions (in B). However, if there are any variables in A that are bound by some quantifier outside of $\lambda x(A)$, then the corresponding variable in $\lambda y(b)$ must be bound by the same operator in order for alphabetic variance to obtain. ($\lambda x(\ldots)$ and $\lambda y(\ldots)$ are alphabetic variants in $(\forall z)$[John, λx(x loves z) & Bill, λy(y loves z)]). Crucially, if $\lambda x(A)$ contains a variable bound outside of $\lambda x(A)$ (for instance, z in $(\forall z)$[John, λx(x loves z)]) and $\lambda y(B)$ contains a corresponding variable bound outside of $\lambda y(B)$ (even one bound by an analogous operator, for instance, w in $(\forall w)$[John, λy(y loves w)]) the two λ-expressions are not alphabetic variants (though here the universally quantified expressions, considered as a whole, would be).[7]

By way of illustration, the following pairs of λ-expressions are alphabetic variants.

(2.1.12)(a) λx(x is happy) = λy(y is happy)

(b) λw(w loves John) = λz(z loves John)

(c) $\lambda w((\forall y)$[w likes y]) = $\lambda z((\forall q)$[z likes q])

(d) $\lambda w((Ez)$[w ate z]) = $\lambda q((Er)$[q ate r])

(e) λx(x said [Mary, λy(y likes x)]) =

λz(z said [Mary, λw(w likes z)])

(f) λx(x loves y) = λz(z loves y) as in

($\forall$y)[John, λx(x loves y) $\rightarrow$ Bill, λz(z loves y)]

Conversely, the pairs of λ-expressions in (2.1.13) are not alphabetic variants.

(2.1.13)(a) λx(x is happy) $\neq$ λy(y is sad)

(b) λw(w loves John) $\neq$ λz(z loves Mary)

(c) λx(x likes y) $\neq$ λw(w likes z) as in

(Ey)[John, λx(x likes y)] & ($\forall$z)[Bill, λw(w likes z)]) or in

John, λy(y said [Mary, λx(x likes y)]) &

Bill, λz(z said [Mary, λw(w likes z)]) [8]

There are many interesting questions about the nature of logical forms. Some of these we will take up in the discussion that follows immediately. We will also have more to say about the nature of logical forms in the last section of this chapter.

Presuming, however, that our rules of SI-1 will be such that the "correspondence problem" alluded to earlier does not arise (i.e. assuming that it makes sense to speak of a correspondence between shallow structure VP's and λ-expressions), we are now in a position to elaborate on our claim that a deletion rule like VPD is subject to recoverability at the level of logical form. We may state our claim precisely as follows.

(2.1.14) With respect to a sentence S, VPD can delete any VP in S whose representation at the level of logical form is a λ-expression that is an alphabetic variant of another λ-expression present in the logical form of S or in the logical form of some other sentence S', which precedes S in discourse.[9]

Note first of all, that this formulation makes the same predictions as a purely syntactic theory in a wide variety of cases. Thus in standard cases of VPD like the following, which can be assigned logical forms indicated in the framework we are developing, the deleted VP's are syntactically identical (hence deletable on the syntactic identity theory) and correspond to λ-expressions that are alphabetic variants (hence deletable according to (2.1.14)).

(2.1.15)(a) Peter likes Betsy, and Sandy does ∅, too.

(b) Peter homered, and (then) Betsy did ∅.

(c) Betsy homered after Peter did ∅.

(d) Betsy homered, and Peter didn't ∅.

(2.1.16)(a) Peter, λx(x likes Betsy) & Sandy, λy(y likes Betsy)

(b) Peter, λw(w homered) & Betsy, λz(z homered)

(c) [Betsy, λx(x homered)] after [Peter, λw(w homered)]

(d) [Betsy, λw(w homered)] & ¬[Peter, λy(y homered)]

We might want to represent some of these sentences differently, of course. If, as is frequently suggested, tense should be treated as a sentence operator (say [-ed] for past tense), (2.1.15)(b) might be represented by any of the following logical forms.

(2.1.17)(a) [-ed][Peter, λw(w homer)] & [-ed][Betsy, λz(z homer)]

(b) [-ed][Peter, λx(x homer) & Betsy, λy(y homer)]

(c) Peter, λu([-ed](u homer)) & Betsy, λx([-ed](x homer))

Any of these representations provides the necessary identity for our theory to guarantee deletability. None of the standard examples in (2.1.15), then, decides between the purely syntactic theory and the logical theory in (2.1.14).

The examples cited in the previous section, however, are more to the point. For ease of presentation we will repeat those examples in full as we reconsider them. (2.1.18) is a good first example.

(2.1.18) Someone hit everyone, and (then) Bill did ø.

Here the syntactic theory and the logical theory make different predictions. The former says nothing about what readings can be assigned to the left conjunct, for on either reading, there exists the appropriate syntactic identity for deletion to take place.

Within the theory of logical form we have been developing, the left conjunct in (2.1.18) has the two logical forms sketched below.

(2.1.19)(a) (Ex)[x, λy(($\forall$z)[y hit z])]

(b) ($\forall$z)(Ex)[x, λv(v hit z)]

(2.1.19)(a) is the result of applying the Derived Verb Phrase Rule, which creates λ-expressions, and subsequent application of quantifier scope assignment rules. Such rules, in most cases, simply move quantifier words to the front of the sentence they are in, creating bound variables as indicated (see Chomsky (1975b) and Reinhart (forthcoming) for more discussion of the workings of such rules). Notice, however, that quantifiers that are inside VP in shallow structure are given what is essentially "VP-scope", i.e. they stay "inside" the λ-expressions created by the Derived Verb Phrase Rule. The second reading, given in (2.1.19)(b), we will take to result from application of an optional "scope jumping" rule, that allows quantifiers to take wider scope under certain conditions. Alternatively, there may be some general principle that allows this scope to be assigned directly, without assuming any optional rule of "scope jumping". Such matters are highly complex, and are further complicated by the fact that certain quantifiers seem to require scope over other quantifiers in spite of their surface order. These questions are well beyond the scope of this investigation. (See Ioup (1976), Bennett (1974), Kroch (1974) for more discussion of these matters.)

Now consider the result of applying these same rules to the (shallow structure) source of the right conjunct of (2.1.18), i.e. to the sentence "Bill (did) hit everyone". Here no multiple scope possibilities exist, because the sentence contains only one quantifier expression. The result then is the following logical representation.

(2.1.20) Bill, $\lambda w((\forall u)[w \text{ hit } u])$

Comparing this formula with the two logical representations in (2.1.19), we find that $\lambda w(\ldots)$ here and $\lambda y(\ldots)$ in (2.1.19)(a) are alphabetic variants. In (2.1.19)(b), however, there is no λ-expression that is an alphabetic variant of $\lambda w(\ldots)$, for the only λ-expression there, $\lambda v(\ldots)$, contains no universal quantifier. Our logical theory of VPD therefore predicts that only when the left conjunct in "Someone hit everyone and (then) Bill hit everyone" is interpreted as in (2.1.19)(a) (where the existential quantifier has wide scope) is deletion possible. Our theory therefore explains precisely the fact we observed earlier.

A second argument for our theory of VPD concerns sentences like this one:

(2.1.21) Sandy greeted everyone when Betsy did.

We observed in Chapter One that the syntactic theory, which derives this sentence from (2.1.22), fails to predict that the former sentence is

(2.1.22) Sandy greeted everyone when Betsy greeted everyone.

ambiguous in a way that the latter is not. In our theory, (2.1.22) will be assigned the logical form in (2.1.23) (again ignoring the problem of the proper representation of <u>when</u>).[10]

(2.1.23) Sandy, $\lambda x((\forall y)[x \text{ greeted } y])$ when Betsy, $\lambda z((\forall w)[z \text{ greeted } w])$

The logical theory therefore predicts, as any purely syntactic theory would also predict, that (2.1.21) is derivable from (2.1.22), for $\lambda x(\ldots)$ and $\lambda z(\ldots)$ in (2.1.23) are alphabetic variants.

Recall, however, that we observed that this deleted sentence (i.e. (2.1.21)) had a second reading, one that it shares with (2.1.24).

(2.1.24) Sandy greeted everyone when Betsy greeted $\left\{\begin{matrix}\text{them}\\ \%\text{him}\end{matrix}\right\}$.

The logical form of (2.1.24) (on this reading) would in our terms be the following.

(2.1.25) $(\forall x)([\text{Sandy}, \lambda y(y \text{ greeted } x)] \text{ when } [\text{Betsy}, \lambda w(w \text{ greeted } x)])$

Since $\lambda y(\ldots)$ and $\lambda w(\ldots)$ contain the <u>same</u> bound variable, they are alphabetic variants. Thus the logical theory of VPD predicts that deletion is possible in (2.1.24). The result? (2.1.21) with the reading in (2.1.25). Note that this is a case where trigger and target VP are quite distinct in form. Such cases are highly problematic for any syntactic theory of VPD. It is of course possible that some other account could be given for the unexpected ambiguity of (2.1.21) within a purely syntactic theory. In the absence of such an account, however, the fact that the logical theory <u>explains</u> the ambiguity of (2.1.21) must be taken as a strong argument for that theory.

More support for the logical theory comes from the interaction of VPD and <u>wh</u>-Movement, which we also discussed in the previous chapter. There we noted that the VP-deletability of fragmented VP's seemed to correlate with whether or not certain variables were bound by the same operator at the level of logical form. The correlation seemed to be that when the logical form of a sentence containing a fragmented VP contains a variable bound by an operator that also binds a corresponding variable in the logical form of the antecedent VP, then deletion is possible.

This is precisely the prediction made by our theory. When the appropriate variables are identical, i.e. bound by the same operator, alphabetic variance is obtained. When the variable binding is diverse, the relevant expressions are not alphabetic variants. Let us reconsider the relevant facts in more detail.

(2.1.26) What did Harry take a picture of? ...*What did Bill $\emptyset$?

[$\emptyset$ = take a picture of]

Here the logical forms are as in (2.1.27)

(2.1.27)(a) (What x)[Harry, λy(y took a picture of x)]

(b) (What z)[Bill, λw(w took a picture of z)]

$\lambda y(\ldots)$ and $\lambda w(\ldots)$ are not alphabetic variants, since x and z are bound diversely outside of the λ-expressions. The logical theory predicts, correctly, that deletion is impossible.

Reconsider, now, this sentence.

(2.1.28) We finally got in touch with John, who my brother Al tried to visit, but couldn't.

It is not clear how best to represent the logical forms of sentences with non-restrictive relative clauses. We would claim, however, that at the level of logical form (in our sense of that term), the representation of (2.1.28) involves variable binding in the manner sketched in (2.1.29).

(2.1.29) $\ldots(\text{who } x)([\text{my brother Al}, \lambda y(y \text{ tried } (y, \lambda z(z \text{ visit } x)))]$
but $[\neg\, \text{could}[\text{my brother Al}, \lambda w(w \text{ visit } x)]])$

Since $\lambda z(\ldots)$ and $\lambda w(\ldots)$ are alphabetic variants, deletion is possible.

This contrasts with (2.1.30), whose logical form we would claim must involve variable binding like that sketched in (2.1.31).

(2.1.30) *We finally got in touch with John, who my brother Al tried to visit, but who he couldn't.

(2.1.31) $\ldots(\text{who } x)([\text{my brother Al}, \lambda y(y \text{ tried } (y, \lambda z(z \text{ visit } x)))])$
but $(\text{who } v)([\neg\, \text{could}[\text{my brother Al}, \lambda r(r \text{ visit } v)]])$

$\lambda z(\ldots)$ and $\lambda r(\ldots)$ are not alphabetic variants. Deletion is impossible. I believe that any theory of the logical form of non-restrictive relatives that treats each <u>wh</u>-word as a variable binding logical operator will provide the appropriate distinction for the logical theory of VPD to make the correct predictions.[11]

The observations concerning pseudo-clefts are another case in point. The contrast between these next two sentences was observed in the last chapter.

(2.1.32) *What Sandy carried was the baseball bat, and what Betsy did $\emptyset$ was the catcher's mit.

(2.1.33) What Sandy wanted to buy, but couldn't $\emptyset$, was the catcher's mit.

We might choose to represent pseudo-clefts logically in terms of set abstraction. This would lead to logical forms like the following for the previous examples (again oversimplifying).

(2.1.34) $\{\text{the baseball bat}\} = \hat{x}(\text{Sandy}, \lambda y(y \text{ carried } x))$ &
$\{\text{the catcher's mit}\} = \hat{z}(\text{Betsy}, \lambda w(w \text{ carried } z))$

(2.1.35) $\{\text{the catcher's mit}\} = \hat{x}([\text{Sandy}, \lambda y(y \text{ wanted } [y, \lambda z(z \text{ buy } x)])]$
$\text{but } \neg[\text{could } [\text{Sandy}, \lambda w(w \text{ buy } x)]])$

(This analysis is preferable to one alternative that comes to mind, namely Russell's iota operator, because of the possibility of plurals as pseudo-cleft foci.) In (2.1.34), $\lambda y(\ldots)$ and $\lambda w(\ldots)$ are not alphabetic variants (x and z are bound diversely); deletion is impossible. In (2.1.35), $\lambda z(\ldots)$ and $\lambda w(\ldots)$ are alphabetic variants, and deletion is possible.

In short, here, as well as in all the previous cases involving fragmented VP's, the purely syntactic theory fails to distinguish between those VP's that can, and those VP's that cannot, undergo VPD. The logical theory, on the other hand, given only one additional assumption (the correctness of treating <u>wh</u>-words as variable-binding logical operators), makes precisely the correct predictions about which fragmented VP's are deletable, and which are not.

Let us now return to the antecedent-contained deletions discussed at the very end of Chapter One. The reader will recall that the syntactic theory floundered badly with respect to those examples, as did Hankamer's "Pied-Wiping" theory. A coherent treatment of those facts would therefore constitute extremely strong support for any theory of VPD. The logical theory we have proposed provides a very coherent account, as we will see.

Sentences like the following ones, for example, certainly involve universal quantification.

(2.1.36) Alan will eat anything you want him to (eat).

(2.1.37) Sandy hit everyone that Bill $\left\{\begin{matrix}\text{hit}\\ \text{did}\end{matrix}\right\}$.

In our theory, we might represent these sentences logically as in (2.1.38)(a) and (b), using restricted quantification, or else as in (2.1.39)(a) and (b), treating them as implicational statements.

(2.1.38)(a) $(\forall x\text{: you}, \lambda y(y \text{ want } [\text{Alan}, \lambda z(z \text{ eat } x)]))\ [\text{Alan}, \lambda w(w \text{ eat } x)]$

(b) $(\forall y\text{: Bill}, \lambda x(x \text{ hit } y))\ [\text{Sandy}, \lambda w(w \text{ hit } y)]$

(2.1.39)(a) $(\forall x)(\text{you}, \lambda y(y \text{ want } [\text{Alan}, \lambda z(z \text{ eat } x)]) \rightarrow \text{Alan}, \lambda w(w \text{ eat } x))$

(b) $(\forall y)(\text{Bill}, \lambda x(x \text{ hit } y) \rightarrow \text{Sandy}, \lambda w(w \text{ hit } y))$

On either account, we find the appropriate identity of λ-expressions. In (2.1.38)(a), $\lambda z(\ldots)$ and $\lambda w(\ldots)$ are alphabetic variants, as are $\lambda x(\ldots)$ and $\lambda w(\ldots)$ in (2.1.38)(b), $\lambda z(\ldots)$ and $\lambda w(\ldots)$ in (2.1.39)(a), and $\lambda x(\ldots)$ and $\lambda w(\ldots)$ in (2.1.39)(b). Thus the logical theory of VPD, with no further modifications correctly predicts that deletion is possible in both (2.1.36) and (2.1.37) above. Furthermore, the mysterious fact we observed about antecedent-contained instances of VPD like (2.1.40), namely that such sentences lose one of the readings that they had before deletion, is also predicted by the logical theory of VPD.

(2.1.40) Betsy's father wants her to read everything her boss does.

One source for this sentence, is the following.

(2.1.41) Betsy's father wants her to read everything her boss reads.

Let us opt (somewhat arbitrarily) to represent such sentences as conditional statements, rather than using restricted quantification. The logical form of (1.2.41) is then

(2.1.42) Betsy's father, λx(x want [($\forall$y)[Betsy's boss, λw(w read y) $\rightarrow$ Betsy, λz(z read y)]])

λw(...) and λz(...) are alphabetic variants, therefore deletion of $[_{VP}$reads], which is represented by λw(...), is possible in (2.1.41). This accounts for why (2.1.42) is one of the possible readings of the elliptical (2.1.40).

The other source for (2.1.40) is this:

(2.1.43) Betsy's father wants her to read everything her boss wants her to read.

As we know, this sentence is ambiguous (see the discussion in Section 1.3 of Chapter One). We represent the two readings as (2.1.44) and (2.1.45).

(2.1.44) ($\forall$x)[Betsy's boss, λy(y want [Betsy, λz(z read x)]) $\rightarrow$ Betsy's father, λw(w want [Betsy, λr(r read x)])]

(2.1.45) Betsy's father, λy(y want [($\forall$x)[Betsy's boss, λz(z want [Betsy, λw(w read x)]) $\rightarrow$ Betsy, λr(r read x)]])

On the first reading the VP: [wants her to read] is represented by $\lambda y(\ldots)$ in (2.1.44) which, to be sure, is an alphabetic variant of $\lambda w(\ldots)$ in (2.1.44) (the final variables are bound by the same operator). That VP is therefore deletable. This accounts for why the elliptical (2.1.40) can be understood as (2.1.44).

If (2.1.43) is assigned its other reading, however (namely the logical form in (2.1.45)), the VP: [wants her to read] corresponds to $\lambda z(\ldots)$. But there is no other λ-expression in (2.1.45) that is an alphabetic variant of $\lambda z(\ldots)$. Deletion of that VP is therefore blocked on this reading, thus accounting for the seemingly mysterious fact that (2.1.40) lacks one of the readings of its syntactic source (2.1.43).[12]

This concludes our discussion of the problems raised in Section 1.3 of Chapter One. We have seen that a general account of those facts can be given in terms of our hypothesis of recoverability of deletion at the level of logical form. Notice that given this hypothesis, we need no longer mention the trigger VP in the structural description of VPD. That is, we may now write that rule simply as the following.

(2.1.46) Verb Phrase Deletion (optional)

	X -	AUX -	VP -	Y
S.D.:	1	2	3	4
S.C.:	1	2	$\emptyset$	4

Indeed such a formulation is independently necessary if we are to account for the "antecedent-contained" deletions such as those in (2.1.36) and (2.1.37) above, for, as we say in Chapter One, no formulation of VPD that "mentions" the trigger VP will allow those sentences to be generated.

Moreover, such a formulation of the rule, given the recoverability condition as we have formulated it, also accounts for cases of VPD across speakers in discourse. We might hypothesize that all rules that operate in discourse as well as sentence-internally have this property, i.e. the property of not specifying the deletion trigger in their structural description. The bi-directionality of VPD (see Chapter One) is also accounted for nicely by our formulation.

Let us return now to those cases observed in Section 2.0 of this chapter which were problematic for the theory of deep structure recoverability. We noted sentences like these, whose deep structure sources

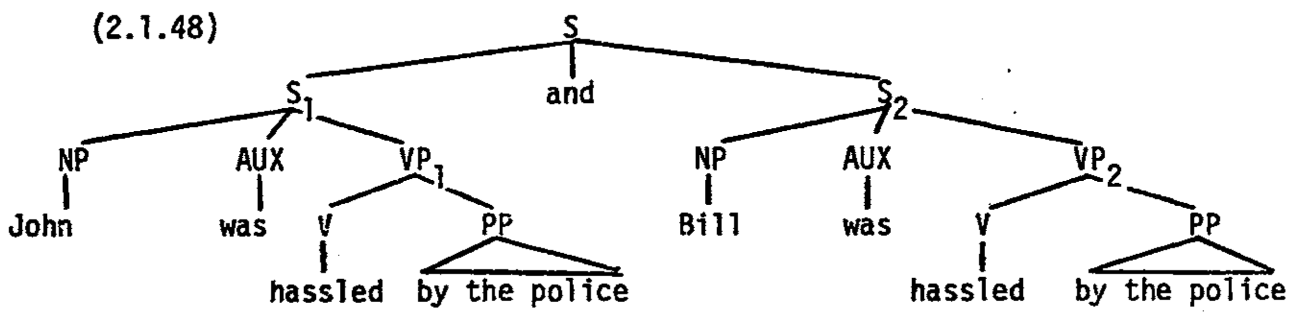
(2.1.47)(a) John was hassled by the police, and Sam was ∅, too.

(b) Betsy seems to me to be unhappy, and Sandy does ∅, too.

(c) Betsy is easy to talk to, and Peter is ∅, too.

would not provide the appropriate identity to predict the possibility of deletion. The problem is actually a form of the "correspondence problem" we noted earlier. To take one example, (2.1.47)(a) would have a standard theory deep structure roughly as in (2.1.48) (the position of AUX is not of concern here).

(2.1.48)



If the notion "deep structure identity" is to be appealed to at all, the problem we encounter is that of establishing a correspondence between the superficial trigger and target VP's and their respective deep structure (non-constituent) ancestors. It is not at all clear how one would go about establishing such a correspondence in an intutive manner.

In a theory of shallow structure interpretation, however, the correspondence between shallow structure VP's and entities of logical form is rather straightforward in these cases. We might follow Bresnan (1976b) and assign (2.1.49) a logical form that looks, in its essentials, like this:

(2.1.50) John, $\lambda x((Ey)[\text{hassle }(y, x) \ \& \ y = \text{the police}]) \ \&$
Bill, $\lambda w((Ez)[\text{hassle }(z, w) \ \& \ z = \text{the police}])$

Here $\lambda x(\ldots)$ and $\lambda w(\ldots)$ are alphabetic variants, and the possibility of deletion is correctly predicted. In a similar manner, the other examples in (2.1.47) do not present correspondence problems for a theory of shallow structure interpretation.

Lakoff's example, which we repeat here in a somewhat more digestable form, is of particular interest.

(2.1.51) The steak is ready to eat, and the chicken is ∅ too.

The standard view of these sentences is essentially that of Lakoff, that the surface VP's here correspond to deep structure VP's which contain object NP's coreferential with the matrix NP's, e.g. $[_{VP}$ready to eat the steak] and $[_{VP}$ready to eat the chicken]. These NP's then undergo an obligatory deletion rule called, variously, Object Deletion (Lasnik and Fiengo (1974))[13] or "Ready Socks" (Hankamer (1971, 1973a)).

There are problems with the standard view. Consider a sentence like (2.1.52), for instance, (2.1.52) The steak which Harry sold to Sue is ready to eat.

the source for which should be (omitting details of the analysis of relative clauses)...

(2.1.52) The steak which Harry sold to Sue is ready to eat the steak which Harry sold to Sue.

What is to prevent optional cyclic rules from applying to the second relative clause, but not to the first one, thus creating non-identity at the level when the deletion rule is to apply? The result will be ungrammatical sequences like the following:[14]

(2.1.53)(a) *The steak which Harry sold to Sue is ready to eat the steak which Harry sold Sue.

(b) *The steak which Harry sold to Sue is ready to eat to eat the steak which was sold to Sue by Harry.

One solution to this dilemma might be to treat the target of such deletion rules as a pronominal element. We might further speculate that such pronominal elements are always to be treated as bound variables, a speculation that possibly receives further support from the existence of sentences like this one:

(2.1.54) Everything is ready to eat.

That is, we would not want to derive this sentence from (in a deep structure theory of interpretation),

(2.1.55) *Everything is ready (for Δ) to eat everything.

for we would not be able to capture the fact that the object of eat is to be treated logically as a bound variable. In other words, we would no more want to treat ready to eat constructions as deletions of a full identical NP than we would standard cases of EQUI.

Given shallow structure interpretation, however, these facts present no particular problem. The fact that the object of eat in such sentences must be interpreted as a bound variable is certainly a property of the ready class of predicates. We might therefore require the λ of the ready predicate to also bind the position of the embedded object pronoun. ready for John to eat would then correspond to the following λ-predicate.

(2.1.54) $\lambda x(\text{ready}(x, \text{eat}(\text{John}, x)))$ or $\lambda x(\text{ready}(x,[\text{John}, \lambda y(y \text{ eat } x)]))$

We then end up with logical forms like the following:

(2.1.55)(a) The steak is ready for you to eat.

(b) The steak, $\lambda x(\text{ready}(x,[\text{you}, \lambda y(y \text{ eat } x)]))$

(2.1.56)(a) Everything is ready for you to eat.

(b) $(\forall x)[x, \lambda y(\text{ready}(y,[\text{you}, \lambda w(w \text{ eat } y)]))]$

Let us now return to the matter of the interaction of ready constructions and VPD. For the purpose of this discussion we will treat the unspecified subject of embedded complements of ready predicates simply as Δ in logical forms. That is, we will write the logical form of ready to eat as $\lambda x(\text{ready}(x,[\Delta, \lambda y(y \text{ eat } x)]))$. Nothing hinges on this decision. Here is the logical form our rules assign to the (pre-deletion) source of (2.1.51).

(2.1.51)' The steak is ready to eat and the chicken is ready to eat, (too).

(2.1.57) The steak, $\lambda x(\text{ready}(x,[\Delta, \lambda y(y \text{ eat } x)]))$ &
the chicken, $\lambda w(\text{ready}(w,[\Delta, \lambda z(z \text{ eat } w)]))$

Clearly, $\lambda x(\ldots)$ and $\lambda w(\ldots)$ are alphabetic variants. We therefore predict that deletion of the VP [ready to eat] in the second conjunct is possible. This accounts for the grammaticality of (2.1.51).

We have not yet indicated how we would represent EQUI sentences like <u>John wants to go</u> or <u>John is ready to go</u>. There are various possibilities, the most straightforward of which is to assume that it is a lexical property of EQUI predicates that their embedded complement subjects are to be treated logically as variables obligatorily bound by the λ of the EQUI predicate, say, as follows.

(2.1.58)(a) John wants to go.
(b) John, $\lambda x(\text{want}(x,[x, \lambda y(\text{go}(y))]))$

(2.1.59)(a) John is ready to go.
(b) John, $\lambda x(\text{ready}(x,[x, \lambda y(\text{go}(y))]))$

Given this, Lakoff's example, where the "crossed" deleted readings are impossible, is easily accounted for, <u>viz</u>.

(2.1.60)(a) The children are ready to eat [i.e. to partake of food] and the chickens are ready to eat [i.e. to be eaten] too.
(b) The children, $\lambda x(\text{ready}(x,[\Delta, \lambda y(y \text{ eat})]))$ &
the chickens, $\lambda x(\text{ready}(x,[\Delta, \lambda y(y \text{ eat } x)]))$

In (2.1.60)(b), the logical form for one of the "crossed" readings, no alphabetic variance obtains.

There are more facts to be predicted here, however, which Lakoff did not observe. Consider the following sentence.

(2.1.61) *The steak is ready to eat, and the chicken is ready to ∅, also.

Notice that there is virtually no way in a deep structure identity theory of recoverability of deletion to explain the ungrammaticality of this sentence, even accepting Lakoff's suggestion that "items that do not appear in the derived structure are completely irrelevant to the question of linguistically significant identity". Why should deletion be possible in (2.1.51) above, where the whole VP: ready to eat has been deleted, but not in (2.1.61), where only the embedded VP: eat has undergone deletion.

Our theory explains this contrast. Consider again the logical form for the (pre-deletion) source of these sentences:

(2.1.57) the steak, λx(ready(x,[Δ, λy(y eat x)])) &
the chicken, λw(ready(w,[Δ, λz(z eat w)]))

As we said earlier, λx(...) and λw(...) are alphabetic variants, thus deletion of $[_{VP}$ready to eat] is predicted. The VP: $[_{VP}$eat] in the second conjunct, however, corresponds to λz(...) in (2.1.57). But there is no other λ-expression in (2.1.57) that is an alphabetic variant of λz(...). This is so because λy(...), the only reasonable candidate for such, contains the variable $\underline{x}$ where λz(...) contains $\underline{w}$, and $\underline{x}$ and $\underline{w}$ are bound diversely. We therefore predict, correctly, that deletion of $[_{VP}$eat] is impossible.

There is more. Consider (2.1.62), where both ready-conjuncts receive an EQUI-interpretation.

(2.1.62) Betsy is ready to give up, and Peter is ready to ∅, also.

Here deletion of $[_{VP}$give up] seems to be possible. Consider our logical form for (the pre-deletion source of) this sentence.

(2.1.63) Betsy, $\lambda x(\text{ready}(x,[x, \lambda y(y \text{ give up})]))$ &
Peter, $\lambda w(\text{ready}(w,[w, \lambda z(z \text{ give up})]))$

Here the embedded VP's correspond to $\lambda y(\ldots)$ and $\lambda z(\ldots)$, which are indeed alphabetic variants. Thus our theory makes precisely the right prediction, namely, that the VP embedded in the complement of ready-type predicates is deletable only when those predicates have an EQUI interpretation.

Predictions of this sort constitute support of the strongest kind for our view of logical form and its relation to the recoverability of deletion. In the next section, we will see that the problem of "sloppy identity" provides further support for our theory.

2.2 The Problem of "Sloppy Identity"

It was Ross (1967a, 1969b) who first observed the ambiguity of sentences like (2.2.1).

(2.2.1) John scratched his arm and Mary did too.

On one reading, this sentence conveys that Mary scratched John's arm. On the other reading, what is conveyed is that Mary scratched her own arm.[15] This latter understanding is what Ross called the "sloppy" reading of (2.2.1).

Ross presumed a purely syntactic theory of VPD, which he attempted to modify, in light of the possibility of "sloppy" interpretations of sentences like (2.2.1), with the following reformulation of the syntactic identity condition.

(2.2.2) [= Ross (1967a, (5.135))]

Constituents are identical if they have the same constituent structure and are identical morpheme-for-morpheme, or if they differ only as to pronouns, where the pronouns in each of the identical constituents are commanded by antecedents in the non-identical portions of the phrase-marker.

Ross himself (Ross, 1969b, nt. 4) points out a serious difficulty with this revised identity condition, namely, that illustrated by this example:

(2.2.3) I told you that you would be famous, and Jack$_i$ told Betty$_j$ (that $\left\{\begin{array}{l}\text{you}\\ \text{she}_j\\ \text{he}_i\end{array}\right\}$ would be famous).

The optionality of the *that* clause in this example, Ross takes to be the result of a rule of S-Deletion. Given the identity condition in (2.2.1), the deleted version of (2.2.3) should have a reading that conveys: *Jack told Betty that Jack would be famous*, which it certainly does not.

In Chapter Four, we will see that the "deleted" version of (2.2.3) does not arise from deletion at all, but rather is an example of Null Complement Anaphora (Sag and Hankamer (1976)), so we need not pursue that matter here. Nevertheless, the inadequacy of Ross's proposal with respect to VPD is readily apparent from examples like the following.

(2.2.4) Norma told Beth's boyfriend to give her a dime, and Judy told Lois's boyfriend to ø.

Given (2.2.2), this sentence should have a reading paraphrasable by (2.2.5).

(2.2.5) Norma told Beth's boyfriend to give Beth a dime, and Judy told Lois's boyfriend to give Judy a dime.

Of course it does not, nor is the class of "sloppy" readings in general as broad as Ross's theory predicts.

Keenan (1971a) draws similar conclusions about the inadequacy of (2.2.2) for Ross's rule of S-Deletion and for cases of so be anaphora.[16] He develops a logical theory of names, descriptions, and quantifiers which allows him to state the conditions on "sloppiness" with considerably greater precision. Keenan's semantic representation for a sentence like the following one is that in (2.2.7).

(2.2.6) John$_i$ was surprised he$_i$ was drunk, and Bill$_j$ was surprised he$_j$ was drunk.

(2.2.7) (John,x)(surprise(x, drunk(x))) & (Bill,y)(surprise(y, drunk(y)))

To account for the paraphrastic alternant of (2.2.6) with so be, namely (2.2.8), a transformation, $T_{\text{so be } 2}$, is posited.

(2.2.8) John was surprised that he was drunk, and so was Bill.

This transformation, Keenan gives as this:

(2.2.9) $T_{\text{so be } 2}$: $\frac{(p,x)S^n_x}{1}$ & $\frac{(p',y)S^n_y}{2}$ $\Rightarrow$ 1 + so be p'

The notation S^n_x denotes the result of replacing each occurrence of n in S by x. The two sentences in (2.2.7) therefore are analyzable as S^n_x and S^n_y respectively with respect to the sentence: surprise (n, drunk(n)). (2.2.7) hence meets the structural description of $T_{\text{so be } 2}$.

Non-"sloppy" deletions in Keenan's theory are handled by a separate transformation, the conventional rule of VPD. VPD interacts with yet another rule, $T_{\text{so be 1}}$. These two rules work together to convert Keenanian intermediate representations like (2.2.10) into transformed structures like (2.2.11).

(2.2.10) (John, x)(Fred, y)(x surprise(x drunk) & y surprise (x drunk))

(2.2.11) (John, x)(Fred, y)(x surprise(x drunk) & so be y)

A further rule of "proper-name-quantifier elimination" produces (2.2.8) from (2.2.11).

The theory of VPD we developed in the last section incorporates the insight of Keenan's approach, namely that "identity of predication" is the relevant parameter that governs VP-deletability.[17] The account of the "sloppy identity" problem offered here, however, will not posit two separate deletion rules as Keenan does. Rather, we will develop the view that a sentence like (2.2.12) has two logical forms, in one of which, the coreferential pronoun is replaced by a bound variable.

(2.2.12) Betsy_i loves her_i dog.

One representation for (2.2.12), we will write as follows:

(2.2.13) Betsy_i, λx(x loves her_i dog)

The other representation will differ only in that the coreferential pronoun her within the λ-expression is replaced by an occurrence of the variable bound by the λ, i.e. as in (2.2.14)[18]

(2.2.14) Betsy_i, λx(x loves x's dog)

The reader will note that we are assuming that referring expressions such as proper names and pronouns appear in logical forms with indices. Some device like indexing is necessary in any theory of VPD, as McCawley (1976) has shown, because in sentences like this next one, no matter what the reference of him may be, a semantic theory must guarantee that both Betsy and Sandy saw the same guy.

(2.2.15) Betsy saw him and Sandy did ∅ too.

Certain previous approaches, e.g. that of Jackendoff (1972), failed to account for this fact. Similarly, any approach to anaphora that eliminates the notion of stipulated coreference entirely, e.g. that of Lasnik (1976), also fails to account for this fact.[19]

We will assign indices to all NP's, including non-referring NP's like some man, everyone, etc. Therefore only in a sub-class of cases will our indices be "referential" indices. I don't see that this presents any particular problem, since semantic rules will obligatorily apply to NP's containing quantifier words, creating quantified formulae with the appropriate bound variables. The only remaining entities with indices in logical forms, therefore, will be those whose indices can be naturally interpreted as referential.[20]

As for the idea that sentences like "Betsy$_i$ loves her$_i$ dog" have more than one (logically equivalent) logical forms, this is not as novel as it might appear. McCawley (1967) and Keenan (1971a) both clearly had this in mind; the former, tacitly, the latter, quite explicitly. Moreover, other recent attempts to be precise about logical representations (Montague (1975), Thomason (1974), Partee (1975)) by necessity posit a plethora of such multiple representations.

We will presume that formulae like (2.2.14) (Betsy$_i$, λx(x loves x's dog)) arise from the application of an optional rule of semantic interpretation (of SI-1, that is) that converts pronouns into bound variables. We may state this rule (call it: PRO$\rightarrow$BV) informally as follows:

(2.2.16) PRO $\rightarrow$ BV

$$NP_i, \lambda x(\ldots \underset{\substack{\Downarrow \\ x}}{PRO_i} \ldots)$$

The interpretive rules thus first produce the formula in (2.2.13) (Betsy$_i$, λx(x loves her$_i$ dog)), which optionally undergoes PRO$\rightarrow$BV to yield (2.2.14).

The rule of PRO$\rightarrow$BV, we will see, is the only additional device we will need for our logical theory to predict all the "sloppy" instances of VPD that have been noted in the literature, and, in addition, several new facts that have heretofore escaped notice. To see this, consider first the following sentence:

(2.2.17) Betsy$_i$ loves her$_i$ dog, and Sandy$_j$ does $\emptyset$, too.

The non-"sloppy" reading of this sentence (where what is being said is that Sandy loves Betsy's dog) has the logical representation shown in (2.2.18).

(2.2.18) Betsy$_i$, λx(x loves her$_i$ dog) & Sandy$_j$, λy(y loves her$_i$ dog)

Clearly, λx(...) and λy(...) are alphabetic variants: deletion is possible.

On the "sloppy" reading (where what is being said is that Sandy loves her own dog), the semantic rules first produce the formula in (2.2.19).

(2.2.19) Betsy_i, λx(x loves her_i dog) & Sandy_j, λy(y loves her_j dog)

Double application of PRO→BV then yields (2.2.20).

(2.2.20) Betsy_i, λx(x loves x's dog) & Sandy_j, λy(y loves y's dog)

$\lambda x(\ldots)$ and $\lambda y(\ldots)$ are again alphabetic variants, allowing deletion. The theory we have presented thus accounts for the simple cases of "sloppy" and non-"sloppy" VPD straightforwardly.

But the formulation of PRO→BV we have given actually has many empirical consequences that are of interest. First, it predicts that, "sloppy" deletions are, in principle at least, unbounded. That is, by allowing any PRO_i within a λ-expression to be replaced by a bound variable, rule (2.2.16) optionally creates "sloppy" predicates no matter how deeply embedded that PRO_i happens to be. We will later return to one qualification that must be made in this regard, but this prediction is certainly in general correct, as can be seen from the following VP-deleted sentences, all of which are ambiguous in the appropriate way.

(2.2.21)(a) John_i said Mary knew he_i was unpopular, and Bill_j did ∅, too.
[∅ = said Mary knew $\text{he}_{i,j}$ was unpopular]

(b) John_i said there was likely to be a good explanation for why Mary hated him_i, and Bill_j did ∅, too.
[∅ = said there likely to be a good explanation for why Mary hated $\text{him}_{i,j}$]

(c) Ford_i knew that it was certain that he_i had won the primary before Jackson_j did ∅.
[∅ = knew it was certain that $\text{he}_{i,j}$ had won the primary]

Secondly, our formulation automatically accounts for an observation made by Witten (1972). Witten observed that if the antecedent VP contained a non-pronominal NP, no "sloppy" reading is possible. Witten's example is the following:

(2.2.22) Sam$_i$ liked Janet's picture of Sam$_i$, and Bill$_j$ did $\emptyset$, too.

Here, $\emptyset$ can be interpreted as <u>liked Janet's picture of Sam$_i$</u>, but not as <u>liked Janet's picture of Bill$_j$</u>. Our theory predicts this because our rule PRO→BV can make "sloppy" predicates only by converting pronouns into bound variables, not full NP's. Not all speakers accept the left conjunct in Witten's example, but the same point can be made by the following, more acceptable sentence:

(2.2.23) His$_i$ mother loves John$_i$, and Bill$_j$'s mother does $\emptyset$, too.
[$\emptyset$ = loves John$_i$, *him$_j$]

A third consequence of our rule of PRO→BV is a proper account of Ross's (1967a, 1969b) observation that "sloppiness" is possible only when the appropriate antecedents command their pronominal anaphors. The relevant examples are the following:

(2.2.24) John$_i$ scratched his$_i$ arm, and [the boy who knew Tom$_k$]$_j$ did $\emptyset$, too. [$\emptyset$ = scratched his$_{i,j,*k}$ arm]

(2.2.25) John$_i$'s sister scratched his$_i$ arm, and Bill$_j$'s sister did $\emptyset$, too
[$\emptyset$ = scratched his$_{i,*j}$ arm]

(2.2.26) John$_i$ scratched his$_i$ arm, and [Bill$_k$'s brother]$_j$ did $\emptyset$, too
[$\emptyset$ = scratched his$_{i,j,*k}$ arm]

Our analysis prevents all the non-occurring readings because our rule of PRO→BV can never create "sloppy" λ-predicates that would produce those readings. The only pronouns that can become bound variables are those that bear the same index as the argument of the λ-predicate, which, in all the cases we have considered so far, is the surface subject. (2.2.27) sketches how the relevant logical forms for the previous examples fail to meet the S.D. of PRO→BV.

(2.2.27)(a) [the boy who know $\text{Tom}_k]_j$, λx(x scratched his_k arm).

(b) $[\text{Bill}_j$'s $\text{sister}]_k$, λx(x scratched his_j arm)

(c) $[\text{Bill}_k$'s $\text{brother}]_j$, λx(x scratched his_k arm)

In similar fashion, we can explain another observation of Witten's (1972, p. 42). He notes that the following sentence is unambiguously non-sloppy.

(2.2.28) Bill_i and Mary are afraid that he_i may be sent to India, and Tom_j and Lucy are ∅, too.
[∅ = are afraid $\text{he}_{i,*j}$ may be sent to India]

PRO→BV can not convert a pronoun bearing the same index as an NP <u>inside</u> the subject into a bound variable, but only one bearing the same index as the subject NP itself. Compare the following minimally different sentence, where a sloppy reading is possible:

(2.2.29) $[\text{Bill and Mary}]_i$ are afraid that they_i may be sent to India, and $[\text{Tom and Lucy}]_j$ are ∅, too. [∅ = are afraid $\text{they}_{i,j}$ may be sent to India]

All these facts fall out of our theory.

A fourth interesting prediction made by our theory concerns the following contrast:

(2.2.30) John_i said Mary_j hit him_i, and Bill_k did ∅, too.
[∅ = said Mary_j hit $\text{him}_{i,k}$]

(2.2.31) John_i said Mary_j hit him_i, and Bill_k said she_j did ∅, too.
[∅ = hit $\text{him}_{i,*k}$]

(2.2.30), where the entire VP in the second conjunct has been deleted, is ambiguous ("sloppy" or non-"sloppy"). (2.2.31), my informants tell me, can have only the non-"sloppy" reading, i.e. the reading that conveys "Bill said she hit John".

This is a surprising contrast. Why should (2.2.31), where only the embedded VP is deleted, not have a "sloppy" reading? All previous treatments of "sloppy" identity that I am aware of (McCawley (1967), Keenan (1971a), Witten (1972), Dahl (1972, 1974), Schiebe (1973)) predict, as far as I can see, that (2.2.31) should also be ambiguous. However, our theory correctly predicts that only (2.2.30) has a "sloppy" reading. To see this, consider (2.2.32), which is the logical representation our rules assign to (the source of) both these sentences on the "sloppy" reading.

(2.2.32) John_i, λx(x said [Mary_j, λy(y hit x)]) &
Bill_k, λz(z said [Mary_j, λw(w hit z)])

The entire VP in the second conjunct (<u>said Mary_j hit him_k</u>) corresponds to λz(...) here. λz(...) is furthermore an alphabetic variant of λx(...). Therefore deletion of the entire VP in the second conjunct (i.e. the deletion in (2.2.30)) is possible. The embedded VP in the second conjunct, on the other hand, corresponds to $\lambda w(\ldots)$ in (2.2.32), which, in order for deletion to be possible, would have to be an alphabetic variant of $\lambda y(\ldots)$. But $\lambda w(\ldots)$ and $\lambda y(\ldots)$ are not alphabetic variants, for the variables x and z are not identical, i.e. they are bound diversely. The deletion in (2.2.31) is therefore predicted on our theory to be possible only on the non-"sloppy" reading.

In the previous discussion, the only cases we have examined are those where our rule of PRO → BV has made bound variables out of pronouns bearing the same index as the surface subject. The "sloppy identity" phenomenon is actually somewhat more pervasive than this. The generalization seems to be that the only pronouns which can undergo PRO → BV are those that bear the same index as a cycle-final subject NP, or an NP that "controls" a subject position. The following contrast illustrates this point.

(2.2.33) Mary$_i$ persuaded John$_j$ to say she$_i$ hit him$_j$, and she$_i$ persuaded Bill$_k$ to ∅, also. [∅ = say she$_i$ hit him$_{j,k}$]

(2.2.34) Mary$_i$ persuaded John$_j$ to say she$_i$ hit him$_j$, and she$_i$ persuaded Bill$_k$'s sister to ∅, also. [∅ = say she$_i$ hit him$_{j,*k}$]

"Sloppy" deletion is possible in the first example only.

Here I think it's reasonable to expect our theory of logical form to help us out. Specifically, it seems perfectly reasonable to assume that the logical forms for sentences with <u>persuade</u>, and the like, will contain embedded propositions which will have the shape, in our theory, NP, $\lambda x(\ldots)$. A sentence like (2.2.33), for example, might have a representation like the following for its left conjunct.

(2.2.35) Mary_i, λx(x persuaded John_j[John_j, λy(y say [she_i, λz(z hit him_j)])]])

Given representations of this sort, PRO $\rightarrow$ BV can apply to create the appropriate "sloppy" predicates in (2.2.33). The discussion of the last few pages furthermore makes clear why (2.2.34) cannot be assigned the appropriate "sloppy" logical form.

Another interesting case is EQUI - constructions like (2.2.36).

(2.2.36) John wants to go, and Bill does too.

As Fodor (1975) points out, deletions like these are necessarily "sloppy". For instance, (2.2.36) must convey that the person Bill wants to go is Bill, not John. (This, by the way, poses another difficulty for the theory of deep structure identity as a recoverability condition, as, say, Lakoff has developed it.)

Again, it seems reasonable to look to our theory of logical forms for a solution to this problem.

We might follow Fodor (and Chomsky) and posit self as the underlying subject of equi-embedded clauses (which he justifies, by the way, on independent grounds).[21] Alternatively, we might (as suggested in the preceding section) simply require a bound variable in the appropriate position in the embedded proposition of the logical forms of such sentences. The logical form of (2.2.36) then might appear as one of the following formulae.

(2.2.37)(a) John, λx(x want[self, λy(y go)]) & Bill, λw(w want[self, λz(z go)])

(b) John, λx(x want[x, λy(y go)]) & Bill, λw(w want[w, λz(z go)])

In either case, our theory of VPD will assign the correct interpretations to sentences like (2.2.36).

This is probably as good a place as any to bring up a seemingly problematic example observed by Schiebe (1973). Schiebe notes that in German, the following sentence can be interpreted as indicated.

(2.2.38) Karl sieht ein, dass er dumm ist, aber Peter nicht, obwohl sogar seine Frau das tut.
'Karl realizes that he (Karl) is stupid, but not Peter [i.e. Peter doesn't realize he (Peter) is stupid] even though his wife does [realize Peter is stupid].'

This is an interesting possible reading for this sentence, though not of immediate relevance, bacause the anaphoric device in the third clause (das tut) involves an overt proform, i.e. it is not VPD. das tut may well be an instance of the German analogue of the English do it. It is known that the it of such anaphoric entities as do it makes free reference to non-linguistic entities (Sag and Hankamer (1976)) or to elements that do not form single entities in syntactic or logical structure, as in the following examples, adapted from Akmajian (1970, 1973) and Chomsky (1972a).

(2.2.39) John turned the hot dog down flat, but it wouldn't have happened with the filet mignon.
[it = John turned down...flat]

(2.2.40) John pounded a hole in each fender of the Pontiac, and Betsy did it to the Mercedes.
[it = pound a hole in each fender of...]

One might propose a free abstraction mechanism, as Akmajian does, to account for such interpretations, but it is quite clear that whatever device is employed, it must assign to sentences containing do it many possible interpretations that an analogous sentence involving VPD would lack.

The point of this digression is the following. Dahl (1972) translates Schiebe's sentence into English, using VPD, and offers it as a counterexample to any analysis that involves VPD and a logical identity condition (such as McCawley's, Keenan's, or ours). His example is the following.

(2.2.41) John realizes that he is a fool, but Bill does not, even though his wife does.

Crucially, Dahl claims this sentence can have the following reading:

(2.2.42) John realizes that he - John - is a fool, but Bill doesn't realize that he - Bill - is a fool, even though his wife realizes that he - Bill - is a fool.

Almost all speakers I have checked with, however, insist that (2.2.41) cannot have this reading. Dahl's factual claim seems to be incorrect.[22]

Why would the possibility of this reading be a problem for our theory? For this reason. In order for the first, "sloppy" deletion to be permissible, PRO $\rightarrow$ BV must apply in the logical derivations of the first two clauses, i.e. the partial logical form would be as in (2.2.43).

(2.2.43) John, λx(x realize [x is a fool]) but $\neg$ Bill, λy(y realize [y is a fool])

In order for the second, non-"sloppy" deletion to occur, however, the last two clauses would have to have the representations in (2.2.44), where

PRO $\rightarrow$ BV has not applied.

(2.2.44) Bill$_i$, λy(y realize [he$_i$ is a fool]) even though his wife, λz(z realize [he$_i$ is a fool])

For both deletions to be legitimate in our theory, the middle clause would have had to have both logical forms at once.

It is therefore an empirical prediction of our theory that the interpretation in question is not possible for (2.2.41), and I interpret my informant's responses as further evidence for its general correctness.

Dahl offers another argument against any analysis that involves a deletion rule that makes reference to logical structures. This concerns sentences like the following:

(2.2.45)(a) John thinks he is smart, and the same is true of Bill.

(b) John thinks he is smart, and Bill suffers from the same delusion.

Dahl concludes (Dahl, 1972, p. 9) that "the existence of cases like [(2.2.45)(a), (b)] shows us that the problem [of "sloppy" identity - I.A.S.] goes beyond what can be explained by deletion and substitution rules". I fail to appreciate Dahl's point. The fact that there exist expressions like <u>the same</u> that make reference to logical entities, like our "sloppy" λ-predicates, but that do not arise by deletion, in no way bears on the question of whether or not deletion rules make use of such entities. As in the case of the <u>it</u> in <u>do it</u>, <u>the same</u> can on occasion refer to abstracted entities that are not entities of logical representations, as in the following example.

(2.2.46) John ate the banana, and the same is true of the hot dog.

The fact that the same sometimes picks out an entity of logical representations, as it also frequently does, is simply irrelevant with respect to evaluating any theory of deletion.

Another of Dahl's (1974) examples is more instructive. Examples like the following, he points out, do not have all the readings one might expect them to.

(2.2.47) Bill believed that he loved his wife, and Harry did, too.

The second conjunct in this example has only the first three of the following four possible interpretations:

(2.2.48) Harry believed that

(a) Harry loved Harry's wife.

(b) Bill loved Bill's wife.

(c) Harry loved Bill's wife.

(d) *Bill loved Harry's wife.

The third possible interpretation, as Dahl also notes, is perhaps marginal. I find it rather plausible, actually, in contexts like the following.

(2.2.49) Who believed that they loved Bill's wife?

- Bill believed that he loved his wife, and Harry did, too.

Nonetheless, the fourth reading in (2.2.48) is quite impossible in any context.

In our theory, this problem reduces to explaining why only the first three of the following four logical formulae are possible.

(2.2.50)(a) Bill_j, λw(w believed [w, λz(z loved w's wife)]) & Harry_i, λx(x believed [x, λy(y loved x's wife)]) Harry loved Harry's wife

(b) Bill_j, λw(w believed [he_j, λz(z loved his_j wife)]) & Harry_i, λx(x believed [he_j,λy(y loved his_j wife)]) Bill loved Bill's wife

(c) Bill_j, λw(w believed [w, λz(z loved his_j wife)]) & Harry_i, λx(x believed [x, λy(y loved his_j wife)]) Harry loved Bil wife

(d) *Bill_j, λw(w believed [he_j, λz(z loved w's wife)]) & Harry_i, λx(x believed [he_j, λy(y loved x's wife)]) Bill loved Harry's wife.

As we have developed it thus far, our theory generates all these formulae: (a) results from "across the board" application of PRO → BV in both conjuncts; (b) results from no applications of PRO → BV; (c) is the result of one application of PRO → BV in each conjunct, as is (d). The difference between (c) and (d) is simply which pronoun becomes a bound variable. The question to be raised then, is whether there is some independent explanation for why (d) is ill-formed. I think there is.

As I mentioned briefly earlier, Chomsky (1975b) has argued that it is possible to give a unified account of certain of Postal's (1971) "crossover" violations, and the impossibility of "backwards" anaphora with quantifier expressions and contrastively stressed NP's. Chomsky's account is in terms of a general constraint barring (logical) structures where a bound variable is preceded by a proform that is related to it anaphorically. The sentences in question are the following.

(2.2.51)(a) *Who_i did his_i mother hate?

(b) *his_i mother hated everyone_i.

(c) *his_i mother hated John_i.

The structures Chomsky has in mind for these sentences are roughly the following (the ι operator is used for the logical form of sentences involving contrastive stress, a matter which we return to in Section 3.5 of Chapter Three).

(2.2.52)(a) $(\text{Who } x_i)(\text{his}_i \text{ mother hated } x_i)$

(b) $(\forall x_i)(\text{his}_i \text{ mother hated } x_i)$

(c) $\iota x_i(\text{his}_i \text{ mother hated } x_i) = \text{John}_i$

The general constraint, Chomsky suggests, bars structures of the form: $\ldots\text{PRO}_i\ldots x_i\ldots$

There are many ways one might formulate this constraint and still retain the essence of Chomsky's idea. We will not explore these possibilities here, but will merely point out that the formula we have given in (2.2.50)(d) is of precisely the same form as Chomsky's examples. That is, the left conjunct in (2.2.50)(d) violates the proposed constraint. Thus whatever version of that constraint turns out to be correct, it is reasonable to presume it will provide an explanation for why (2.2.50)(d) is not a well-formed logical representation, and hence for why the deleted sentence in (2.2.47) cannot be so interpreted. Our proposed theory thus interacts with independently motivated constraints on logical forms in a principled, if not elegant way.[23]

The last matter we will take up in this section is the problem of reflexives. The first relevant cases are these.

(2.2.53)(a) John liked himself, and Bill did, too.

(b) John liked himself before Bill did.

Some informants find such cases unambiguously "sloppy", i.e. they interpret the second conjunct of (2.2.53)(a) only as Bill liked himself. For other speakers, sentences like these are always ambiguous, except, of course, for sentences like the following, which, for lexico-semantic, or perhaps for purely pragmatic reasons, are unambiguously "sloppy".

(2.2.54) John perjured himself, and Bill did, too.

Other sentences involving reflexives show the possible non-"sloppy" readings more clearly.

(2.2.55) Betsy couldn't imagine herself dating Bernie, but Sandy could.

Here it is clear that either a "sloppy" or a non-"sloppy" interpretation is possible.

The judgements are not entirely clear in the above cases. However, in cases like this, where some people claim a sentence lacks a reading that others find perfectly natural for it, more often than not, it seems to me, there is no "dialect variation" involved (see Hindle and Sag(1975), Labov (1972) for more discussion of this point). Rather, the prudent conclusion in many such cases is that extraneous factors affect people's introspective judgements in a way that compels them to reject interpretations that are actually possible. This being the case, we should be reluctant to conclude that sentences like (2.2.53)(a) and (b) are unambiguous for any speakers, and we will presume that our rules should assign them "sloppy" and non-"sloppy" readings.

This could be done in the following manner. Let us assume reflexive pronouns are treated just as non-reflexive pronouns are (allowing the syntactic distribution of the morpheme self to be handled by a reflexive transformation (Lees and Klima (1963)) or by an interpretive mechanism (Jackendoff (1968, 1972)), or whatever turns out to be correct). The logical form for a simple reflexive sentence is then as in (2.2.56).

(2.2.56) John_i shot himself.

John_i, $\lambda x(x \text{ shot } \text{him}_i)$

This will also allow us to obtain certain inferences straightforwardly ((2.2.56), for example, entails: <u>Someone hit him_i</u>), and does not seem to create any difficulties elsewhere in the grammar.

PRO$\rightarrow$BV can then apply to logical representations like that in (2.2.56). Therefore the ambiguity of (2.2.53)(a) and (b) is automatically accounted for. Note further that the following sentence is also assigned only the correct possible readings.

(2.2.57) Betsy talked to Sandy about herself, and Myra did, too.

If the first conjunct of this sentence is interpreted as <u>Betsy talked to Sandy about Betsy</u>, then it has the logical representation shown in (2.2.58).

(2.2.58) Betsy_i, $\lambda x(x \text{ talked to } \text{Sandy}_j \text{ about } \text{her}_i)$

If the right conjunct is interpreted analogously, its logical form contains the same λ-predicate and a non-"sloppy" deletion is possible. Alternatively, (2.2.58) can undergo PRO$\rightarrow$BV. If the appropriate representation for the right conjunct also undergoes PRO$\rightarrow$BV, a "sloppy" deletion is also possible.

Suppose, however, the first conjunct of (2.2.57) is interpreted as <u>Betsy talked to Sandy about Sandy</u>. This interpretation is represented as the following.

(2.2.59) Betsy$_i$, λx(x talked to Sandy$_j$ about her$_j$)

PRO$\rightarrow$BV cannot apply, since $\lambda x(\ldots)$ contains no occurrence of PRO$_i$. Therefore the only λ-expression occurring in the right conjunct of the logical form of (2.2.57) (on this reading) which would allow deletion is this one:

(2.2.60) Myra$_k$, λx(x talked to Sandy$_j$ about her$_j$)

Our theory therefore predicts that when the left conjunct of (2.2.57) is interpreted as in (2.2.59), i.e. as ...to Sandy about Sandy, then the only possible deletion is a non-"sloppy" one. This is a correct prediction.

2.3 Residual Matters

In Section 2.0 we noted that VPD was not sensitive to such distinctions as that of generic vs. specific interpretations of the indefinite article. In terms of our hypothesis about the nature of recoverability of deletion, this observation leads to the conclusion that this distinction should not be represented at the level of logical form, i.e. that the indefinite article should be treated univocally there. But, as we have seen in the previous sections, we have been able to give a coherent account of the interaction of quantifier scope assignment and VPD by assuming that certain quantifier-bound variable relations are specified in representations at the level of logical form. It is then of some interest to ask what sorts of "differences in understandings" should be represented at the level of logical form, and which should not be. If our hypothesis about the nature of recoverability of deletion is correct, then we have a litmus (to use a term of Harris's) to help us answer this question, namely, whether or not deletion rules are respectful of the difference (in understanding) in question.

2.3.1 Plurals

Let us take up first the matter of sentences involving plural NP's. It seems that, as regards deletion, VP's that contain plural NP's are to be considered non-distinct from analogous VP's containing singular NP's. Consider the following example.

(2.3.1) John's uncles are bachelors, but Betsy claims her uncle isn't $\emptyset$. [$\emptyset$ = a bachelor]

This example is similar to one noted by Chomsky (1965, p. 180) for comparative sentences. Chomsky's solution to the problem posed by such sentences (see the discussion in Section 2.0 above) was to weaken the notion of identity between deletion target and deletion trigger to a notion of non-distinctness, where two such entities that differ only with regard to transformationally-inserted features (such as plurality) are to be considered sufficiently non-distinct to allow deletion to take place. Given our theory, however, we cannot make use of Chomsky's suggestion to account for the grammaticality of (2.3.1), because our rule of VPD does not make use of the notion of syntactic identity at all. We require of a target VP only that it correspond to a λ-predicate in logical form which is an alphabetic variant of some other λ-predicate in the logical form of that sentence or of some preceding sentence. Therefore we would draw a somewhat different conclusion from such sentences.

We would take the grammaticality of (2.3.1) as providing support for the view that plural is to be treated as a sentence operator at the level of logical form, as suggested in Chomsky (1975c). That is, a sentence like the left conjunct of (2.3.1) (John's uncles are bachelors) should have a logical form like the following, where [PL] designates a plural sentential operator.

(2.3.2) [PL](John's uncle, λx(x is a bachelor)) or
[PL](John's uncle, λx(bachelor(x)))

In this way, the logical form of (2.3.1) would be the following, where $\lambda x(\ldots)$ and $\lambda y(\ldots)$ are alphabetic variants, thus predicting that deletion is possible.

(2.3.3) [PL](John's uncle, λx(bachelor(x))) but (Betsy, λz(z claim $\neg$ [her uncle, λy(bachelor(y))]))

There are some further observations to be made about the interaction of VPD and plurality. Consider, for instance, the following sentence containing three plural NP's.

(2.3.4) The women gave lectures at museums.

There are several understandings such a sentence might take on. One such understanding would represent a situation where, say, each woman in question gave one lecture at one museum. We represent this situation as follows, where the letters are to be understood in the obvious mnemonic way.

(2.3.5)

$$\begin{tikzcd}
W_1 \arrow[r] & L_1 \arrow[r] & M_1 \\
W_2 \arrow[r] & L_2 \arrow[r] & M_2 \\
W_3 \arrow[r] & L_3 \arrow[r] & M_3 \\
\vdots & \vdots & \vdots
\end{tikzcd}$$

Alternatively, a hearer of (2.3.4) might understand a different message, say, where each woman gave several lectures at given museums, this picture:

(2.3.6)

$$\begin{tikzcd}
 & L_1 \arrow[dr] & \\
W_1 \arrow[ur] \arrow[r] \arrow[dr] & L_2 \arrow[r] & M_1 \\
 & L_3 \arrow[ur] & \\
 & L_4 \arrow[dr] & \\
W_2 \arrow[ur] \arrow[r] \arrow[dr] & L_5 \arrow[r] & M_2 \\
 & L_6 \arrow[ur] & \\
\vdots & \vdots & \vdots
\end{tikzcd}$$

A third possibility is an understanding where each woman in the domain of discourse gave one lecture at several museums:

(2.3.7)

$$\begin{tikzcd}
 & & M_1 \\
W_1 \arrow[r] & L_1 \arrow[ur] \arrow[r] \arrow[dr] & M_2 \\
 & & M_3 \\
 & & M_4 \\
W_2 \arrow[r] & L_2 \arrow[ur] \arrow[r] \arrow[dr] & M_5 \\
 & & M_6 \\
\vdots & \vdots & \vdots
\end{tikzcd}$$

One might suppose that these three understandings correspond to three different logical forms for (2.3.4), i.e. that sentence might be taken to be three-ways ambiguous. The following sentence is therefore of some interest.

(2.3.8) The women gave lectures at museums, and Sam voluteered to $\emptyset$, also.

The judgements involved in interpreting this sentence are admittedly difficult, but several grueling hours with a few informants has convinced me that (2.3.8) is utterable (truthfully) even in a situation where Sam only volunteered to give one lecture at one museum. What's more, it seems that at the same time that the right conjunct of (2.3.8) is construed in this way, the left conjunct may be assigned any of the interpretations just noted.

Now the women gave lectures at museums may indeed be ambiguous. Notice that if we represent one of these readings as the following

(2.3.9) [PL](the woman, λx(x give a lecture at a museum)) ,

then this reading is minimally satisfied by the state of affairs in the first picture above (i.e. (2.3.5)). That is (2.3.9) requires only that for each woman in question there is a member (call it *l*) of some set of lectures L and there is a member (call it *m*) of some set of museums M, such that that woman gave *l* at *m*. It further requires that the cardinality of each set (the set of women in question, L, and M) be greater than 1. Therefore all three pictures sketched above are consistent with the truth of (2.3.9).

(2.3.8), then (our deleted sentence) can be derived from the following sentence, whose logical form is as indicated.

(2.3.10) The women gave lectures at museums, and Sam volunteered to give a lecture at a museum, (also).

(2.3.11) [PL](the woman, λx(x five a lecture at a museum)) & Sam, λy(y volunteered [y, λz(z give a lecture at a museum)])

$\lambda y(\ldots)$ and $\lambda x(\ldots)$ are alphabetic variants, correctly predicting the possibility of deletion. Since any of the situations sketched in the above pictures are consistent with the truth of the left conjunct of this last formula [= (2.3.9) above], we have therefore accounted for the fact that our deleted sentence [= (2.3.8)] could be true if Sam volunteered to give only one lecture at a museum, while each woman might have been giving any number of lectures at any number of museums.

Plural sentences should not always be represented in this way, however. These "sentential plural" readings, in fact, seem to arise only when there is a plural subject and a plural entity elsewhere (later) in the sentence. We might propose that the interpretive rule which "extracts" plurals "across-the-board" to produce these logical forms operates only when there is a plural subject. Let us further assume that this rule is optional. Given this, the sentence <u>The women gave lectures at museums</u> would have a second logical form, where, intuitively, the property of giving lecture<u>s</u> at museum<u>s</u> is being predicated of the women in question. This we might write as the following.

(2.3.12) The women, λx([PL](x give a lecture at a museum))

This we would take to correspond to another reading of the sentence in question, perhaps even the most natural one.

Accepting this, we can account for another reading of (2.3.8)

(The women gave lectures at museums and Sam volunteered to also), which is the reading where Sam did indeed volunteer to give several lectures at several museums. The source on this reading would be (2.3.13) whose logical form is as given.

(2.3.13) The women gave lectures at museums and Sam volunteered to give lectures at museums (also).

(2.3.14) the women, $\lambda x([PL](x \text{ give a lecture at a museum}))$ & Sam, $\lambda y(y \text{ volunteered } [y, \lambda z([PL](z \text{ give a lecture at a museum}))])$

Here $\lambda z(\ldots)$ and $\lambda x(\ldots)$ are alphabetic variants, and the possibility of deletion is predicted, thus allowing for the reading of (2.3.8) where what is understood is that Sam volunteered to give more than one lecture.

Notice, however, that our decision to regard this sentence as ambiguous is by no means forced upon us by the facts just considered, for, as we noted above, the multiple lecture-museum understandings are perfectly consistent with the "sentential" plural understandings. That is, if we derived (2.3.8) only from (2.3.10) where the deleted VP is (Sam volunteered to) $[_{VP}$give a lecture at a museum], the logical form we would assign to (2.3.8) (i.e. (2.3.11)) would always be consistent with the state of affairs where Sam actually volunteered to give more than one lecture at more than one museum. In other words, since Sam volunteered to give lectures at museums entails Sam volunteered to give a lecture at a museum, it is not clear whether our treatment of (2.3.8) as ambiguous is well-motivated, or whether we should simply consider it "vague" as to how many lectures Sam actually volunteered to give.

There is further evidence however that the decision we have made is the right one. That evidence comes from sentences where both trigger and target VP contain plural elements, but where the subject of each sentence is singular. Since we have hypothesized that "sentential plurality" arises only when the subject NP is also plural, then a sentence like John has living parents can have only the logical form in (2.3.15), not one like (2.3.16) where [PL] has been extracted as a sentential operator.

(2.3.15)(a) John has living parents.

(b) John, λx([PL](x has a living parent))

(2.3.16) *[PL](John, λx(x has a living parent))

This predicts that a sentence like John has living parents and Bill does too cannot be derived from (2.3.17) whose logical form is as in (2.3.18).

(2.3.17) John has living parents and Bill has a living parent too.

(2.3.18) John, λx([PL](x has a living parent)) &

Bill, λy(y has a living parent)

λx(...),which contains [PL], is not an alphabetic variant of λy(...). In other words, we predict that John has living parents and Bill does too would be false if Bill has only one living parent, because we can derive it only from the following sentence, whose logical form is as given.

(2.3.19) John has living parents, and Bill has living parents too.

(2.3.20) John, λx([PL](x has a living parent)) &

Bill, λy([PL](y has a living parent))

λx(...) and λy(...) are alphabetic variants. Our theory therefore predicts that John has living parents and Bill does too is false if Bill has only one living parent just as Bill has living parents would be false in those circumstances. This seems to be a correct prediction.

2.3.2 Comparative Ellipsis

We might bring our theory of VPD to bear on another syntactic problem that has been discussed in the literature, namely, the matter of reduction in comparative clauses. A common view is that all of the following sentences are derived from the same syntactic source.

(2.3.21)(a) John ate more potatoes than Harry ate.

(b) John ate more potatoes than Harry did.

(c) John ate more potatoes than Harry.

The first of these is generally taken (Hankamer (1971, 1973a), Bresnan (1972,1973, 1975)) to result from the application of a rule of Comparative Deletion (CD) to an underlying structure roughly like

(2.3.22) John ate more potatoes than Harry ate x-many potatoes.

The only precise formulation of CD that I am aware of is that of Bresnan (1975). An alternative view, due to Chomsky (1973) is that (2.3.21)(a) results from relativization and obligatory deletion of the *wh*-word. Chomsky cites the following sorts of sentence as justification for that analysis.

(2.3.23) John ate more potatoes than what Harry ate.

There are some dialects where such sentences are grammatical (Hankamer's, for instance - see Hankamer (1971)). The dialect difference given the relativization analysis would be neatly characterizable as a difference between an optional and an obligatory rule deleting the fronted relative pronoun.[24] Thus two views of the derivation of (2.3.21)(a) have been formulated quite explicitly.

(2.3.21)(b) and (c) have been taken to be the result of two further rules which perform ellipsis within comparative clauses. However, no one has ever given an explicit formulation of these rules. Indeed they seem rather difficult to write. Notice also that the hypothesized rules of Comparative Ellipsis (1 and 2) are rather like the rule of Auxiliary Ellipsis proposed by Akmajian and Wasow which we showed in Chapter One to be both undesirable and unnecessary.

There is a further problem, for as Bach, Bresnan and Wasow (1974) note, CD does not seem to allow "sloppy identity", but Comparative Ellipsis (CE) does. They cite the following examples.

(2.3.24)(a) John lost more of his hair than George lost. (CD)
(b) John lost more of his hair than George (did). (CE)

(2.3.24)(a) does not seem to admit the reading where George lost "x much of his own hair", which is the only plausible understanding given the peculiar context one would have to imagine in order for George to lose John's hair at all (let alone less of it than John lost himself). (2.3.24)(b) (in either version), on the other hand, freely permits the natural"sloppy" reading. Notice that this is a rather peculiar state of affairs, if (2.3.24)(b) is to be derived from (2.3.24)(a) by an optional deletion rule. Why should an extra reading pop up?

What I would like to suggest, is that perhaps sentences like "John lost more of his hair than Bill did", or "John ate more potatoes than Bill did" do not involve CE at all. Given our analysis, it's quite possible that these sentences can be derived by VPD. In order for this to be the case, we would have to assume that the logical forms for comparative sentences are rather simpler than what has sometimes been proposed. This may be independently motivated. Recent work on the semantics of comparative sentences (Bartsch and Vennemann (1972), Moravcsik and Gabbay (forthcoming)) points to the conclusion that computing what is actually being compared in such sentences is actually far more complex than people have imagined (consider what is being compared in a sentence like "More dentists use Lavoris than any other mouthwash") and that knowledge of the real world is often inextricably involved in interpreting them. Thus the rather simplistic proposals for the logical forms of comparative sentences (e.g. that of Postal (1974b)) which specify operators like EXCEED or SAME, which are double-variable-binding, are unlikely to be sufficient for representing the meanings of such sentences.

Now in any case it does seem to make sense to use a quantifier-like operator to represent comparative sentences logically. Let us write this operator [ER]. It may be the case, however, that the logical forms, in our sense, of comparatives need not specify the particular parameter of comparison. It may be perfectly sufficient to give logical forms like the following, and leave the further specification of what is being compared to the rules of SI-2, with the possible intrusion of other (perhaps pragmatic) systems as well (the ultimate semantic representations of comparative may well involve definite description operators or perhaps rather complex measure functions (as Bartsch and Vennemann suggest)).[25]

(2.3.25)(a) John ate more potatoes than Harry ate.

(b) $[ER]^\circ([\text{John}, \lambda x(x \text{ ate } \Delta^\circ \text{ many potatoes})], [\text{Harry}, \lambda y(y \text{ ate } \Delta^\circ \text{ many potatoes})])$

or John, $\lambda x([ER]([x, \lambda y(y \text{ ate } \Delta° \text{ many potatoes}], [Harry, \lambda z(z \text{ ate } \Delta°$ $\text{many potatoes})]))$

If something like this can be maintained, then Bach, Bresnan and Wasow's examples might have a logical form essentially like...

(2.3.26) $[ER]°([John_i, \lambda x(x \text{ lost } \Delta° \text{ much of } his_i \text{ hair})],$
$[George_j, \lambda z(z \text{ lost } \Delta° \text{ much of } his_{i,j} \text{ hair})])$

VPD could then apply, as $\lambda x(\ldots)$ and $\lambda z(\ldots)$ are alphabetic variants. If PRO→BV applied in each sentence in (2.3.26), we would furthermore have the possibility of the "sloppy" reading.

Further support for this analysis might come from sentences like the following (due to Edwin Williams), which have all the properties observed

(2.3.27) Mary's father wanted her to work harder than her boss did.

in Chapter One regarding sentences like (2.3.28):

(2.3.28) Betsy's father wanted her to read everything her boss did.

Most importantly, (2.3.27) cannot be used to report a situation where Mary's father says: "Mary, I want you to work harder than your boss wants you to work". Thus, quite analogous to our treatment of (2.3.28) (see Section 2.1 above), VPD would be impossible on the following reading, where $\lambda z(\ldots)$ has no alphabetic variant.

(2.3.29) Mary's father, $\lambda x(x \text{ want}([ER]°([Mary, \lambda y(y \text{ work } \Delta° \text{ hard})],$
$[Mary\text{'s boss}, \lambda z(z \text{ want}(Mary, \lambda w(w \text{ work } \Delta° \text{ hard})))])))$

A possible objection to the reanalysis of Comparative Ellipsis

(CE_1, actually) as VPD, might be raised from sentences like these next two, as several people have pointed out to me.

(2.3.30)(a) John ate more beans than Mary did bananas.

(b) Fred took more of his brothers to the movies than Bill did to the circus.

The problem here is that in these sentences, less than a VP has been deleted, which is, in general, not possible with VPD (but see the remarks at the end of Section 1.2 in Chapter One).

These sentences are generally presumed to arise from the application of (first) Subdeletion and (then) CE_1. This, I would argue, is indeed a possible analysis of them. If we reserve CE_1 for cases that have undergone Subdeletion, then we have reduced the number of environments CE_1 must apply in, and have perhaps rendered it possible to formulate (although the difficulties that remain in writing the rule of CE_1 still appear formidable). An alternative might be to collapse CE_1 and Subdeletion, treating CD as Relativization, perhaps, as Chomsky suggests.

There are two bits of evidence I have been able to find that point to the correctness of isolating cases of CE_1 which are reanalyzable as VPD. First, the type of ellipsis occurring in (2.3.30) seems to be very restricted. Unlike the other cases of CE_1, this type of ellipsis rapidly becomes unintelligible as the depth of embedding increases. Consider the following minimal pair, for instance.

(2.3.31)(a) *?Fred took more of his brothers to the movies than Jane claimed Sandy said that he did ø (ø) to the circus.

(b) Fred took more of his brothers to the movies than Jane claimed Sandy said that he did ø.

Subdeletion itself decreases in acceptability as the depth of embedding increases, of course, as Bresnan notes. The difference between (2.3.31)(a) and (b), however, is particularly striking, indicating that "unbounded" CE_1 contributes to the deviance of such sentences. But if (2.3.31)(b) also involves CE_1, then why is there no decay of acceptability? The non-deviance of (2.3.21)(b) follows if it is derived by VPD, which, as we have seen, freely applies over "unbounded" stretches of tree.

Another piece of evidence that cases like those in (2.3.20) should not be derived by VPD is the fact that they do not allow "sloppy" readings. Consider the following minimal pair, for instance.

(2.3.32) Fred took more of his brothers to the movies than Bill did.

(2.3.33) Fred took more of his brothers to the movies than Bill did to the circus.

(2.3.32), it seems, where an entire VP has been deleted, is ambiguous. Bill could have taken either Fred's or his own brothers to the movies. (2.3.33), however, seems to permit only the reading where Bill took Fred's brothers to the circus. That is, (2.3.33) lacks a "sloppy" reading, thus providing support for treating such sentences not on a par with the cases reanalyzable as VPD (e.g. (2.3.32)), but rather as instances of the separate rule of Subdeletion (with the possible further application of CE_1).

2.3.3 Some and Any

A matter of much dispute in the recent literature (both philosophical and linguistic) concerns the proper treatment of the quantifier word any (see Quine (1960), Vendler (1967), Horn (1972), Kamp (1973) for some initial discussion). One view, most notably that of Quine, is that _any_ is to be regarded always as a universal quantifier. This is a popular view among philosophers. The contrary view, popular among linguists (at least until lately) originates with Klima (1964), who proposed a syntactic rule converting _some_ to _any_ in certain (he called them "affective") environments (see R. Lakoff (1969) for interesting problems with this proposal). This latter position involves the claim that _any_ is to be treated as an existential quantifier, at least in some circumstances.

A standard argument that "every _any_ means _every_" (to use the slogan of Savin (1971)) involves sentences like the following one.

(2.3.34) If anyone leaves, he will be unhappy.

The claim is that in order to treat the pronoun _he_ as a bound variable, one must analyze a sentence like this as involving a universal quantifier with scope over the conditional, say as the following

(2.3.35) $(\forall x)(x \text{ leaves} \rightarrow x \text{ will be unhappy})$.

So, the argument goes, _any_ must be treated as a universal quantifier (with widest possible scope) if we are to assign reasonable logical forms to sentences like (2.3.34).

Kamp (1973) points out a flaw in this line of reasoning, however. He notes the possibility of sentences like this next one.

(2.3.36) If someone leaves, he will be unhappy.

Here, there is no natural account one can give by treating _someone_ as having wide scope. Thus this next logical form does not accurately represent the meaning of (2.3.36).

(2.3.37) (Ex)(x leaves $\rightarrow$ x will be happy)

(2.3.37) would be true, for instance, if there is somebody who leaves and is not happy, just as long as there is at least one other person who leaves and is happy. In this same situation, (2.3.36) would be false.

Clearly some further device is necessary to "link up" quantifiers with pronouns that are not (by virtue of the syntax) within their scope. Further evidence for the necessity of some such device, also pointed out by Kamp, involves quantifier-pronoun "linkings" ("linkages"?) in discourse:

(2.3.38) Speaker A: Someone left.
Speaker B: Do you know why he left?

It's hard to see how one could give an account of such cases in terms of an existential quantifier whose scope extends across sentences and speakers in discourse.

But if some such device is independently necessarily, as Kamp argues, then there is no necessity to treat <u>any</u> in examples like (2.3.34) above as a universal quantifier. Whatever device accounts for sentences like (2.3.36), which clearly involves an existential quantifier, will also handle analogous sentences with <u>any</u>, if <u>any</u> is treated as an existential quantifier.

The evidence from VPD provides support for Kamp's position, that is, for the position that <u>any</u> (at least sometimes) is to be treated as an existential rather than a universal quantifier. The examples are well-known, actually. They are cases like the following:

(2.3.39) John doesn't see anyone, but Bill does.

This clearly means the same thing as the following sentence.

(2.3.40) John doesn't see anyone, but Bill does see someone.

If we assume that both <u>someone</u> and <u>anyone</u> are represented by existential quantification at the level of logical form, then we can account for this sort of deletion with no difficulty. (2.3.40) will give rise to (2.3.39) by virtue of its logical form, which we write as this:

(2.3.41) $\neg$John, $\lambda x((Ey)(x \text{ see } y))$ but Bill, $\lambda w((Er)(w \text{ see } r))$

$\lambda x(\ldots)$ and $\lambda w(\ldots)$ are alphabetic variants. Crucially, observe that treating <u>any</u> as a universal quantifier here will predict that no deletion is possible in (2.3.40):

(2.3.42) $(\forall x)$ $(\neg$ John, $\lambda y(y \text{ see } x))$ but Bill, $\lambda w((Ez)(w \text{ see } z))$

$\lambda w(\ldots)$ has no alphabetic variant. Our theory thus makes an interesting empirical prediction about the logical representations of certain quantifier words.

2.3.4 <u>Until</u>

There's been quite a controversy over this word and its relation to VPD. Actually, the controversy has all been about the following sentence, which Grinder and Postal (1971) (and about one third of my informants) find ungrammatical.

(2.3.43) %John won't leave until midnight, but Bill will.

Chomsky (1972b), however, finding this sentence acceptable, uses it as an argument against a deletion theory of VPD, for its source (where <u>until</u> occurs in an inappropriate environment in the second conjunct) is certainly ungrammatical in all dialects.

(2.3.44) *John won't leave until midnight, but Bill will leave until midnight.

Chomsky does not make clear how an interpretive theory would do any better than a deletion theory with regard to (2.3.43), however.

With respect to our theory, this raises an interesting question about the nature of dialect variation. Suppose, for instance, that Karttunen (1974) is right about the semantic representation of until-sentences. He proposes that a sentence like John won't leave until midnight should have a semantic representation ("logical form" in his sense) like the following.

(2.3.45) NOT(John will leave before midnight)

The implication that John will leave at midnight (or shortly afterwards) arises through pragmatic presupposition in Karttunen's theory. For our purposes, actually, all that is relevant is that until be decomposed into before.

If we accept any such decompositional analysis of until, we then can derive (2.3.43) from this next sentence (which, if you think about it, means exactly what (2.3.43) means), whose logical form is as given.

(2.3.46) John won't leave until midnight, but Bill will leave before midnight.

(2.3.47) $\neg$ WILL(John, λx(x leave before midnight)) but WILL(Bill, λy(y leave before midnight))

Here $\lambda x(\ldots)$ and $\lambda y(\ldots)$ are alphabetic variants, and deletion is possible.

The dialect variation then (if indeed that's what it is) would be described in terms of when the decomposition of until into before (and whatever else) takes place. In a dialect that allows the sentences in question, the rule that effects the decomposition of until is a rule of SI-1. In dialects that don't, it is a rule of SI-2. In the latter dialects then, at the level of logical form, undecomposed until will prevent the alphabetic variance necessary for deletion.

It should be noted that, in general, the cases of lexical decomposition that have been proposed in the literature (e.g. by McCawley (1973), Lakoff (1965, 1972), Katz (1972), etc.) have no place whatsoever in our logical forms. For example, if transitive verbs like, say, melt are decomposed into CAUSE plus (intransitive) melt by rules of SI-1, then we would expect such horrendous deletions as the following to be grammatical.

(2.3.48) *Bill melted the copper vase, and the magnesium vase did ø, too.

[ø = melt]

Therefore, if lexical decomposition of verbs has any place at all in linguistic theory (which I am dubious of), that place is certainly not in logical forms.

Determining whether there is some principled bases for allowing decomposition with connectives like before but not with other lexical items, I leave as an exercise for the reader, to whom I also apologize for the extremely sloppy use in this discussion of the term "dialect".[26]

2.3.5 Proper Names and Definite Descriptions

Let us backtrack for a moment and reconsider an example we treated in some detail earlier. In Section 2.1 we accounted for the ambiguity of sentences like (2.3.49).

(2.3.49) Bill wanted Betsy to read everything that Sam wanted her to read.

We assigned this sentence two logical forms. One reading, which seems plausibly regarded as "opaque" in Quine's (1960) sense, we gave as (2.3.50).[27]

(2.3.50) Bill, λx(x want [($\forall$y:Sam, λz(z want [Betsy, λw(w read y)]))
[Betsy, λq(q read y)]])

This is crudely paraphrasable as: <u>What Bill wants is that Betsy read everything Sam wants her to read</u>.

The other reading, intuitively "transparent" in Quine's sense, is paraphrasable roughly as: <u>Everything that Sam wants Betsy to read is also such that Bill wants Betsy to read it</u>. This we represented as (2.3.51).

(2.3.51) ($\forall$x:Sam, λy(y want [Betsy, λz(z read x)]))
[Bill, λw(w want [Betsy, λr(r read x)])]

Now this is not quite right, for we must account for the possibility of the following sentence, which seems to allow a transparent reading.

(2.3.52) Bill wanted Betsy to read everything that Sam wanted her to read, and Peter did ∅ too.
[∅ = wanted Betsy to read everything that Sam wanted her to read]

This is a problem, because if we assign the right conjunct in (2.3.52) a logical form analogous to (2.3.51), i.e. this:

(2.3.53) $(\forall q{:}\mathrm{Sam},\ \lambda s(s\ \mathrm{want}\ [\mathrm{Betsy},\ \lambda m(m\ \mathrm{read}\ q)]))$
$[\mathrm{Peter},\ \lambda o(o\ \mathrm{want}\ [\mathrm{Betsy},\ \lambda k(k\ \mathrm{read}\ q)])]$,

then the appropriate λ-expressions (λo(...) in the latter representation and λw(...) in the former) will not be alphabetic variants, because the variables q and x are bound from outside the λ-expressions.

This situation is easily remedied by allowing double λ-abstraction in such cases. That is, we could write, instead of (2.3.53), the logically equivalent (2.3.54).

(2.3.54) $\mathrm{Peter},\ \lambda p((\forall q{:}\mathrm{Sam},\ \lambda s(s\ \mathrm{want}\ [\mathrm{Betsy},\ \lambda m(m\ \mathrm{read}\ q)]))$
$[p,\ \lambda o(o\ \mathrm{want}\ [\mathrm{Betsy},\ \lambda k(k\ \mathrm{read}\ q)])])$

Allowing double λ-abstraction does not present any particular difficulty for our theory. It will mean, however, that in some cases we will assign multiple logical representations to certain sentences, though those representations will be logically equivalent.

This will allow us to account not only for the deletion in (2.3.52), but also for such double deletion possibilities as the following (on the transparent reading),

(2.3.55) Bill wanted Betsy to read everything that Sam did $\emptyset_1$, and Peter did $\emptyset_2$, too.
[$\emptyset_1$ = wanted Betsy to read; $\emptyset_2$ = wanted Betsy to read everything that Sam wanted Betsy to read]

The logical form here will be:

(2.3.56) Bill, $\lambda t((\forall x{:}Sam, \lambda y(y \text{ want } [Betsy, \lambda z(z \text{ read } x)]))$
$[t, \lambda w(w \text{ want } [Betsy, \lambda r(r \text{ read } x)])])$ &
Peter, $\lambda p((\forall q{:}Sam\ \lambda s(s \text{ want } [Betsy, \lambda m(m \text{ read } q)]))$
$[p, \lambda o(o \text{ want } [Betsy, \lambda k(k \text{ read } q)])])$

Alphabetic variance obtains between $\lambda y(\ldots)$ and $\lambda w(\ldots)$, and between $\lambda t(\ldots)$ and $\lambda p(\ldots)$. Thus both deletions in (2.3.55) are predicted to be possible, correctly. [28]

Now that we have cleared up this matter, notice that there is another interesting fact to be accounted for here. The following sentence, which one might expect to be derivable from the same source as (2.3.55), but where only the embedded VP: [read everything that Sam wanted Betsy to] in the second conjunct has been deleted, seems to lack the transparent reading.

(2.3.57) Bill wanted Betsy to read everything that Sam wanted her to read, and Peter wanted her to ∅, too.
[∅ = read everything that Sam wanted Betsy to read]

This sentence can be interpreted only opaquely, i.e. as ...<u>and what Peter wanted Betsy to do was to read everything that Sam wanted her to read.</u>

This follows from our theory, because the deleted embedded VP in (2.3.57) corresponds to $\lambda k(\ldots)$ in (2.3.56), the logical representation for the transparent reading. $\lambda k(\ldots)$ has no alphabetic variant in (2.3.56), however, because the only reasonable candidate for such, $\lambda r(\ldots)$, contains <u>x</u> bound from outside where $\lambda k(\ldots)$ contains <u>q</u> bound from outside. [29] This is a very nice empirical prediction that our theory makes.

In fact, this is a characteristic prediction of our theory. In cases where a quantifier has wide scope, i.e. scope over two VP's, we will always predict that the "higher" VP is deletable and that the "lower" one is not. We saw this earlier with regard to the object deletion cases (ready to eat) in Section 2.1, and with respect to the possibility of "sloppy" readings (John said Mary hit him and Bill did too vs. ...and Bill said she did too) in Section 2.2. Those cases all involved variations in alphabetic variance that concerned variables bound by our λ-operators. Thus the fact that quantifiers like everything work the same way, as in the examples we have just considered, provides further support for our entire λ-machinery, i.e. for the correctness of the λ-calculus with regard to capturing linguistically significant generalizations.[30]

What then of proper names and definite descriptions? These are often analyzed as logical operators admitting of scopal variation (the position is essentially due to Russell (1905); recent specific proposals have been made by Keenan (1970,1971a, 1972), McCawley (1971), and others). Take proper names, first. It is well-known that they too admit of transparent-opaque differences in understanding in sentences like the following.

(2.3.58) Norma thinks that she likes Ronald Reagan.

The position to the right of likes in (2.3.58) may be treated as purely referential or not. If it is so treated, then a truthful utterance of (2.3.58) taken together with the truth of (2.3.59)

(2.3.59) The guy who mugged Norma last night is Ronald Reagan.

allows us to draw the conclusion (by the well-known principle of the substitutivity of identity) that...

(2.3.60) Norma thinks that she likes the guy who mugged her last night.

"thinks that", in this case, is transparent.

Alternatively, (2.3.58) may be understood in such a way that, given that (2.3.59) is true, we might still be prepared to deny (2.3.60) (suppose, say, it was too dark for Norma to get a good look at the guy who mugged her). In this case, we would say that "thinks that" is opaque.

Now if we were to follow Keenan, we would assign two different logical representations to (2.3.58) to account for these two different understandings. Proper names, on Keenan's account, are treated as variable-binding operators, on a par with $\underline{\forall}$, and $\underline{E}$. Adopting his proposal would lead us to represent the transparent understanding of (2.3.58) roughly as follows (we will write (Proper Name-variable), which is analogous to $(\forall y)$ or (Ex)).

(2.3.61) (Norma-x)(Ronald Reagan-y)[x, λw(w think [w, λz(z like y)])],
or (Norma-x)[x, λw((Ronald Reagan-y)[w, λr(r think
[r, λz(z like y)])])]

Transparent and opaque differences in understanding, in Keenan's system, correspond to differences in scope of the Proper Name operators. The opaque reading of (2.3.58) would be represented by a formula like (2.3.62), where (Ronald Reagan-y) is inside the scope of think.

(2.3.62) (Norma-x)[x, λw(w think [(Ronald Reagan-y)[w, λz(z like y)]])]

This proposal makes some incorrect predictions about deletion, which we will turn to in a moment. First, however, it should be pointed out that there are some fundamental problems with this entire approach.

Notice that precisely the same difference in understanding as the one just described for sentences containing proper names seems to exist for other sentences as well. For instance, given that both of these next two sentences are true,

(2.3.63) Nixon signed Document 12.

(2.3.64) Signing Document 12 caused the slaughter of thousands of Asians.

it follows that...

(2.3.65) Nixon caused the slaughter of thousands of Asians.

If we embed (2.3.63) in the complement of believe, however, an opaque understanding is possible. Thus, given that both (2.3.66) and (2.3.64) are true,

(2.3.66) Nixon believes he signed Document 12,

it does not necessarily follow that...

(2.3.67) Nixon believes he caused the slaughter of thousands of Asians.

Facts like these pose grave difficulties for a semantic theory like Keenan's. Would one introduce predicate quantifiers that bind variably that range over VP's? I think not. Moreover, it seems that transparent-opaque differences in understanding can involve referential vs. non-referential position of almost any part of speech, rendering rather dubious any approach like Keenan's, which must quantify into all such positions. Surely all such cases must be given a unified account.

Aside from this objection, it is clear that Keenan's theory makes incorrect predictions about the possibility of VPD. A sentence like the following one, for instance, would have a Keenanian logical representation roughly like that in (2.3.69) on the transparent reading, and like that in (2.3.70) on the opaque reading.

(2.3.68) Norma thinks she likes Ronald Reagan, and Lois thinks she likes Ronald Reagan, too.

(2.3.69) (Norma-r)[r, λx((Ronald Reagan-y)[x thinks [x, λz(z like y)]])] & (Lois-m)[m, λw((Ronald Reagan-q)[w think [w, λp(p like q)]])]

(2.3.70) (Norma-r)[r, λx(x think [(Ronald Reagan-y)[x, λz(z like y)]])] & (Lois-m)[m, λw(w think [(Ronald Reagan-q)[w, λp(p like q)]])]

Now given this, we would predict that the following deleted sentence is possible on either reading.

(2.3.71) Norma thinks she likes Ronald Reagan, and Lois does ∅, too. [∅ = thinks she likes Ronald Reagan]

In both logical representations, λw(...) and λx(...) are alphabetic variants.

This prediction would seem to be correct. The further prediction, that no "crossed" reading is possible, is difficult to verify. I have not been able to construct convincing examples to decide whether or not in general such "crossed" readings exist. If such readings do exist, of course, then we have an independent argument that transparent vs. opaque differences in understanding should not be represented at the level of logical form at all.

Notice, however, that if we adopt a Keenanian approach to this problem, then we would predict that deletion of just the embedded VP: likes in the second conjunct should be possible only on the opaque reading:

(2.3.72) Norma thinks she likes Ronald Reagan, and Lois thinks she does too.

That is, on the transparent reading (in (2.3.69)), the relevant λ-expressions (λp(...) and λz(...)) are not alphabetic variants because they contain non-identical variables (y and q), which are bound from outside. In other words, if proper names are to be treated on a par with quantifiers, admitting of scopal variation, then we would expect them to behave like quantifiers which, as we saw earlier, do not permit deletion of embedded VP's when they have wide scope.

Crucially, the prediction seems to be incorrect. (2.3.72) allows a transparent reading with no apparent difficulty. This would seem then to provide further evidence that proper names should not be given a scopal treatment at the level of logical form.

This conclusion is not quite forced upon us, however. We would account for the descrepancy between the deletion facts of sentences containing proper names and those of sentences containing quantifier words by letting proper names take wide scope over (in the last sentence we looked at) both conjuncts. Thus we might assume that the logical representation for the source of (2.3.72) on the transparent reading is the following.

(2.3.73)(a) Norma thinks she likes Ronald Reagan, and Lois thinks she likes Ronald Reagan, too.

(b) (Ronald Reagan-y)[Norma, λx(x think [x, λz(z like y)]) & Lois, λw(w think [w, λp(p like y)])]

Given this representation, we predict correctly that both VP's in the second conjunct are deletable (i.e. that (2.3.71) and (2.3.72) above both allow transparent readings), because in this logical representation $\lambda w(\ldots)$ and $\lambda x(\ldots)$ are alphabetic variants, as are $\lambda p(\ldots)$ and $\lambda x(\ldots)$.

This proposal, which is essentially Keenan's, would pose some problems for our theory of logical form as the output of sentence grammar. This is so, because in sentences like this next one, where the first utterance is interpreted transparently,

(2.3.74) Speaker A: John was sad that he hit Bill.
Speaker B: Are you sure he did ∅?
[∅ = hit Bill]

would force us to construct a logical form where 'Bill' has scope over two sentences in a discourse. We might do this as follows (where "?" is to be interpreted as a question operator):

(2.3.75) (Bill-y)[John, λz(z was sad that [z, λx(x hit y)]).
?(you're sure that [John, λw(w hit y)])]

Since this would involve revising our entire theory of logical form, which, as we have been developing it, is part of an autonomous sentence grammar, it's worth asking whether there is any other evidence that might bear on this question. There is, it seems, a certain amount of evidence that other discourse matters that pertain to semantic interpretation should not be resolved at the level of logical form.

Recall the examples we discussed in Section 2.3.3 above (due to Kamp), where certain pronouns which do not, strictly speaking, fall within the scope of certain quantifiers must be anaphorically linked to those quantifiers at some level of representation. We observed such examples as the following.

(2.3.76) Speaker A: Someone left.

Speaker B: Do you know why he left?

We hypothesized that a rule of SI-2 (or a rule of discourse grammar) might operate to convert such pronouns as <u>he</u> in this example into bound variables.

Consider now the following discourses.

(2.3.77) Fortune Teller: Mary will meet someone.

Client: Do you think I will meet him too?

(2.3.78) Fortune Teller: Mary will meet someone.

Client: Does that mean I will meet someone too?

In the first discourse, <u>him</u> is anaphorically linked to someone in the previous sentence. Notice that if we allowed the existential quantifier to have wide scope at the level of logical form, allowing also that <u>him</u> be represented as a variable bound by that quantifier, then we would have a logical form for the discourse like the following, where $\lambda w(\ldots)$ and $\lambda y(\ldots)$ are alphabetic variants.

(2.3.79) (Ex)[WILL(Mary, λy(y meet x)).
?(you think [WILL [I, λw(w meet x)]])]

Accepting this, we would expect VPD to be possible. This would allow (2.3.77) to become (2.3.80).

(2.3.80) Fortune Teller: Mary will meet someone.

Client: Do you think I will $\emptyset$ too?

But, as far as I can tell, this last discourse is not interpretable in this way.[31] Rather, it is only interpretable as elliptical for (2.3.78), i.e.

where [$_{VP}$meet someone] has been deleted. The possibility of that interpretation is, of course, unproblematic, for the logical forms of the two sentences in (2.3.78) would be the following, where $\lambda x(\ldots)$ and $\lambda y(\ldots)$ are alphabetic variants.

(2.3.81)(a) WILL(Mary, λy((Ez)[y meet z]))

(b) ?(you, λw(w think [WILL(I, λx((Es)[x meet s]))]))

Precisely these facts would be accounted for if the (discourse) rule that operates in examples like (2.3.77) does not apply until after the level of logical form. Thus our hypothesis that only rules of sentence grammar apply to map syntactic objects into logical forms actually does some work for us in these cases. For this reason, it would seem to be preferable not to treat proper names scopally at the level of logical form in the way we speculated a moment ago (if at all), and to continue to represent them as indexed constants.

Now definite descriptions may be another matter, depending on whether or not sentences like the following are taken to be grammatical.

(2.3.82) John read the books that Bill did ∅.

[∅ = read]

Presuming that such sentences are acceptable (most people seem to find them so), it's easy to show that definite descriptions must be treated scopally at the level of logical form. Treating NP's with the as complex (but non-scopal) entities leads us to logical forms like (2.3.83).

(2.3.83) John, λx(x read{(the y)[books(y) & Bill, λw(w read y)]})

This leaves us with no way of accounting for the possibility of deletion in (2.3.82), for $\lambda w(\ldots)$, which corresponds to the deleted $[_{VP}\underline{\text{read}}]$ in (2.3.82), has no alphabetic variant in (2.3.83).

One approach we might take to the logical form of sentences containing definite descriptions is the one suggested by Chomsky (1975c). Chomsky proposes to represent a sentence like (2.3.84), as indicated, where B, O, and A denote the class of books, the things we ordered, and the things that arrived, respectively, and c[X] represents the cardinality of (the set) X.

(2.3.84) The book we ordered arrived.

(2.3.85) $B \cap O \subset A \; ; \quad c\,[\,B \cap O\,] = 1$

Plural definite descriptions, such as the one in (2.3.85) differ only with respect to cardinality, as is indicated.

(2.3.86) The books we ordered arrived.

(2.3.87) $B \cap O \subset A \; ; \quad c\,[\,B \cap O\,] \geq 2$

Chomsky goes on to suggest that an analysis of this type allows the word <u>the</u> to be treated uniformly as a universal quantifier. How might this proposal fit into our theory of logical form?

First of all, it seems that Chomsky's proposal must be modified slightly. Notice that if we take something like (2.3.87) as the logical representation of (2.3.86) (with plural definite description), then in the event that we only ordered one book and that book arrived, (2.3.86) will be false. This appears to be incorrect. That is, imagine a situation where the set of books under discussion is of a cardinality greater than 1, but (say, on this particular occasion) the number of books we actually ordered is only one (perhaps on every other day we ordered more than one book).

In that situation, it seems to me, (2.3.86) would be true. One might argue that in such a situation (2.3.86) would be less appropriate than, say, (2.3.84), but this difference in appropriateness, I feel, is entirely pragmatic (one might invoke Grice's (1975) maxim of quantity). In other words, the inappropriateness of (2.3.86) in the context under discussion is rather like the inappropriateness of uttering (2.3.88) if the utterer knows that Mary in fact has four daughters:

(2.3.88) Mary has three daughters.

In this situation, we would not accuse the utterer of (2.3.88) of having told a lie, but rather of not having been as informative as he might have been. In this same situation, notice, (2.3.89) would, by anyone's standards, be false.

(2.3.89) Mary has five daughters.

(2.3.86) uttered in a situation where only one book was ordered (and that book arrived) seems to pattern with (2.3.88), not with (2.3.89), i.e. it would be true, but not maximally informative.

Given this, we might propose that (2.3.86) be assigned essentially the logical representation in (2.3.90), and (2.3.84) the one in (2.3.91).

(2.3.90) $B \cap O \subset A \; ; \; c\,[\,B\,] \geq 2$

(2.3.91) $B \cap O \subset A \; ; \; c\,[\,B\,] = 1$

In terms of our notation, we then have the following logical forms for the two sentences in question (($\forall$x:...) again represents a restricted quantifier).

(2.3.92) $(\forall x\colon [\text{we}, \lambda y(y \text{ ordered } x)] \;\&\; x \in \hat{z}(\text{book}(z)) \;\&$
$c\,[\hat{z}] \geq 2)\,[x, \lambda w(\,w \text{ arrived})]$

(2.3.93) $(\forall x\colon [\text{we}, \lambda y(y \text{ ordered } x)] \;\&\; x \in \hat{z}(\text{book}(z)) \;\&$
$c\,[\hat{z}] = 1)[x, \lambda w(w \text{ arrived})]$

Other notations are of course possible.

Notice that a treatment of the definite article along these lines also solves the problem we were worying about a moment ago re VPD. John read the books that Bill read would be assigned a logical form as in (2.3.94).

(2.3.94) $(\forall x\colon [\text{Bill}, \lambda y(y \text{ read } x)] \;\&\; x \in \hat{z}(\text{book }(z)) \;\&$
$c\,[\hat{z}] \geq 2)[\text{John}, \lambda w(w \text{ read } x)]$ or perhaps

$\text{John}, \lambda p((\forall x\colon [\text{Bill}, \lambda y(y \text{ read } x)] \;\&\; x \in \hat{z}(\text{book}(z)) \;\&$
$c\,[\hat{z}] \geq 2)[p, \lambda w(w \text{ read } x)])$

in either case, $\lambda w(\ldots)$ is an alphabetic variant of $\lambda y(\ldots)$ (since they both contain the variable x and are otherwise identical). The deletion in (2.3.82) above (John read the books that Bill did) is therefore accounted for.[32]

It is important to note, however, that definite descriptions otherwise work like proper names with respect to deletion. In this next example, for instance, we find two possible deletions.

(2.3.95) John was sure he saw the man Sue liked.

(a) Bill was ∅, too. [∅ = sure he saw the man Sue liked]

(b) Bill was sure he did ∅, too. [∅ = see the man Sue liked]

These discourses seem quite possible, when the first sentence is interpreted transparently or opaquely. This means that we would not want to let definite descriptions take variable scope (the argument is quite analogous to the one made above for proper names), unless, of course, we allow them to have scope over two (or more) sentences in discourse. In light of the above discussion, however, that would seem unadvisable. The tentative conclusion we will draw then, is that definite descriptions should be treated as quantifiers, but, as far as logical form is concerned anyway, their scope should be determined by the clause they are in (as in (2.3.94) above for instance).

Now given that something like this is a correct treatment of the logical forms of the-sentences, it's quite possible that a similar treatment of related sentences with the same can be brought to bear on another problem that has been noted in the literature. Liddell (1975) faults Postal (1974b), correctly, in my opinion, for the latter's claim that the ungrammaticality of (2.3.96) is due to a general prohibition against lowering predicates (Postal analyzes SAME as a two-place predicate in logical structure) into island contexts.

(2.3.96) *Jill bought Melvin's picture of the same orgy Arthur did.

Liddell points out that examples like the following one, which, in Postal's system, would also involve lowering the "predicate" SAME into an island context.

(2.3.97) Jill bought Melvin's picture of the same orgy Arthur bought a picture of.

The problem with (2.3.96), Liddell argues, "seems to be the replacement of bought a picture of with did". This intuition seems to be correct.

It is not at all clear what the right logical representations for sentences involving words like the same should be, but notice that the word same doesn't really add very much to the meaning of such sentences. It's difficult, for instance, to decide what the difference in meaning is between say, (2.3.97) and a sentence like (2.3.98).

(2.3.98) Jill bought Melvin's picture of the orgy that Arthur bought a picture of.

Same almost seems to be entirely "emphatic" in nature in sentences like (2.3.98) (very behaves similarly).

We might propose then that sentences with the same or the very (in pro-nominal position when a restrictive relative clause follows) are assigned logical forms which are not essentially different from those of corresponding sentences with just the definite article. The logical form of (2.3.97) then would be very close to the logical form of (2.3.98), which, given the approach we have been developing, would be roughly as follows.

(2.3.99) $(\forall x$: [Arthur, λy(y bought a picture of x)] & $x \in \hat{z}(orgy(z))$ & $c\ [\hat{z}] = 1)$[Jill, λw(w bought Melvin's picture of x)]

Let us assume that this is at least a fair approximation of the logical form of (2.3.97). Given this, it's easy to see why deletion is not possible: $\lambda y(\ldots)$ and $\lambda w(\ldots)$ contain instances of the same bound variable all right, but they are not otherwise identical, i.e. they are not alphabetic variants.

Suppose we modify (2.3.99) minimally so as to induce alphabetic variance. This could be done in two ways. We could modify either $\lambda w(\ldots)$ or $\lambda y(\ldots)$. Suppose we modify $\lambda w(\ldots)$. The result is the logical formula in (2.3.100).

(2.3.100) $(\forall x$: [Arthur, λy(y bought a picture of x)] &
$x \in \hat{z}(\text{orgy}(z))$ & $c\ [\hat{z}] = 1$)[Jill, λw(w bought a picture of x)]

This representation corresponds to the following sentence.

(2.3.101) Jill bought a picture of the same orgy Arthur bought a picture of.

Our theory predicts deletion should be possible (we have induced alphabetic variance),and it is.

(2.3.102) Jill bought a picture of the same orgy Arthur did.

On the other hand, if we take the other tack, and minimally modify λy(...) in (2.3.99), the result is this:

(2.3.103) $(\forall x$: [Arthur, λy(y bought Melvin's picture of x)] &
$x \in \hat{z}(\text{orgy}(z))$& $c\ [\hat{z}] = 1$)[Jill, λw(w bought Melvin's picture of x)]

Now we might expect this to give rise to Postal's example (2.3.96), for λy(...) and λw(...) are indeed alphabetic variants. But notice what the source for the deletion would have to be:

(2.3.104) *Jill bought Melvin's picture of the same orgy Arthur bought Melvin's picture of.

(2.3.104) would provide an appropriate source for (2.3.96) in our theory, but it is a case of extraction out of an island, illegitimate in anyone's theory. Therefore, for purely syntactic reasons, Postal's example has no source. Our theory of VPD thus interacts with well-motivated syntactic constraints to predict exactly the right array of facts.

2.4 Summary of Chapter Two

In Chapter Two we have developed a theory of logical form and a theory of deletion. Having examined previous discussions of the recoverability of deletion, we concluded that "identity of logical form" (a notion we gave considerable substance to) is not only a necessary condition for deletion, but a sufficient one as well. The theory of logical form we have developed has drawn heavily both from Chomsky's Extended Standard Theory and from Montague Grammar. We have incorporated insights from both of these approaches. This has enabled us to give a coherent account not only of long-standing problems, e.g. the problem of "sloppy identity", but also of many new facts we have discovered concerning the interaction of quantifier interpretation and VPD.

Footnotes to Chapter Two

1. The entire notion of recoverability of deletion is, I believe, due to Harris. An explicit statement of his views can be found in Harris (1968, p. 78 ff.).

2. Hankamer (1971) considers this problem also, but in less detail than Lakoff.

3. Many of the observations of the last section of Chapter One and certain of the explanations proposed in this chapter (particularly certain applications of the lambda calculus) have been discovered independently by Edwin Williams. His ideas are developed in Williams (1975, 1976)(q.v.).

4. Note that this amounts to the claim that there are no cyclic rules that perform deletion <u>under identity</u>. The fact that there may be cyclic conditions on such rules (as shown, for instance, for Super Equi by Jacobson and Neubauer (forthcoming)) does not contradict this.

5. Or the alternative analysis suggested in Note 9 of Chapter One, where <u>Do</u>-Support and Affix Hopping apply before VPD.

6. Wasow's (1972) discussion differs somewhat from Chomsky's, actually. Wasow attempts to motivate a notion of "determinateness". Of the three understandings of a sentence like (*i*), only the (non-generic) non-specific understanding is claimed to be non-determinate.

(*i*) A Cro Magnon man lumbered along.

The specific and generic understandings are claimed to be determinate in virtue of the fact that both refer: the former to an individual and the latter to a set (the set of Cro Magnon men).

Only non-determinate entities are treated as variable-binding under Wasow's proposal. Curiously, the only examples he gives of illegitimate "backwards" variable binding with (indeterminate) non-specific singular indefinite NP's (i.e. NP's that begin with the indefinite article) are (*ii*) and (*iii*), which, as Ken Hale pointed out to Wasow (p. 84, nt. 8), are counterexamples to his claim that all non-specific indefinite NP's are subject to the constraint against "backwards" variable binding.

(*ii*) If you ask for it nicely, you can have an ice cream.

(*iii*) If you really want it, you can have a lollipop.

In short, Wasow presents no evidence that any NP's containing the indefinite article should be treated as variable-binding entities at the level of logical form.

This means of course, that we must treat such indefinite NP's univocally at the level of logical form, and that some account of the generic readings of sentences containing indefinite NP's must be given in terms of rules of SI-2, or in terms of some other system entirely. For some intriguing suggestions as to how a purely pragmatic theory of such matters might proceed, see Nunberg and Pan (1975) and Nunberg (forthcoming).

7. For a more formal discussion of this notion, see van Fraassen (1971, pp. 102-104), Hughes and Cresswell (1968), Kleene (1952), and especially Kalish and Montague (1964). λ is to be considered on a par with $\forall$ or $\exists$ with respect to alphabetic variance.

8. It is important to point out that here, and in all that follows, we will not repeat variables, even where repetition is allowed in standard formal logic. It is crucial, for example, that we write (*i*) and not (*ii*).

(i) $(\forall x)(f(x))$ & $(Ez)(g(z))$

(ii) $(\forall x)(f(x))$ & $(Ex)(f(x))$

9. Subject of course to the Backwards Anaphora Constraint. See the discussion in Chapter Four and especially that in Sag and Hankamer (1976).

10. This will be modified slightly in Section 2.3 of this chapter.

11. The tentative nature of this analysis of relative clauses should be reemphasized. Especially in light of Larry Horn's sentences cited in note 26 of Chapter One.

12. Similar facts are observable in the case of comparatives, e.g. sentences like (i) (due to Edwin Williams):

(i) Mary's father told her to work harder than her boss did.

These will be discussed in Section 2.3.

13. Fiengo and Lasnik argue that cases of Tough-Movement should also be treated as Object Deletion. On this see Jackendoff (1975).

14. This type of argument was pointed out to me by Geoff Pullum (personal correspondence - June 4, 1975), who attributes it to Michael Brame.

15. Not all speakers find (2.2.1) ambiguous. For some of those speakers, it is the discrepancy of gender that blocks the "sloppy" reading. Others object to similar examples involving discrepancy of person and/or number. Virtually all speakers find (i) ambiguous, however.

(i) John scratched his arm, and Bill did, too.

To placate the masses, I will henceforth skirt these problems by using examples like (i). I have observed, however, several instances of people performing gender-violating deletions of the sort they previously claimed they would not say. I am therefore inclined to allow all such sentences to be generated by the syntax.

16. So-anaphora may or may not be a subcase of VPD. See Sag and Hankamer (1976) for more discussion.

17. Keenan's theory of proper names is examined critically in Section 2.3 of this chapter.

18. Representations like (2.2.13), where the pronoun bears the same index as the subject NP, are a special case, of course. The important distinction to draw is that between pronoun (referring to whatever) and bound-variable.

19. The essence of Lasnik's proposal, which draws upon Dougherty's (1969) "anaporn" relation is perhaps not threatened by facts like these. The crucial point made by both Dougherty and Lasnik, that pronouns behave like NP's, is surely correct with respect to deletion. Thus the required (intended) coreference of (2.2.15) is also a property of cases like (*i*), where a non-pronominal NP inside a VP has been deleted.

(*i*) Betsy saw Peter, and Sandy did ∅, too. [∅ = saw Peter]

For (*i*) to be true, Betsy and Sandy must also have seen the same guy.

Thus, we might simply posit a general principle that requires coreference between referring entities that are inside the deletion target and corresponding referring entities inside the deletion trigger. Given this, indices of the sort we are using may be supplanted by some other device.

20. The problem of opaque contexts is discussed in the next section.

21. We have not dealt with Fodor's theory of "sloppy" identity in any detail. This is because Fodor makes the incorrect assumption that the phenomenon is restricted to genitive pronouns. The range of examples considered in this section makes clear why that assumption is incorrect.

22. In Section 3.5 of Chapter Three, however, we will offer an hypothesis, motivated on independent grounds, that accounts for why some speakers find such readings possible.

23. The problem is actually somewhat more complicated. In particular, the constraint on "backwards" variable binding cannot be stated merely in terms of preceding anaphoric proform. This is because, as Ross (frequently delivered unpublished lecture on "Primacy") points out, sentences like (*i*) have a reading paraphrasable as (*ii*), but not one paraphrasable by (*iii*) (underlining means coreference).

(*i*) <u>Edith</u> said that finding <u>her</u> husband nude had upset <u>her</u>, and Martha did ∅, too.

(*ii*) ...Martha said that finding Edith's husband nude had upset Martha.

(*iii*) ...*Martha said that finding Martha's husband nude had upset Edith.

A proper version of this constraint must therefore involve a notion something like Ross's notion of primacy. It is important to note that similar facts can be observed for cases of primacy-violating quantifier-bound variable linkages:

(*iv*) The woman who saw everyone talked to him.

(*iv*) cannot mean (*v*).

(*v*) $(\forall x)$(the woman who saw x talked to x)

This provides further support for our claim that the non-existing "sloppy"

21. We have not dealt with Fodor's theory of "sloppy" identity in any detail. This is because Fodor makes the incorrect assumption that the phenomenon is restricted to genitive pronouns. The range of examples considered in this section makes clear why that assumption is incorrect.

22. In Section 3.5 of Chapter Three, however, we will offer an hypothesis, motivated on independent grounds, that accounts for why some speakers find such readings possible.

23. The problem is actually somewhat more complicated. In particular, the constraint on "backwards" variable binding cannot be stated merely in terms of preceding anaphoric proform. This is because, as Ross (frequently delivered unpublished lecture on "Primacy") points out, sentences like (*i*) have a reading paraphrasable as (*ii*), but not one paraphrasable by (*iii*) (underlining means coreference).

(*i*) <u>Edith</u> said that finding <u>her</u> husband nude had upset <u>her</u>, and Martha did ∅, too.

(*ii*) ...Martha said that finding Edith's husband nude had upset Martha.

(*iii*) ...*Martha said that finding Martha's husband nude had upset Edith.

A proper version of this constraint must therefore involve a notion something like Ross's notion of primacy. It is important to note that similar facts can be observed for cases of primacy-violating quantifier-bound variable linkages:

(*iv*) The woman who saw everyone talked to him.

(*iv*) cannot mean (*v*).

(*v*) $(\forall x)$(the woman who saw x talked to x)

This provides further support for our claim that the non-existing "sloppy" predicates are ruled out by the same constraint that rules out cases of illegitimate variable binding with ordinary quantifier words.

24. There are further arguments in favor of Chomsky's analysis, actually. (*i*), for instance, has an optional variant with overt wh-word in all dialects.

(*i*) John is the same as he was yesterday.

(*i*)' John is the same as what he was yesterday.

Moreover, sentences like (*i*) seem to have no source in a CD analysis, i.e. they lack a deletion trigger.

Further problematic cases for Bresnan's analysis of CD are discussed in Section 3.3 of Chapter Three.

25. We use notation [ER]°...Δ° as a shorthand to relate the [ER] operator to the positions it binds. Other notations are of course possible. Δ° should be thought of as a bound variable.

26. Other cases of VPD involving polarity items are not discussed here in any detail. For whatever it's worth, however, notice that the sources for what are commonly taken to be the problematic cases in this area are in general not so bad:

(*i*) John didn't budge an inch but Bill did ∅.

(*ii*) John didn't budge an inch, but Bill did budge an inch.

Whatever problems actually lurk in this domain are no more problematic for our theory than for any other theory that I am aware of.

27. In this section we will employ restricted quantification, instead of representing sentences like (2.3.49) with conditionals as we chose to earlier. In (2.3.50), (∀y: Sam, λz(z want [Betsy, λw (w read y)])) is a restricted

universal quantifier to be thought of roughly as in (*i*).

(*i*) (for all y, y such that Sam wanted Betsy to read y)

28. It might be objected that our semantic rules will not assign sentences like (*i*) and (*ii*) an appropriate logical form (this objection was pointed out to me by Larry Horn).

(*i*) John can eat everything.

(*ii*) John can eat most things.

The problem is that <u>can</u> must be within the scope of the quantifier in these examples. (*ii*), for instance, can be paraphrased roughly as "most things are such that John can eat them". We would therefore not want to represent (*ii*) as, say, (*iii*).

(*iii*) CAN [John, λx((Most y: things(y))[x eat y])]

An acceptable alternative within our framework would be to assign (*ii*) a logical form like the one shown in (*iv*).

(*iv*) John, λx((Most y: things(y))[CAN [x, λz(z eat y)]])

We now have a peculiar case of a VP (<u>eat most things</u>) which corresponds to a λ-predicate which contains material which corresponds to a syntactic entity outside of that VP. λx(...), which corresponds to [$_{VP}$eat most things], contains CAN, which corresponds to the syntactic entity <u>can</u> which is outside of that VP.

Oddly enough, this peculiar state of affairs makes some empirical predictions. It should be the case then that the only time [$_{VP}$eat most things] in an example like (*ii*) can be deleted, is when the trigger VP is also preceded by <u>can</u>. Only in such a situation will the necessary alphabetic variance obtain between the relevant λ-expressions.

Unless I'm mistaken, this prediction seems to be correct. (*v*), of course, is a legitimate instance of VPD.

(*v*) Bill can eat most things, and John can ∅, too.

[∅ = eat most things]

(*vi*), on the other hand, has no reading which renders it synonymous with (*vii*).

(*vi*) John can eat most things, but when he does ∅, he gets sick.

(*vii*) John can eat most things, but when he eats most things, he gets sick.

Although I have the feeling there are some real problems lurking in this domain of facts, I can't seem to find them.

29. The opaque reading for the source of (2.3.57) will, however, be represented as in (*i*).

(*i*) Bill, λx(x want [(∀y: Sam, λw(w want [Betsy, λs(s read y)]))
[Betsy, λt(t read y)]]) & Peter, λm(m want [(∀z: Sam, λr(r want
[Betsy, λp(p read z)])) [Betsy, λo(o read z)]])

Here the higher VP corresponds to λm(...), which is an alphabetic variant of λx(...), and the embedded VP(which is the one deleted in (2.3.57)) corresponds to λo(...), which is an alphabetic variant of λp(...). On the opaque reading then, deletion of either VP is permitted.

30. Notice that comparatives work the same way. Consider the following sentence, for instance.

(*i*) John claimed he was taller than he is, and Mary did too.

(*i*) allows either a transparent (= sensible) or an opaque (= contradictory or, to use the technical term of Postal (1974b), "stupid") reading.

(*ii*), on the other hand, allows only an opaque reading where both John and Mary claimed a contradiction.

(*ii*) John claimed he was taller than he is, and Mary claimed he was, also.

The fact that deletion of only the embedded VP is impossible on the transparent reading follows from our theory, given the assumption that comparatives are represented scopally at the level of logical form. The logical form for the transparent reading is roughly as in (*iii*).

(*iii*) John, $\lambda x([ER]^{\circ}$ ([x, λy(y claimed [John, λz(z be Δ° tall)])], [John, λw(w be Δ° tall)])) & Mary, $\lambda q([ER]^{1}$ ([q, r(r claimed [John, λs(s be Δ^{1} tall)])], [John, λt(t be Δ^{1} tall)]))

The higher VP, which was deleted in (*i*), corresponds to λq(...), which is an alphabetic variant of λx(...) in (*iii*). The embedded VP, however, which was deleted in (*ii*), corresponds to λs(...), which, because it contains the variable Δ^{1} bound from outside, is not an alphabetic variant of λz(...), which contains Δ° bound from out_ide.

The notation here becomes somewhat baroque. I hope that some future investigator will improve upon it.

31. Note that the failure of (2.3.80) to be derivable from (2.3.77) provides strong evidence against explaining some of the facts we observed in Chapters One and Two in a syntactic identity theory by weakening the notion of syntactic identity to allow a VP containing a quantifier to count as non-distinct from a VP (otherwise identical) containing a pronoun anaphoric to that pronoun. That is an alternative we failed to mention in Chapter One as possible dodge which might enable a syntactic identity theory of VPD to derive (*i*) from (*ii*) (on one reading).

(*i*) Sandy greeted everyone when Betsy did ∅.

(*ii*) Sandy greeted everyone$_i$ when Betsy greeted $\{\begin{smallmatrix}\text{them}_i \\ \%\text{him}_i\end{smallmatrix}\}$.

32. Notice that with respect to examples like this, the remarks just made about cardinality of sets are particularly apt. Thus _John read the books that Bill did_ is unquestionably true in a situation where Bill read only one book (and John read it too).

If the discrepancy between this example and the previous one (_the books we ordered arrived_) is very striking to the reader (it's not to me), then perhaps we should adopt Chomsky's proposal for some of the cases, and the proposal we have just given for others (those where deletion is possible). The data do not seem to be sufficiently clear to decide this matter.

CHAPTER THREE

Gapping

3.1 Syntactic Overview

3.1.0 Introduction

This chapter will deal with the well-known elliptical phenomenon illustrated by the sentences in (3.1.1).

(3.1.1)(a) Sandy played shortstop, and Betsy ∅ first base.
[∅ = played]

(b) Alan ran to second base, and Betsy ∅ to first base.
[∅ = ran]

(c) Betsy plays first base for the first four innings, and Peter ∅ for the rest of the game.
[∅ = plays first base]

(d) Sandy wanted to begin to write a novel, and Betsy ∅ a short story. [∅ = wanted to begin to write]

Harris (1957, 1964) was, I believe, the first to propose a deletion transformation to account for such sentences. His discussion of the phenomenon, however, is not comprehensive. The first detailed examination of this transformation that I am aware of is by Ross (1967b). We will follow Ross's terminology, which has become standard, and refer to this rule as Gapping.

An overview of this phenomenon is somewhat difficult to present because of certain cases which have been subsumed under Gapping by some, and not by others. Proposals have been made which collapse Gapping with virtually every other elliptical process in English (the interpretive theory of Fiengo (1974)), while others regard only cases of sentence-internal ellipsis of (strings of) verbs as Gapping (Stillings (1975)). In the next section we will examine the various proposals that have been made in detail. Here we will only sketch salient syntactic properties. Our exposition owes much to the excellent discussion in Lust (ms.).

3.1.1 Where to Gap:

Gapping is restricted to coordinate structures:

(3.1.2) *Sandy played the guitar while (after, before, since, although...)
Betsy ∅ the recorder.

Furthermore, within coordinate structures, only the highest S can undergo Gapping.

(3.1.3) *Alan went to New York, and
(a) I know (that)
(b) it seems (that)
(c) Bill met a man who claimed (that)
Betsy ∅ to Boston.

The range of coordinating conjunctions is restricted. Almost all speakers accept Gapped sentences with <u>and</u>, <u>or</u> and <u>nor</u>, but most people seem to find analogous sentences with <u>but</u> to be unacceptable:

(3.1.4)(a) Peter hit a double, and Betsy ∅ a triple.
(b) Either Peter hit a double, or Betsy ∅ a triple.

(c) John didn't go to the laundromat, nor Betsy ø to the grocery store.

(d) (%)*Joan likes Richard, but Betsy ø Peter.

Under certain circumstances, Gapping is possible with no overt conjunction:

(3.1.4)(e) Peter plays first base; Betsy ø second base.

There is another reduction process which applies in comparative clauses to produce maimed clauses similar to those derived by Gapping:

(3.1.5) Betsy likes swimming more than Sandy Ø softball.

It seems preferable to keep this process distinct from Gapping (see Jackendoff (1971) for more discussion and also the discussion in Section 2.3 of the preceding chapter)

Backwards Gapping does not occur in English:

(3.1.6)(a) *Peter ø first base, and Alan played left field.

(b) *Either Norma ø Beth, or Lois saw Norma.

In some languages, backwards ellipsis of this type seems to be allowed. Ross (1967b) argued that in languages where both forward ellipsis (=Gapping) and backwards ellipsis occurs, both ellipses should be effected by a single rule (i.e. a bi-directional Gapping rule). Arguments against this position have been given by Hankamer (1971, 1972a) and Maling (1972).

It's an interesting question whether or not Gapping applies in discourse. In Sag and Hankamer (1976), examples like the following were noted, and the conclusion was drawn that Gapping can indeed apply in discourse, at least sometimes.

(3.1.7) Speaker A: Jorge is peeling an apple.

Speaker B: And Ivan ∅ an orange.

It's not at all clear, however, what to make of examples like this. The discourse in (3.1.7) seems to some people to be a peculiar case of two people collaborating on what is actually a single sentence, in which case Gapping should perhaps be restricted to single sentences. Alternatively, one might argue that Gapping is a rule of discourse grammar, but that since Gapped clauses must begin with conjunctions, it is only in peculiar situations like (3.1.7) that Gapping can apply inter-sententially. We will return to this point in the next chapter.

Another interesting property, which we will have quite a bit to say about in the last section of this chapter, is that elements left behind in a Gapped clause are intonationally marked. Typically, each Gapping remnant carries a pitch accent (Bolinger (1958, 1965)) of its own (separated by a pause) though the relative pitch prominence in a Gapped clause, it seems to me, is not fixed.

(3.1.8)(a) Alan played poker, and Betsy ∅ canasta.

(b) (Did Gwendolyn play poker and Alan canasta?)

No, Alan played poker, and Gwendolyn ∅ canasta.

Gapping remnants must also, in some poorly understood sense, be parallel to corresponding elements in the left conjunct. Thus we do not have:

(3.1.9)(a) *Sam hates reptiles, and Sandy ∅ to talk to Oh.

(b) *Beth ate yogurt, and Norma ∅ at midnight.

The only attempt I am aware of to make this notion precise is due to Hankamer (1971), p. 25) who says "the conjoined sentences must be structurally identical simplexes" where "structurally identical simplexes" are structurally identical down to highest NP's (and embedded S's).

Gapping can apply recursively in structures containing more than two conjuncts:

(3.1.10) Ray plays the clarinet, Lois ∅ the oboe, John ∅ the piano, Sandy ∅ the guitar...

But, as Lust (ms., sec. 11 A-§7) phrases it, Gapping is not <u>linearly recursive</u>, for we do not have:

(3.1.11) *Max wanted Ted to persuade Alex to see Mary, and Sam ∅ Walt ∅ Tanya ∅ Tom.

3.1.2 When to Gap

Gapping can be shown to apply after cyclic rules:

(3.1.12)(a) Betsy was hassled by the police, and Norma ∅ by the FBI. (Passive)

(b) Norma seemed to be eating less, and Beth ∅ to be eating more (Subject-Subject Raising)

(c) Yesterday it was obvious that you were happy, and the day before, ∅ ∅ ∅ that you were quite unhappy. (Extraposition)

(d) Tomorrow there will be a riot, and the day after ∅ ∅ ∅ a holocaust. (There-Insertion)

Moreover, as Perlmutter and Soames have pointed out (1975 Class lectures), certain difficulties could conceivably arise if Gapping were allowed to apply before certain cyclic rules. For instance the following sort of sentence might be generated, presuming a common deep structure source for both conjuncts:

(3.1.13) *John is likely to give up, and it ø that Max will persist.

Conversely, arguments that Gapping must apply <u>before</u> some other syntactic rule are rare. One such argument, due to Jackendoff (1971, p. 23, nt. 2), concerns sentences like these:

(3.1.14) Did Bill eat the peaches, or Harry the grapes?

In this sentence, Jackendoff argues, Gapping must be assumed to apply before Subject Auxiliary Inversion, i.e. to the structure in (3.1.15).

(3.1.15) (Q) - Bill - past - eat - the peaches or

(Q) - Harry - $\underbrace{\text{past - eat}}_{\Downarrow\ \emptyset}$ - the grapes

Unless Gapping applies at this level, Subject-AUX Inversion will apply, moving, in this case, only the tense morpheme in each conjunct leftward over the subject. <u>Do</u>-Support would then apply in each conjunct. The resulting sentence, prior to Gapping would be this:

(3.1.16) Did Bill eat the peaches or did Harry eat the grapes?

Gapping then applies, yielding:

(3.1.17) *Did Bill eat the peaches, or did Harry ø the grapes?

Ordering Gapping before Subject-Auxiliary Inversion solves this problem, for tense is gapped along with the verb in the second conjunct, hence blocking Subject-Auxiliary Inversion and Do-Support there. We will return to this problem in Section 3.4.

3.1.3 What to Gap

It is clear that simple verbs are the most likely deletion targets for Gapping, as we have seen in most of the above examples. Auxiliary verbs, if present, also undergo Gapping:

(3.1.18)(a) Alan should have been upset, and Peter ∅ overjoyed. [∅ = should have been]

(b) Alan has talked to his editor, and Janis ∅ to her graduate advisor. [∅ = has talked]

One perplexing restriction, pointed out by Ross (1967b) is that a negation in the auxiliary cannot Gap:

(3.1.19)(a) I didn't eat fish and Bill didn't eat ice cream

(b) *I didn't eat fish and Bill ∅ ice cream.

However, if the negation inheres in the conjunction (i.e. nor), Gapping is possible.

(3.1.20) I didn't eat fish, nor Bill ∅ ice cream.

This is a curious property of Gapping which remains mysterious (see Jackendoff (1971) and Stillings (1975) for some discussion).

Identical pre-verbal adverbs are also Gappable, as the following sentences show.

(3.1.21)(a) Betsy quickly dropped the frying pan, and Peter ø the soup bowl. [ø = quickly dropped]

(b) Betsy sometimes sleeps with Duke, and Peter ø with his teddy bear. [ø = sometimes sleeps]

However, as Ross and Jackendoff have pointed out, when the pre-verbal adverbs are not identical, Gapping is impossible:

(3.1.22)(a) *Betsy quickly dropped the frying pan, and Peter suddenly ø the soup bowl.

(b) *Betsy sometimes sleeps with Duke, and Peter frequently ø with his teddy bear.

Hankamer (1972a) observes similar restrictions on (forward) Gapping in Turkish. He seems to be suggesting that a "like-adverb restriction" might be a general condition on Gapping (universally). Alternatively, one might expect a proper formulation of the rule to predict these facts.

A standard presumption about remnants in Gapped clauses is that there can only be two of them. Some standard examples are...

(3.1.22)(a) *Alan gave Sandy a book, and Peter ø Betsy a magazine.

(b) *Alan told Harry that the sky was falling, and Sam ø Betsy that Chicken Little was right.

(c) *Arizona elected Goldwater Senator, and Massachusetts ø McCormack Congressman.

No mention has been made in the literature of cases like the following (which virtually all speakers find acceptable) where the Gapped clause contains three remnants (NP - PP - PP).

(3.1.23)(a) Peter talked to his boss on Tuesday, and Betsy ∅ to her supervisor on Wednesday.

(b) John talked to his supervisor about this thesis, and Erich ∅ to the dean about departmental policies.

Judgements concerning other sentences containing Gapped clauses with three remnants seem to vary considerably from speaker to speaker. Jackendoff (1971, p. 26) finds a contrast in acceptability depending on whether or not the third remnant is a PP that is strictly subcategorized by the (Gapped) verb. He cites these examples.

(3.1.24)(a) ?*Willy put the flowers in a vase, and Charlie on the table.

(b) ??Charlie entered the bedroom at 5:30, and Vera ∅ the kitchen at 6:00.

Other speakers find this contrast even more striking than Jackendoff does, accepting (3.1.24)(b), but totally rejecting (3.1.24)(a).

In the most common cases, where only the verb has been Gapped and the first remnant is the subject of the second conjunct, the second remnant may be an NP, an Adjective (Phrase), an Adverb (Phrase), a <u>that</u>-clause, a <u>for-to</u> clause, or a Prepositional Phrase:

(3.1.25)(a) Peter loves Betsy, and Betsy ∅ Peter. (NP)

(b) Alan seemed happy, and Sandy ∅ sad. (AP)

(c) Tom ran slowly, and Alex ∅ quickly. (Adv. Ph.)

(d) Alan claimed that he was cheated, and Sandy ∅ that she was the one who cheated him. (<u>that</u>-S)

(e) Alan prefers for Tom to do it, and Sandy ∅ for Alan to do it. (<u>for-to</u> clause)

(f) Betsy stood in left field, and Sandy ∅ in right field. (PP)

As Ross (1976) notes, auxiliaries by themselves are not very good Gapping targets. Put somewhat differently, VP is an unacceptable second remnant in a Gapped clause. Ross cites these examples, inter alia (the judgements are his):

(3.1.26)(a) ?*He may stay inside, and she ∅ go to the beach.

(b) ?*He has taken the Star of Pittsburgh, and she ∅ stolen the Moon of Altoona.

(c) ??He was squeezing a tennis ball, and she ∅ greasing a shoe.

(d) ?He was driven to Aix, and she ∅ taken to Ghent.

Not all speakers find these facts quite as gradient as Ross finds them to be. Nevertheless most speakers find all the examples in (3.1.26) less acceptable that those in (3.1.25) where something other than VP is the second remnant in the Gapped clause.

Notice also, that if the second Gapping remnant is an S (or $\overline{S}$ in the sense of Bresnan (1970, 1972)), then a complementizer must be present, even if that complementizer is otherwise deletable:

(3.1.27)(a) Sandy said (that) he was a fool, and Betsy ∅ that he was out of his mind.

(b) *Sandy said (that) he was a fool, and Betsy ∅ he was out of his mind.

(3.1.28)(a) Sandy preferred for Alan to do it, and Betsy ∅ for Peter to do it.

(b) *Sandy preferred for Alan to do it, and Betsy ∅ Peter to do it.

These facts might be attributed to a general principle blocking clauses without complementizers in certain positions (as was suggested by Lasnik; see Postal (1974a, p. 128, nt. 35; p. 131, nt. 39) and Lust (ms.) for some discussion). Alternatively, one might expect an appropriate formulation of the rule of Gapping itself, perhaps interacting with some other general principle (the A-over-A principle comes to mind), to account for these facts. We will return to this problem.

There are many other constraints on what elements can be Gapped that have been pointed out in the literature. Jackendoff observed that the following sentences are unacceptable, where an NP immediately following the verb has also been Gapped.

(3.1.29) *Mary gave a nickel to Sally, and Maxine ∅ to John.
[∅ = gave a nickel]

Hankamer (1973a) subsequently pointed out that in similar sentences, an NP that does not immediately follow the verb cannot be Gapped either:

(3.1.30) *Max gave Sally a nickel, and Harvey $\emptyset_1$ Susan $\emptyset_2$.
[$\emptyset_1$ = gave; ∅ = a nickel]

However, if no NP is being Gapped, there are sometimes multiple output possibilities. The following illustration of that is due to Ross (1969 Linguistic Institute lecture and mimeo).

(3.1.31) John tried to begin to write a play, and Harry

(a) ø to begin to write a novel.

(b) ø to write a novel.

(c) ø a novel.

Precisely these and no other collection of remnants are possible:

(3.1.32) *John tried to begin to write a play, and Harry

(a) ø begin to write a novel. [ø = tried to]

(b) ø write a novel. [ø = tried to begin to]

(c) ø novel. [ø = tried to begin to write a]

(3.1.32)(c) illustrates a more general property of Gapping of course (parts of NP's cannot Gap by themselves), which is also illustrated by sentences like the following:

(3.1.33) *Bill read his father's exciting new book, and Alan

(a) ø mother's boring old article. [ø = read his]

(b) ø boring old article. [ø = read his father's]

(c) ø old article. [ø = read his father's exciting]

(d) ø article. (ø = read his father's exciting new]

The only attempt to be precise about specifying what a well-formed Gapping remnant (that I am aware of) is due to Hankamer (1973a), who states that they must be "major constituents", which he defines this way: "A 'major constituent' of a given sentence S_0 is a constituent either immediately dominated by S_0, or immediately dominated by VP which is immediately dominated be S_0" (Hankamer (1973a, p. 18, nt. 2))

Such a condition, as it stands, is neither a sufficient nor even a necessary condition for Gapping remnants. Hankamer offers a proposal, as we will see, to account for certain unacceptable instances of Gapping where the Gapped-clause remnants do staisfy his definition of major constituent. It should be noted, however, that in most syntactic frameworks (certainly in Hankamer's), Ross's examples in (3.1.31)(c) contains a remnant which is not "immediately dominated by a VP which is immediately dominated by S_0", for an S-node intervenes. One can imagine various ways of restating Hankamer's condition. We will see, however, that there is no proposal for Gapping to date that accounts for all the facts that have previously been observed. We will present further facts that are also problematic for existing theories of Gapping.

To pursue those matters now, however, would.result only in a series of unrelated facts. Let us therefore now turn to the various proposals that have been made _re_ the formulation of the rule, and examine these further facts as they become relevant.

3.2 Previous Proposals and Problems

3.2.1 Ross (1967b)

As mentioned earlier, Ross was the first to discuss Gapping in any detail. He gave no explicit formulation of the rule. However, he clearly had in mind a rule roughly as sketched in (3.2.1) (this formulation of what Ross had in mind is due to Hankamer (1973a)).

(3.2.1) Gapping (Ross)

NP_1 - X - A - Y - _and_ - NP_2 - X - B - Y $\Rightarrow$

NP_1 - X - A - Y - _and_ - NP_2 - B

where A and B are non-identical major constituents (see above)

This rule has a couple of noteworthy properties. First, it deletes non-constituents, i.e. a variable. This was, as far as I know, the first proposal that rules could delete non-constituents. Other deletion rules, e.g. Equi-NP-Deletion, Super-Equi-NP-Deletion, Agent Deletion, VP-Deletion, have usually been assumed to delete only constituents. Deletion of a variable is certainly a coherent notion, however. Furthermore, rules that delete a variable are statable even in a very restricted theory of transformations (say even in a version of transformational grammar that does not allow arbitrary Boolian conditions on analyzability). Let us adopt the convenient terminology of Hankamer (1971, 1973a) and refer to rules that delete a variable as ellipsis rules. Ross's formulation of Gapping is then an ellipsis rule.

A second noteworthy property of Ross's rule is that it effects only clause-internal and right-peripheral ellipsis. In Ross's analysis then, both deletions in the following sentence would be the result of a single application of Gapping.

(3.2.2) Betsy talked to Peter on Sunday, and Alan $\emptyset_1$ to Sandy $\emptyset_2$.

[$\emptyset_1$ = talked; $\emptyset_2$ = on Sunday]

Left-peripheral ellipsis, such as that in the next two sentences, however, would not be the result of Gapping in Ross's proposal, but, presumably, some other rule (perhaps, at least in the first case, Conjunction-Reduction --see Stockwell, et al. (1972), Sjoblom (1975a, 1975b)).

(3.2.3)(a) John talked to Bill and $\emptyset$ seemed upset.

[$\emptyset$ = he = John]

(b) John persuaded Sandy to stand up, and $\emptyset$ Peter to sit down.

[$\emptyset$ = John persuaded]

None of the problems noted in the previous section are discussed by Ross.

3.2.2 Hankamer (1971, 1973a)

Hankamer's proposal differs from Ross's first in that left peripheral deletion, such as those in (3.2.3) are included within the domain of Gapping. Hankamer conceives of Gapping as a rule that deletes variable strings freely, leaving just two remnants in the right conjunct. His rule looks like this, roughly:

(3.2.4) Gapping (Hankamer)

$$X - A_1 - Y - B_1 - Z - \underline{\text{and}} - X - A_2 - Y - B_2 - Z \Rightarrow$$
$$X - A_1 - Y - B_1 - Z - \underline{\text{and}} - A_2 - B_2$$

Hankamer provides no measurements for collapsing the left peripheral ellipsis with Gapping. He simply remarks that "there seems to be no reason to assume that a rule other than Gapping is responsible" (1973a, p. 27).

Hankamer's formulation thus amounts to an empirical hypothesis that the left peripheral ellipsis in sentences like those in (3.2.3) has the same properties as sentence internal and right peripheral ellipsis, i.e., Gapping. There is a certain amount of evidence, however, that such left-peripheral ellipsis has strikingly different properties.

For instance, the Gapping deletion target cannot be a sequence-like verb-preposition or (auxiliary-)adjective-preposition if the preposition is the head of a prepositional phrase:

(3.2.5)(a) *John spoke to Harry, and Bill ø Mike.[1]

[ø = spoke to]

(b) *John was happy with his girlfriend, and Betsy ∅ her boyfriend. [∅ = was happy with]
(cf...and Betsy ∅ with her boyfriend)

Notice that quite independently of the particular mechanism one employs, some constraint must be stated that prevents term B_2 in Hankamer's rule from analyzing a NP when that NP occurs in the configuration: $[_{PP}P\_\_\_\_]$. A similar constraint holds for term A_2 for cases of Gapping, as the following contrast shows:

(3.2.6)(a) At our house, we play poker, and at Betsy's house ∅ bridge. [A_2 = PP, ∅ = we play]
(b) *At our house, we play poker, and $\emptyset_1$ Betsy's house $\emptyset_2$ bridge. [$\emptyset_1$ = At; A_2 = NP, $\emptyset_2$ = we play]

Whether some modification of Hankamer's notion of major constituent is sufficient to account for these facts, or whether some independent constraint (say, the A-over-A principle) is involved, is not of concern here. The point is simply that for the Gapping cases some such constraint on both terms (A_2 and B_2) is in effect.

Under the assumption that left-peripheral deletions result from the same rule (namely (3.2.4)), then we would expect the same restriction to be in effect when that rule deletes only identical left-peripheral material, i.e. where Y and Z are null. This is simply not the case, as the following examples show:

(3.2.7)(a) My mother met with the principal on Thursday, and ∅ the dean on Friday. [∅ = my mother met with]

(b) My sister spoke to Mrs. Wimble on Friday and ø the dean on Saturday. [ø = my sister spoke to]

Here we have left-peripheral ellipsis that includes a preposition. In terms of Hankamer's rule, term A_2 has successfully analyzed an NP in the configuration $[_{PP}P\_\_\_\_\_]$; precisely what we observed is impossible in cases of Gapping. It is difficult to see how Hankamer's collapsed rule can be constrained so as to take these facts into account.

A second argument against collapsing these ellipses into one rule is rather simplistic. As noted earlier, most people reject Gapped sentences when the conjunction involved is <u>but</u>.

(3.2.8) %*Sandy ate the bagels, but Betsy ø the creamcheese.

However all speakers accept sentences involving left-peripheral ellipsis with <u>but</u>:

(3.2.9)(a) Betsy gave the bagels to Tommy, but ø the creamcheese to Mike. [ø = Betsy gave]

(b) I told my brother to wait there, but ø my sister to go on ahead. [ø = I told]

Since it seems that all rules that effect ellipsis in coordinate structures must mention the appropriate conjunctions, these facts would present no particular difficulty if Gapping and the process that effects left-peripheral ellipsis are two distinct rules. Under Hankamer's theory, however, these facts appear indescribable (unless, of course, there is some extra-grammatical explanation available for the relevant contrasts, a hypothesis which I have been unable to give substance to, but which I realize is at least conceivable).

As a third argument, consider the following ungrammatical instances of Gapping.

(3.2.10)(a) *That Alan was late annoyed Betsy, and that Sandy was late ∅ Bernie. [∅ = annoyed]

(b) *For us to appoint Alan would infuriate Betsy, and for us to appoint Sandy ∅ Bernie. [∅ = would infuriate]

These examples illustrate a constraint on Gapping that has escaped notice in the literature, namely...

(3.2.11) Gapping cannot apply after a sentential subject.

This constraint is easily accomodated by modifying the structural description of the rule. In Hankamer's theory, this amounts to saying that term A_2 cannot analyze an S (or $\overline{S}$).

But observe the consequences of imposing such a restriction on A_2. It is then predicted that in cases of left-peripheral ellipsis, the first remnant in the right conjunct cannot be a _that_-clause or a _for to_ clause either. This prediction is incorrect:

(3.2.12)(a) I proved that the sky is blue last Tuesday, and ∅ that the earth is round last Wednesday. [∅ = I proved]

(b) I prefer for you to get the job when I'm drunk, and ∅ for Betsy to get the job when I'm sober. [∅ = I prefer]

In order to generate sentences like these, term A_2 of Hankamer's rule must be permitted to analyze S (or $\overline{S}$). But again, this is precisely what A_2 must be prevented from doing in cases of Gapping.

A fourth difference between these two processes, it seems to me, can be heard in their intonations. As we noted earlier, remnants in Gapped clauses must in general be separated by a pause. Most speakers however, do not require such pauses in cases of left-peripheral ellipsis.(This is left to the reader to verify.) It's of course possible that this has some independent explanation (say within the intonational system), but I take this prosodic disparity to be highly suggestive that two distinct processes are at work here.

A fifth curious difference between left peripheral ellipsis and Gapping has to do with the conjunction as well as (not to be confused with comparatives of the adjective (or adverb) well). Sentences containing this much-ignored conjunction are discussed by Fiengo (1974), who points out that sentences like (3.2.13) are ungrammatical.

(3.2.13) *Tom is happy as well as Dick is sad.

It seems that as well as cannot conjoin sentences, but only sub-constituents:

(3.2.14)(a) I walk as well as talk.
(b) Betsy is tired as well as hungry.
(c) He spoke to Alan as well as to Bill.

Notice however, that embedded sentences with complementizers can be conjoined with as well as, but the same sentences without complementizers are impossible:

(3.2.15)(a) I know that John is a fool as well as that Bill is a fool.
(b) *I know John is a fool as well as Bill is a fool.

Presumably then, as well as can conjoin two $\bar{S}$'s, but not two S's.

From this it follows, nothing more being said, that Gapping is impossible with as well as. This seems to be correct.[2]

(3.2.16)(a) ??John likes Susan, as well as Peter ø Melinda.

(b) *Peter is happy as well as Betsy ø sad.

Left-peripheral ellipsis however, is possible with as well as:

(3.2.17)(a) I spoke to Betsy on Friday as well as ø to Sandy on Saturday. [ø = I spoke]

(b) I read a limerick to my sister as well as ø a fairy tale to my brother. [ø = I read]

This difference between left-peripheral ellipsis and Gapping is very telling. The fact that such sentences as (3.2.17)(a) and (b) are grammatical suggests that left-peripheral ellipsis is not a sentence-ellipsis process at all, for these sentences have no source involving sentential conjunction:

(3.2.18)(a) *I spoke to Betsy on Friday as well as I spoke to Sandy on Saturday.

(b) *I read a limerick to my sister as well as I read a fairy tale to my brother.

It is quite possible then that cases like these are to be treated as base-conjoined VP's, with subsequent application of Conjunction Reduction. The derivation of, say, (3.2.17)(a), would proceed roughly as follows.

(3.2.19)(a) I spoke to Betsy on Friday as well as spoke to Sandy on Saturday. (base-conjoined VP's) $\Rightarrow$ (3.2.17)(a) (Conjunction Reduction).

Our hypothesis then is that the restrictions on as well as are statable at the level of deep structure. This predicts that cases of derived VP conjunction will be impossible with as well as, for even in a system that assumes free conjunction in the base (cf. Dougherty (1970, 1971)), sentences like the following are taken to be derived via sentential conjunction.

(3.2.20) ?*John seems to Harry to be sad as well as to Mike to be happy.

The prediction would seem to be correct, but other sentences, such as this next one, are not as bad as they should be.

(3.2.21) John was hassled by the police on Friday as well as by the FBI on Saturday.

I confess to having no comprehensive account of these facts. However the hypothesis that all cases of left-peripheral ellipsis arise from Conjunction Reduction (with the standard notions of rebracketting and relabeling - see Stockwell et al. (1972) for discussion), has much to recommend it, for it would account for most of the discrepancies we have observed between left-peripheral ellipsis and Gapping. That is, that hypothesis would lead us to expect that the range of conjunctions was different, that further reductions would be possible (because of rebracketting and subsequent reapplication of Conjunction Reduction), that pauses are not obligatory (also because of rebracketting), and that no restrictions vis à vis that-clauses and for-to clauses are in force.[3]

It should be noted that Hankamer does offer an argument for treating left-peripheral deletions, when the appropriate pauses occur, as the result of Gapping. That argument rests crucially on the ungrammaticality of such sentences as this (with comma pause).

(3.2.22) You have raised a question which is an important one, and will continue to be asked until an answer is found.

Hankamer regards such sentences as "queer" and invokes the Coordinate Structure Constraint as the reason for their queerness.

Hankamer's judgement is highly idiosyncratic. Moreover, if such sentences were generally judged to be unacceptable, one would be hard-pressed to explain why, for whatever process produces "across the board" extractions like the one in this next sentence would also produce (3.2.22), with the pause, even if no (Subject) Gapping had taken place:

(3.2.23) You have raised a question which he loves to talk about, but he can never find an answer for.

Furthermore, compare whatever deviance might inhere for some speakers in these last two sentences with the ungrammaticality of the following sentence, where extraction has been attempted out of a coordinate structure wherein Gapping has applied.

(3.2.24) *You have brought up the matter of the Hindenburg, which is an important question, and the question of the lost continent ø a stupid question. [ø = is]

Let us now turn to Hankamer's proposed constraint on Gapping, bearing in mind that his proposal is couched in terms that presuppose the correctness of including left-peripheral ellipsis within the domain of Gapping.

Hankamer examines certain sentences which could arise, given his generalized formulation of Gapping, in two distinct ways. One such case is the following sentence, which, Hankamer remarks, can be interpreted only as (i.e. derived only from) (a), not (b).

(3.2.25) Bill expects Harry to find the way to the party, and Sue to find the way home.

(a) ---and [Bill expects] Sue to find the way home.

(b) ---and Sue *[expects Harry] to find the way home.

Related examples, all due to Hankamer, are all of the following, which, he claims, are acceptable only if interpreted as indicated.

(3.2.26) Jack calls Joe Mike, and Sam Harry.

(a) ---and [Jack calls] Sam Harry.

(b) ---and Sam *[calls Joe] Harry.

(3.2.27) Max wanted Ted to persuade Alex to get lost and Walt Ira.

(a) ---and [Max wanted] Walt [to persuade] Ira [to get lost].

(b) ---and Walt *[wanted Ted to persuade] Ira [to get lost].

(c) ---and Walt *[wanted] Ira [to persuade Alex to get lost].

Hankamer proposes the following "No-Ambiguity Condition" to account for these facts (1973a, p. 29).

(3.2.28) The No-Ambiguity Condition (NAC)

Any application of Gapping which would yield an output structure identical to a structure derivable by Gapping from another source, but with the "gap" at the left extremity, is disallowed.

The kind of "ambiguity" Hankamer has in mind is purely structural. Thus sometimes a particular Gapping derivation must be blocked even if the blocking derivation, i.e. the one with a left-peripheral gap, produces an ungrammatical sentence. For instance this next example, as Hankamer notes, is blocked in the (b) interpretation, even though the source for the (a) interpretation is ungrammatical, as indicated.

(3.2.29) Jack asked Mike to wash himself, and Sue to shave himself.

(a)*---and [Jack asked] Sue to shave himself.

(b) ---and Sue *[asked Mike] to shave himself.

Several obvious conterexamples to the NAC are in fact discussed by Hankamer. Jackendoff's sentences, for instance, like the following,

(3.2.30) Massachusetts elected McCormack Congressman, and Pennsylvania, Schweicker.

(a)*---and [Massachusetts elected] Pennsylvania, Schweicker.

(b) ---and Pennsylvania [elected] Schweicker [Congressman].

should be unacceptable on the (b) reading by the NAC if titular NP's like congressman are indeed NP's (i.e. if the constituent structure of elected McCormack Congressman is V NP NP). The argument is then made, not unconvincingly, that such titular NP's are in fact not NP's at all (and hence do not interact with the NAC as real NP's do). We will not reproduce those arguments here.

Similarly, sentences like this one:

(3.2.31) Paul Schachter has informed me that the basis order in Tagalog and related languages if VOS; Ives Goddard that the unmarked order in Algonkian is OVS; and Guy Cardin that the basic order in Aleut is OSV. (Ross (1967b)).

are potential counterexamples to the NAC. Hankamer argues that here the pronoun me has been cliticized to the preceding verb, and has hence lost its

NP status. To support this, he compares a similar sentence with a full NP instead of the (cliticized) pronoun me, where the NAC in fact seems to make the correct prediction that only the (a) reading is possible:

(3.2.32) Paul Schachter has informed Haj Ross that the basic order in Tagalog and realted languages is VOS; Ives Goddard that the unmarked order in Algonkian is OVS; and Guy Cardin that the basic order in Aleut is OSV.

(a) --- [Paul Schachter has informed] Guy Carden that the basic order in Aleut is OSV.

(b) ---Guy Cardin *[has informed Haj Ross] that the basic order in Aleut is OSV.[4]

Hankamer has no explanation for the acceptability of sentences like this next one, noted by Jackendoff, where the NAC is clearly violated.

(3.2.33) Max writes plays in the bedroom, and Harvey in the basement.

(a) *---and [Max writes] Harvey in the basement.

(b) --- and Harvey [writes plays] in the basement.

Hankamer in fact finds this sentence questionable. Here too, his judgement is highly idiosyncratic. Leaving aside, for a moment, the question of the empirical adequacy of the NAC, it is important to notice that Hankamer views it as a "transderivational constraint" and as "an explicit formulation of what has been called a 'perceptual strategy'" (1973a, p. 36, nt. 12). In fact, the NAC is viewed as a particular case of a more general, universal constraint which he terms "The Structural Recoverability Hypothesis":

(3.2.34) The Structural Recoverability Hypothesis

> Deletion rules involving variables are universally subject to a transderivational condition which prevents them from applying in such a way as to introduce structural ambiguity. (Hankamer, 1973a, p. 40)

At this point, several comments are in order. First, many of the crucial predictions made by the NAC are sentences which informants do not agree about. In particular, many informants accept certain sentences which are ruled out by the NAC. It is of course possible that this is simply an instance of people not having direct access via introspection to their internalized grammars, or that certain principles extraneous to grammar proper render certain ungrammatical sentences acceptable. It seems to me, however, that even if <u>one</u> of these instances of informant variation represents true dialect variation, then a "universal transderivational constraint" is hardly an appropriate account of the facts in the dialects that reject the sentences in question.

It should further be noted that Hankamer's factual claims are in many cases extremely controversial. Quirk et al. (1972), for instance cite the following examples as fully acceptable (as indeed they seem to be).

(3.2.35)(a) Joan will cook the meals today, and Barbara ø tomorrow. [ø = will cook the meals]

(b) Peter is playing football for his school and Paul ø for his club. [ø = is playing football]

Both these examples are glaring counterexamples to the NAC as Hankamer has formulated it.

The following passage is even more striking.

> Notice that in certain contexts there can be ambiguity as to whether the subject and verb are ellipted or the verb and object are ellipted. For example, the sentence
>
> Bob will interview some candidates this morning and Peter this afternoon.
>
> can be interpreted as having either of these two kinds of ellipsis:
>
> Bob will interview some candidates this morning and (he will interview) Peter this afternoon.
>
> Bob will interview some candidates this morning and Peter (will interview some candidates) this afternoon. (Quirk et al., 1972, p. 580)

All my informants are in agreement with the judgements given by Quirk et al.

Secondly, the facts in English, especially those like (3.2.29) above, show that semantic and/or morphological factors do not interact with the NAC. That is, in order for the NAC to account for many of the facts Hankamer addresses himself to, morphologically and semantically (selectionally) induced violations must be disregarded. This taken together with the supposed universality of the NAC would entail that, even in languages with highly intricate morphologies, where, for example, case relations are morphologically distinguished sharply, one would expect that NAC violations would still produce unacceptable sentences.

Two such languages have in fact been examined recently with respect to precisely the phenomenon in question. In both cases, the same negative conclusion has been reached. Channon (1974, 1975) and Fedorowicz-Bacz (1973) conclude that with respect to Russian and Polish (respectively), violations of the NAC are, with one or two exceptions, entirely grammatical. This throws the universality of the NAC, as formulated, into question, even if it can be shown to be adequate for English.

In summary, Hankamer's theory is rather dubious, though the problems he raises are certainly in need of explanation. In a sense, all the subsequent literature on Gapping has been an attempt to provide alternative explanations for those problems. Let us now turn to some of those attempts.

3.2.3 Langendoen

Langendoen objects to Hankamer's claim that the appropriate constraint on Gapping must be transderivational. His point, which is well-taken, is that in examples like (3.2.29) and (3.2.30) above, the NAC must be brought to bear to reject one of two possible derivations even when the unblocked derivation produces an ungrammatical sentence. Therefore, he concludes, the constraint must be grammatical in nature, not transderivational. He then attempts to motivate an alternative account in terms of the "standard" theory (Chomsky (1965)).

Langendoen points out certain additional data that the NAC is incapable of accounting for. The following are two of Langendoen's examples.

(3.2.36) Max sent Sally the messenger last week, and Susan yesterday.

(a) ---and [Max sent Sally] Susan yesterday.

(b) ---and *[Max sent] Susan [the messenger] yesterday.

(3.2.37) Max wanted Ted to persuade Alex to see Mary, and Walt, Ira.

(a) ---and Walt *[wanted] Ira [to persuade Alex to see Mary]

(b) ---and Walt *[wanted Ted to persuade] Ira [to see Mary]

(c) ---and Walt *[wanted Ted to persuade Alex to see] Ira

(d) ---and *[Max wanted] Walt [to persuade] Ira [to see Mary]

(e) ---and *[Max wanted] Walt [to persuade Alex to see] Ira

(f) ---and [Max wanted Ted to persuade] Walt [to see] Ira

These facts, Langendoen argues, suggest a generalization different from

Hankamer's, namely, what he calls the "non-left-peripheral Noun Phrase Constraint" (NLPNPC). NLPNPC states that no NP can be deleted by Gapping unless it is either left-peripheral or else involved in a continuous left-peripheral gap. NLPNPC is supposed to account for all of Hankamer's data as well as for the facts of (3.2.36) and (3.2.37). Moreover, Langendoen claims NLPNPC is independently motivated because of the ungrammaticality of sentences like the following:

(3.2.38) *Mary sang the Bach cantata, and Sam played $\emptyset$.

Note in passing that Langendoen's facts are not as clear as he claims they are. Thus (3.2.37)(e) is a possible interpretation for many speakers. It is true, however, that non-sentences like (3.2.38) have no explanation in Hankamer's analysis as it stands.

Langendoen then goes on to collapse Gapping with Conjunction Reduction. He proposes the following rule.

(3.2.39) Conjunction Reduction (Langendoen)

$$X_1 - Z_1 - X_2 - Z_2 - X_3 - (\text{and} - X_1 - Z_1' - X_2 - Z_2' - X_3)^*$$

$$1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 - 11$$

opt. $\Rightarrow$

(A) $1 - 2 - 3 - 4 - 5 - 6 - \emptyset - 8 - \emptyset - 10 - \emptyset$

(B) $1 - 2 \theta 6 + 8 - 3 - 4 \theta 6 + 10 - 5 - \emptyset - \emptyset - \emptyset - \emptyset - \emptyset - \emptyset$

respectively

Conditions: (1) 2 and 8; 4 and 10 are the same major category, and $2 \neq 8$; $4 \neq 10$

(2) A is precluded if:

(a) $4\ 5 = \emptyset$ and 2 is not VP, or

(b) 3 or 5 contain NP other than a clitic pronoun adjoined to its V.

This rule performs Gapping, Conjunction Reduction, Rebracketing and "respectively insertion" in one fell swoop (more precisely, two fell swoops). Condition (2)(b) is NLPNPC.

There are two main arguments against Langendoen's proposal. First, since his analysis collapses Gapping with the rule that effects left-peripheral ellipsis, it is subject to all of the same objections raised in the previous section. This analysis, like Hankamer's, fails to account for why left-peripheral ellipsis and Gapping differ in the ways we have noted. Secondly, NLPNPC does not seem to be an appropriate generalization. Thus Langendoen's proposal fails to account for the grammaticality of all the following sentences (many of which we have noted already) where non-left-peripheral NP's have been Gapped.

(3.2.40)(a) Max writes plays in the bedroom, and Harvey $\emptyset$ in the basement. [$\emptyset$ = writes plays] (Jackendoff (1971))

(b) Joan will cook the meals today, and Barbara $\emptyset$ tomorrow. [$\emptyset$ = will cook the meals] (Quirk et al. (1972))

(c) Peter is playing football for his school and Paul $\emptyset$ for his club. [$\emptyset$ = is playing football] (Quirk et al. (1972))

(d) At Betsy's house we play bridge, and at our house $\emptyset$ poker. [$\emptyset$ = we play]

(e) John took Harry to the movies, and Bill $\emptyset_1$ Mike $\emptyset_2$. [$\emptyset_1$ = took; $\emptyset_2$ = to the movies]

In short, although Langendoen's proposal may be metatheoretically preferable to Hankamer's, it shares with Hankamer's analysis the drawback of empirical inadequacy.

3.2.4 Stillings (1975)

Stillings's solution to the problems raised by Hankamer is to restrict the deletion target of the rule of Gapping. The "no-ambiguity" violations, taken together with other examples where deletion of a sequence [V - NP] results in ungrammatical sentences, she argues, leads "very clearly" to the conclusion that "direct object nouns are never permitted to gap along with the verbs that precede them" (p. 260). The following formulation of the rule is offered to account for this "fact".

(3.2.41) Gapping (Stillings)

$$NP \quad V^* \; C \; \begin{Bmatrix} \text{and} \\ \text{or} \end{Bmatrix} \; NP \quad \underset{\substack{\Downarrow \\ \emptyset}}{V^*} \quad C$$

<u>V</u>* in Stillings's rule is a string variable which can be expanded as any string of elements dominated by V. It is intended that this notation account for sentences of the type we saw earlier, where multiple Gapping outputs are possible. Sentences like these:

(3.2.42) John wanted to begin to write a play, and Harry

(a) ∅ to begin to write a novel.

(b) ∅ to write a novel.

(c) ∅ a novel

Stillings does not discuss how V* is able to analyze the infinitive marker <u>to</u>, as it would have to in (3.2.42)(b) and (3.2.42)(c) (though this is surely a minor objection).

<u>C</u> is taken to be a "constituent variable", i.e. a variable which ranges solely over single constituents. One of Stillings's major conclusions, in

fact, is that a proper formulation of Gapping (hers) requires that linguistic theory countenance such constituent variables. The C variables in (3.2.41) are necessary, Stillings argues, to account for the "fact" that "it is not possible to gap a verb unless what remains to the right of the gap is a single constituent" (p. 249).

It should be noted that this analysis is not subject to the objections we raised above about left-peripheral ellipsis. Stillings argues that all such cases are to be derived by Conjunction Reduction. Her argument is that left-peripheral ellipsis does not require that only a single constituent follow the gap, and must therefore be the result of a rule other than Gapping.

Now for the reasons we discussed earlier, we must agree with Stillings's conclusion about the analysis of left-peripheral ellipsis. Her argument for that conclusion, however, is faulty, because her claim that no more than one constituent can ever follow the gap is incorrect. In support of that claim she offers the following example (the judgement is Stillings's).

(3.2.43) *He mumbled to the gumbo of the function at the junction, and she ø [to the tea] [of the latest chivaree].

This example, of course, sounds like a perfectly fine Lewis Carroll line (as many people have pointed out to me). Since few people have intuitions about gumbos, chivarees, and mumbling to teas, however, the example can hardly be taken to show very much about constraints on Gapping.

But some of the examples we considered earlier are precisely of the form Stillings's analysis does not permit. The following ones, for instance:

(3.2.44)(a) Peter talked to his boss on Tuesday, and Betsy, ø [to her supervisor], [on Wednesday].

(b) John talked to his supervisor about his thesis, and Erich, ø [to the dean], [about departmental policies].

It's true that many sentences where a gapped clause contains three or more remnants are unacceptable (Jackendoff, Hankamer, and Stillings all give examples of these), but the fact that three-remnant gapped clauses are almost always fully acceptable when the last two remnants are both prepositional phrases argues against a formulation of Gapping like Stillings's. Moreover, one example attributed by Stillings to Bach, where the gapped clause contains the sequence [NP - ADV - NP], is judged to be perfectly acceptable (though perhaps awkward) by almost all of my informants:

(3.2.45) Monk probably enjoyed epistrophy, and Albert Ayler, almost certainly, ghosts.

In Section 3.4 of this chapter, we examine these cases in more detail. We will offer an account of some of these facts in terms of a surface constraint. The conclusion for the moment, however, must be that, in principle, Gapping has no constraint of the sort suggested by Stillings. Thus Stillings has presented no convincing evidence that linguistic theory must countenance "constituent variables".

Consider now her claim that the deletion target of Gapping must be V* rather than an X-variable, as had previously been assumed. Stillings considers, albeit briefly, certain obvious counterexamples to her claim, such as the following two familiar sentences (due to Jackendoff).

(3.2.46) Simon quickly dropped the gold, and Jack, ø the diamonds.

[ø = quickly dropped]

(3.2.47) Paul Schachter has informed me that the basic order in Tagalog and related languages is VOS; Ives Goddard $\emptyset_1$ that the unmarked order in Algonkian is OVS, and Guy Cardin $\emptyset_2$ that the basic order in Aleut is OSV. [$\emptyset_1$ = $\emptyset_2$ = has informed me]

As regards these sentences, Stillings cites Hankamer (1971, 1973a) as having "argue[d] on independent grounds that in one case <u>the preverbal adverb has become cliticized to the verb prior to Gapping</u> [emphasis added-I.A.S.], and that in the other case the postverbal pronoun has become cliticized" (p. 263). When Gapping applies, therefore, the sequences ADV-V or V-Pronoun will be dominated by V, and hence analyzable as deletion targets.

Now we mentioned earlier Hankamer's argument that post-verbal pronouns are cliticized. That argument, which seems quite plausible, has to do with the impossibility of Dative movement over direct-object pronouns (e.g. *He gave John it). However, nowhere in Hankamer (1971) or (1973a) is it argued that pre-verbal adverbs are cliticized onto following verbs.[5] Nor does Stillings offer any arguments for that position. Until such arguments can be found, then, the fact that pre-verbal adverbs can undergo Gapping must be taken as strong counterevidence to Stilling's claim that only (strings of) verbs undergo Gapping.

Note further that all the examples quoted earlier as counterexamples to Hankamer's and Langendoen's proposals are also unaccounted for in Stillings's analysis, for they are mostly cases where sequences like [V-NP] or [AUX-V-NP] have successfully been gapped (e.g. the sentences in (3.2.40) above).

There is a further objection to be made against Stillings's analysis. Notice that her rule makes no provision for right-peripheral ellipsis. Thus the right-peripheral ellipsis ($\emptyset_2$) in examples like this next one, are taken to be the result of an independently-motivated rule of (right-peripheral) Truncation (which Stillings attributes to Bresnan).

(3.2.48) Jack begged Elsie to get married and Wilfred $\emptyset_1$ Pheobe $\emptyset_2$.
[$\emptyset_1$ = begged; $\emptyset_2$ = to get married]

As an argument for the correctness of this two-rule view of the derivation of (3.2.48), the following minimally different example is offered, where only Truncation has applied.

(3.2.49) Jack begged Elsie to get married and Wilfred begged Pheobe $\emptyset$.

This argument is fallacious, however, for it is not the case that all right-peripheral ellipsis in gapped clauses can be attributed to application of an independent rule (though that rule (i.e. Truncation) may very well exist). Thus consider the following.

(3.2.50) Betsy believed Peter to be sexy, and
(a) Alan $\emptyset_1$ Barbara $\emptyset_2$. [$\emptyset_1$ = believed; $\emptyset_2$ = to be sexy]
(b) *Alan believed Barbara $\emptyset$.

In this example, it is clear that right-peripheral ellipsis can apply only if sentence-internal ellipsis (i.e. Gapping) has also applied. From this I conclude that a correct formulation of Gapping must also include a second (right-peripheral) deletion target, which Stillings's rule does not. Stillings's conjecture with regard to similar cases that "the two deleted variables are so closely connected semantically that it is impossible to delete one of them without the other" (p. 271) is at best a hand-wave at this serious defect of her analysis.

In summary, Stillings's analysis has little to recommend it. Her proposal is empirically inadequate, and therefore her claims about the necessity of incorporating "constituent variables" into linguistic theory, though interesting, remain unsupported.

From this discussion, it seems clear that Gapping must be formulated so as to delete not only one, but two X-variables. Any attempt to write the rule in this way, however, will overgenerate quite a bit. In Section 3.4 we will offer an analysis of this sort and we will argue that certain unacceptable outputs are to be explained by means of surface constraints. Our analysis will not give a grammatical explanation for Hankamer's "no-ambiguity" violations discussed earlier. This will be unnecessary, as we will see in what follows immediately, because a plausible extra-grammatical explanation of those cases has already been proposed by Kuno.

3.2.5 Kuno (1976)

Kuno points out many counterexamples to Hankamer's NAC and Langendoen's NLPNPC. Among them are the following.

(3.2.51)(a) Some people live in this city because they like living here, and others ∅ because they don't have means to move to the suburbs. [∅ = live in this city]

(b) One of the muggers hit Mary with a baseball bat, and another ∅ with a bicycle chain. [∅ = hit Mary]

(c) Some people go to Europe every year, and others ∅ every other year. [∅ = go to Europe]

(d) 50% of his constituents asked the Senator to vote for the bill, and 25% ∅ to vote against it. [∅ = asked the Senator]

(e) Two days ago John took Mary out to dinner, and this afternoon ∅ to the movies. [∅ = John took Mary out]

In all of these sentences, the remnants in the gapped clauses are something other than proper names. One reason why previous investigators were led to their (incorrect) conclusions about Gapping, Kuno argues, is that they considered only examples involving proper names, which have a strong tendency to be interpreted as "old information". Kuno argues that elements left in a gapped clause must represent "new information".[6] Therefore, sentences involving proper names, when considered in isolation, are sometimes difficult to assign the required interpretations.

To emphasize this point, he offers the following discourses, in which violations of the NAC and NLPNPC are possible even with proper names. The preceding wh-questions make clear that the proper names represent "new information".

(3.2.52)(a) Q: Who persuaded who to examine Mary?

(b) A: John persuaded Dr. Thomas to examine Mary, and Bill $\emptyset_1$ Dr. Jones $\emptyset_2$.

[$\emptyset_1$ = persuaded; $\emptyset_2$ = to examine Mary]

(3.2.53)(a) Q: With what did John and Bill hit Mary?

(b) A: John hit Mary with a stick, and Bill $\emptyset$ with a belt.

[$\emptyset$ = hit Mary]

Now to account for the tendency of many speakers to interpret certain gapped sentences in a way that is in accordance with Hankamer's NAC, Kuno posits a perceptual principle. This Minimal Distance Principle (due originally to Langendoen and Grosu in their course: "Interactions Among Systems of Verbal Ability", given at the 1974 Linguistic Institute of the L.S.A.) is as follows.

(3.2.54) The Minimal Distance Principle (MDP)

The two constituents left behind by Gapping can be most readily coupled with the constituents (of the same structures) in the first conjunct that were processed last of all.

This principle interacts with several other principles Kuno develops to account for the complex array of acceptability judgements vis à vis gapped sentences.

One other principle Kuno posits is the following:

(3.2.55) The Tendency for Subject-Predicate Interpretation (TSPI)

When Gapping leaves an NP and a VP behind, the two constituents are readily interpreted as constituting a sentential pattern, with the NP representing the subject of the VP.

The following sentence of Hankamer's is deviant because it violates TSPI.

(3.2.56) *Jack asked Mike to wash himself, and Sue ø to shave himself.
[ø = asked Mike]

The two Gapping remnants [$_{NP}$Sue], [$_{VP}$to shave himself] are not in a subject-predicate relationship.

As further evidence for the correctness of TSPI, Kuno offers the following examples.

(3.2.57) Bill was persuaded by John to donate $200, and Tom ø to donate $400. [ø = was persuaded by John]

(3.2.58) *John persuaded Bill to donate $200, and Tom ø to donate $400.
[ø = persuaded Bill]

Only in (3.2.58) is TSPI violated. Moreover, the acceptability of the following sentence is in accord with TSPI.

(3.2.59) John promised Bill to donate $200, and Tom $\emptyset$ to donate $400.
[$\emptyset$ = promised Bill]

$[_{NP}$Tom] and $[_{VP}$to donate $400] are in a subject-predicate relationship. (Notice that Hankamer's NAC fails to distinguish between (3.2.58) and (3.2.59).)

The overall acceptability of a gapped sentence then is argued to be a complex function involving various perceptually-based factors. (3.2.56) and (3.2.58) violate both MDP and TSPI, and are therefore consistently judged by speakers to be the least acceptable. (3.2.57) and (3.2.59) violate MDP but not TSPI, and hence are judged to be acceptable by most speakers. Moreover, the following interpretations of (3.2.58) and (3.2.59), which do not violate MDP, are judged completely acceptable by all speakers.

(3.2.60) John persuaded Bill to donate $200, and $\emptyset$ Tom to donate $400.
[$\emptyset$ = John persuaded]

(3.2.61) John promised Bill to donate $200, and $\emptyset$ Tom to donate $400.
[$\emptyset$ = John promised]

These next ungrammatical examples are also considered by Kuno:

(3.2.62) *John persuaded Dr. Thomas to examine Jane, and Bill $\emptyset$ Martha.
[$\emptyset$ = persuaded Dr. Thomas to examine]

(3.2.63) *Dr. Jones was persuaded by Bill to examine Jane, and $\emptyset_1$ by John $\emptyset_2$ Martha. [$\emptyset_1$ = Dr. Jones was persuaded; $\emptyset_2$ = to examine]

(3.2.64) *John hugged his sister to please his mother, and $\emptyset_1$ his wife $\emptyset_2$ his father. [$\emptyset_1$ = John hugged; $\emptyset_2$ = to please]

Sentences like these, it is argued, violate yet another perceptually motivated constraint, which Kuno refers to as The Requirement for Simplex-Sentential Relationship:

(3.2.65) The Requirement for Simplex-Sentential Relationship (RSSR) The two constituents left over by Gapping are most readily interpretable as entering into a simplex-sentential relationship. The intelligibility of gapped sentences declines drastically if there is no such relationship between the two constituents.

In each of the last three examples, the two Gapping remnants bear no simplex-sentential relationship to each other, i.e. the first remnant in each case originates in the matrix clause while the second remnant originates in an embedded clause (and moreover the first remnant does not control any position in the embedded clause). Therefore all these exmples are in violation of RSSR. RSSR, as Kuno notes, "is a very strong and nearly inviolatable constraint".

Now the extreme variation from speaker to speaker concerning acceptability judgements of gapped sentences should come as no surprise if Kuno's principles, or anything like them, are correct. One would expect such principles to vary in strength from individual to individual. Given that the facts in this area are almost all of this nature, idiolectally varying weightings of extra-syntactic (primarily perceptual) factors seems to be a rather intuitive approach to the problem. This is not to say that the specific formulations that Kuno has given for these various principles are precisely correct as they stand. There are many particular problems that are not quite worked out in sufficient detail. For example, (3.2.61), which is clearly acceptable to all speakers, obeys MDP and RSSR, but not TSPI. There is surely more that needs to be said about the interaction of the various principles (although obeying two out of three principles comes pretty close to being a sufficient condition for acceptability). We will not attempt to refine Kuno's analysis here. Rather, we will assume that something on the order of Kuno's principles is in fact at work, and we will proceed to a formalization of the rule of Gapping that will interact with those principles appropriately.

Summarizing now, we have examined five previous analyses of Gapping. The proposals of Ross, Hankamer, Langendoen, and Stillings have all been shown to be empirically inadequate. The hypothesis of Kuno's that various extra-syntactic factors influence the acceptability of gapped sentences has been seen to provide a fairly successful account of previously-observed problematic data.

While we have accepted Kuno's arguments, we cannot agree with his conclusion that "these non-syntactic factors leave, in the domain of pure syntax, a set of rather uninteresting constraints on Gapping". In Section 3.4 of this chapter we will see that a proper formulation of the rule of Gapping poses many interesting problems, and, in fact, interacts in intriguing ways with general syntactic principles. Certain problems concerning the nature of general syntactic principles must be cleared up first, however, and this is the object of the next section.

3.3 RAOAP Reconsidered

At the very end of Section 1.2 of Chapter One, we put forth a view of the nature of the interaction of constraints on rules of grammar. We offered an account of certain problematic cases where VPD is unable to apply in terms of a hypothesis that the recoverability of deletion functions independently of other constraints on rule application.

Bresnan (1975, 1976a) takes a rather different view of the nature of constraint interaction. She proposes that RAOAP is dependent upon the principle of recoverability of deletion. That is, she views RAOAP as a principle that guarantees, inter alia, that the target of a deletion rule will be the maximal constituent that is recoverable. This is to say that the principle of recoverability of deletion first delimits the class of possible proper analyses of a given sentence with respect to a given transformation, and that RAOAP then ensures that, of the proper analyses that are not ruled out by the recoverability condition, those that are non-maximal are ruled out. Her position then is essentially the same as Grosu's (see the discussion in Section 1.2 of Chapter One).

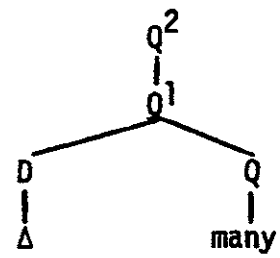
Bresnan buttresses her argument for this view with an in depth proposal for Comparative Deletion (CD). The deletion target of CD is a phrasal node (of the type X^2 in Bresnan's (1976a) system) that begins with a measure phrase (Q^2 in Bresnan's system). The following formulation is given.

(3.3.1) Comparative Deletion (Bresnan (1976a))

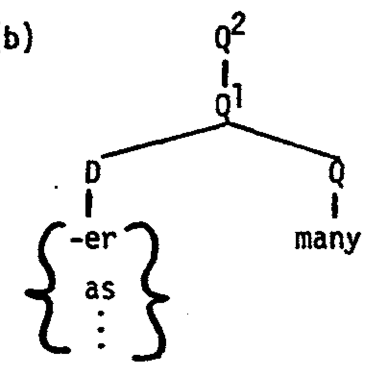
$$
\begin{array}{lccccccc}
[_{X^2}\ [_{X^2} & Q^2 - & W_1\ ] - & W_2]\ [_{\bar{S}} & W_3 - [_{X^2} & Q^2 - & W_4\ ] - & W_5] \\
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
\Rightarrow & 1 & 2 & 3 & 4 & & \emptyset & 7
\end{array}
$$

Bresnan's view of the recoverability of deletion involves a notion of "non-distinctness". Term 5 of this rule must be able to analyze Q^2's like the one in (3.3.2)(a), which are to be deleted under identity with Q^2's like those in (3.3.2)(b), even though they are not identical in their overt syntactic structure.

(3.3.2)(a)



(b)



A more detailed account of her view of the syntax of the comparative clause can be found in Bresnan (1973), which I will presuppose familiarity with here.

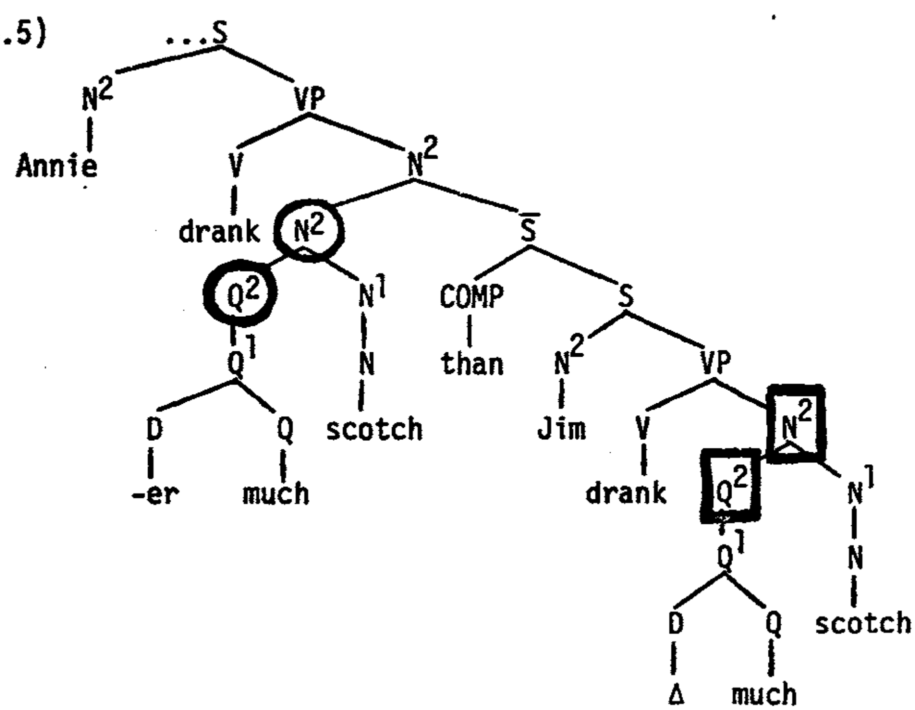
RAOAP interacts with CD to account for such contrasts as the following.

(3.3.3) Annie drank more scotch than Jim drank ø. [ø = Δ much scotch]

(3.3.4) *Annie drank more scotch than Jim drank ø scotch. [ø = Δ much]

The deletion in (3.3.4), where only the Q^2 has been deleted (Bresnan refers to this phenomenon as Subdeletion), is ruled out by RAOAP, because that Q^2 is not the maximal identical Q^2 initial X^2. That is, assuming the structure in (3.3.5) prior to deletion,

(3.3.5)



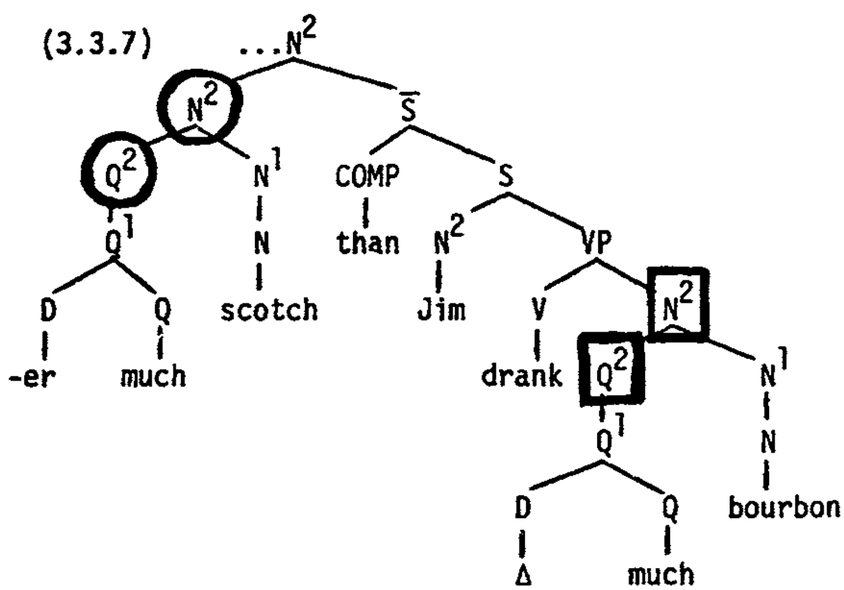
two proper analyses should be possible. $\boxed{Q^2}$ and $\boxed{N^2}$ both meet the identity condition and are possible deletion targets for CD (analyzable as terms 5 and 6 of rule (3.3.1) above). Since $\boxed{Q^2}$ is not the maximal deletion target, however, deleting only $\boxed{Q^2}$ is ruled out by RAOAP, thus accounting for the deviance of (3.3.4).

Crucially, CD (subdeletion, actually) is allowed to apply in examples like this next one.

(3.3.6) Annie drank more scotch than Jim drank bourbon. [ø = Δ much]

This illustrates Bresnan's view of the interaction of the recoverability of deletion and RAOAP. The pre-deletion structure for this sentence is as given:

(3.3.7)



Here $\boxed{N^2}$ does not meet the identity condition, for, as anyone knows, bourbon ain't scotch. The maximal deletion target that meets the identity condition is $\boxed{Q^2}$. If RAOAP functioned as we suggested in Chapter One, we would predict that no deletion is possible here. On Bresnan's view, however, RAOAP serves to rule out one of two possible proper analyses only when both meet the recoverability condition. That view therefore correctly allows the non-maximal deletion in (3.3.6) (i.e. the maximal recoverable deletion) to occur. How can these facts be reconciled with our earlier proposal?

The first thing to say, I think, is that there is a certain unclarity in the data. Reconsider (3.3.4):

(3.3.4) Annie drank more scotch than Jim drank scotch.

The fact of the matter is, only about sixty percent of my informants reject this sentence (when presented with it in isolation). Moreover, most of those that do, find it more "awkward" or "redundant" than "ungrammatical". But if this is the case, it seems to me rather unlikely that RAOAP is an appropriate explanation for whatever deviance inheres in (3.3.4). If RAOAP is to be accepted, it will be the only thing responsible for the deviance of the following examples, as Bresnan herself argues:

(3.3.8)(a) *How difficult do you want to solve a problem? (cf. How difficult a problem do you want to solve?) [Wh-Q-Movement]

(b) *How many is the girl feet tall? (cf. How many feet tall is the girl?) [Wh-Q-Movement]

(c) *He considers many of stupid my best friends. (cf. He considers stupid many of my best friends) [Complex NP Shift]

Here there is no informant variation. All speakers concede that such non-sentences are aptly characterized as "word salad". Why should there be such a difference between the status of (3.3.4) and that of the sentences in (3.3.8) if only RAOAP is being violated in each case?

Edwin Williams (personal communication) has pointed out to me that there is some further evidence that RAOAP is the wrong explanation for the deviance of (3.3.4). He observes the following pair of examples:

(3.3.9) Annie drank more water than Jim drank water.

(3.3.10) Annie drank more water than Jim drank H_2O.

If the deviance of (3.3.4) is due to the fact that a non-maximal identical target was deleted, then (3.3.10) should be grammatical, for the two N^2's water and H_2O are not identical in form, and the deleted element is indeed the maximal identical target.

Now (3.3.10), with accent on Jim and reduced stress thereafter is judged by informants to be on a par with (3.3.9) (i.e. "redundant"), or else to be even less acceptable. With accent on H_2O, the sentence is consistently judged to be peculiar, but because it would seem to be saying that H_2O is something different from water.

In order to give an account of these facts, we must posit some kind of principle that remnants in a Subdeleted clause must be accented. I am not at all certain whether the proper account should be in terms of accent itself, in terms of the functional conditions that determine that accent (say new or contrastive information, as Kuno suggests for remnants in Gapped clauses - see Kuno (1972, 1976)), or in terms of more general considerations having to do with contrastive accent, logical form, and the recoverability of deletion (as will be motivated in Section 3.5 of this chapter). In any case, however, we would then have an independent explanation for all the facts in question. (3.3.10) with the accent on H_2O would not violate the principle, but would be pragmatically anomolous: items bearing contrastive accent must, at the very least, convey new information (though this is not as easy to make precise as some (e.g. Kuno) think it is). H_2O (accented) in (3.3.10) is then a Gapping remnant that conveys old information, but which, by virtue of its accent, must convey new information; hence, the pragmatic anomoly. (3.3.10) with reduced stress on $\underline{H_2O}$, (3.2.9), (3.2.8), and furthermore most of Bresnan's examples illegitimate non-maximal Subdeletion violate the proposed principle, for anaphoric material in subordinate clauses (whether pronominal, lexically identical, or paraphrastic (e.g. $\underline{H_2O}$ for water)) requires destressing (see Carrier (1974)). The requirement of destressing is in conflict with the principle that such remnants must be contrastively accented, and the conflict induces reduced acceptability.

Since the proposed explanation is in terms of extra-grammatical factors, it is not at all surprising that there is variation among speakers. Such factors may very well vary from speaker to speaker. RAOAP, on the other hand (or the modification of it we will ultimately suggest), is an immutable metatheoretical principle not subject to inter-subjectual variation.

A further argument for the correctness of this approach to the general problem of non-maximal Subdeletion has been pointed out to me by Larry Horn. He observes that the acceptability of such sentences is significantly enhanced if the accent on the repeated identical element is motivated in discourse. The following example illustrates Horn's point:

(3.3.11) Speaker A: Annie drank more scotch than Jim drank bourbon.
Speaker B: No, you've got it all wrong. Annie drank more scotch than Jim drank scotch (not bourbon).

Indeed, virtually all of my informants find such examples to be perfectly acceptable.

There also seem to be cases where (when the discourse permits) non-maximal Subdeletion is possible even without an accent on the Gapping remnant. These all seem to involve contrast between the comparative head and an appropriate entity in the preceding discourse:

(3.3.12) Speaker A: The table is wider than the desk is long.

Speaker B: No, you've got it all wrong. The table is longer than the desk is long.

Notice that we find exactly analogous facts in sentences not involving comparatives, suggesting that there is an independent explanation (having nothing to do with maximal application of a deletion rule) that accounts for all the Subdeletion facts we have looked at so far:

(3.3.13)(a) John's height exceeds Bill's height.

(b) John's height exceeds Bill's height. [OK as a correction of: John's height exceeds Bill's weight]

(3.3.14)(a) John's height exceeds Bill's weight.

(b) John's height exceeds Bill's weight. [OK as a correction of: John's weight exceeds Bill's weight]

All the above facts are irreconcilable with Bresnan's theory. If CD and Subdeletion are collapsed into a single rule, which applies subject to RAOAP, there is no way to account for the grammaticality of (3.3.11) and (3.3.12). But surely all such sentences must be generated by the syntax. There are essentially two ways out of this dilemma. Either we separate CD and Subdeletion into two separate processes (perhaps following Chomsky in treating CD as Relativization), or we modify RAOAP in such a way that a combined CD-Subdeletion rule is allowed to apply in (3.3.11) and (3.3.12). Although in what follows we will see that the first hypothesis is preferable, it is nevertheless also the case that RAOAP is in need of revision.

Notice that if we accept Bresnan's collapsed CD-Subdeletion rule, her formulation of RAOAP, and the particular details of her analysis of the syntax of the comparative clause, the view we put forth at the end of Chapter One vis à vis the interaction of RAOAP and the recoverability of deletion makes some incorrect predictions. In particular, if we assume structures like those in (3.3.5) and (3.3.7) above, then independent application of the two principles in question will incorrectly predict that no deletion is possible (in either (3.3.4) or (3.3.6)). Since assuming independence of constraints on rules enabled us to explain certain problematic cases of VPD in Chapter One, it is highly desirable that we reconcile that assumption with the facts at hand.

I would like to suggest first of all that RAOAP should be replaced by the following metatheoretical principle, which we will call the Immediate Domination Principle:[7]

(3.3.15) Immediate Domination Principle (IDP)

Given 2 proper analyses, PA_1 and PA_2, of a sentence S with respect to a transformation T which differ minimally with respect to the value of some predicate P in the structural description of T, if (1) $/P/_{PA_1}$(the value of P under PA_1) is some terminal substring t_1,
and (2) $/P/_{PA_2}$ is some terminal substring t_2,
and (3) t_1 is analyzable as A_1 ($A_1 \varepsilon V_{NT}$),
and (4) t_2 is analyzable as A_2 ($A_2 \varepsilon V_{NT}$),
and (5) A_1 immediately dominates A_2,
then PA_2 is an inadmissable proper analysis of S with respect to T.

Let us further clarify the notion "differ minimally". Intuitively, two proper analyses differ minimally if for every predicate in the structural description, except one, the value of the predicate is the same under both proper analyses. Since variables are not predicates, this means that in all cases of two minimally differing proper analyses, a variable term will "cover" a longer substring in one than in the other. More precisely,

(3.3.16) Two proper analyses PA_1 and PA_2, of a given sentence S with respect to a transformation T <u>differ minimally</u> if

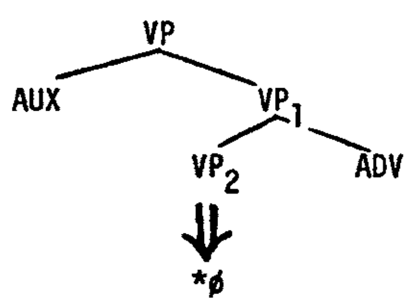
(1) There is some predicate P_i in the structural description of T such that $/P_i/_{PA_1} \neq /P_i/_{PA_2}$, and

(2) For all other predicates P_n (in the structural description of T), $/P_n/_{PA_1} = /P_n/_{PA_2}$.

Notice that IDP rules out fewer proper analyses (in any particular situation) than Bresnan's RAOAP. It thus makes weaker predictions. We will see, however, that this is preferable on empirical grounds.

Reconsider for a moment the cases of illegitimate VPD discussed at the end of Chapter One. We invoked RAOAP to explain why VPD could not apply to delete VP's in the position of VP_2 in (3.3.17).

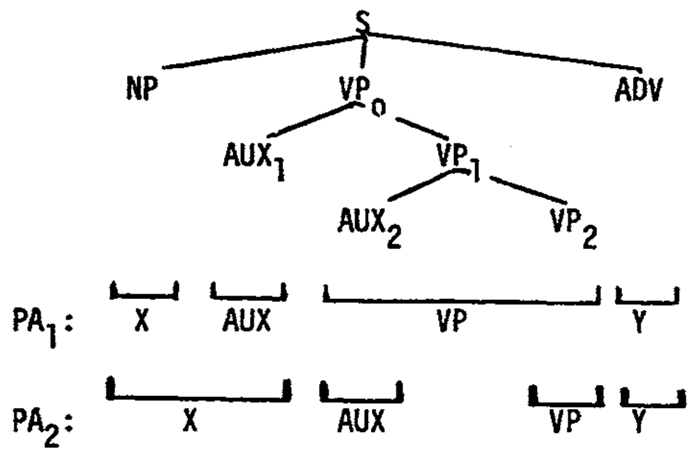
(3.3.17)



IDP also prevents this deletion, because VP_1 immediately dominates VP_2, and the two proper analyses in question differ minimally.

It's important to note that IDP does not block deletion of VP_2 in a configuration like the following one.

(3.3.18)



Here PA_1 and PA_2 do not differ minimally, for there are two predicates in the SD of VPD which have different values. Recall that this is precisely the right prediction (assuming the independence of IDP and the recoverability of deletion). VPD is possible in (3.3.18), but not in (3.3.17).

IDP, as we have formulated it, accounts for many facts that are generally subsumed under the A-over-A Principle, the Left Branch Condition, or the Coordinate Structure Constraint. This will become clear in the discussion that follows.

Ross (1967a) argued that the ungrammaticality of examples like those in (3.3.19), which are presumed to derive from the corresponding structures in (3.3.20), is to be explained by a single general principle.

(3.3.19)(a) *Whose did you read book?

(b) *The boy who(m) she expected who(m)[8] I knew to show up...

(c) *How Bob is muscular!

(3.3.20)(a) [You read Wh + someone's book]

(b) [the boy_i [she expected the boy_i [whom I knew to show up]]]...

(c) [Bob is to wh-some extent muscular]

This principle, the Left Branch Condition, he formulated as in (3.3.21).

(3.3.21) The Left Branch Condition (LBC)

No NP which is the left-most constituent of a larger NP can be reordered out of this NP by a transformational rule.

Ross also noted that LBC could not be universal, because some languages allow such reorderings. The following Russian examples (due to Ross) illustrate this:

(3.3.22)(a) Č'oju knigu ty čitageš?
whose book you read
'Whose book are you reading?'

(b) Č'ju ty čitaješ' knigu?
whose you read book?
*'Whose are you reading book?'

Although examples like (3.3.22)(b) seem to be highly marked, nevertheless it is clear that Russian allows such sentences, whereas English does not. Ross argued that it is an idiosyncratic fact whether or not a particular language obeys LBC.

However, as Grosu (1974) points out, not all of the cases ruled out by LBC are grammatical in Russian, as sentences like the following show.[9]

(3.3.23) *Čto ja postal čto vesilo piat'
what I sent which weighed five
kilogram moemu drugu v Ameriku?
kilograms to my friend from America
'What did I send to my friend from America which weighed five kilos?'

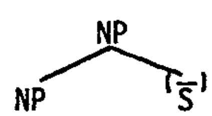
Here extraction of the head of a relative clause results in ungrammaticality.

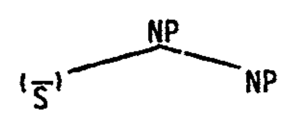
Grosu examines various facts from English, Russian, Turkish, Japanese, and Rumanian. His conclusion is that Ross's LBC "constitutes a spurious generalization". He offers the following tentative generalizations:

(3.3.24) A. The freezing of the heads of complex NP's seems to be a linguistic universal; and

B. The behavior of modifying NP's (of which "possessor" and/or genitive NP's constitute subinstances) and of adjectival or nominal quantifiers with respect to reordering processes must be specified - with different degrees of idiosyncracy - for every language.

This state of affairs is quite consistent with IDP. That is, as a meta-principle, we would expect it to be universal. The only one of Grosu's generalizations which falls under IDP, I would claim, is (3.3.24)A (we will have more to say about this in a moment). Specifically, in any language where heads of relative clauses appear in one of the configurations in (3.3.25), IDP correctly predicts that any rule that reorders NP must reorder the higher NP, not just the head.

(3.3.25)(a) 

(b) 

Thus IDP accounts for Grosu's Turkish examples like the following, where right-branching heads are unreorderable.

(3.3.26) *Mary şapka giyen ø Bille gösterdi çocuk
Mary hat wearing to Bill showing boy
*'The boy who Mary showed who was wearing a hat to Bill.'

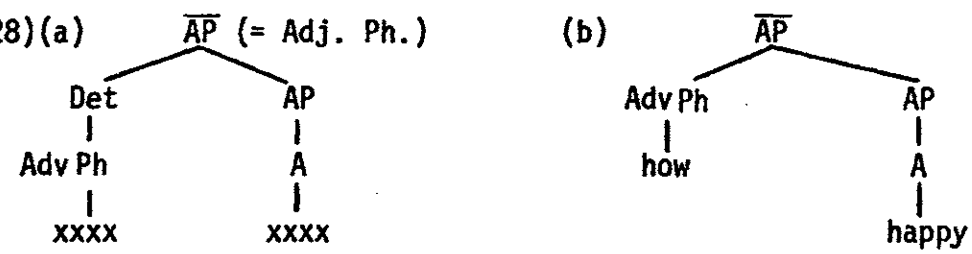
*(3.3.19)(b) above also falls under IDP, for it is an illegal reordering of a left-branching head of a relative clause (in English). Since IDP, unlike LBC, is formulated in terms of immediate domination, not left-to-right order, it accounts for precisely the generalization that Grosu proposes.

What about adverbial modifiers? If Ross's and Grosu's claim is correct that there are some languages in which sentences like the following ungrammatical English example are grammatical (Ross's Russian example which Grosu cites is of dubious grammaticality, as Morris Halle informs me), then what is the explanation for the English facts?

(3.3.27) *How is John happy? (≠ How happy is John?)

One possibility is that the difference between a language in which sentences like (3.3.27) are grammatical and English, where they are clearly not, is simply a difference in the constituent structure of the Adjective Phrases. Perhaps the following difference:

(3.3.28)(a) (b)



However, if, as I suspect, there are no languages where sentences like (3.3.27) are grammatical, then we can suppose that no languages have the $\overline{AP}$ structure of (3.3.28)(a).[10]

In English, at any rate, it's clear that *Wh*-(Q)-Movement reorders both $\overline{AP}$ and $\overline{Adv\ Ph}$ (cf. *When does John talk?*). Therefore, assuming the structure in (3.3.28)(b), IDP correctly predicts that the only $\overline{AP}$ can be reordered, for $\overline{AP}$ immediately dominates $\overline{AdvPh}$, and the two proper analyses in question differ minimally.

As for genitive NP's, Grosu cites sentences from Hankamer (1971) which show that in Turkish, genitive NP's can be Relativized and Topicalized.

(3.3.29)(a) Türkyede ∅ film *i* yasak olan roman...
in Turkey film + possessive forbidden being book

'The book whose film is forbidden in Turkey...'

(b) Kardeşin *i* gördüm Johnun
brother + possessive I saw John's

'Speaking of John, I saw his brother.'

(Topicalized NP's are sentence-final in Turkish)

The ungrammaticality of such sentences in English, i.e. the ungrammaticality of examples like the following, must be due to a language-particular constraint against moving genitive NP's.

(3.3.30)(a) *The man whose I liked ∅ book...

(b) *John's, I liked ∅ book.

This explanation is further supported by examples like (3.3.30)(c) which one of Grosu's anonymous reviewers pointed out to him.

(3.3.30)(c) *Your wife's, I met an uncle of

English genitive NP's are frozen, no matter what phrase-structure configuration they occur in.

But just what is the phrase structure of pre-nominal genitives in English? Notice that the following examples are all ungrammatical.

(3.3.31)(a) *the John's book

(b) *a John's book

(c) *John's a book

(d) *John's the book

These facts constitute an argument that possessive NP's are dominated by Det (determiner). The non-occurrence of articles and possessives would follow if we posit the usual phrase structure rules:

(3.3.32)

$$\text{NP} \rightarrow \text{Det} - \overline{\text{N}}\ldots$$
$$\text{Det} \rightarrow \left\{ \begin{array}{c} \text{Article} \\ \text{NP} \\ \vdots \end{array} \right\}$$

This being the case, the examples in (3.3.30) are ruled out only by the genitive constraint, for the Det node that intervenes between the two possible NP movement targets blocks IDP. Further, arguments for a Det node over possessive NP's can be found in Siegel (1974).

By now it should be clear that is also the genitive constraint, and not IDP that will be held responsible for the ungrammaticality of examples like the following (due to Bresnan) where only a genitive NP has undergone Subdeletion:

(3.3.33) *John was as many boys' favorite teacher as he was ø ('s) favorite advisor.

Now as for the scotch and bourbon examples we looked at earlier, it seems reasonable to assume that pre-nominal quantifier phrases are also dominated by Det. Thus we do not have...

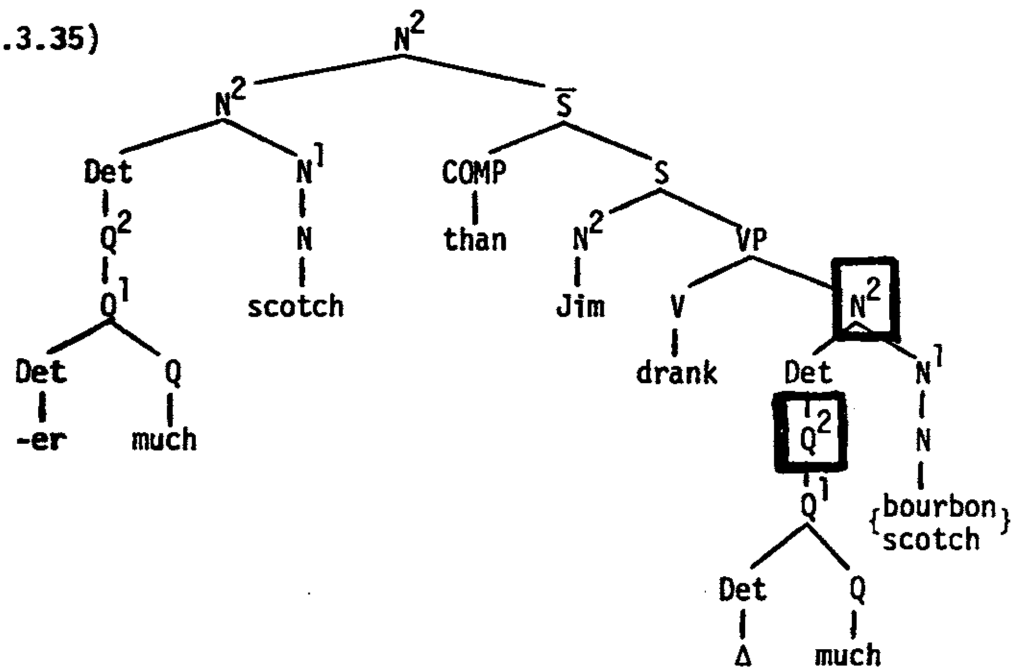
(3.3.34)(a) *the more books than John read.

(b) *the as many books as John read

(c) *some more books than John read. etc.

Examples like (3.3.4) and (3.3.6) above then will have constituent structures of the following general shape.

(3.3.35)



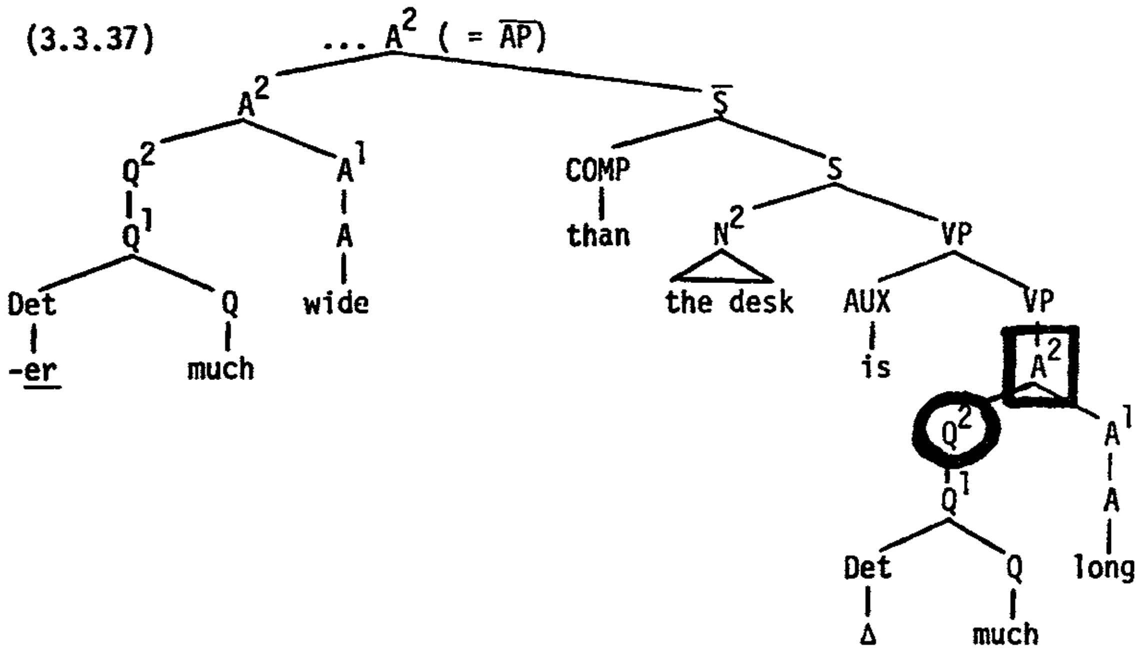
Now if we follow Bresnan in collapsing CD and Subdeletion, both $\boxed{N^2}$ and $\boxed{Q^2}$ are possible deletion targets. $\boxed{N^2}$ does not immediately dominate $\boxed{Q^2}$ so deletion of just $\boxed{Q^2}$ is permitted by IDP. Deletion of $[_{N^2}\ Q^2$ - bourbon] would of course be non-recoverable.

Since we have just argued that Adjective Phrases do not have Determiners, examples like the following which we also considered earlier, are of considerable interest.

(3.3.36) The table is wider than the desk is long.

We must assume the following structure for this example (prior to deletion).

(3.3.37)



But here $\boxed{A^2}$ immediately dominates Q^2. Thus if CD-Subdeletion is a single rule, IDP incorrectly predicts that (Sub-)deletion of Q^2 is unallowed. This leads naturally to the conclusion that CD and Subdeletion should not be collapsed. If the target term of Subdeletion is simply Q^2 (= $\overline{QP}$), then in (3.3.35) and in (3.3.37) above there is only one Subdeletion target. The question of IDP does not even arise. Cases of CD then may well be cases of Relativization, as Chomsky has suggested.

It's interesting to note that almost all of Bresnan's examples of RAOAP maximization are at least arguably cases involving immediate domination. Take Complex NP Shift [= Heavy NP Shift], for instance. This rule moves NP's and PP's rightward within VP. Bresnan cites the following paradigm.

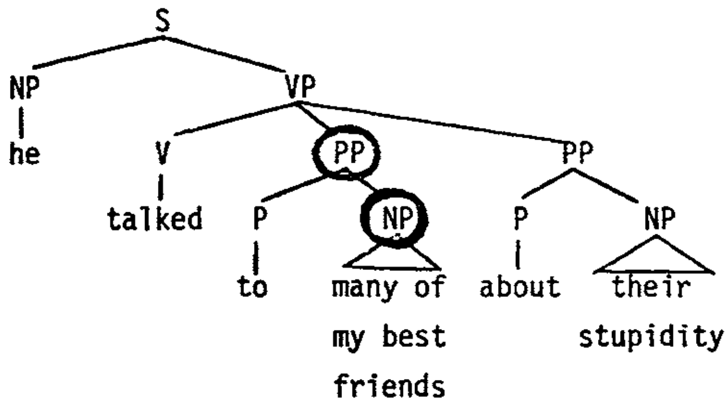
(3.3.38)(a) He talked to many of my best friends about their stupidity.

(b) He talked about their stupidity to many of my best friends.

(c) *He talked to about their stupidity many of my best friends.

Assuming the constituent structure in (3.3.39), these facts follow from RAOAP,

(3.3.39)



whereas both PP and NP here should be able to shift rightward. RAOAP rules out the non-maximal shifting of NP. But IDP also does the job, because, as is almost completely uncontroversial (<u>pace</u> Jackendoff (1976)), PP immediately dominates NP.

Bresnan also cites the following paradigm in her discussion of Complex NP Shift:

(3.3.40)(a) He considers many of my best friends stupid.

(b) He considers stupid many of my best friends.

(c) *He considers many stupid of my best friends.

I know of no decisive arguments in the literature concerning the internal structure of the complex NP's in these examples. I note, however, that these sequences: <u>of NP</u> do not undergo such rules as Extraposition of PP (Ross (1967a), Jacobson and Neubauer (to appear)):

(3.3.41)(a) *Many arrived of my best friends.

(b) *Several came in of my best friends.

(c) *All arrived of my best friends.

These examples suggest that these sequences: <u>of NP</u> are not PP's.

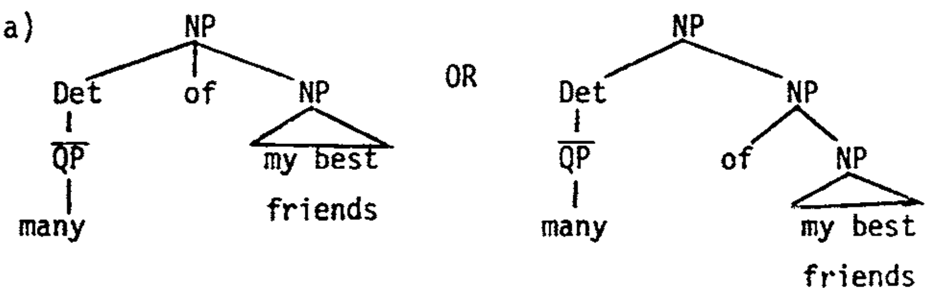
This is rather surprising, I would say. Notice that other sequences of <u>of</u> <u>NP</u> extrapose normally:

(3.3.42)(a) A professor arrived of considerable fame.

(b) A package was brought in of considerable bulk.

The facts of (3.3.41) would follow if we posited one of the following constituent structures for the complex NP in question. The structure in (3.3.43)(b) would be the natural result of an <u>of</u>-Insertion analysis, as has been suggested.

(3.3.43)(a)



But if some structure like this is correct, then we have an explanation for Bresnan's facts in (3.3.40) also. One NP (or N^2) immediately dominates the other. Thus (3.3.40)(c), but not (3.3.40)(b), is ruled out by IDP.

Let us now turn to Bresnan's analysis of Relativization. She posits the following rule

(3.3.44) Relativization (Bresnan (1976a))

$$\begin{array}{llllllllll} & NP & - [_{\overline{S}}\ COMP & - W_1 & - (P) & - [_{\substack{X^2 \\ [-V]}} W_2 & - rel & - W_3\ ] & - W_4\ ] \\ & 1 & 2 & 3 & 4 & \underbrace{5 \quad\quad 6 \quad\quad 7} & & & 8 \\ \Rightarrow & 1 & \overbrace{5\ 6\ 7} & 3 & 4 & \emptyset & & & 8 \end{array}$$

The Relativization target (terms 5, 6, 7) is what Woisetschlaeger (1975) calls a "mixed term". Mixed terms are a way of specifying that a rule applies to constituents that contain some designated element. [-V] is in Bresnan's system the feature specification that NP's and PP's have in common. Relativization then moves either NP's or PP's. RAOAP will here impose cross-categorical maximization effects. The parenthesized P in the structural description of (3.3.44) is the device Bresnan uses to account for optional preposition stranding (this was first suggested, I believe, by Rodman (1972)). When the P option is taken, i.e. when the value of the context predicate (P) is non-null, then RAOAP maximizes the Relativization target relative to the context predicate.

The interaction of Bresnan's Relativization rule and RAOAP accounts for all the following data (taken from Ross (1967a)):

(3.3.45)(a) Reports which the government prescribes the height of the lettering on the covers of are invariably boring.

(b) Reports the covers of which the government prescribes the height of the lettering on almost always put me to sleep.

(c) Reports the lettering on the covers of which the government prescribes the height of are a shocking waste of public funds.

(d) Reports the height of the lettering on the covers of which the government prescribes should be abolished.

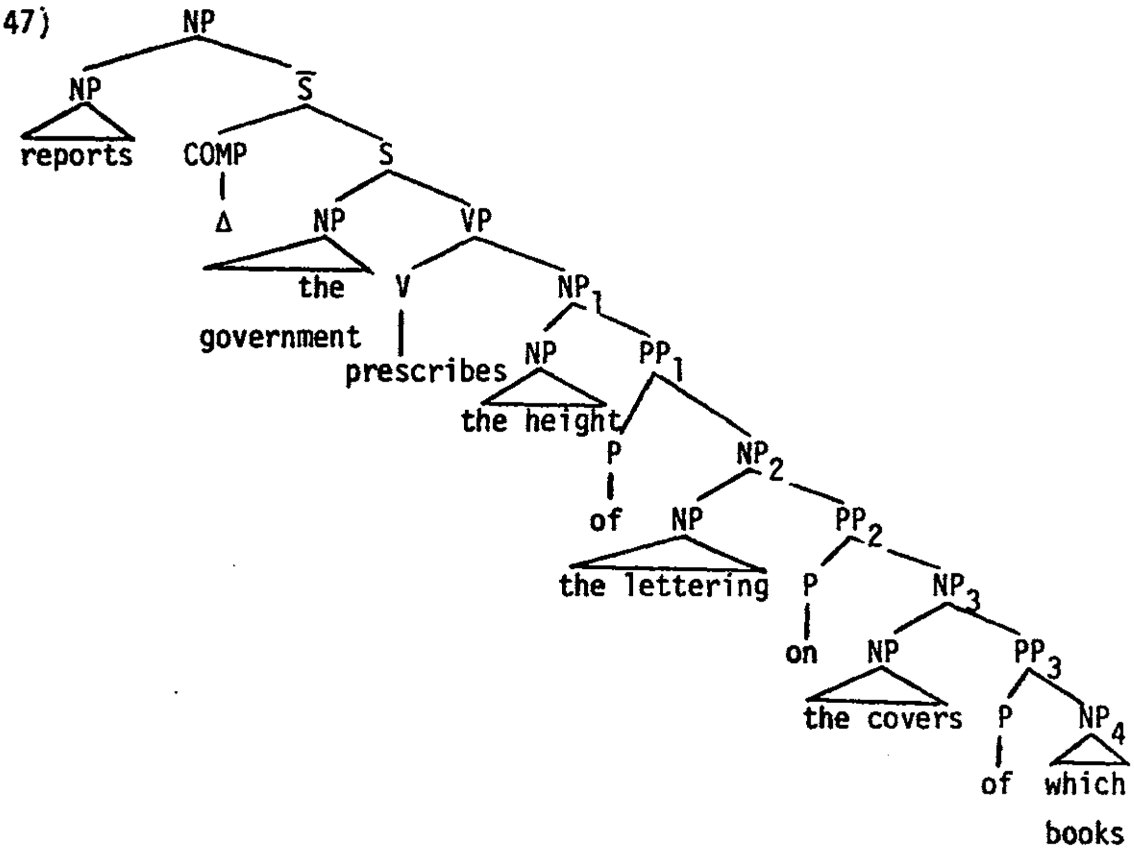
(3.3.46)(a) *Reports of which the government prescribes the height of the lettering on the covers are invariably boring.

(b) *Reports on the covers of which the government prescribes the height of the lettering almost always put me to sleep.

(c) *Reports of the lettering on the covers of which the government prescribes the height are a shocking waste of public funds.

To see this, consider (3.3.47), which is the underlying structure posited for these examples.

(3.3.47)



(P) in the SD of rule (3.3.44) can optionally analyze any of the prepositions in (3.3.47). RAOAP then guarantees that the moved element, for each choice of P, will be the maximal NP to its right. This accounts for the grammticality of (3.3.45)(a)-(c). If no P is so analyzed, then the Relativization target must be the maximal NP or PP containing the <u>wh</u>-word (= <u>rel</u> in (3.3.44). The maximal such entity is NP_1. This accounts for the grammaticality of (3.3.45)(d). Moreover, the ungrammatical examples in (3.3.46) are all in violation of RAOAP. In each case, no P has been analyzed by term 4 of the rule, and the moved constituent is not the maximal NP or PP containing the <u>wh</u>-word.

Notice, however, that IDP makes precisely the same predictions in all these cases. No constituent is a legitimate target term (relative to the given context term) if there is some other constituent that immediately dominates it that is also a possible target term. Thus assuming the structure in (3.3.47), (3.3.46)(a), (b), and (c) are the result of analyzing PP_3, PP_2, and PP_1 as the Relativization target. But each of these targets is immediately dominated by another possible target, thus violating IDP (the relevant proper analyses differ minimally). The only proper analysis of (3.3.47) (assuming no P has been analyzed) that is consistent with IDP is the one where NP_1 is the Relativization target. That proper analysis yields (3.3.45)(d). All Ross's facts in (3.3.45) - (3.3.46) above are therefore correctly accounted for by IDP.

Now as Bresnan notes, judgements on these "pied-piping" sentences vary considerably from speaker to speaker. Bresnan gives no indication of how her proposal can be modified to account for any array of facts other than Ross's, except to offer the speculation that "slightly different formulations of rules give different results".

The matter is really more serious than Bresnan's comment would imply. Some speakers, for example, accept sentences like the following one.

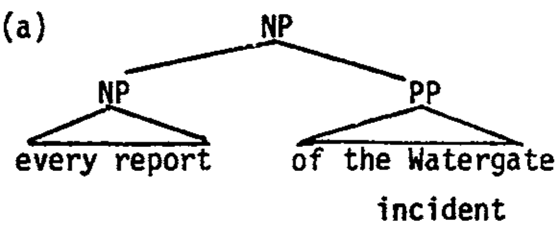
(3.3.48) %The Watergate incident of which I have carefully examined virtually every report...

This type of idiosyncratic variation seems to me rather difficult to reconcile with Bresnan's analysis. If the locus of the inter-subjectual variation is lexical rather than structural, it's hard to see how giving the Relativization rule a slightly different formulation will do the job. Assuming IDP, however, a rather intuitive characterization of the alternative set of facts seems to be possible. The variation is simply one of constituent structure.

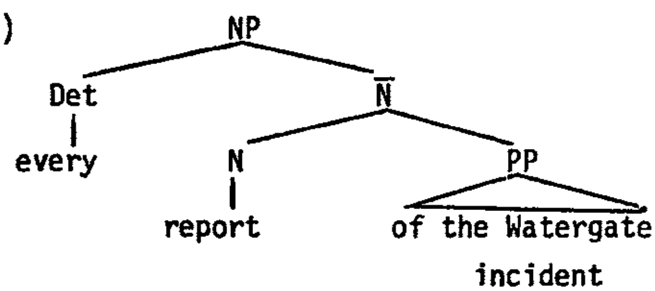
The whole question of the constituent structure of the English NP is an extremely vexed one. Many pages have been written about the problem, with very few convincing conclusions (see, for instance, Harris (1946, 1951), Wells (1947), Chomsky (1955, 1965, 1970), Newmeyer (1975), and Jackendoff (1976)). Frequently the evidence is simply contradictory and crucial judgements difficult to make. One reason for this, it seems to me, is that there is a certain amount of variation from speaker to speaker regarding the phrase structure configuration assigned to a given NP. This, I think, is precisely what's going on with the NP's in sentences like those in (3.3.48) and (3.3.49).

In some idiolects, an NP like every report of the Watergate incident has a constituent structure something like the one in (3.3.50)(a). For other idiolects, this same NP is assigned the structure in (3.3.50)(b).

(3.3.50)(a)

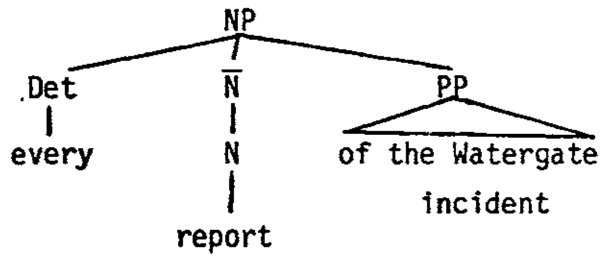


(b)



It's even conceivable that the structure in (3.3.51) exists for some idiolects.

(3.3.51)



This variation in constituent structure interacts with IDP. Assuming the structure in (3.3.50)(b), IDP will not prevent the derivation of (3.3.48) or (3.3.49), for $\bar{N}$, which immediately dominates the PP, is not a possible Relativization target. Assuming either the structure in (3.3.50)(a) or that in (3.3.51), however, the NP immediately dominating PP is itself a possible Relativization target, and derivations like those in (3.3.48) and (3.3.49) are therefore blocked by IDP. That various prepositions occur in different structures (e.g. *on* vs. *of*) in some idiolects is not at all implausible. Thus with respect to Relativization also, IDP can explain the same facts that RAOAP can explain, and perhaps even a few more.

Wh-Movement in questions is another interesting case in point. We will follow Bresnan in referring to the rule in question as Question Movement, though it may very well be the case that there is only one *wh*-Movement rule that operates in both questions and relatives, as Chomsky (1964, 1973, 1975a, 1975b) has suggested. Nothing hinges on this decision in what follows.

Bresnan argues that Question Movement applies to Noun Phrases, Adjective Phrases, Adverb Phrases and Quantifier Phrases that begin with an interrogative morpheme. She cites the following examples:

(3.3.52)(a) What book did you read?

(b) How long is it?

(c) How quickly did you read it?

(d) How much did it cost?

The target term of the rule is therefore formulated as...

(3.3.53) $[_{X^2}\ \underline{wh} - W_1\ ]$

Given this, Bresnan shows that RAOAP will account for such facts as the following:

(3.3.54)(a) How difficult a problem do you want to solve?

(b) *How difficult do you want to solve a problem?

(c) *How do you want to solve a difficult problem (≠ (a))

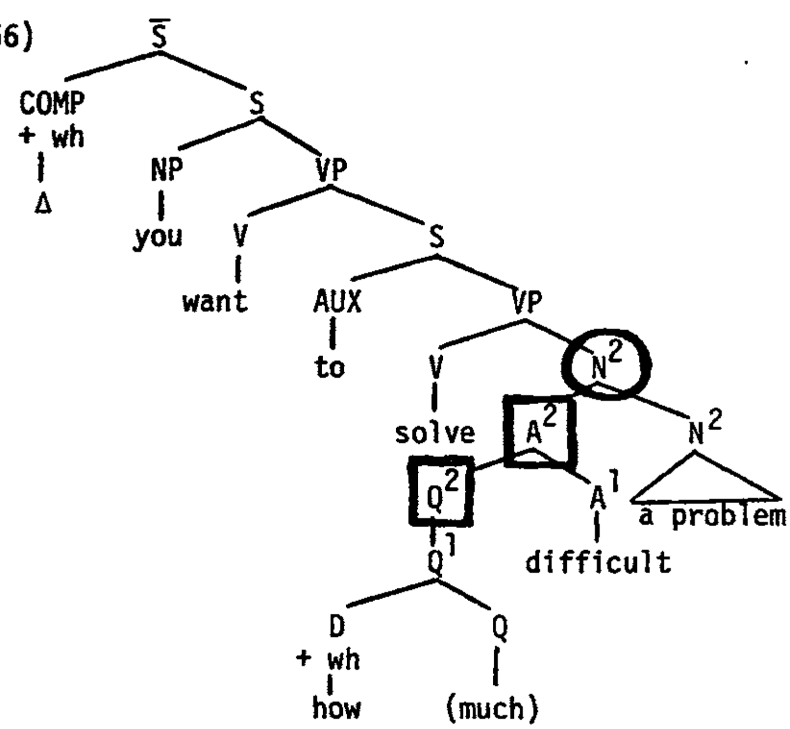
(3.3.55)(a) How many feet tall is the girl?

(b) *How many feet is the girl tall?

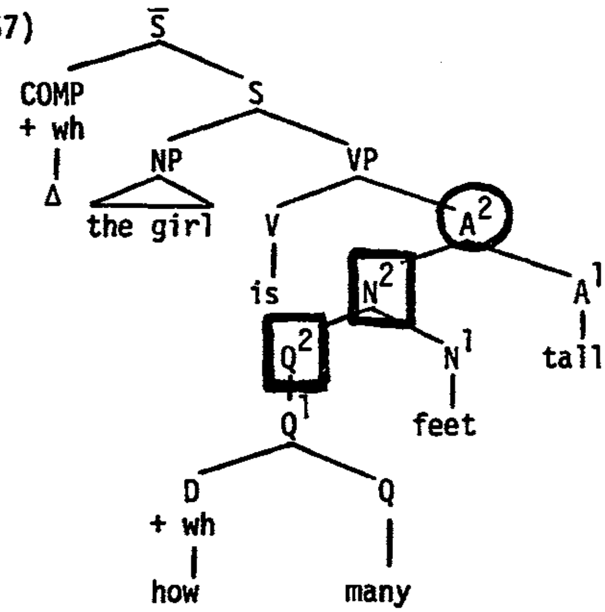
(c) *How many is the girl feet tall?

The relevant structures are the following ones. The boxed nodes are the targets that are ruled out by RAOAP. (3.3.54)(a) and (3.3.55)(a) result from moving (the terminal strings exhaustively dominated by) the circled nodes, which are the only targets allowed by RAOAP. [11]

(3.3.56)



(3.3.57)



Now granting all the premises of Bresnan's argument (i.e. granting her formulation of the target term and granting her constituent structures), it's clear that IDP accounts for precisely the same facts as RAOAP: Every inadmissible target in (3.3.56) and (3.3.57) is immediately dominated by another possible target.

One of Bresnan's claims, however, seems to me rather dubious, namely, the claim that Question Movement applies to Quantifier Phrases (Q^2). Her argument for this is based on the existence of sentences like (3.3.52)(d) above (<u>How much did it cost</u>?). But surely the phrase <u>how much</u> is an NP (it may be a Q^2 also). NP's can in general occur as the complement of verbs like <u>cost</u>:

(3.3.58)(a) It cost five dollars

(b) It cost an exorbitant amount of money.

(c) It cost a bundle.

True these NP's cannot passivize when the verb is <u>cost</u>, as is well-known (Chomsky (1965), Lakoff (1965), Bresnan (1976b)). Nevertheless, with other verbs how much is clearly passivizable, betraying its true NP nature:

(3.3.59)(a) How much was brought into the house?

(b) How much was taken by the crooks?

The following examples show that phrases like how much can also undergo There-Insertion and Raising.

(3.3.60)(a) How much was there?

(b) How much is likely to bother you?

This allows us to eliminate Quantifier Phrases as possible targets of Question Movement. Notice that this means that the ungrammaticality of (3.3.54)(c) and (3.3.55)(c) above is no longer to be explained by IDP. Rather, these examples are ungenerated because the movement target is a Quantifier Phrase, which Question movement does not apply to.

Further refinements are necessary, for PP's can also undergo Question Movement:

(3.3.61) In which cities did the Exorcist play?

We might then propose that the target term of Question Movement is the following, where X^2 is restricted to NP, PP, Adj. P, and Adv. P.

(3.3.62) $[_{X^2}$ (P) - wh - W_1]

Bresnan also considers the paradigm in (3.3.63).

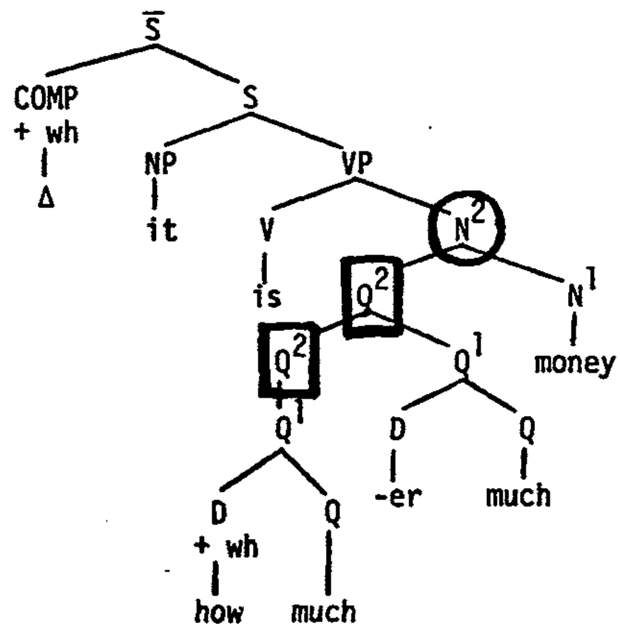
(3.3.63)(a) How much more money is it?

(b) *How much more is it money?

(c) *How much is it more money?

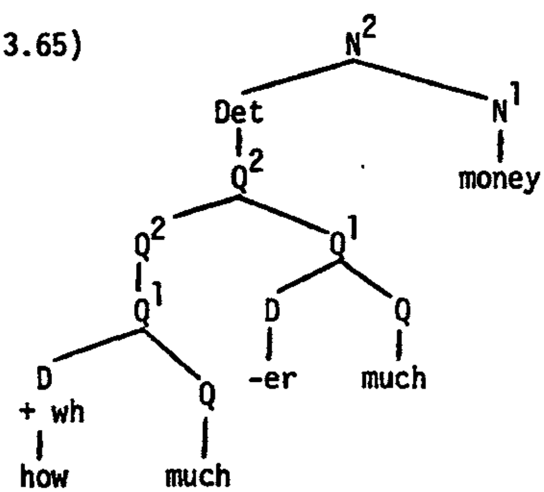
She gives the following constituent structure. RAOAP guarantees that only N^2 can be moved.

(3.3.64)



Now if our earlier suggestion that pre-nominal Quantifier Phrases are dominated by Det is correct, then N^2 should have the structure in (3.3.65) rather than the one that Bresnan indicates.

(3.3.65)



How then can IDP account for *(3.3.63)(b)? The answer is straightforward: In both (3.3.63)(b) and (3.3.63)(c) the target is a Quantifier Phrase; but

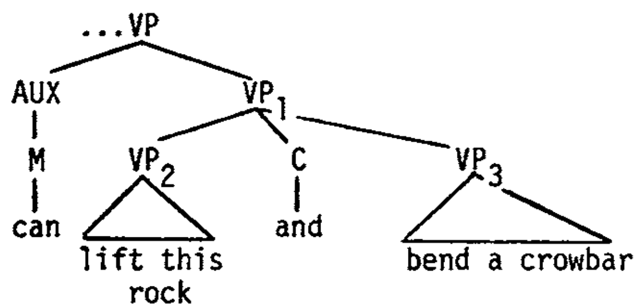
Question Movement does not apply to Quantifier Phrases, as we have seen.

There is one very important final point to be made in this section regarding the comparison of RAOAP and IDP. Given RAOAP and Bresnan's view of its interaction with the principle of recoverability of deletion, there is no explanation to be found for the observation made in Section 1.1 of Chapter One that VPD obeys the Coordinate Structure Constraint (more precisely Grosu's Conjunct Constraint). We observed the following example (due to Grosu):

(3.3.66) *I couldn't lift this rock, but I know a boy who can ∅ and bend a crowbar too. [∅ = lift this rock]

The relevant structure is given in (3.3.67).

(3.3.67)



Assuming either Bresnan's formulation of VPD or the one we developed in Chapters One and Two, VP_1 and VP_2 are both possible deletion targets. VP_1, however, is non-recoverable. On Bresnan's view, RAOAP maximizes only with respect to recoverable deletion targets, thus, ohne weiteres, *(3.3.66) will be generated. On our view, however, each constraint has an independent domain of application. Thus VP_2 is not a possible VPD target because of IDP, and VP_1 isn't either, because of the principle of recoverability of deletion.

In order to account for these facts, Bresnan must assume a Coordinate

Structure Constraint in addition to RAOAP. But surely giving an account of Conjunct Constraint violations is one of the central motivations for any A-over-A Principle. Thus unless we assume that recoverability of deletion functions independently from other constraints on rules (quite apart from the quesion of whether RAOAP or IDP (or either) is the correct formulation of the A-over-A Principle), we have no way of capturing the significant generalization that the various instances of illegitimate non-maximal proper analyses examined in this chapter and Conjunct Constraint violations are the same phenomenon.

It seems, in fact, that given IDP, we can eliminate the Coordinate Structure Constraint entirely. IDP, together with our view of its interaction with the recoverability of deletion, clearly accounts for all violations of the Conjunct Constraint. Moreover, Element Constraint violations, which, as Grosu (1973) and Ross (1967a) note, are not ill-formed in all languages (or even in English in some circumstances), quite probably have a different explanation entirely (perhaps subjacency, in the sense of Chomsky (1973)). Apart from these, the only troublesome cases involving coordinate structures that remain are those like the following.

(3.3.68) You and which army did you say would beat me up?

These are generated under Bresnan's proposal and under ours. Curiously, sentences like this do not seem to be ungrammatical, but rather they can only be interpreted as "echo" or "accosting" questions. A reasonable approach then would be to allow the syntax to generate these sentences freely. The interpretive rules that convert _wh_-words in COMP into quantifier-like entities ("questioners" in Hiż's (1962) sense) cannot apply to (i.e.interpret) a _wh_-word in a coordinate structure. Sentences containing no _wh_-words that are interpreted as "questioners" are always interpreted as "echo" questions (or "accosting" questions, or REF-questions (see Pope (1972) and Chomsky (1973)). Sentences like (3.3.68) then present no particular difficulties for our analysis.

As this section has been somewhat of a digression, a few summary remarks are perhaps in order. We have looked more closely at Bresnan's Relativized A-over-A Principle and found it to be unsatisfactory in certain respects. We have argued on empirical grounds that a different constraint, which we have called The Immediate Domination Principle is to be preferred. Detailed justification for this claim has been given on the basis of facts from various languages and, in particular, on the basis of such phenomena in English as relativization, wh-movement in questions, Complex NP Shift, Verb Phrase Deletion, and deletion in comparative clauses.

3.4 A Formulation of Gapping

In this section we will offer a formal treatment of Gapping that is in keeping with all (or almost all) the observations made in the first two sections of this chapter, and which is not subject to the criticisms we have raised against previous proposals. Recall that we have accepted Kuno's proposal that there are independent, primarily perceptual factors which drastically affect the acceptability of Gapped clauses. The formulation we will propose will consequently overgenerate to a considerable extent.

Summarizing previous observations now, we have seen first (in Section 3.2.2) that left-peripheral ellipsis should not be collapsed with Gapping. Secondly, as we noted in Section 3.2.4, any attempt to restrict the deletion target to something other than an X-variable encounters insurmountable difficulties. Thirdly, certain cases of right-peripheral ellipsis seem to be possible only when (sentence-internal) Gapping has applied, suggesting that these be accounted for by adding a second deletion target to the structural description of the Gapping rule (this was also discussed in 3.2.4). A good first approximation to the rule then would be the following: [12]

$$(3.4.1)\quad A_1 - X - B_1 - Y - \left\{\begin{matrix}\text{and}\\ \text{or}\end{matrix}\right\} A_2 - \underset{\emptyset}{\underset{\Downarrow}{X}} - B_2 - \underset{\emptyset}{\underset{\Downarrow}{Y}}$$

Now unlike previous discussions, which have been concerned with imposing restrictions on the deletion target(s), we will be concerned here with a further specification of the context terms, i.e. the Gapping remnants. Recall that in Section 3.1 we observed that the remnant to the right of (the first) gap (i.e. B_2 in (3.4.1)) must be either an NP, an Adjective Phrase, an Adverb Phrase, a That-clause, a for-to clause, or a Prepositional Phrase. The examples were sentences like these:

(3.4.2)(a) Peter loves Betsy, and Betsy $\emptyset$ Peter (B_2 = NP)

(b) Alan seemed more reluctant than Peter, and Peter $\emptyset$ more reluctant than Betsy. (B_2 = Adj. Phrase)

(c) Tom ran extremely quickly, and Alex $\emptyset$ more slowly than anyone I'd ever seen. (B_2 = Adv. Phrase)

(d) Alan claimed that he was cheated , and Sandy $\emptyset$ that she was the one who cheated him. (B_2 = that-S)

(e) Alan prefers for Tom to do it, and Sandy $\emptyset$ for Alan to do it. (B_2 = for-to clause)

(f) Betsy stood in left field, and Sandy $\emptyset$ in right field. (B_2 = PP)

We also cited Ross's (1967) observation that Gapping only an auxiliary leads to unacceptability:

(3.4.3) *?He may stay inside, and she ∅ go to the beach.

We noted that such examples may be viewed as showing that B_2 cannot be VP. Note furthermore that all the examples we have cited in this chapter are in conformity with the pattern of (3.4.2) -*(3.4.3).

Now it could be the case that we simply have to list an arbitrary disjunction of possiblities for B_2 in the structural description of our Gapping rule, say...

(3.4.4) $$\ldots\left\{\begin{matrix}\text{and}\\ \text{or}\end{matrix}\right\} A_2 - X - \left\{\begin{matrix}\text{Adv P}\\ \text{NP}\\ (\overline{\text{S}})\\ \text{Adj P}\\ \text{PP}\end{matrix}\right\} - Y$$

but surely it would be preferable to have some way of referring to this collection of categories (presuming, of course, that it is in some sense a "natural class"). One very obvious condidate that comes to mind in this regard is the $\overline{X}$ -theory.

There has been a considerable amount of research trying to work out an $\overline{X}$-theory for the syntax of several languages (but primarily English). There has actually been very little agreement, however, as to how many levels should be posited, and what the actual feature system that projects syntactic categories should be (see for example Harris (1946, 1951), Chomsky (1970), Bowers (1969), Selkirk (1970), Jackendoff (1974, 1976), Hornstein (1975), Halitsky (1975), and Bresnan (1976a)). One clear objective of the $\overline{X}$-theory, which everyone would agree to, I think, is to solve a problem which, as Bresnan points out, was observed by Lyons (1968: 330-332) who writes:

> ...As far as the formalization of phrase-structure grammars is concerned, it is a matter of 'accidental' coincidence that linguists will include in their grammars of different languages rules which always expand NP into a string of symbols containing N and rules which always expand VP into a string of symbols containing V... what is required, and what was assumed in traditional grammar, is some way of relating sentence-constituents of the form XP to X (where X is any major category: N, V, etc.).

One solution to the Lyons problem is the one offered by Chomsky (1970) who suggests that the base rules introducint N, A, and V should be replaced by a schema like the following.

(3.4.5) $X^n \rightarrow [\ \text{Spec},\ X^{n-1}\ ]\ X^{n-1}$

Chomsky proposes the following sketch of the initial rule of the (English) base grammar.

(3.4.6) $S \rightarrow N^2 - V^2$

X^2, in fact, seems to be the maximum $\overline{X}$ level that can be motivated. The various proposals that have been made which involve "higher" expansions than this end up proliferating nodes without justification in the expansion of minor categories (Jackendoff's (1976) Det^3 and Deg^3, for instance). This is not a desirable result. We will therefore adopt Chomsky's "2-bar" system here.

Notice that it would be perfectly natureal to treat $\overline{S}$ (in Bresnan's (1972) sense) as S^2, and S as S^1. This is in keeping with the fact that S has no lexical categore, i.e. has no X^0 expansion. We will thus follow Harris (1946, 1951) in distinguishing S from the X-bar system (though we do not follow Harris in the number of exponents). The base grammar will then contain the following two rules.

(3.4.7) $S^2 \rightarrow$ COMP - S^1 ...

(3.4.8) $S^1 \rightarrow N^2 - V^2$ ...

We might further propose to accomodate the analyses of the Verb Phrase we proposed in Chapter One by treating the "highest VP" (the VP which immediately dominates that AUX which expands to tense (M)(have-en)) as V^2, and the other VP's as recursive V^1. This would require phrase-structure rules like the following:

$$(3.4.9)\quad \left.\begin{array}{l} V^2 \rightarrow \text{AUX} - V^1 \\ V^1 \rightarrow \text{AUX} - V^1 \end{array}\right\} \text{ or } \left\{ V^{1,2} \rightarrow \text{AUX} - V^1 \right.$$

$$V^1 \rightarrow V - (\text{NP}) - (\text{PP}) \ldots$$

VPD would then presumably delete only V^1.

This phrase structure system will also allow us to capture the generalization we need for term B_2 of our Gapping rule. The only possible B_2's are X^2's, i.e. the major phrasal categories. Accepting this, we will modify the structural description of the rule in the way sketched in (3.4.10) (to avoid confusion, we will change the target variables to W_1 and W_2).

$$(3.4.10)\quad \ldots\left\{\begin{matrix}\text{and}\\\text{or}\end{matrix}\right\}\ A_2 - \underset{\substack{\Downarrow\\ \emptyset}}{W_1} - X^2 - \underset{\substack{\Downarrow\\ \emptyset}}{W_2}$$

Term A_2 must also be restricted. In almost all the examples of Gapping we have considered, A_2 is an NP. However we have also seen cases like the following ones, where a preposed Adverb or PP serves as A_2.

(3.4.11)(a) At our house, we play poker, and at Betsy's house, ∅ bridge. [A_2 = PP]

(b) Yesterday we went to the movies, and last Thursday, ∅ to the circus. [A_2 = Adv (P)]

It would seem reasonable then to replace A_2 by X^2 as well. However it may be necessary to restrict this term so that it cannot analyze S's, for as we saw in Section 3.2.2 above, Gapping does not seem to be possible when the second clause contains a sentential subject:

(3.4.12) *That Harry is a fool bothers Dick, and that Bill is a fool ∅ Sam.

Alternatively, this restriction may follow from a more general constraint, say, the one argued for by Zaenen and Pinkham (to appear), who observe that clauses containing sentential subjects are "islands" with respect to many rules. We will not resolve that matter here, except to observe that sentences like (3.4.12) are not ill-formed for all speakers, suggesting that appealing to a more general constraint may be inappropriate.

Replacing A_2 by X^2, we now have this revision of the rule:

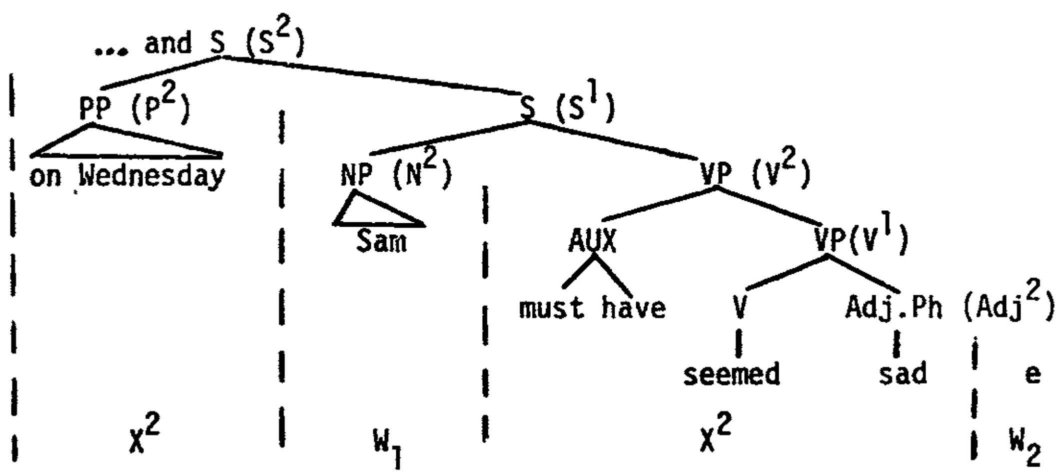
(3.4.13) $\ldots\left\{\begin{matrix}\text{and}\\\text{or}\end{matrix}\right\} - X^2 - \underset{\emptyset}{\underset{\Downarrow}{W_1}} - X^2 - \underset{\emptyset}{\underset{\Downarrow}{W_2}}$

Notice that the existence of the preposing cases (like those in (3.4.11)) force us to impose a further restriction on term B_2. That is, if the only candidates for term A_2 were subject NP's, then no further modification would be necessary to account for the impossibility of Gapping just AUX's. Presuming that only the highest VP is a V^2, if term B_2 were to analyze V^2, it would necessarily be the case that the constituents analyzed as A_2 and B_2 were adjacent, i.e. that there was no deletion target. Once we subsume the preposing cases under Gapping, however, something must be done to prevent cases like this:

(3.4.14) *On Tuesday, Sam must have seemed happy and on Wednesday ∅ must have seemed sad.

That (3.4.14) can in fact be generated by the formulation in (3.4.13) is shown more clearly in the following tree diagram.[13]

(3.4.15)



Thus the second X^2 in (3.4.13) must be restricted so that it cannot analyze V^2. This would presumably be done by means of syntactic features, but we will not pursue that matter here. The proposed revision, however, must guarantee that the only good Gapping output from the structure underlying *(3.4.14) is (3.4.16).

(3.4.16) On Tuesday, Sam must have seemed happy, and on Wednesday ∅ sad.

There is only one refinement left to be made, I believe. If we are to maintain our hypothesis that all deletion rules apply after the level of shallow structure (see the discussion in Chapter Two), then Gapping is certainly non-cyclic. Therefore sentences like the following, where Gapping applies entirely within an embedded cyclic domain, are of considerable interest.

(3.4.17)(a) Bill said that Betsy played shortstop, and Alan ∅ 1st base. [∅ = played]

(b) That Alan played 1st base and Betsy ∅ shortstop, is not surprising.

These sentences show that our formulation of Gapping must be revised so as to include "end-variables" (on both ends). Let us then formulate Gapping as follows:

(3.4.18) Gapping (first formulation)

$$\begin{array}{lcccccccccc} W_3 & - [_S X^2 & - W_1 & - X^2 & - W_2] & - \left\{\begin{matrix}\text{and}\\ \text{or}\end{matrix}\right\} & - [_S X^2 & - W_1 & - X^2 & - W_2] & W_4 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ \Rightarrow 1 & 2 & 3 & 4 & 5 & 6 & 7 & \emptyset & 9 & \emptyset & 11 \end{array}$$

It's also worth pointing out that the existence of sentences like those in (3.4.17) in the context of the proposed analysis bears on the proper formulation of the "Strict Cycle Condition" (Chomsky (1973)).[14] Certain rules (perhaps only non-cyclic rules) must be allowed to apply so as to "affect" only a proper subdomain of a cyclic domain which has already been "passed". Similar examples can be constructed with VPD, Sluicing, and other deletion rules to make the same point. This matter is of considerable interest, but we will not pursue it here.

Now that we have given a precise formulation of the Gapping rule, let us turn to the matter of its interaction with general constraints on rules.

Consider first the following examples, similar to cases we noted in Section 3.1.

(3.4.19)(a) Beth left after the first act, and Norma $\emptyset$ after the second act. [$\emptyset$ = left]

(b) *Beth left after the first act, and Norma $\emptyset$ the second act. [$\emptyset$ = left after]

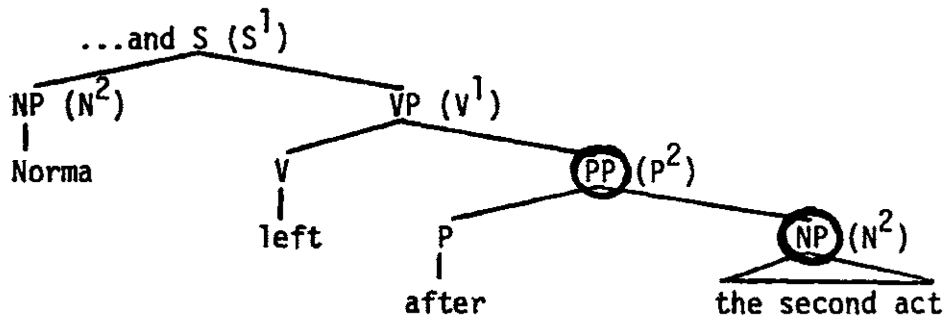
(3.4.20)(a) Jim was hassled on Winthrop Street, and Norma $\emptyset$ on Hooker Street. [$\emptyset$ = was hassled]

(b) *Jim was hassled on Winthrop Street, and Norma $\emptyset$ Hooker Street. [$\emptyset$ = was hassled on]

If Gapping deletes an X-variable, as we have argued, and if the remnants to the right of the gap in all of these examples are X^2's (i.e. NP's or PP's), as is surely the case, why are only the (a) examples grammatical?

The Immediate Domination Principle developed in the last section provides an answer to this question. Assuming the structure in (3.4.21), NP cannot be analyzed as term 9 of our rule (3.4.18), because PP is also thus analyzable, and PP immediately dominates NP, invoking IDP.

(3.4.21)



For some speakers, the Gapping of certain prepositions along with a preceding verb is marginally possible, say, in examples like the following:

(3.4.22) %*Myra talked to Harry, and Sandy ∅ Betsy.

[∅ = talked to]

As we noted earlier, this seems to correlate with the "lexicizability" of these sequences of V - P, i.e. the possibility of analyzing such sequences as: [$_V$ V - P]. Notice that this parallels pretty closely the "lexicizability" of such sequences with respect to other rules, like Passive. Thus while the passive in (3.4.23) is possible, the sequence V-*after* never allows passivization, as (3.4.24) illustrates.

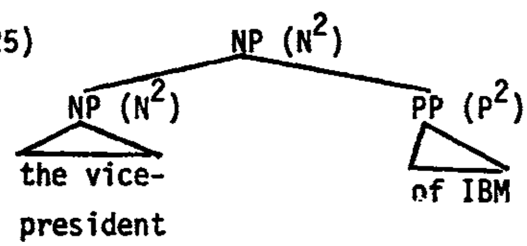
(3.4.23) Harry was talked to (?by Myra).

(3.4.24) *The first act was left after (by Norma).

If such sequences are indeed lexicizable then there should be the possibility of Gapping in (3.4.22), for IDP will not be violated.

Consider next Noun Phrases like *the vice-president of IBM*. These arguably have the constituent structure shown in (3.4.25).

(3.4.25)



One argument for this structure would be the grammaticality of sentences like the following:

(3.4.26) The former vice-president and the current secretary-treasurer of IBM met in Peoria.

Here there is no obligatory pause before the PP: *of IBM*, hence a derivation via Right-Node-Raising, (that is a derivation from a source containing a second identical PP which gets deleted) which is always associated with such pauses, seems untenable. This sentence seems best regarded as base conjunction. But the conjoined entities seem to allow the full array of NP-recursion (the sequence Det-Adj-N), suggesting the structure in (3.4.25).

If this structure is correct, we can explain the following contrasts <u>re</u> Gapping.

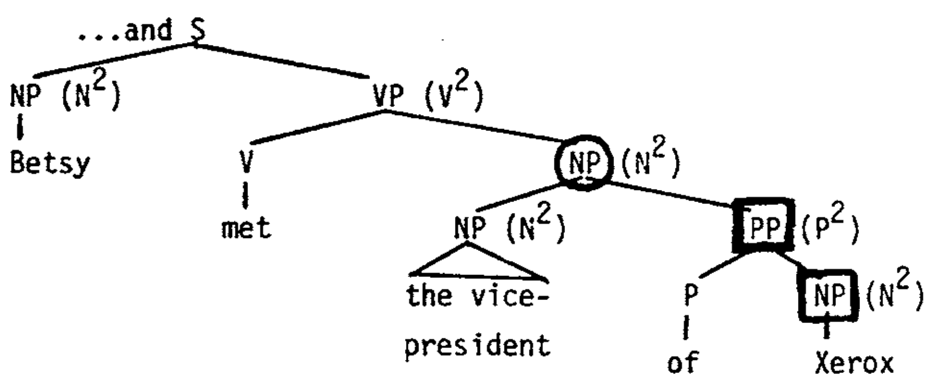
(3.4.27)(a) John met the vice-president of IBM, and Betsy ø the vice president of Xerox. [ø = met]

(b) *John met the vice-president of IBM, and Betsy ø of Xerox. [ø = met the vice-president]

(c) *John met the vice-president of IBM, and Betsy ø Xerox. [ø = met the vice-president of]

The structure prior to Gapping is the one given in (3.4.28).

(3.4.28)



There are three potential right-remnants in this example: NP, $\boxed{\text{PP}}$, and $\boxed{\text{NP}}$. Since they are in an immediate-domination relationship, however, IDP correctly predicts that Gapping can apply only so as to leave NP behind, thus accounting for the facts of (3.4.27). It should be noted that exactly analogous facts hold with respect to Pied-Piping in relativization. That is, the following facts are also accounted for by IDP, assuming roughly Bresnan's analysis and the structure in (3.4.28) (see the discussion in Section 3.3 above).

(3.4.29)(a) We drove past Xerox, the vice-president of which Betsy knows very well.

(b) *We drove past Xerox, of which Betsy knows the vice-president very well.

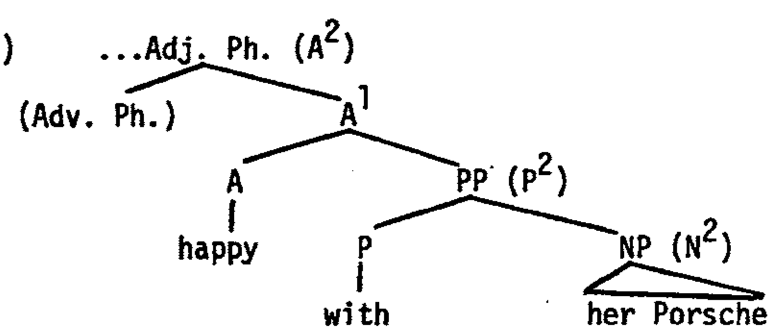
Interestingly, Adjective Phrases (A^2's) seem not to immediately dominate PP adjuncts. Given a sequence: Adj. - PP in an appropriate input to Gapping, the sequence itself (i.e. the A^2) or just the PP are both possible right-remnants, as shown by the following pair of examples.

(3.4.30)(a) Carol was happy with her Oldsmobile, and Margie ø upset with her Porsche.

(b) Carol was upset with her Oldsmobile, and Margie ø with her Porsche.

These facts suggest the following constituent structure for Adjective Phrases.

(3.4.31)



Since A^2 does not immediately dominate P^2 in (3.4.31), term 9 of our Gapping rule can freely choose either A^2 or P^2. IDP does not allow N^2 in (3.4.31) to be so analyzed, however, because P^2, another possible value for term 9, immediately dominates it. This correctly accounts for the ungrammaticality of the following example.

(3.4.32) *Carol was happy with her Oldsmobile, and Margie ø her Porsche.

Recall that we observed in Section 3.1 of this chapter that occasionally there are triple-output possibilities when Gapping applies. We noted these examples (due to Ross):

(3.4.33) John tried to begin to write a play, and Harry

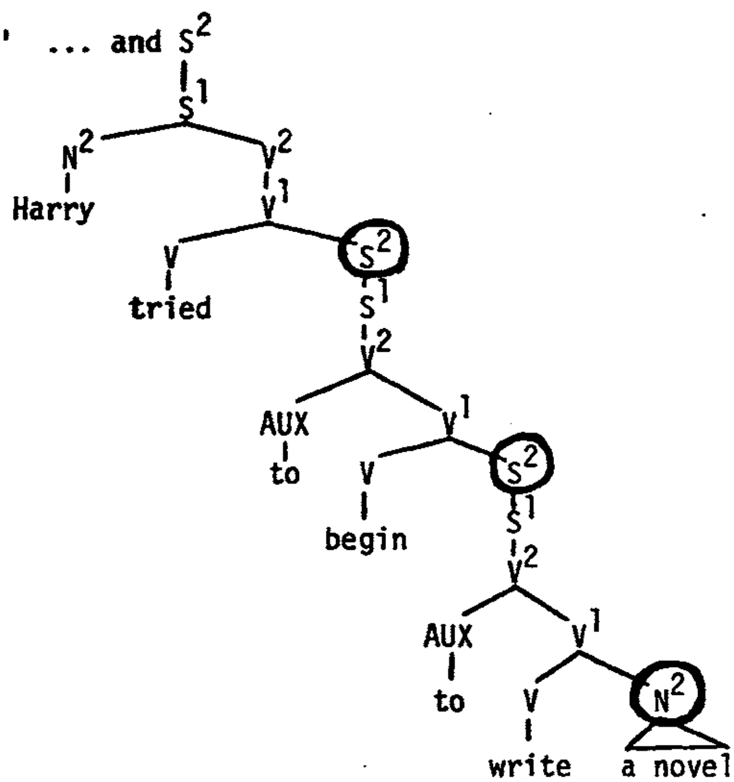
(a) ø to begin to write a novel

(b) ø to write a novel

(c) ø a novel

A fairly standard view of the syntax of examples like these would involve a pre-deletion structure like the one sketched in (3.4.33).

(3.4.33)'



Any of the circled nodes in this structure is a potential value for term 9 of our Gapping rule.[15] Moreover, no proper analyses are ruled out by IDP. Therefore all three sentences in (3.4.32) are legitimate instances of Gapping.

Our analysis also accounts for why those sentences are the only possible Gapping outputs, i.e. for why none of the following examples is possible.

(3.4.34) *John tried to begin to write a play and Harry

(a) ∅ begin to write a novel

(b) ∅ write a novel

(c) ∅ novel

In each of these cases, something other than an X^2 has been left to the right of the gap, which is unallowed, given our formulation of the rule.

The fact that term 9 of our Gapping rule must analyze X^2's also accounts for why complementizers cannot be Gapped along with a preceding verb, as the following contrast, which was also observed earlier, shows.

(3.4.35) Sandy said (that) he was a fool, and Betsy

(a) ∅ that he was out of his mind.

(b) *∅ he was out of his mind.

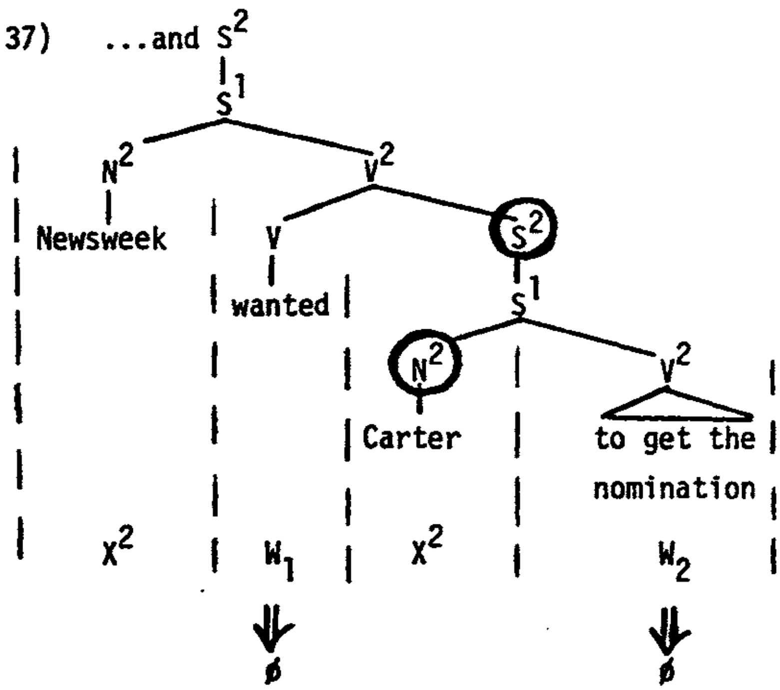
The <u>that</u> clause is an S^2, and hence a possible value for term 9, but the same clause without the <u>that</u> is an S^1, and hence not a possible right-remnant. These facts, though interesting, may not provide any further support for our analysis, for there may very well be an independent principle requiring the presence of complementizers in such situations, as Lasnik and Chomsky have suggested.

Thus far we have considered only cases of clause-internal ellipsis. Both deletion targets will be utilized in the derivation of sentences like this next one.

(3.4.36) Time wanted Udall to get the nomination, and Newsweek $\emptyset_1$ Carter $\emptyset_2$.

Such sentences are unproblematic, as the following analyzed tree diagram shows:

(3.4.37)



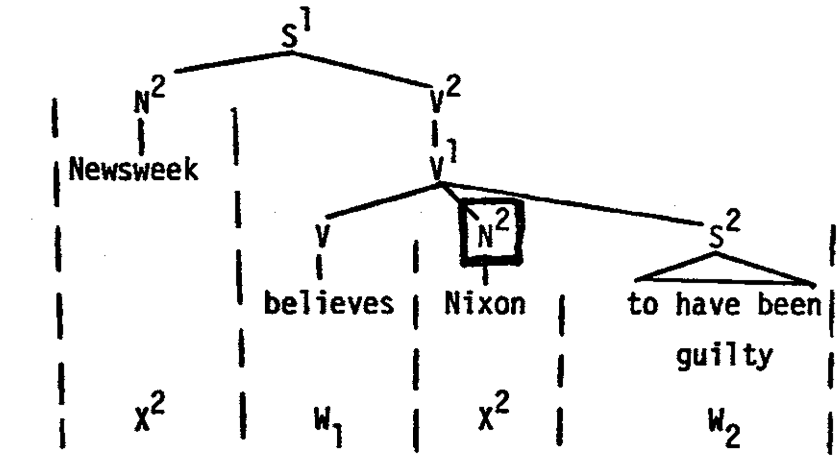
Since S^2 does not immediately dominate N^2, IDP does not rule out the indicated proper analysis.

Note further that similar cases with <u>believe</u> pose no particular problem for our analysis either:

(3.4.38) Time believes Agnew to have been guilty, and Newsweek $\emptyset_1$ Nixon $\emptyset_2$.

If, as Chomsky argues, there is no rule of Raising to Object Position, then the structure for such sentences will be just as in (3.4.37). Alternatively, if the derivation of sentences like (3.4.38) does involve Raising, then the pre-deletion structure will be that in (3.4.39).

(3.4.39) ...and S^2



$\boxed{N^2}$ is not immediately dominated by any X^2, hence IDP is not in force. Thus whether or not there is a rule of Object-Raising, our Gapping rule and IDP interact to make the correct predictions in all these cases.

We are also in a position to explain an observation of Jackendoff's mentioned in Section 3.1, namely, the contrast between these next two examples.

(3.4.40) Did Bill eat the peaches, or Harry ∅ the grapes?

(3.4.41) *Did Bill eat the peaches, or did Harry ∅ the grapes?

The source for these examples is the one underlying (3.4.42) which, of course, results if no Gapping takes place.

(3.4.42) Did Bill eat the peaches, or did Harry eat the grapes?

Left-peripheral ellipsis, which we have argued is a separate rule (the same rule as Conjunction Reduction) can now apply to yield (3.4.43).

(3.4.43) Did Bill eat the peaches, or Harry eat the grapes?

Only after this rule has applied is the structural description of Gapping met, for the X^2 that is the first Gapping remnant (i.e. the subject NP) must be adjacent to the conjunction. Therefore only (3.4.40) will be generated. (3.4.41), where <u>did</u> stands between <u>or</u> and the N^2 <u>Harry</u>, is not a suitable input for Gapping.

Thus we need not accept Jackendoff's alternative solution of ordering Gapping before Subject-Auxiliary Inversion. This is a very desirable result with respect to our ordering hypothesis, for Subject-Auxiliary Inversion, which clearly affects semantic interpretation (see Klima (1964), Lasnik (1972), Liberman (1974)), must apply before the level of shallow structure. Gapping applies after the level of shallow structure.

A similar explanation obtains with regard to contrasts of the following type, pointed out by Fiengo (1974).

(3.4.44) Betsy said that Alan went to the ballgame, and Betsy ∅ to the movies. [∅ = went]

(3.4.45) *Betsy said that Alan went to the ballgame, and that Betsy ∅ to the movies. [∅ = went]

When the complementizer <u>that</u> stands between the conjunction and the first Gapping remnant, the structural description of our Gapping rule is not met.

We have saved for last a discussion of those cases where more than two remnants can be left behind in a Gapped clause. The relevant cases, the reader will recall from Section 3.1, are those like the following.

(3.4.46)(a) Peter talked to his boss on Tuesday, and Betsy ∅ to her supervisor on Wednesday.

(b) John talked to his supervisor about his thesis, and Erich ∅ to the dean about departmental policies.

These examples, where two PP's occur to the right of the gap, are by far the best of the multiple-remnant possibilities. Still, many people feel reluctant to reject examples like the next one, where a sequence NP-PP occurs to the right of the gap.

(3.4.47) Charlie entered the bedroom at 5:30, and Vera ø the kitchen at 6:00.

All my informants, however, agree with Jackendoff in finding a contrast between (3.4.47) and (3.4.48).

(3.4.48) ?*Willy put the flowers in a vase, and Charlie ø the book on the table.

The acceptability of such sentences seems to depend on whether or not the third remnant is dominated by the VP. It seems reasonable to assume that PP's like the ones in (3.4.47) and (3.4.46a) are at least sometimes generated as daughters of S. The PP's in (3.4.48) on the other hand, are strictly sub-categorized by the verb, and, moreover, obligatory. It is therefore usually assumed that they are always dominated by VP (V^1).

Notice that this difference correlates with the possibility of leaving the PP behind after VPD:

(3.4.49)(a) Charlie didn't enter the bedroom at 5:30, but he did ø at 6:00.

(b) ?*Willy didn't put the flowers in a vase, but he did ø on the table.

Furthermore, <u>about</u>-Phrases like the one in (3.4.46)(b) pattern with the time adverbs in this regard:

(3.4.50) John didn't talk to his supervisor about his thesis, but he did ø about departmental policies.

What all this suggests is that the possibility of some constituent being a third remnant in a Gapped clause correlates with the possibility of its being outside of VP (V^1).

A reasonable way to account for such facts, it seems to me, is to allow our Gapping rule to generate any number of right remnants, and to rule out the unacceptable outputs independently, say, via surface constraint. This is further supported by examples like the following, where even more than two remnants to the right of the gap are possible.

(3.4.51) Betsy dances with a parasol in the living room on Fridays, and Peter ø with a meat cleaver in the bar on Saturday nights.

Let us therefore modify our Gapping rule as follows, where $[X^2]^*$ can be expanded as any number of X^2's. [16]

(3.4.52) Gapping (final formulation) [17]

$$\begin{array}{lccccccccccc} & W_3 & - [_S\ X^2 & - W_1 & - [X^2]^* & - W_2\ ] & - \{^{\text{and}}_{\text{or}}\} & - [_S\ X^2 & - W_1 & - [X^2]^* & - W_2] & - W_4 \\ & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ \Rightarrow & 1 & 2 & 3 & 4 & 5 & 6 & 7 & \emptyset & 9 & \emptyset & 11 \end{array}$$

Much of the overgeneration of this rule can be ruled out by a surface constraint of the sort illustrated in (3.4.53)

(3.4.53) $*[_S\ X^2\ [_{V^1}\ X^2 - C^*]\ ]$, where C* stands for any sequence of constituent (within V^1 (VP)). As previously mentioned, judgements vary considerably among speakers. Thus some people accept sentences like (3.4.54), whose Gapped right conjunct has the structure: $[_S\ N^2\ [_{V^1}\ N^2\ S^2]\ ]$.

(3.4.54) %Janis told Alan that he was crazy, and Betsy ø Peter that he should go take a bath.

Our Gapping rule generates such sentences. The idiolectal variation is handled by positing slightly different surface constraints. This, I would claim, is a very natural way to handle what seem to be completely idiosyncratic differences in acceptability judgements.

We have now developed a formulation of Gapping that accounts for various observations made in the preceding sections of this chapter. Our Gapping rule interacts with the Immediate Domination Principle developed in Section 3.3 in a principled fashion. The proposed analysis overgenerates considerably, which, we have argued, is a desirable result. Many of the sentences generated by our Gapping rule will be ruled out by the (perceptual) constraints proposed by Kuno (see Section 3.2), or else by surface constraints. The extreme variation among speakers with regard to the acceptability of Gapped sentences provides strong justification for the view we have developed.

3.5 Gapping and Logical Form

In Chapter Two we put forth a hypothesis as to the nature of the recoverability of deletion. Our argument was that identity of logical form (a notion which we gave considerable substance to) was sufficient, in and of itself, to determine the applicability of VPD. In this section we will justify having couched that entire discussion in extremely general terms by giving an outline of our view of the general interaction of Logical Form and Gapping.

Gapped sentences always involve contrast. In virtually all cases, this is reflected prosodically, with placement of contrastive accent not only on remnants in the Gapped clause, but also on corresponding entities in the immediately preceding conjoined clause (as was noted in Section 3.1). This has usually been regarded as an incidental fact. Hankamer (1971) makes the obviously correct observation that it must be the case that rules that assign accent must apply before ellipsis rules. Kuno (1976) envisages a general functional principle that remnants in elliptical clauses must be new (or contrastive) information.

The prosodic facts seem to me to be far from incidental. As we noted in Section 2.2 of Chapter Two, there is a certain amount of evidence that contrastive accent drastically alters the logical form of a sentence, as many people have suggested (see for instance Wilson (1926)). Dretske (1972) makes much the same point, arguing, rather convincingly in my opinion, that the presence and position of such accents can alter the truth conditions of a given sentence. [18]

To take just one of Dretske's cases, consider the following contrast.

(3.5.1) If Clyde hadn't made the last shot, we would have lost.
(neutral intonation)

(3.5.2) If Clýde hadn't made the last shot, we would have lost.

Imagine a situation in which Basketball Player A utters (3.5.1) to Basketball Player B after their team has just won a game (121-120) owing to their teammate Clyde's incredible hook shot from mid-court in the last two seconds of play. In such a situation we are likely to concede that A's utterance is true. If Player B were to respond with (3.5.3)...

(3.5.3) Oh, I don't agree. We would have won no matter which of us had made the last shot.

we would accuse him of purposely having distorted what A had said. (3.5.3) is not a legitimate denial of (3.5.1). If Player A had uttered (3.5.2), on the other hand, then B's response would indeed count as a denial. In the situation we have imagined, then, (3.5.1) is true (or so we may assume), and (3.5.2) is false.

Similarly, we may want to consider well-formedness conditions on question-answer pairs as a matter of logical form; accentual position is clearly relevant here as well. Thus, just as discourses like the one in (3.5.4) are anomalous, so are similar discourses where the accent is simply in the wrong place, as (3.5.5) shows.

(3.5.4) Q: What did the strike force you to call off?

#A_1: What forced us to call off the lecture was the strike.[19]

#A_2: It was the strike that forced us to call off the lecture.

(cf. What the strike forced us to call off was the lecture).

(3.5.5) Q: What did the strike force you to call off?

#A_1: The strike forced us to call off the lecture.

#A_2: The strike forced us to call off the lecture., etc.

(cf.The strike forced us to call off the lecture)

A standard view of questions and corresponding well-formed direct answers is that they must share an appropriate presupposition. Without becoming mired in the obscurities surrounding the entire notion of presupposition, we will simply point out that a standard view of the presuppositions involved in the pseudo-clefts of (3.5.4) is that they are logical presuppositions rather than pragmatic (see Keenan (1971b), Keenan and Hull (1972), but also Soames (1976)). If pseudo-clefts and direct questions enter into well-formed question-answer pairs by virtue of shared logical presuppositions, then the fact that contrastive statements exhibit the same behavior suggests that contrast also is a matter of logical presupposition.

This conclusion is further supported by the observation made by Higgens (1974) (and perhaps others) that in pseudo-cleft sentences the grammatical focus (i.e. the element in post-copular position) is the only natural recipient of focal accent:

(3.5.6)(a) What Harry ate was the bagel.

(b) #What Harry ate was the bagel.

(c) #What Harry ate was the bagel.

(3.5.6)(b) and (c) are not ungrammatical, of course. However, those cases seem best regarded as what Pike (1945) calls "hypostasis". That is, they are almost necessarily corrective, on a par with cases like (3.5.7).

(3.5.7) I didn't say [ɛ]conomics, I said [iy]conomics!

Contrastive accent in cases like this, I think it's fair to say, is an isolable phenomenon, quite apart from the other cases we have seen ((3.5.6)(a) included) where it's arguably the case that logical matters are involved.

What then is the logical form of contrastive sentences like those in (3.5.5) above? Chomsky (1975b) has suggested Russell's iota-operator as an appropriate device for representing the logical forms of sentences with focal accent. A sentence like (3.5.8), under Chomsky's proposal, would have a logical representation roughly as indicated.

(3.5.8) Betsy loves Peter.

(3.5.8)' ιx(x love Peter) = Betsy

(3.5.8)' is to be interpreted in the standard Russellian way, i.e. as: 'There is one and only one x such that x loves Peter and that x is Betsy.'

This is not quite right, however. As Baker (1968) notes, sentences with with focal accent differ from corresponding (it-) cleft sentences precisely in that the latter necessarily imply uniqueness, whereas the former do not. Baker offers examples like the following.

(3.5.9) #It was the play that Betsy went to. She went to the movie, too.

(3.5.10) Betsy went to the play. She went to the movie, too.

The discourse in (3.5.9) is contradictory. The cleft-sentence requires that the play be the unique element (in the relevant domain of discourse) that is such that Betsy went to it. (3.5.10) is non-contradictory, because sentences with focal accent do not have this requirement of uniqueness. Therefore the iota-operator is not really the device we want to represent the logical form of sentences with focal accent. If we accept Chomsky's argument about the relation of such sentences to sentences with overt quantifier expressions, however (see the discussion in Section 2.2 of Chapter Two), then we do want to represent focal accent by means of some kind of variable-binding operator at the level of logical form.

This can be accomplished very easily, it seems to me, by a rather straightforward use of set abstraction. Recall that in Section 2.1 of Chapter Two we proposed to represent the logical form of pseudo-clefts in this way also. A sentence like (3.5.6)(a) above (What Harry ate was the bagel), we gave a representation like the following.

(3.5.11) $\{\text{the bagel}\} = \hat{x}[\text{Harry}, \lambda y(y \text{ ate } x)]$

Notice that this logical form correctly indicates that the bagel is the only thing John ate (in the relevant domain of discourse). That is, (3.5.11) is a plausible logical form for (3.5.6)(a), because a discourse like the following one is contradictory.

(3.5.12) #What John ate yesterday was the bagel. He also ate the banana yesterday.

A very simple modification of (3.5.11) will give us an intuitive representation for (3.5.13), thus accented.

(3.5.13) Harry ate the bagel.

We can accomodate Chomsky's hypothesis about variable binding and at the same time avoid the problem of non-uniqueness by representing (3.5.13) as in (3.5.14).

(3.5.14) $\{\text{the bagel}\} \subset \hat{x}[\text{Harry}, \lambda y(y \text{ ate } x)]$

When the subject is accented, as in (3.4.15), a different set is being picked out. The logical form in (3.5.16) reflects this.

(3.5.15) Harry ate the bagel.

(3.5.16) $\{\text{Harry}\} \subset \hat{x}[x, \lambda y(y \text{ ate the bagel})]$

The general constraint on bound variables with preceding anaphoric pro-form interacts with this proposal in an appropriate manner. Thus (3.5.17) is ill-formed because of the position of the bound variable in its logical form, as sketched in (3.5.18).

(3.5.17) *His_i mother loves John_i.

(3.5.18) $\{\text{John}_i\} \subset \hat{x}[\text{his}_i \text{ mother}, \lambda y(y \text{ love } x)]$

Moreover, the general approach we are taking encounters no difficulties (as an iota-operator theory would) with plural noun phrases, which can be Clefted, Pseudo-Clefted, or accented.

Notice that a system of this sort may very well suffice to account for the prosodic constraints on the well-formedness of question-answer pairs we observed a moment ago. To take a fresh example, consider the discourse in (3.5.19).

(3.5.19) Q: What did Betsy eat?

A: Betsy ate the bagel.

The logical forms of this question and answer, given what we have said so far, would be as follows.

(3.5.20)(a) (For what x) [Betsy, $\lambda y(y \text{ eat } x)$]

(b) $\{\text{the bagel}\} \subset \hat{z}[\text{Betsy}, \lambda w(w \text{ eat } z)]$

We might characterize well-formed pairs in terms of shared open sentences in their logical forms. Thus as open sentences, [Betsy, $\lambda y(y \text{ eat } x)$] and [Betsy, $\lambda w(w \text{ eat } z)$] in (3.5.20) are alphabetic variants, hence (3.5.20)(b) is a well-formed direct answer to (3.5.20)(a).

(3.5.21), on the other hand, with accent on <u>Betsy</u>, has the logical form as given in (3.5.22).

(3.5.21) Betsy ate the bagel.

(3.5.22) $\{\text{Betsy}\} \subset \hat{z}[z, \lambda w(w \text{ eat the bagel})]$

Here the open sentence is [z, λw(w eat the bagel)], which is not an alphabetic variant of the open sentence in the logical form of the question in (3.5.20)(a). Consequently (3.5.31) is not a well-formed direct answer of that question, i.e. the following discourse is ill-formed.

(3.5.23) Q: What did Betsy eat?
#A: Bétsy ate the bagel.

This approach can be naturally extended to multiple wh-question. Consider (3.5.24), for example.

(3.5.24) Who gave the book to who?

Treating each question word as a variable-binding operator, we arrive at roughly the following logical form.

(3.5.25) (for what (person) x) (for what (person) y)
(x, λz(z give the book to y))

Direct answers to multiple wh-questions must have multiple accents, and those accents must furthermore be in the right place. Thus (3.5.26), but not (3.5.27) is a well-formed direct answer to (3.5.25).

(3.5.26) Bétsy gave the book to Péter.
(3.4.27) Bétsy gave the bóok to Peter.

A natural approach to the problem of sentences containing multiple accents would be to simply allow multiple abstraction. The resulting system, which is not unlike that of Jackendoff (1972), would lead us to represent (3.5.26) as the following.[20]

(3.5.28) {Betsy, Peter} $\subset \hat{w}\,\hat{s}$[w, λr(r give the book to s)]

Notice that in this formula we have the open sentence: [w, λr(r gave the book to s)]. This is an alphavetic variant of the open sentence in the logical form of the multiple <u>wh</u>-question in (3.5.24), i.e. of [x, λz(z give the book to y)] in (3.5.25). In this way we can account for the fact that (3.5.25) is a well-formed direct answer to (3.5.24).

Conversely, (3.5.27) with an accent on <u>the book</u>, but not on <u>Peter</u>, will have the following logical form.

(3.5.29) $\{\text{Betsy, the book}\} \subset \hat{w}\ \hat{r}[w, \lambda s(s \text{ give } r \text{ to Peter})]$

But the open sentence here, [w, λs(s give r to Peter)] is not an alphabetic variant of [x, λz(z give the book to y)]. Therefore the account we have given correctly predicts that (3.5.27) is not admissable as a direct answer to (3.5.24), i.e. that discourses like this one are ill-formed:

(3.5.30) Q: Who gave the book to who?
A: Bétsy gave the bóok to Peter.

This brief sketch of the interaction of accent and logical form provides, I think, a fairly intuitive account of question-answer pair well-formedness. Let us now consider how our proposal bears on our rule of Gapping.

As we have mentioned, Gapping remnants and corresponding constituents in the left conjuncts must be accented. Under the proposal we have just made, this will be systematically reflected at the level of logical form. (3.5.31), for example, is a possible input to Gapping.

(3.5.31) Jím likes Júdy and Péter likes Bétsy.

The logical form of this sentence is shown in (3.5.32).

(3.5.32) $\{Jim, Judy\} \subset \hat{x}\,\hat{y}[x, \lambda z(z \text{ like } y)]$ &

$\{Peter, Betsy\} \subset \hat{r}\,\hat{s}[r, \lambda w(w \text{ like } s)]$

It follows from the theory we have just outlined that sentences which can undergo Gapping, i.e. which are appropriately accented, are such that the material to be Gapped "corresponds" in a very intuitive sense to an entity at the level of logical form. Gapping remnants, by this same token, also "correspond" to identifiable entities.

Note that this provides a straightforward solution to what might have appeared to be a rather formidable correspondence problem. That is, given our hypothesis of the general nature of the recoverability of deletion and its relation to logical form, we must always be able to associate syntactic entities with corresponding logical entities, as we mentioned in Chapter Two with respect to VPD. That association was somewhat less problematic in the case of VPD once we constructed our logical system in such a way that every surface Verb Phrase corresponds to a λ-predicate in logical form. As we have seen, however, Gapping must be formulated so as to delete X-variables. Therefore, given the theory of logical form proposed in Chapter Two, Gapping deletion targets do not correspond to any single logical entity. The proposal we have just made, for the logical form of contrastive sentences, however, not only provides a natural account of question-answer pair well-formedness, but also provides logical entities (of the form $\hat{v}_1\,\hat{v}_2[\ldots v_1 \ldots v_2 \ldots]$) which "correspond" to Gapping deletion targets in a straightforward and intuitive way. Thus by providing a uniform account of the logical form of contrastive sentences, we have offered a solution to the correspondence problem with respect to Gapping.

Note further that this "solution" is actually a strong empirical hypothesis that we have formulated here with one of many possible notations. This hypothesis, more generally put, is that there is a fundamental relation between ellipsis rules (rules which delete X-variables) and the prosody of the remnants they leave behind. Contrastive accent alters the logical form of a given sentence in such a way as to induce a correspondence between arbitrary sequences of syntactic entities and single entities of logical formulae. THe proposal seems quite naturally extended to cases of Stripping (Hankamer (1971)), Comparative Ellipsis, and many of the cases of ellipsis observed by Morgan (1973), but a detailed examination of those cases is well beyond the scope of this enterprise.[21]

One might well ask whether there is any independent evidence that the recoverability of Gapping must make reference to logical form, and not merely to syntactic identity. I think there is. Consider the following example.

(3.5.33) My brothers have all gone to the circus, and my sister ∅ to the carnival. [∅ = has gone]

This sentence, which all my informants accept, is of considerable interest. If syntactic identity were the appropriate condition for Gapping, (3.5.33) would be derivable only from the following ungrammatical source.

(3.5.34) *My brothers have all gone to the circus, and my sister have all gone to the carnival.

One possible solution to this problem, of course, would be to presume a rule of Quantifier Float ordered after Gapping (we will not concern ourselves here with the singular-plural problem). However, as we have mentioned (cf. Chapter One) that is a dubious rule. Furthermore, if Quantifier Float were a rule, it would clearly affect semantic interpretation (for instance if a quantifier were "floated" over the negative particle not). Hence it would have to apply before the level of Shallow Structure (given our ordering hypothesis), and would therefore have to precede Gapping. Any such solution therefore seems unlikely.

However our theory of the recoverability of deletion, nothing more being said, predicts that sentences like (3.5.33) should be possible. The logical form for (3.5.33) would be roughly as follows.

(3.5.35) $(\forall x\text{: my brother}(x))[\{x, \text{the circus}\} \in \hat{y}\,\hat{z}[y, \lambda w(w \text{ go to } z)]] \;\&$
$\{\text{my sister, the carnival}\} \in \hat{r}\,\hat{s}[r, \lambda t(t \text{ go to } s)]$

Quantifying into a position that represents a member of a set in the left conjunct does not affect the possibility of Gapping at all. $\hat{y}\,\hat{z}[\ldots]$ and $\hat{r}\,\hat{s}[\ldots]$ are alphabetic variants. The existence of such sentences as (3.5.35) therefore provides further support for our proposal for Gapping in particular, and for our more general hypothesis about the relation between deletion and logical form.

It should be pointed out that any proposal to represent contrastive accent at the level of logical form must, at the very least, involve second order logic. This is because of the simple fact that syntactic entities corresponding to logical predicates are possible contrastive foci. Consider (3.5.36), for instance.

(3.4.36) John is happy, and Bill ∅ sad.

There are many potential problems in this area, and many possible solutions. The one solution we will suggest here is the introduction variables ranging over predicates (we will write P and Q for such variables). The logical form of (3.5.36), then, may be given as (3.5.37).

(3.5.37) $\{John,(\lambda x(happy(x)))\} \subset \hat{y}\ \hat{p}[y, P]$ &

$\{Bill,(\lambda z(happy(z)))\} \subset \hat{w}\ \hat{q}[w, Q]$

$\hat{y}\ \hat{P}[...]$ and $\hat{w}\ \hat{Q}[...]$ are alphabetic variants, predicting the possibility of deletion in (3.5.36). We will have no more to say about this interesting and rather complex matter here.

Another interesting problem concerns the interaction of VPD and Gapping. Some informants accept examples like the following one, where Gapping and VPD have both applied.

(3.5.38) Alan likes Sandy, and Betsy $\emptyset_1$ Peter$_j$, although she doesn't know why she does $\emptyset_2$. [$\emptyset_1$ = likes; $\emptyset_2$ = like Peter$_j$ (or like him$_j$)]

It seems that general considerations of complexity are relevant here, for less complex examples where Gapping and VPD have both applied seem much less acceptable:

(3.5.39) ??Alan likes Sandy, and Betsy $\emptyset_1$ Peter, and Lois does $\emptyset_2$ too.

Why are these cases a problem? Because the logical forms we would posit for such sentences are roughly as follows (details omitted):

(3.5.40) $\{Alan, Sandy\} \subset \hat{x}\ \hat{y}[x, \lambda r(r\ like\ y)]$ &

$\{Betsy, Peter_j\} \subset \hat{w}\ \hat{z}[w, \lambda s(s\ like\ z)]$

$...[Lois, \lambda t(t\ like\ \{\begin{matrix} Peter_j \\ him_j \end{matrix}\})]$

$\hat{x}\ \hat{y}[...]$ and $\hat{w}\ \hat{z}[...]$ are alphabetic variants, predicting the possibility of Gapping, but $\lambda s(s\ like\ z)$ and $\lambda t(t\ like\ him_j)$ are not alphabetic variants, which should mean that VPD is impossible.

The marginal nature of some of the examples (e.g. (3.5.39)) and the variation among speakers makes it difficult to resolve this problem. One possible solution, however, is in terms of our indexing mechanism. We might speculate that the rules of SI-1 that create λ-expressions assign an index to the variable bound by the λ which is the same index as that of the NP which is the argument of the λ-predicate. On this view, a sentence like (3.5.41) would have a logical form like that in (3.5.42).

(3.5.41) Betsy$_b$ loves Peter$_p$.

(3.5.42) Betsy$_b$, $\lambda x_b(x_b$ loves Peter$_p)$

By the same token, with respect to Gapping we have logical forms like the following.[22]

(3.5.43) Alan$_a$ likes Sandy$_s$, and Betsy$_b$ $\emptyset$ Peter$_p$.

(3.5.44) {Alan$_a$, Sandy$_s$} $\subset \hat{x}_z\ \hat{x}_a[x_a, \lambda x_a(x_a$ like $x_s)]$ &

{Betsy$_b$, Peter$_p$} $\subset \hat{x}_b\ \hat{x}_p\ [x_b, \lambda x_b(x_b$ like $x_p)]$

We might then allow for the examples in (3.5.38) and (3.5.39) by a weakening of the notion of alphabetic variance, allowing, say, $\lambda x_i(f(x_i))$ to count as an alphabetic variant of $\lambda x_i(f\ (PRO_i))$. Thus (3.5.39) might have the following logical form.

(3.5.45) {Alan$_a$, Sandy$_s$} $\subset \hat{x}_a\ \hat{x}_s[x_a, \lambda x_a(x_a$ like $x_s)]$ &

{Betsy$_b$, Peter$_p$} $\subset \hat{x}_b\ \hat{x}_p[x_b, \lambda x_b(x_b$ like $x_p)]$ &

[Lois$_1$, $\lambda x_{.}(x_1$ like him$_p)]$

If $\lambda x_1(x_1$ like him$_p)$ counts as an alphabetic variant of $\lambda x_b(x_b$ like $x_p)$ then we have accounted for the possibility of VPD (together with Gapping) in examples like these.

The proposed weakening of the notion of alphabetic variance, it should be noted, if viewed as a speaker-variable process, ancillary to the central grammar (rendering such sentences in effect derivatively generated) would have further consequences. It might, for example, account for why some speakers accept sentences like the following one (which we discussed in Section 2.2 of Chapter Two) on the reading indicated.

(3.5.46) John thinks he's unpopular, and Bill does $\emptyset_1$ too, although his wife doesn't $\emptyset_2$.

(3.5.47) John thinks John is unpopular, and Bill thinks Bill is unpopular, although Bill's wife doesn't think Bill is unpopular.

Such a reading would be accounted for under this proposal, for $\lambda x_w(x_w$ think [he_b, $\lambda x_b(x_b$ be unpopular)]) and $\lambda x_b(x_b$ think [x_b, $\lambda x_b(x_b$ be unpopular)]) would count as alphabetic variants in the following logical formula.

(3.5.48) $John_j$, $\lambda x_j(x_j$ think [x_j, $\lambda x_j(x_j$ be unpopular)]) &
$Bill_b$, $\lambda x_b(x_b$ think [x_b, $\lambda x_b(x_b$ be unpopular)]) although
[his_b $wife_w$, $\lambda x_w(x_w$ think [he_b, $\lambda x_b(x_b$ be unpopular)])]

Again, the facts are simply not clear enough to resolve this matter, although the proposed account, I think, is not at all implausible.

Finally, it should be noted that the general theory outlined in this section bears on another frequently-observed fact concerning the general relation of deletion and prosody. It seems to be a fact, in English at least, that for any sentence that has undergone an optional deletion rule, there exists another sentence where the same material is not deleted, but is destressed. This has led some, in particular Chomsky, to suggest a principle that only destressed material can be deleted.

No arguments for such a principle appear in the literature (aside from the observation just noted). For whatever it's worth, then, I'll give two arguments that I've managed to find for such a principle. The first, due to Haj Ross (personal communication), concerns the deletion rule involved in sentences like the following.[23]

(3.5.49) ø stand on this wire, and you'll get a shock. [ø = you]

Under some circumstances, the subjects in the left conjuncts in sentences like this are contrastively accented:

(3.5.50) You hit yourself, and I'll hit myself.

In just those circumstances, Ross observes, deletion of those subjects is not allowed:

(3.5.51) *ø hit yourself, and I'll hit myself.

This fact would follow if there were some principle that prohibited deletion of accented material.

A second argument concerns the rule of Sluicing (Ross (1969b), Sag and Hankamer (1976)). Sluicing can apply as in (3.5.52), where the material to be deleted is destressed.

(3.5.52) I know someone left, but I don't know who $\left\{\begin{matrix}\text{left}\\ ø\end{matrix}\right\}$

When the material to be deleted contains something accented, however, as in (3.5.53), Sluicing cannot apply.

(3.3.53) I know Betsy didn't leave, but I wonder who $\left\{\begin{matrix}\text{díd leăve} \\ \text{*}\emptyset\end{matrix}\right\}$.

The only possible deletion here is VPD, which can delete the unaccented VP to give:

(3.5.54) I know Betsy didn't leave, but I wonder who díd $\emptyset$.

Again, these facts would follow if there were some principle prohibiting deletion of accented material.

In the theory we have developed, accented entities correspond to bound variables at the level of logical form. In many cases then, this will bring it about that the relevant logical entities are not alphabetic variants, hence preventing deletion. It's quite possible then that the principle prohibiting deletion of accented material will result as a theorem. In the absence of a coherent proposal for the treatment of such matters as stressed do (as in (3.5.53)), however, this remains an interesting problem for future research.

3.6 Summary of Chapter Three

This chapter has dealt primarily with the rule of Gapping. We have critically examined all previous treatments of this rule and provided an alternative formulation. In the process of doing that, we travelled rather far afield in order to reformulate Bresnan's Relativized A-over-A Principle. Our alternative to that, the Immediate Domination Principle has been seen to interact with our Gapping rule to provide an account of several well-known, but previously unexplained facts. A hypothesis as to the nature of the particular relation between Gapping and logical form and of the more general relation between ellipsis rules and the recoverability of deletion has also been offered.

Footnotes to Chapter Three

1. Some informants accept this sentence. I think they are reanalyzing speak with as a verb. This is supported by the possibility of passivization.

(*i*) Don't talk unless you're spoken to.

An "unlexicizable" sequence of verb-preposition neither Gaps nor passivizes:

(*ii*) *Jane left after the first act, and Mary ∅ the second act.

(*iii*) *This act will be left after (by hundreds of people).

Similarly, sequences of adjective-preposition are not lexicizable.

2. Fiengo (1974, Chapter Four) finds sentences like these to be acceptable. His judgements, however, are highly idiosyncratic.

3. Needless to say, the observations just made speak against the theories or Tai (1969) and Koutsoudas (1971), who collapse all rules performing deletion in coordinate structures into one coordinate deletion rule.

4. Many of my informants accept this interpretation of (3.2.32).

5. This has been kindly verified for me by Jorge Hankamer (personal communication of May 3, 1976), who knows of no arguments whatever for that position. Neither do I.

6. See Kuno (1972, 1975b) for a more detailed account of his ideas about such notions. For a more general review of the work on "functional sentence perspective", see Garvin (1964), Firbas (1964), Halliday (1967), Vachek (1964), and the references cited there.

7. IDP should not be confused with Ross's (1974) Immediate Self-Domination Principle, which is something quite different.

IDP says, less formally, simply that if a rule applies ambiguously to two nodes A and B, and if A immediately dominates B, the rule cannot apply to B. This formulation is quite independent of any assumptions about the nature of syntactic categories.

8. It might be objected that (3.3.19)(b) is independently ruled out because its source would require a (+wh) pronominal head (pronominal heads of relative clauses are in general inadmissible). One might even raise the further question of whether heads of relative clauses are NP's at all, and hence the question of whether they can ever serve as values for target terms of rules that apply to NP's.

This matter is by no means clear, but there seems to be a fair amount of evidence that heads of relative clauses are indeed NP's. They seem to allow full NP recursion:

(*i*) [The man from Peoria] who I like...

(*ii*) [My younger sister] that you don't like is coming to dinner.

They also admit of what would appear to be base conjunction:

(*iii*) [The man and the woman] who met in Peoria arrived yesterday.

Moreover, PP's and relative clauses seem to occur in either order:

(*iv*) [The men that I liked] in this room...

(*v*) [The men in this room] that I liked...

Moreover, the ban against pronouns as heads of relative clauses seems to be a recent innovation of the language:

(*vi*) Can a man be profitable unto God, as he that is wise may be profitable unto himself? [Job, 22.2]

(*vii*) For he whom God hath sent speaketh the words of God... [St.John, 3.34]

All these facts would follow if we assumed a phrase structure rule as in (*viii*).

(*viii*) $NP \rightarrow NP - (\{ \begin{matrix} \overline{S} \\ PP \end{matrix} \})$

Other analyses may be possible, obviating some of the following discussion, but the view that heads of relative clauses are indeed NP's has quite a bit to recommend it.

9. Thanks here go to Morris Halle for correcting Grosu's Russian.

10. I find the arguments given by Jackendoff (1976) for a Det node in Adjective Phrases rather unconvincing.

11. We will follow Woisetschlaeger (1975, 1976) in assuming that $\boxed{A^2}$ in (3.3.56) has been permuted to N^2-initial position from its base position, which is after Det.

12. Notice that this is not very different from Hankamer's conception of Ross's (1967b) formulation of the rule. Ross (1976), however, has a different conception of Ross's (1967b) formulation of the rule. Personally, I can't figure out what Ross (1967b) had in mind.

13. This tree diagram may be somewhat inaccurate. The preposed PP may be outside of S^2. Nothing hinges on this.

14. For further elaboration of this notion in the domain of phonology, see Kean (1974).

15. Notice that here it is crucial that we assume no "pruning" conventions of the sort frequently suggested in the literature. Alternatively, we could

presume pruning of non-branching nodes if the circled S nodes in (3.4.33) are still branching when Gapping applies. There are several formulations of EQUI that this would be consistent with.

16. This rule must be constrained somehow to prevent Term 8 from being null. Gapping must effect right-peripheral ellipsis only when it is concommitant with sentence-internal ellipsis.

17. Notice that Hankamer's "unlike adverb" constraint, mentioned in Section 3.1, is an automatic consequence of our formulation of Gapping. Sentences like (*i*) cannot be generated, because only one remnant to the left of the (first) Gap is allowed.

(*i*) *John certainly likes Sue, and Joan possibly ø George.

[ø = likes]

18. Dretske actually couched his discussion in terms of "contrast", which he claims is usually, but not always reflected by accentual prominence. This may in fact be a better way to look at the matter.

19. "#", following Sag and Hankamer (1976), indicates a sentence that cannot be appropriately uttered in the context indicated.

20. Further refinements may be necessary to guarantee the proper order within the set: {Betsy, Peter} . This might be done by means of ordered pairs (n-tuples, actually). Further intonational facts, i.e. relative pitch prominence between Topic and Focus in answers (incorrectly associated with Bolinger's (1958) A and B pitch accents by Jackendoff (1972)) might also be treated in a logical system of the type we are developing. We will not pursue that matter here.

21. For instance the ellipsis rule that produces short answers to questions might be formulated simply as (*i*).

$$(i) \quad \underset{\emptyset}{\underset{\Downarrow}{W_1}} - X^2 - \underset{\emptyset}{\underset{\Downarrow}{W_2}}$$

General considerations of logical form will then account for most of the restrictions on the application of this rule. An approach along these lines has the potential to provide a solution to many of the problems raised by Morgan (1973), which are mostly problems only for a purely syntactic theory of deletion.

22. Making the indices on these variables work out right is no trivial task.

23. The deletion rule involved in such sentences has been studied rather extensively by Thrasher (1974).

Chapter Four

A "Mixed" Theory of Anaphora

4.0 Introduction

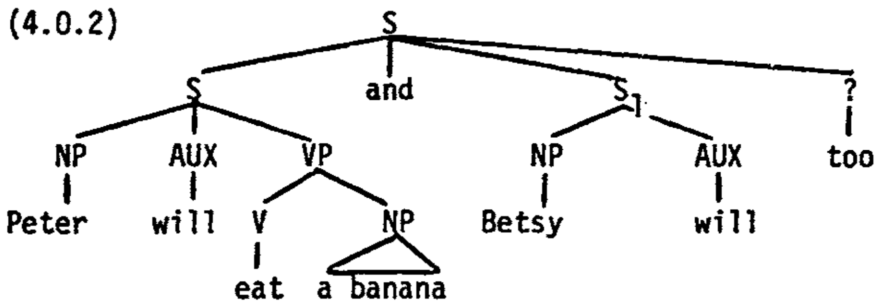
In the preceding chapters we have accepted the Harrisian position that "elliptical" sentences are derived from "non-elliptical" entities by deletion transformations. Deletion rules, on our view, are only one type of anaphoric rule. Such phenomena as do it Anaphora, Sentential-it Anaphora, and definite pronominalization are assumed to involve interpretive rules which associate the pronominal anaphors in question with their controllers. Ours is then a "mixed" theory of anaphora. For a discussion of the different "monolothic" theories that have been proposed, and for a summary of arguments in favor of such a "mixed" theory, the reader is referred to Sag and Hankamer (1976).

Though our use of deletion rules is rather traditional, there has been a certain amount of research by various people in recent years (e.g. Akmajian (1970), Shopen (1972), Chomsky (1972b), Jackendoff (1972), Wasow (1972)), that has challenged the correctness of this general approach. These people have suggested that deletion rules should be replaced by rules of semantic interpretation that operate to supply interpretations (in a way that has never been made precise) for syntactic objects that are generated by the base grammar in elliptical form. On this view, all anaphoric processes are handled with the same, or at least similar, mechanisms.

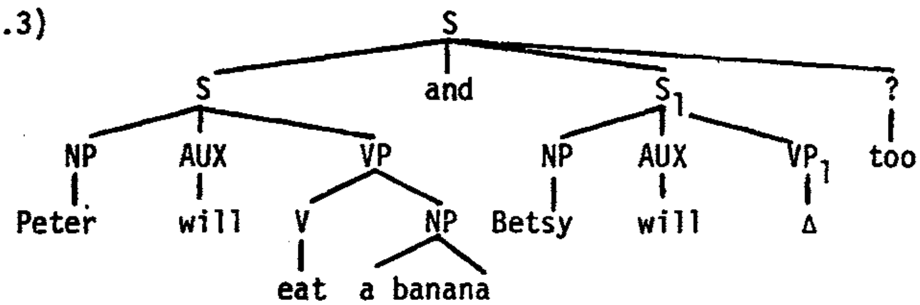
There has been some response to these proposals in the literature (e.g. Ross (1969b), Grinder and Postal (1971), McCawley (1976) - see the summary of this response in Sag and Hankamer (1976)). Virtually all published arguments against non-deletion theories, however, have been directed against specific proposals, some of them straw-men, which either assumed that elliptical sentences have no syntactic structure beyond that determined by lexically-present material, or else posited the existence of syntactically non-complex null anaphors (i.e. Δ). The interpretive theories that have been argued against then are all such that for a sentence like (4.0.1), they would posit either a base structure like (4.0.2) or one like (4.0.3), which would be in all essential respects equivalent to the surface structure.

(4.0.1) Peter will eat a banana, and Betsy will, too.

(4.0.2)



(4.0.3)



In these theories, an interpretive rule would "fill in" a "reading" either for S_1 or for VP_1. The rule of VPD would be eliminated.

The arguments against such theories I think are generally regarded as conclusive. We will therefore neither summarize those arguments here, nor dwell further on those theories (see Hankamer (ms.) for further remarks on this matter). The nature of many of those arguments will in fact become clear in the discussion that follows.

There is one recent ("monolithic") interpretive theory, however, which has not been included in prior criticisms of interpretive theories. That is the Empty Structures Hypothesis proposed by Wasow (1972). Wasow's proposal is in fact considerably more sophisticated than its predecessors, and is consequently not subject to the anti-interpretive arguments of Ross and of Grinder and Postal.

The Empty Structures Hypothesis (ESH) would then appear to be a possible alternative to the "mixed" theory we have offered. Wasow, in fact, argues for the superiority of ESH over a theory like ours. More precisely, he has argued that certain phenomena, such as Comparative Deletion should be regarded as deletion rules, but that such processes as VPD and Sluicing (Ross (1969b)) should be treated on a par with definite pronouns, i.e. interpretively. His arguments have, as far as I know, gone unchallenged. In fact, there have been at least three subsequent investigations that would appear to have simply accepted Wasow's arguments against a deletion analysis of VPD and Sluicing (Fiengo (1974), Williams (1975, 1976)).

In this chapter we will be concerned with several issues. First, we will examine Wasow's arguments against deletion theories. We will conclude that none of those is an argument against the deletion theory we have developed in the previous chapters. Secondly, we will summarize the discussion in Sag and Hankamer (1976), who argue for a fundamental dichotomy of anaphoric processes. Thirdly, in the course of discussing Wasow's arguments for the "unity of anaphora", we will observe certain discrepancies in the behavior of various anaphoric processes. These and other discrepancies we will point out, taken together with the fundamental dichotomy of anaphoric processes discussed by

Hankamer and myself, will be argued to favor the "mixed" theory of anaphora we have adopted over Wasow's ESH.

4.1 Interpretation vs. Deletion

4.1.1 The Empty Structures Hypothesis

The essence of ESH is that lexical insertion is optional. The base grammar then generates structures with empty nodes in any position or combination of positions. Under certain conditions (precisely the same conditions where deletion can take place in a deletion theory, in fact), anaphora rules are allowed to "associate...an empty structure with an antecedent" and "then the reading of the antecedent can be associated with the empty node" (Wasow (1972, p. 98)). The syntax then generates strings like those in (4.1.1).

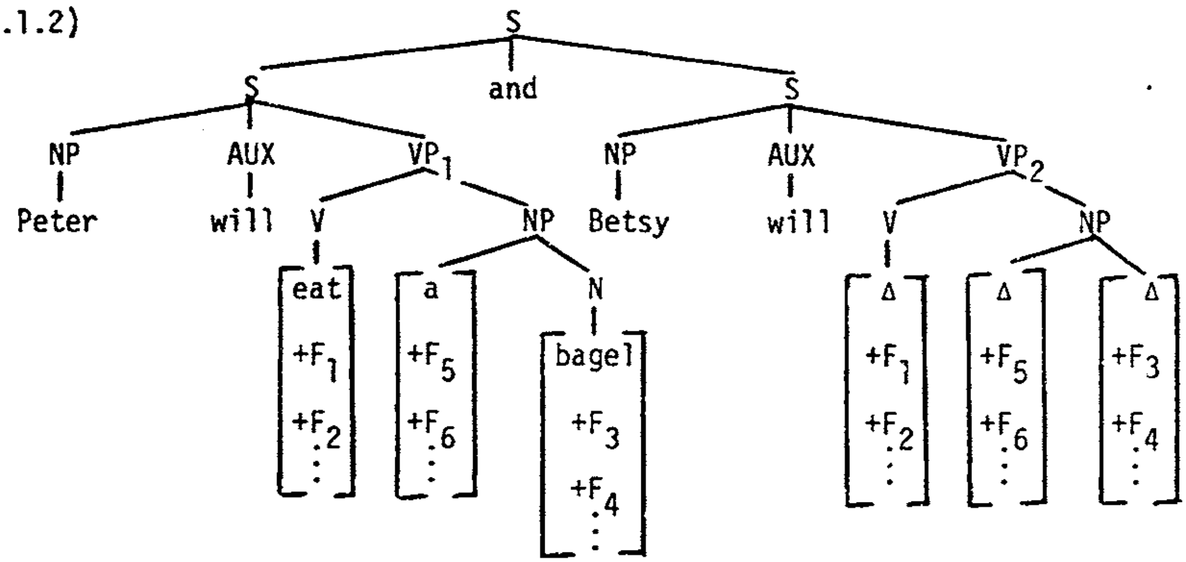
(4.1.1)(a) Δ Δ John Δ to the Δ Δ Δ

(b) A Δ Δ know Δ Δ does Δ Δ bananas.

Examples like these, to which no interpretive rule can apply, are filtered out by a general convention that "surface structures containing uninterpreted empty nodes would be regarded as semantically anomalous" (p. 98).[1] In Wasow's theory, then, null anaphors are sequences of dummy nodes which "have all the structure of their antecedents, lacking only phonetic material". This makes it possible to "generate structures with all of their normal syntactic properties, but lacking any phonological or semantic material" (p. 98).

As an illustration of how ESH differs from previous interpretive theories of anaphora, reconsider the example of the last section, which under ESH would be generated as the following (where F_1, F_2... represent syntactic features).

(4.1.2)



Since VP_1 and VP_2 are identical (more precisely, non-distinct; we will return to this), except for the presence of the actual lexical items eat, a, and bagel, an interpretive rule can apply to this structure to "associate" the "reading" of each lexical item in VP_1 with the corresponding empty node in VP_2.

Dummy nodes in ESH undergo transformations just like any lexical entities, as indeed they must in order to avoid Ross's anti-interpretive arguments concerning examples like the following, where VPD applies to the output of passive and There-Insertion (Δ's indicate ESH dummy nodes that have undergone transformations).

(4.1.3)(a) Betsy was hassled by the police and Peter was Δ Δ Δ, too.

(b) Sandy said there was a mouse in the bedroom, and there was Δ Δ Δ Δ Δ.

Notice that ESH is really not very different from a deletion theory. One might look at the difference as simply one of where the feature [+ phonetically null] is assigned. In a deletion theory, it is assigned at a rather superficial level (after the level of shallow structure, in fact, as we have seen in the preceding chapters). In ESH, it is assigned at the level of deep structure, but indiscriminately, and the rules of interpretation, which determine which phonetically-null sequences are allowable and which are not, also apply at a rather superficial level (perhaps also at the level of shallow structure).

One might then ask whether ESH and a deletion theory are empirically distinguishable. Wasow claims that they are. There are various facts, he argues, which can be accounted for in ESH but not in a deletion theory. Let us now examine those arguments.

4.1.2 Arguments Against Deletion

1. Wasow's first argument concerns Ross's (1967a) Complex Noun Phrase Constraint, which Ross formulated as follows:

(4.1.4) No element contained in a sentence dominated by an NP with a lexical head noun may be moved out of that NP by a transformation.

Deletion rules like Comparative Deletion and the rule involved in the derivation of sentences like (4.1.7) obey this constraint.

(4.1.5) John is taller than Mary believes that Bill said he is ∅.
(4.1.6) *John is taller than Mary believes the claim that Bill is ∅.
(4.1.7) John is not the doctor that you claim his father is ∅.
(4.1.8) *John is not the doctor that you made the claim that he is ∅.

If these examples are to be derived by deletion, as Wasow assumes, then Ross's constraint must be extended (as Ross noted) to prevent deletion (under identity) of elements contained in sentences dominated by an NP with a lexical head noun (as Ross noted).

VPD and Sluicing, however, do not obey Ross's constraint:

(4.1.9) John didn't take LSD, but Bill believed the claim that he did $\emptyset$.

(4.1.10) John takes LSD, but I don't know the reason why $\emptyset$.

Nor does pronominal anaphora obey the Complex Noun Phrase Constraint:

(4.1.11) John_i believes the prediction that he_i will win.

(4.1.12) John has taken LSD, but most of the people who know it won't talk about it. [it = that John has taken LSD]

Wasow makes the (reasonable) assumption that pronouns are not derived by deletion rules, and that they are associated with their antecedents by rules of semantic interpretation. The above facts, he argues, therefore provide evidence that VPD and Sluicing should be treated on a par with pronominal anaphora, i.e. the null anaphors in (4.1.9) and (4.1.10) should be regarded as base-generated dummy nodes which are then associated with their antecedents by interpretive rule(s). Since he regards the cases in (4.1.5)-(4.1.8) as being derived by deletion, the generalization can be stated that deletion rules (and movement rules) are the only rules that obey Ross's Constraint.

Obviously the formulation of the Complex Noun Phrase Constraint in (4.1.4) must be modified (as Ross noted) if it is going to generalize to deletion rules. Otherwise sentences like the following will be ruled out, where two island-obeying deletion rules have applied entirely within a Complex NP environment.

(4.1.13)(a) John denied the claim that Bill is smarter than Harry is $\emptyset$.

(b) John finally came around to accepting our claim that Bill is not the doctor that his father is.

This problem is trivially solved, of course, by stating the constraint in terms such as "no rule can involve two positions A and B which are separated by a Complex NP environment" (see also the final remarks at the end of Chapter Six of Ross (1967a) which are in terms of "no rule can cross...", etc.).

Note that even accepting Ross's formulation, thus modified, the theory of VPD we have put forth accounts for why VPD does not obey the Complex Noun Phrase Constraint. The rule itself, whose structural description is X - AUX - VP - Y, does not "involve" two VP positions at all. The relation between target and trigger VP's is an indirect one, established only through the correspondence of entities of the syntax and entities of logical form.

Sluicing moreover would have to work the same way, i.e. it would be (in Edwin Williams's sense[2]) a "free" deletion. This is so because it applies bidirectionally and in discourse. We would therefore expect Sluicing to take place in Complex NP environments,as it does.

The facts regarding the Complex Noun Phrase Constraint then do not choose between our deletion theory and ESH. Both are consistent with the fact that VPD and Sluicing do not obey that constraint.

2. Wasow's second argument concerns the fact that VPD and Sluicing, like rules of pronominal anaphora, are bidirectional, and hence, given the standard formalism for writing transformations, "very costly in terms of the evaluation metric". Wasow proposes to eliminate bidirectionally anyway, e.g. for pronominal anaphora.

Given our fomulation of VPD (as well as Sluicing) as a "free" deletion rule, however, no complex machinery is required in the actual grammar. The bidirectionality of "free" deletions is a consequence of the fact that these rules do not mention the position of the deletion controller at all. Intra-sentential bidirectional control is predicted to be possible whenever it is also possible for the deletion controller to be in the previous discourse. Whatever "complexity" is involved in this is thus relegated to the metatheory, i.e. to the principle of the recoverability of deletion.

3. Wasow constructs a third argument against a deletion approach to VPD which involves Chomsky's "Strict Cycle Condition" and Grinder and Postal's "missing antecedent" phenomenon. On the basis of sentences like (4.1.14), Wasow argues that VPD must be cyclic, or else it would violate the "Strict Cycle Condition".

(4.1.14) The newpapers reported that reliable sources claim that Tricia won't come unless Ed does ∅. [∅ = come]

But if VPD is cyclic, he continues, then in the derivation of a sentence like (4.1.15), VPD will have applied before the rule of pronominal anaphora.

(4.1.15) The man who claimed he didn't have a car actually did ∅, and it was a convertible. [∅ = have a car_i; it = (the) car_i]

Thus when the rule of pronominal anaphora applies, there is no antecedent that it can associate with the pronoun _it_. Under ESH, the antecedent of _it_ would be a dummy node present throughout the entire derivation of (4.1.15). The pronominal anaphora rule can associate _it_ with its antecedent at any level.

This argument has very little to recommend it. First of all, the "Strict Cycle Condition" is sensibly construed only as a condition on cyclic rules (although the formulation in Chomsky (1973) does not make this clear).[3]

If VPD is not a cyclic rule, it is then not subject to that principle. The position that VPD is non-cyclic is further supported by the fact that no arguments can be constructed to show that it is cyclic.

But if VPD is post-cyclic, as the theory we have developed requires it to be, no problem of the sort Wasow anticipates arises. Rules of semantic interpretation apply to shallow structures as do deletion rules. The pronoun _it_ in (4.1.15) can be associated with its antecedent prior to deletion. There is no argument against our deletion theory here.

4. Another of Wasow's arguments concerns the interaction of VPD and _Do_-Support. If VPD is a deletion rule, then it must precede _Do_-Support in the derivation of sentences like the following.

(4.1.16) John will come if Bill does ø.

_Do_-Support would then have to apply in violation of the "Strict Cycle Condition". As Wasow himself notes, however, this argument is inconclusive if that condition does not apply to "housekeeping" rules (in the sense of Bach (1965,1971)), as Chomsky (1973) suggests, i.e., if that condition is interpreted as a condition on the application of cyclic rules, as is natural.[4]

5. The existence of examples like (4.1.17)(a) are argued to provide additional evidence against a deletion theory (of Sluicing), for (what Wasow takes to be) their sources are ungrammatical.

(4.1.17)(a) John can't come along, but he won't say why not.

(b)*John can't come along, but he won't say why not he can('t) come along.

He further argues that the synonymy of (4.1.17)(a) and (4.1.18) follows (under ESH) from the fact that "anaphora rules may optionally ignore negatives" (p. 107).

(4.1.18) John can't come along, but he won't say why.

Note first of all that the claim that the interpretive analogues of deletion rules "optionally ignore negation" has intolerable consequences. That claim incorrectly predicts that the following instances of VPD are all ambiguous depending on whether or not negation has been "ignored".

(4.1.19) John didn't go to school after he found out that Betsy did ∅.

(4.1.20) John didn't go to school after he found out that Betsy didn't ∅.

This is of course not the case. These two sentences are under no circumstances synonymous.

Even if some independent remedy could be found for this oversight in Wasow's reasoning, the claim that Sluiced sentences "optionally ignore negation" is itself dubious. If this were a general property of Sluiced sentences, we would expect all the following examples to be grammatical, which they are not.

(4.1.21)(a) *Joan didn't like one of my arguments, but she didn't say which one not.

(b) *Either Goldwater didn't win in Ohio, or he didn't win in New York; Bill wouldn't tell me where not.

(c) *Either Jerry didn't like Myra, or he didn't like Dick; Henry wouldn't tell me who not.

The example Wasow gives ((4.1.17)(a)) is thus entirely idiosyncratic. Perhaps <u>why not</u> is lexicalized (if so it would have the peculiar property of requiring null material after it in surface structure). Whatever might be the proper treatment of why not, it's clear that any theory that predicts that VPD or Sluicing "optionally ignores negation" is grossly defective.

The only convincing cases Wasow (or anybody else, to my knowledge) has given where an anaphor is assigned a non-negative interpretation when its antecedent contains an overt negative are those like (4.1.22), due to Lakoff.

(4.1.22) John didn't marry Mary, although the fortune-teller predicted it. [it = that John would (?) marry Mary]

These remain somewhat mysterious, though given the range of possible interpretations for sentential-it anaphora, the fact that these pronouns can refer to non-negative entities (which are after all entities at the level of logical form) is not all that surprising.

Wasow also discusses cases like (4.1.23), but since we have already made a proposal concerning these cases (cf. Section 2.3 of Chapter Two), we will not deal with them here.

(4.1.23) %Although John will trust nobody over 30, Bill will ∅.

To summarize, if examples like (4.1.22) show anything significant at all, they illustrate a striking difference between the behavior of pronominal anaphors and null anaphors. Thus the interaction with negation does not at all support ESH, as Wasow claims, but rather provides evidence against it.

6. Wasow's last argument against a deletion theory of null anaphors concerns the problematic cases of Sluicing like (4.1.24), which as Ross (1969b) pointed out, come from ungrammatical sources (cf. (4.1.25)) in a deletion theory.

(4.1.24) John believes their claims about some products, but I don't know which products ∅.

(4.1.25) *John believes their claims about some products, but I don't know which products he believes their claims about.

The grammaticality of (4.1.24), it is argued, follows under ESH, if Ross's formulation of the Complex Noun Phrase Constraint in (4.1.4) above is accepted. That is, since the empty nodes through which the wh-phrase in (4.1.24) has moved have the structure of a complex NP, but lack a lexical head (since no lexical insertion has taken place there), (4.1.24) can be generated. Ross's formulation of the constraint together with ESH correctly predict, it is claimed, that precisely in those environments where no lexical insertion has taken place, extraction out of a complex NP environment is possible.

Wasow is again relying on the specifics of Ross's formulation of the Complex Noun Phrase Constraint. Interestingly, Ross's formulation and ESH seem to be incompatible. Thus compare the following two examples.

(4.1.26) Joan believed Bill's claim that he liked Sue, but who did she believe (Sam said) that Mike liked?

(4.1.27) *Joan believed Bill's claim that he liked Sue, but who did she believe Mike's Δ that he didn't like? [Δ = claim]

This minimal pair indicates that (4.1.27) should be ruled out as a violation of the Complex Noun Phrase Constraint. But in this example, the wh-Movement was extracted from an NP that has no lexically-present head (the head is presumably Δ under ESH). Thus Wasow cannot invoke Ross's formulation of the constraint without giving up the obvious explanation for the deviance of examples like (4.1.27).[5]

In short, examples like (4.1.24) are just as problematic for ESH as they are for a deletion theory. Hence they do not bear on the correctness of either. I don't have much to say about this general problem of sourceless Sluices except to note that the one proposed explanation that I am aware of, which is Chomsky's (1972b) suggestion that clauses containing violations of constraints be marked with "#", which (if left undeleted) doom a sentence to ungrammaticality at the level of surface structure, makes a very general empirical claim that seems to me to be incorrect. To see this, consider (4.1.28) and (4.1.29), which involve the interaction of VPD and Topicalization.

(4.1.28) ?Betsy liked Peter, but Larry, she claimed she didn't ∅.
[∅ = like]

(4.1.29) *Betsy liked the guy who told her about Peter, but Larry, she didn't ∅. [∅ = like the guy who told her about]

Chomsky's proposal is tantamount to the claim that deletion will always restore grammaticality to sentences which otherwise would violate island constraints. The contrast between (4.1.28) and (4.1.29) shows, I think, that no such general principle can be correct. Somebody should certainly solve this problem.

This exhausts Wasow's arguments against deletion theories. Aside from the brief argument made by Chomsky (1972b), (which has been dealt with in Section 2.3 of Chapter Two), there are no other arguments against such theories in the literature that I am aware of. Wasow's further arguments for "the unity of anaphora", will be taken up in Section 4.3. For the moment, however, it seems fair to conclude that Wasow has presented no convincing arguments against the general theory of deletion we have developed in the last three chapters.

4.2 Deep and Surface Anaphora

This section summarizes with some critical comment the major conclusions about English anaphoric processes drawn in Sag and Hankamer (1976).

As has long been known, certain anaphoric expressions[6] which typically receive their interpretation on the basis of some linguistic "antecedent", can under certain circumstances be assigned an interpretation purely on the basis of some aspect of the non-linguistic (or pragmatic) environment. This is the case, for instance, with definite third person pronouns. In (4.2.1), the underlined pronouns are interpreted on the basis of the underlined non-pronominal NP's. The examples in (4.2.2) may be uttered in the absence of (relevant) previous linguistic context if the extra-linguistic context is sufficient to supply a (more or less) unambiguous interpretation for the underlined pronominal anaphors.

(4.2.1)(a) My brother's a doctor, and he says your hair will fall out.

(b) Anyone who eats that will lose his hair.

(c) If the unicorn were actually an animal, it would certainly be a herbivore.

(4.2.2)(a) He's saying that your hair will fall out.

(b) Her hands are trembling.

(c) I hope it's a herbivore.

We will speak of the anaphors in (4.2.1) (when they are assigned their interpretation solely on the basis of the underlined antecedents) as being grammatically controlled, and the (deictic) anaphors in (4.2.2) as being pragmatically controlled.

Not all anaphoric processes can be pragmatically controlled. The following utterance-context event, for instance, where VPD has been attempted under pragmatic control, is ill-formed.

(4.2.3) [Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop]

Sag: It's not clear that you'll be able to ø.

Compare this with the minimally different event which transpires if (4.2.4), which involves <u>do it</u> Anaphora instead of VPD, is uttered in the very same (extra-linguistic) context:

(4.2.4) Sag: It's not clear that you'll be able to do it.

(4.2.5)(a),(b) is a minimal pair of exactly the same type.

(4.2.5) [Hankamer produces a cleaver and prepares to hack off his left hand]

Sag: Don't be alarmed, ladies and gentlemen. We've rehearsed this act dozens of times and...

(a) # he never actually does.

(b) he never actually does it.

<u>Do it</u> Anaphora, but not VPD, may be pragmatically controlled. The latter process requires the presence of a <u>linguistic</u> antecedent (= controller = deletion trigger) either in the sentence containing the null anaphor, or else in the preceding discourse. The fact that anaphoric processes separate into two classes, those that can be pragmatically controlled and those that cannot, is actually rather surprising. Given existing theories of anaphora, there is no particular reason to expect that any such dichotomy exists.

Notice that this difference between the behavior of VPD and do it Anaphora corresponds to another discrepancy between those two processes that has been pointed out in the literature. Grinder and Postal (1971) observe that the controller of a definite pronoun can sometimes be an NP that is part of a VP that has been deleted by VPD. This "missing antecedent" phenomenon is illustrated in the following examples.

(4.2.6) I've never ridden a camel, but Jorge has ∅, and he said it_i stank horribly. [∅ = ridden a camel_i]

(4.2.7) I don't keep gerbils in my office, although Judy does ∅, and she says they_i eat holes in her books. [∅ = keep gerbils_i in her office]

The ungrammaticality of the following examples shows that it must in fact be the missing NP's that are the antecedents of the underlined pronouns in (4.2.6) and (4.2.7).

(4.2.8) *I've never ridden a camel_i and Jorge says it_i stank horribly.

(4.2.9) *I don't keep gerbils_i in my office, and Judy says they_i eat eat holes in her books.

As Bresnan (1971) points out, however, do it Anaphora does not exhibit the "missing antecedent" phenomenon. This is shown by the following example.

(4.2.10) *Jack didn't slash the tent with a knife; Norma did it, and it was rusty.
[it = the knife Norma slashed the tent with]

Thus VPD requires grammatical control and VPD anaphors may contain "missing antecedents". Do it anaphors may be pragmatically controlled, and contain no "missing antecedents".

Bresnan goes on to show that Sentential-it Anaphora, of which do it Anaphora may very well be a special case, also fails to exhibit the "missing antecedent" phenomenon. This is illustrated by examples like these:

(4.2.11) *Mary didn't get picked off by a throw to first, but it happened to Bill, and it singed his ear.
[it = the throw to first that picked Bill off]

(4.2.12) *My uncle has never ridden a camel, but his brother managed it, and it was lame.
[it = the camel my uncle's brother managed to ride]

Sentential-it anaphors, like do it, may be pragmatically controlled:

(4.2.13) [Sag successfully rips Boston phone book in half]
Hankamer: I don't believe it.

(4.2.14) [same circumstance]
Sag: It's not easy.

Now judgements concerning "missing antecedent" sentences are notoriously labile. They vary from speaker to speaker as well as from moment to moment.[7] It is extremely difficult in many cases to know just what conclusions can be drawn. Notice, for instance, that the (minimally-different) analogues of (4.2.11) and (4.2.12) with VPD are not appreciably improved:

(4.2.15) ?*Mary didn't get picked off by a throw to first, but Bill did, and it singed his ear.
[it = the throw to first that picked Bill off]

(4.2.16) ?*My uncle has never ridden a camel, but his brother managed to, and it was lame.
[it = the camel my uncle managed to ride]

In Bresnan (1971) and in Sag and Hankamer (1976), the difference between VPD and Sentential-it Anaphora with respect to the "missing antecedent" phenomenon is argued to be crucial. At this writing, I am not convinced of that crucialness. Note further that even if the facts were quite clearly the opposite of what has been claimed (although I don't think the facts are "quite clearly" anything), it is difficult to see what would follow. The afore-mentioned authors claim, following Grinder and Postal (1971), that those anaphoric processes which exhibit the "missing antecedent" phenomenon should be derived by deletion and furthermore that those which do not should involve syntactically non-complex anaphors present in the base.

This does not necessarily follow. If it could be shown that, say, Sentential-it Anaphora sometimes does exhibit the phenomenon in question, all that would follow, assuming sentential-it pronouns are base-generated, is that interpretive rules can sometimes apply to the output of other interpretive rules. That is, if the rule that interprets sentential-it pronouns applies first, then in an example like (4.2.11), the rule that associates it with its antecedent can apply to the output of that rule. The extreme case of this, of course, would be the well-known cases of "anaphoric island" violations like (4.2.17) (see Postal (1969), Lakoff and Ross (1972), Ross (1971), Corum (1973)).

(4.2.17) John is a guitarist because he thinks it's a nice instrument.
[it = the guitar]

Sag and Hankamer (1976) do not consider this possibility and argue furthermore that there is a clear "clustering of properties" argument that leads to the conclusion that there is a fundamental dichotomy of anaphoric processes in English. Those processes which exhibit the "missing antecedent" phenomenon and which cannot undergo pragmatic control are termed Surface Anaphora. Those which do not exhibit the "missing antecedent" phenomenon and whose anaphors may be pragmatically controlled are termed Deep Anaphora. Surface anaphors are argued to arise at a relatively superficial syntactic level via the application of deletion transformations.

The crucial difference between the framework assumed by Sag and Hankamer and the one argued for in the previous chapters of this essay concerns where semantic interpretation takes place. In the former framework, the bulk of semantic interpretation takes place at the level of deep structure. In the latter, all semantic interpretation takes place at the level of shallow structure. Thus in the framework we have developed, the terminology "Deep" and "Surface" Anaphora is somewhat inappropriate.

Nevertheless, in spite of the skeptical remarks made above, there does indeed seem to be a dichotomy of anaphoric processes along the lines suggested by Sag and Hankamer. Those processes requiring grammatical control do indeed appear to be the processes which most readily exhibit the "missing antecedent" phenomenon. We will now examine various other anaphoric processes discussed by Sag and Hankamer. For the present discussion, we will continue to use the terminology "Deep" and "Surface" Anaphora.

For many anaphoric processes, e.g. Sluicing and Gapping, it is difficult to construct examples to show convincingly that they exhibit the "missing antecedent" phenomenon, for in virtually all cases, an NP in the controlling clause contains a possible controller for the pronoun in question. (4.2.18) is an attempt to demonstrate that Gapping does exhibit the phenomenon, though this example is already of reduced acceptability because it violates, among other things, the Minimal Distance Principle (see Section 3.2 of Chapter Three).

(4.2.18) Betsy$_i$ took her$_i$ clothes to the laundromat, and Harry$_j$ ∅ to the dry cleaner's, even though he knew very well <u>they</u> were all wash-and-wear.
[∅ = took his$_j$ clothes; <u>they</u> = the clothes Harry took to the dry cleaner's]

In spite of the unclarity of this example, it's nevertheless clear that Gapping is a case of Surface Anaphora. Gapping under pragmatic control is exceedingly bizarre:

(4.2.19) [Hankamer produces an orange, proceeds to peel it, and just as Sag produces an apple...]
Hankamer: And Ivan ∅ an apple.[8]

Sluicing is also an instance of Surface Anaphora. Compare the following context-utterance events which differ minimally in terms of whether the controller is an linguistic, or a non-linguistic entity.

(4.2.20) Hankamer: Someone's just been shot.
Sag: Yeah, I wonder who ∅.

(4.2.21) [Hankamer produces a gun, points it offstage and fires, whereupon a scream is heard]
Sag: Jesus, I wonder who ∅.

As a fourth instance of Surface Anaphora, consider Stripping, which, as (4.2.22) shows, can take place when the controller is (in) a previous sentence.

(4.2.22) Hankamer: Listen, Ivan; he's playing the William Tell Overture on the recorder.

Sag: Yeah, but ∅ not ∅ very well.

The controller must be a linguistic entity however. Consider the following bizarre context-utterance event where Stripping has been attempted under pragmatic control.

(4.2.23) [Sag plays William Tell Overture on recorder]

Hankamer: Yeah, but not very well.

Stripping anaphors may also contain missing antecedents, as (4.2.24) illustrates.[9]

(4.2.24) Alan took his hat off, but not Betsy ∅; she always refuses to take <u>it</u> off. [<u>it</u> = Betsy's hat]

Stripping is therefore quite clearly Surface Anaphora.

Conjunction Reduction (including Left-Peripheral Ellipsis - see Section 3.1 of Chapter Three) is virtually impossible to test for the "missing antecedent" phenomenon, for much the same reasons as Sluicing (see above). Nonetheless, Conjunction Reduction quite clearly is unacceptable under pragmatic control. The following examples illustrate this.

(4.2.25) [Hankamer has been peeling his orange for ten minutes]

Sag: And ∅ is dropping orange peels all over my foot.

(4.2.26) [Hankamer, Sag and Timberlake are playing volleyball. Timberlake, the setter, sets the ball too high for Sag to spike it. On the next volley, his set is too low for Hankamer, excellent spiker though he is, to get a good swipe at it, whereupon Sag utters...]

And ∅ too low for Jorge.

Conjunction Reduction is therefore a fifth instance of Surface Anaphora.

Let us turn now to one(s) Pronominalization. This anaphoric process, contrary to the claims of Grinder and Postal (1971), does not exhibit the "missing antecedent" phenomenon, as the next example shows.

(4.2.27) *Harry didn't sink a boat carrying a gorilla, but George sank one, and it drowned.

[it = the gorilla in the boat George sank]

Here it must be interpreted as the boat, if anything. One(s) Pronominalization is quite clearly possible under pragmatic control. Consider the following examples, for instance.

(4.2.28) [Sag produces an apple]

Hankamer: Did you bring one for me?

(4.2.29) [Betsy drives up in her brand new yellow Porsche]

Peter: The one she used to have was green.

One(s) Pronominalization therefore patterns with do it Anaphora and Sentential-it Pronominalization. It is an instance of Deep Anaphora.

At this point, one might be tempted to conclude that all pronominal anaphora is Deep Anaphora, and that those processes which involve null anaphors are all Surface Anaphora. Bresnan (1971) in fact suggests this generalization vis à vis the "missing antecedent" phenomenon. Curiously, no such conclusion is possible that can be seen by considering the phenomenon of do so Anaphora.

Do so cannot be pragmatically controlled:

(4.2.30) [Hankamer attempts to pass a 12-inch ball through a 6-inch noop]

Sag: I don't think you can do so.

(4.2.31) [Sag successfully rips Boston phone book in half]

Hankamer: I'm amazed that you can do so.

Moreover, contrary to what Wasow (1972, p. 113) claims, do so Anaphora does exhibit the "missing antecedent" phenomenon. The following example illustrates this point.

(4.2.32) I didn't ride a camel, but Jorge must have done so, and now our office is infested with its fleas.

[it = the camel Jorge must have ridden]

Note further that other kinds of so anaphora work similarly. When so occurs as the complement of verbs of saying or believing, for instance, it cannot be pragmatically controlled:

(4.2.33) [Hankamer succeeds in ripping Boston phone book in half]

Sag: I don't believe so.

But lurking somewhere in the structures (or derivations) of such sentences, are missing antecedents:

(4.2.34) Hankamer: Ivan, have you ever ridden a camel?

Sag: I believe you might say so; at least I sat on _its_ back while _it_ walked.

[_it_ = the camel Sag believes you might say he had ridden]

These cases of _so_ anaphora, then, are clear instances of _pronominal_ Surface Anaphora.

Interestingly, it seems that it is not even possible to maintain the weaker hypothesis that all anaphoric processes involving null anaphors are Surface Anaphora. This can be seen by considering a process termed by Sag and Hankamer Null Complement Anaphora (NCA). NCA is the anaphoric process involved in sentences like the following.

(4.2.35)(a) I asked Bill to leave, but he refused.

(b) Sue was attempting to kiss a gorilla, and Harry didn't approve.

(c) We needed somebody to carry the oats down to the bin, but nobody volunteered.

This process has been mentioned by Jackendoff (1972) and discussed extensively by Shopen (1972). Neither of them, however, distinguishes cases like (4.2.35), which involve null anaphors in the object position after verbs like _refuse_, _approve_, _complain_, _volunteer_, etc., from VPD. Shopen demonstrates, however, that these anaphors can be pragmatically controlled, as in these next examples.

(4.2.36) [Indulgent father feeds baby chocolate bar for dinner]

Mother: I don't approve $\emptyset$.

(4.2.37) [Two people are disturbed by loud noises of popcorn-eating in adjacent row]

One to the Other: Don't you think we should complain ∅?

"Missing antecedent" judgements involving NCA are admittedly rather delicate (and as noted above, difficult to draw conclusions from, even when clear), but it in fact seems to be the case that most informants reject examples like the following.

(4.2.38) *He said that one of us had to give up his seat, so Sue volunteered ∅, because <u>it</u> was too narrow for her anyway.

[<u>it</u> = Sue's seat]

Compare this to the following minimally different example involving VPD, which all my informants accept.

(4.2.39) He said that one of us had to give up his seat, so Sue volunteered to ∅, because <u>it</u> was too narrow for her anyway.

[<u>it</u> = Sue's seat]

NCA then is a process involving null anaphors, which is nonetheless a case of Deep Anaphora.

There are further arguments presented by Sag and Hankamer for the essentially "deep" nature of Deep Anaphora, and the essentially "surface" nature of Surface Anaphora. Consider (4.2.40) and (4.2.41), for instance.

(4.2.40) Nobody else would take the oats down to the bin, so

(a) Bill did ∅. [VPD (Surface Anaphora)]

(b) Bill did it ∅. [do it Anaphora (Deep Anaphora)]

(c) Bill volunteered ∅. [NCA (Deep Anaphora)]

(4.2.41) The oats had to be taken down to the bin, so

(a) *Bill did ∅. [VPD (Surface Anaphora)]

(b) Bill did it. [do it Anaphora (Deep Anaphora)]

(c) Bill volunteered. [NCA (Deep Anaphora)]

They conclude that Deep anaphors are interpreted at the level of deep structure and that clauses containing controllers of Deep anaphors are then free to undergo cyclic transformations (e.g. Passive, as in (4.2.41)). Surface anaphors must have the same superficial syntactic shape as their controllers, as the ungrammaticality of (4.2.41)(a) shows.

The conclusion that do it anaphors and Sentential-it anaphors can always be associated with a single entity present at the level of deep structure, however, is surely incorrect. In examples like the following there are no appropriate deep structure entities for the pronominal entities in question to be associated with in any straightforward way.[10]

(4.2.42)(a) I don't see why you shouldn't be forced to take the exam without a dictionary; they did it to the rest of us.

(b) John got stopped by a state trooper (and so did Bill), but if it ever happens to me, I'm gonna write my congressman.

There are many constraints on the interpretation of these pronominal anaphors which are mysterious for the moment. At the very least, however (as we remarked in Chapter Two), some sort of abstraction mechanism must be introduced to allow the it in, say, (4.2.42)(b) to be assigned an interpretation something like [x get stopped by a state trooper]. Such a mechanism would of course be compatible with either a theory of deep structure interpretation or the theory of shallow structure interpretation we have developed.

Whatever the precise nature of this mechanism turns out to be, it can be incorporated into the rule which interprets do it Anaphora and Sentential-it Anaphora at the level of shallow structure.

The most important conclusion then of Sag and Hankamer (1976) is that an adequate theory of anaphora must provide a principled account of why some anaphoric processes allow pragmatic control which others do not. The extent to which the possibility of the "missing antecedent" phenomenon correlates with this is not completely clear, yet I think it is reasonable to say that there is still a "clustering of properties" argument to be made for a fundamental dichotomy of anaphoric processes. This dichotomy, summarized in (4.2.43), must be explained by a theory of anaphora.

(4.2.43) Deep Anaphora

1. Do it Anaphora
2. Sentential-it Pronominalization
3. Null Complement Anaphora (NCA)
4. one(s) Pronominalization

Surface Anaphora

1. Verb Phrase Deletion (VPD)
2. Gapping
3. Sluicing
4. Stripping
5. So Anaphora
6. Conjunction Reduction

In the next section, we will consider how this dichotomy might be accounted for in our "mixed" theory or in Wasow's ESH.

4.3 The Non-Unity of Anaphora

The fundamental dichotomy of anaphoric processes observed in the last section provides to my mind the strongest justification for a "mixed" theory of anaphora of the sort we have adopted. Even granting the unclarity of the facts (and the conclusions that can be drawn from them) with respect to the "missing antecedent" phenomenon, the fact that some anaphors but not others can be pragmatically controlled is surely an important fact that a theory of anaphora must explain.

In a theory which distinguishes deletion rules from interpretive rules, the existence of such a dichotomy is in fact explained. The theory of deletion, which, as we have seen in the preceding chapters, must stipulate that deletions are recoverable at the level of logical form, has as a trivial consequence that deletion anaphors cannot be pragmatically controlled. Since non-linguistic entities have no representation at the level of logical form, no deletion anaphors con ever be pragmatically controlled.

In a mixed theory of the sort we have proposed, the further generalization can be stated that interpretive rules apply to interpret anaphoric expressions on the basis of either the linguistic context or the extra-linguistic context. The existence of two classes of anaphoric processes follows from the existence of two classes of anaphoric mechanisms: deletion rules and interpretive rules. The account is straightforward and intuitive.

How then, we might ask, would the existence of a dichotomy of anaphoric processes be reconciled with ESH? Edwin Williams (personal communication) has suggested a way. He suggests that, accepting ESH, one can simply stipulate that null anaphors require grammatical control, i.e. that the inability to be pragmatically controlled is a property of empty nodes.

Observe that any such proposal encounters difficulty with the process we termed Null Complement Anaphora (NCA). NCA, we argued, is a case of Deep Anaphora, and hence, as Shopen (1972) showed, may be pragmatically controlled (see the discussion in Section 4.2 above).

Williams (1975, 1976) is aware of this difficulty, and therefore argues that NCA is not anaphora at all, but rather simply a case of various verbs occurring intransitively. As evidence for this claim, Williams offers examples like the following one, the grammaticality of which, he argues, shows that NCA does not obey the "fundamental law of anaphora", the "Backwards Anaphora Constraint" (BAC), and hence cannot possibly be anaphora.

(4.3.1) John refused (ø) when we asked him to leave. [ø = to leave]

He finds a crucial difference between (4.3.1) and (4.3.2) which involves VP-Anaphora in his system (=VPD).

(4.3.2) (*)John refused to ø when we asked him to leave. [ø = leave]

Frankly, I don't see much difference between the two cases. Both sound peculiar in isolation if the accent falls on <u>leave</u>. In the right context, however, either sentence seems possible in the intended interpretation:

(4.3.3) When did John refuse to leave?
He refused (to) ø when we asked him to leave.

The well-formedness of this last discourse exhibits a general property of certain anaphoric processes, which we will have more to say about shortly. For the moment, though, I think it's fair to say that there's no striking discrepancy here between VPD and NCA. In short, though the facts are not particularly clear with respect to NCA, the following examples show that NCA does indeed obey BAC, about which we will also have more to say in a moment.

(4.3.4)(a) Harry went to the store, but I refused ø.

[ø = go to the store]

(b) *I refused ø when Harry went to the store.

[ø = to go to the store]

(4.3.5)(a) I don't approve of her driving a truck, though she's old enough ø. [ø = to drive a truck]

(b) Though she's old enough ø, I don't approve of her driving a truck. [ø = to drive a truck]

(c) Though she's old enough to drive a truck, I don't approve ø.

[ø = of her driving a truck]

(d) *I don't approve ø, though she's old enough to drive a truck.

[ø = of her driving a truck]

Williams' example above ((4.3.1)), I think, is acceptable only when ø (the null anaphor) is controlled from elsewhere in the linguistic or extra-linguistic environment.

Notice that there is another argument to be made against analyzing the above examples as intransitive verbs. Almost all other cases of intransitive uses of verbs that are ordinarily transitive require an "unspecified object" sort of interpretation. Consider (4.3.6), for instance (example due to Sag and Hankamer (1976)).

(4.3.6) I bring him soup and potatoes, but he won't eat.

This sentence does not mean that he specifically won't eat soup and potatoes, but rather that he won't eat anything at all. A sentence like (4.3.7), on the other hand, means specifically that my wife doesn't approve of my playing cards and shooting dice, not that she doesn't approve of anything in general.

(4.3.7) I play cards and shoot dice, and my wife doesn't approve ∅.

This is in fact characteristic of cases of NCA. This interpretational difference provides further evidence aginast Williams' suggestion that these are simply intransitive usages.

Leaving NCA aside for the moment, notice that there are other anaphoric processes which involve phonetically null anaphors that can be pragmatically controlled. Consider the process involved in sentences like the following:

(4.3.8)(a) Betsy's paintings are good, and Peter's ∅ are bad.

[∅ = paintings]

(b) I like Bill's wine, but Max's ∅ turns me off.

[∅ = wine]

This is the anaphoric process referred to by Jackendoff (1971) as $\bar{N}$-Deletion. As (4.3.8)(b) shows, the standard analysis of these cases (e.g. the one of Perlmutter (1970)) as <u>one(s)</u> Deletion is defective, for the deleted element is a mass noun which cannot be pronominalized by <u>one</u> (see Jackendoff (1971) for some discussion):

(4.3.9) *Bill's red wine is good, but Max's white wine <u>one</u> is bad.

[<u>one</u> = wine]

The null anaphors of sentences like those in (4.3.8) can, however, be pragmatically controlled. Consider the following example.

(4.3.10) [We walk into a room full of Harry's paintings, whereupon I exclaim...]

Betsy's ø must be in the next room.

Moreover, this process is surely anaphora. The following examples show it obeys BAC.

(4.3.11)(a) Peter's mother knows why Betsy's ø left. [ø = mother]

(b) *Peter's ø knows why Betsy's mother left. [ø = mother]

In (4.3.11)(b), ø can of course be controlled by mother somewhere in the preceding discourse, which is immaterial to our present concerns (see the above discussion). What Jackendoff called $\overline{N}$-Deletion then is another case of Deep Anaphora involving null anaphors.[11]

Another case in point is the process involved in sentences like the next two.

(4.3.12)(a) You've given me two alternatives, but both ø are unsatisfactory. [ø = alternatives]

(b) Many people are smart, but few ø are brave. [ø = people]

This anaphoric process also allows pragmatic control, as the following example shows.

(4.3.13) [You present me with two presents, whereupon I say...]

Both ø are unsatisfactory.

And, as one might expect, we have a case of real anaphora on our hands, for this process obeys BAC:

(4.3.14)(a) My father and my mother have told me that both ø are unhappy with my progress. [ø = my father and my mother]

(b) *Both ø have told me that my father and my mother are unhappy with my progress. [ø = my father and my mother]

The process illustrated in (4.3.12)-(4.3.14) is therefore another case of Deep Anaphora involving null anaphors.

This process and the one Jackendoff called $\bar{N}$-Deletion may in fact be a single process. Both involve NP's whose heads are null. Let us therefore tentatively refer to the entire class of cases as Null NP-Head Anaphora (N^2PHA).

For some speakers, apparently, N^2PHA is more general than the sentences we have considered so far would suggest. Thus Harris (1965, 1968) and Quirk et al. (1972, p. 590) cite examples like the following.

(4.3.15)(a) We wanted fried fish, but they gave us boiled ø. [ø = fish]

(b) She wore the red dress, but the blue ø suits her better. [ø = dress]

(c) He prefers Dutch cheese, and I prefer Danish ø. [ø = cheese]

The acceptability of such examples varies considerably from speaker to speaker. Nevertheless it would seem reasonable to treat the examples o. (4.3.15) as N^2PHA.

One might attempt to handle the above cases of N^2PHA by positing a rule that optionally deletes <u>of them</u>. I don't think this is really a viable alternative. In examples like the following, it would have to be assumed that of it was the deleted element.

(4.3.16)(a) I've eaten a lot of the cake, but any more ø would make me sick.

(b) Harry at a third of the cake, and Betsy ate a quarter ø.

Moreover, in examples like (4.3.8)(b), and (4.3.15)(c) above and in examples like the following one, there is no source in such a deletion analysis.

(4.3.17) Much happiness can be found in our religion, but very little $\{{}^{\emptyset}_{\text{*of it}}\}$ can be found in your religion.

N^2PHA anaphors appear to be best treated as base-generated empty nodes, which receive their interpretation by interpretive rules.[12]

N^2PHA and NCA therefore show that Williams' suggestion is untenable. Within ESH, one cannot simply stipulate that empty nodes cannot be pragmatically controlled. To account for the facts of N^2PHA and NCA within ESH, one would have to posit two kinds of empty nodes, those that can be pragmatically controlled and those that cannot. This, I think, constitutes a strong metatheoretical argument that ESH has no natural way of accounting for the dichotomy between Deep and Surface Anaphora.

The existence of that dichotomy, however, follows directly from the "mixed" theory of anaphora we have proposed, as noted above. All semantic interpretation takes place at the level of shallow structure. All anaphoric elements present there may be pragmatically controlled. Deletion rules apply to convert shallow structures into surface structures. Since deletion is subject to recoverability at the level of logical form, the null anaphor produced by a deletion rule arises only when there is some linguistic entity (the deletion controller) which induces that logical recoverability. Also as previously noted, it is a trivial consequence of the theory of deletion that those null anaphors that are the result of deletion rules cannot be pragmatically controlled.

This of course leaves open the possibility that there are other null anaphors that do not arise by deletion, but are present at the level of shallow structure and hence susceptible to pragmatic control. NCA and N^2PHA would be examples of this latter type of null anaphor.

The "mixed" theory of anaphora we have proposed therefore predicts that there is a principled distinction between two kinds of anaphoric processes. This prediction is confirmed by the existence of the dichotomy between Deep and Surface Anaphora. The fact that this is the case provides a strong argument for accepting our "mixed" theory and rejecting ESH.

Another crucial argument for a distinction between deletion rules and interpretive rules can be made by examining the behavior of various anaphoric processes with respect to coordinate structures. Immediately conjoined VP's, we have observed, cannot undergo VPD:

(4.3.18) *I couldn't lift this rock, but I know a boy who can ∅ and bend a crowbar, too. [∅ = lift this rock]

We offered an explanation of this fact in Chapter Three in terms of our Immediate Domination Principle (IDP), which, as *(4.3.18) shows, VPD obeys.

It is easily demonstrated that other deletion rules, e.g. Sluicing, also obey IDP. Compare these next two examples.

(4.3.19) John ate a bagel, but I don't know why he ate a bagel and Betsy didn't (eat a bagel).

(4.3.20) *John ate a bagel, but I don't know why ∅ and Betsy didn't (eat a bagel).

This behavior of VPD and Sluicing is accounted for by the view we put forth as to the nature of the interaction of the principle of recoverability of deletion and IDP (see Chapters One and Three). It appears that the following generalization can be stated:

(4.3.21) All deletion rules obey IDP.

It is well-known, however, that pronominal anaphora does not work this way. Definite pronouns, for instance, may be associated with their antecedents even when those pronouns are immediately conjoined with another NP:

(4.3.22)(a) Peter said he and Betsy were going to be late.

(b) Peter said Betsy and he were going to be late.

If definite pronouns are associated with their antecedents by the same rule which associates null VP anaphors are associated with their controllers, as Wasow suggests, then the observed discrepancy remains quite problematic. Thus the fact that deletion rules obey IDP and interpretive rules do not constitutes a second argument for a "mixed" theory, which distinguishes between the two types of rules, and a second argument against a theory like ESH, which does not.

Now whereas the two arguments just given provide evidence for distinguishing two kinds of anaphoric processes, this is not to deny that there may be properties shared by anaphoric processes of both types. Wasow in fact presents five arguments (corresponding to five properties shared by anaphoric processes of various types) which he takes as showing that all anaphoric processes should be handled by "a single mechanism or a single type of mechanism". We will examine those arguments with care, pointing out various additional facts (not discussed by Wasow) which, as we will see, suggest rather different conclusions.

One of Wasow's arguments (which is due to Akmajian) is that "the antecedent in an anaphoric relation may contain a negative element not included in the interpretation of the anaphor." We have already dealt with this claim in Section 4.1, where we saw the severe difficulties that would be encountered by any theory of anaphora that allows processes like VPD and Sluicing to "optionally ignore negation". The only convincing case of an anaphoric process that may be assigned an interpretation that corresponds to the "interpretation" of its controller "minus negation" is Sentential-it Anaphora. But in light of the remarks made in Section 4.2 above, it must be concluded that this is a specific case of a more general property of the interpretation of Sentential-it anaphors.

The interaction of negation and anaphora provides no argument at all for the "unity of anaphora", and we will have nothing more to say about that here.

Another argument offered by Wasow concerns the observation (again following Akmajian) that "certain adjectival and adverbial modifiers may sometimes be ignored by anaphora rules" (p. 91). This is illustrated by the following examples. Each of the (a) sentences is ambiguous between a reading equivalent to that of the corresponding (b) sentence and one equivalent to that of the corresponding (c) sentence.

(4.3.23)(a) John has a big fancy car, but Bill doesn't have one.

(b) John has a big fancy car, but Bill doesn't have a big fancy car.

(c) John has a big fancy car, but Bill doesn't have a car.

(4.3.24)(a) John beats Mary because he hates her, and Bill does

$$\left\{\begin{array}{l}\text{it too}\\ \text{likewise}\\ \text{the same thing}\end{array}\right\}$$

(b) John beats Mary because he hates her, and Bill beats Mary because he hates her.

(c) John beats Mary because he hates her, and Bill beats Mary too.

(4.3.25)(a) John has been approached by strange women in New York, and it also happened to Bill.

(b) ...and Bill has also been approached by strange women in New York.

(c) ... and Bill has also been approached by strange women.

(4.3.26)(a) I suspect that the DA accidently suppressed evidence, and even Perry believes it.

(b) ...and even Perry believes that the DA accidently suppressed evidence.

(c) ...and even Perry believes that the DA suppressed evidence.

(4.3.27)(a) Yesterday, John jogged a mile in spite of the rain, and today Mary did so.

(b) ...and today Mary jogged a mile in spite of the rain.

(c) ...and today Mary jogged a mile.

As Wasow himself notes (p. 123, nt.3), however, null anaphors do not have this property. He cites the following examples, where the (a) sentences cannot be interpreted as synonymous with the corresponding (b) sentences.

(4.3.28)(a) Yesterday, John jogged a mile in spite of the rain, and today, Mary did ∅.

(b) Yesterday, John jogged a mile in spite of the rain, and today, Mary jogged a mile.

(4.3.29)(a) Mary beats John because she loves him, and Sara does ∅, too.

(b) Mary beats John because she loves him, and Sara beats John, too.[13]

The facts of (4.3.27)-(4.3.29) are not at all clear to me. If Wasow's observation concerning (4.3.28)-(4.3.29) is correct, I think it also holds of (4.3.27). The Sluicing facts may be a little clearer. I don't think (4.3.30)(a) can be synonymous with (4.3.30)(b).

(4.3.30)(a) John eats apples in New York City, but I don't know why ∅.

(b) John eats apples in New York City, but I don't know why he eats apples.

Note that the anaphoric processes that freely allow modifiers to be "ignored" seem to be all cases of Deep Anaphora. The judgements involving the do so and VPD cases are sufficiently unclear, however, that no conclusions can be drawn with much confidence.

On the other hand, given the theory of deletion developed in Chapter Two, in particular the dependence of deletion rules on matters of logical form, it's hard to see what would follow about a theory of anaphora even if VPD and Sluicing worked exactly like other anaphoric processes in this regard. In short, none of the facts considered provides evidence for choosing between ESH and the deletion theory we have proposed.

The other three arguments offered by Wasow concern constraints that all anaphoric processes are argued to obey. We will first give a brief sketch of those arguments, and then consider some further data.

Wasow takes as "the most compelling piece of evidence for treating different anaphoric relations uniformly" Ross's (1967a) observation that all bidirectional anaphoric processes (VPD and Sluicing included) are subject to (the same) BAC. This constraint which Wasow states (with certain qualifications), roughly as (4.3.31),[14] is illustrated in the examples that follow [= Wasow's (1)-(8) (pp. 89-90); underlined elements are in an anaphoric relation].

(4.3.31) If an anaphor precedes its antecedent (= its controller) then it must also be more deeply embedded than its antecedent.

(4.3.32)(a) John dropped out after he tried LSD.

(b) After John tried LSD, he dropped out.

(c) After he tried LSD, John dropped out.

(d) *He dropped out after John tried LSD.

(4.3.33)(a) John tried LSD after Bill did ∅.

(b) After Bill tried LSD, John did ∅.

(c) After Bill did ∅, John tried LSD.

(d) *John did ∅ after Bill tried LSD.

(4.3.34)(a) John tried LSD after Bill had done so.

(b) After Bill had tried LSD, John did so.

(c) After Bill had done so, John tried LSD.

(d) *John did so after Bill had tried LSD.

(4.3.35)(a) John believes that Bill takes LSD, although no one else believes it.

(b) Although no one else believes that Bill takes LSD, John believes it.

(c) Although no one else believes it, John believes that Bill takes LSD.

(d) *John believes it, although no one else believes that Bill takes LSD.

(4.3.36)(a) John will take LSD if Bill does it.

(b) If Bill takes LSD, John will do it.

(c) If Bill does it, John will take LSD.

(d) *John will do it, if Bill takes LSD.

(4.3.37)(a) John takes LSD, although I don't know why ø.

(b) Although John takes LSD, I don't know why ø.

(c) Although I don't know why ø, John takes LSD.

(d) *I don't know why ø, although John takes LSD.

(4.3.38)(a) John dropped a capsule of LSD after Bill took one.

(b) After Bill took a capsule of LSD, John dropped one.

(c) After Bill took one, John dropped a capsule of LSD.

(d) *John dropped one after Bill took a capsule of LSD.

(4.3.39)(a) John freaked out, although it wouldn't have happened to Bill.

(b) Although John freaked out, it wouldn't have happened to Bill.

(c) Although it wouldn't have happened to Bill, John freaked out.

(d) *It wouldn't have happened to Bill, although John freaked out.

[all of the above judgments are Wasow's]

Wasow interprets the fact that all the above processes obey BAC as showing either that there is only one anaphora rule that applies (subject to BAC) or that BAC is a constraint only on a "formally distinguishable class" of anaphora rules. ESH is consistent with either alternative.

Wasow seems to be assuming that a "mixed" theory which treats pronouns interpretively but which derives null anaphors via deletion transformations is inconsistent with the (alleged) uniform behavior of both types of processes with respect to BAC. Although he does not come out and say this, it is evident that he is making some such assumption.

Another closely related argument concerns the "Transitivity Condition", which is formulated by Jackendoff as the following.

(4.3.40) If A, B, and C are three elements in a sentence such that an anaphoric relation holds between A and B and an anaphoric relation holds between B and C, then the sentence is marked ungrammatical unless an anaphoric relation holds between A and C.

Perhaps the clearest justification for this condition concerns the battery of examples in (4.3.41).

(4.3.41)(a) The woman <u>he</u> loved said that <u>John</u> was a jerk.

(b) The woman <u>John</u> loved hurt <u>him</u>.

(c) *Mary told <u>him</u> that <u>John</u> was a jerk.

(d) *The woman <u>he</u> loved told <u>him</u> that <u>John</u> was a jerk.

Assume none of the pronouns in (4.3.41) is deictic and, therefore, that in each of these sentences every pronoun must be associated with some antecedent in the same sentence. The problem is (4.3.41)(d). <u>him</u> and <u>John</u> cannot be in an anaphoric relation (as (4.3.41)(c) shows) but he and him can be (as (4.3.41)(b) shows) and he and John can be (as (4.3.41)(a) shows). There is a legitimate antecedent for each pronoun, yet the example is ill-formed. The Transitivity Condition accounts for this.

Wasow (following Williams (1971)) argues that VPD also obeys the Transitivity Condition. The next two examples, Williams argues, are grammatical.

(4.3.42) Because Mary didn't ∅, James didn't want to join the party.

(4.3.43) Because nothing happened until some time after Sam joined the party, James didn't want to ∅.

Therefore, unless VPD obeys the Transitivity Condition, we should expect (4.3.44) also to be grammatical, which it is not.

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

The fact that VPD and definite pronominalization both obey the Transitivity Condition, Wasow argues, is evidence in support of a theory like ESH, which treats both uniformly.

The last of Wasow's arguments for the "unity of anaphora" concerns a constituent he states as the following.

(4.3.45) No part of the complement of the specifier of a cyclic node may be anaphorically related to the head of that node (where the head of S is VP, and the head of NP is $\overline{N}$).

This constraint (the ad hoc nature of which Wasow laments) would account for facts like the following.

(4.3.46)(a) *A proof that God exists does $\emptyset$.

(b) *A proof that God does $\emptyset$, exists.

(4.3.47)(a) *Learning that vitamin C improves people's health does so.

(b) *Learning that vitamin C does so improves people's health.

(4.3.48)(a) *The fact that LSD causes people to freak out does $\left\{\begin{array}{l}\text{it (too)}\\ \text{likewise}\\ \text{the same thing}\end{array}\right\}$

(b) *The fact that LSD does $\left\{\begin{array}{l}\text{it (too)}\\ \text{likewise}\\ \text{the same thing}\end{array}\right\}$ causes people to freak out.

(4.3.49)(a) *A trainer of horses' ones are generally healthier than mustangs.

(b) *The winner of the game's one was off today.

These curious facts, Wasow argues, illustrate another property shared by various anaphoric processes, and hence constitute a further argument for the correstness of ESH, which treats all such processes uniformly.

Now these last three arguments of Wasow's are of considerable interest. If indeed there are general constraints on all anaphoric processes of the sort Wasow suggests, it is certainly the case that an adequate theory of anaphora should be able to express the generalization involved, if not explain it.

But just how uniformly do these various anaphoric processes behave with respect to constraints like BAC, if indeed that constraint is properly construed as a prohibition against "backwards" anaphora (see Reinhart (1974, forthcoming) for a different view)? There are various discrepancies to be observed which Wasow does not consider.

Notice first of all that in an example like (4.3.50), he and John may not be anaphorically related even in a context like (4.3.51), where he has another possible antecedent in the preceding discourse (this point is made by Lasnik (1976)).

(4.3.50) *He thinks John is unpopular.

(4.3.51) *John has problems. He thinks John is unpopular.

This is a property of definite pronouns that is not shared by VPD or Sluicing. Thus whereas (4.3.52) and (4.3.53) appear to be deviant in isolation, the discourse in (4.3.54) and (4.3.55) are well-formed, though perhaps somewhat redundant.[15]

(4.3.52) *He did ∅ when they asked him to leave. [∅ = leave]

(4.3.53) *I don't know why ∅, although it seems clear that he left.
[∅ = he left]

(4.3.54) Did Harry leave? He did ∅ when they asked him to leave.

(4.3.55) Why did Harry leave, if indeed he did? I don't know why ∅, although it seems clear that he left.

Notice that similar behavior can be observed with respect to other anaphoric processes:

(4.3.56)(a) *Bill did so after Harry raised his hand.

(b) Who raised his hand? Bill did so after Harry raised his hand.

(4.3.57)(a) *John believes it, although no one else believes that Bill takes LSD.

(b) Who believes that Bill takes LSD? John believes it, although no one else believes that Bill takes LSD.

(4.3.58)(a) *John will do it after Bill takes LSD.

(b) Who will take LSD? John will do it after Bill takes LSD.

(4.3.59)(a) *John dropped one, although Bill's never taken a capsule of LSD.

(b) Who dropped a capsule of LSD? John dropped one, although Bill's never taken a capsule of LSD.

(4.3.60)(a) *It happened to Bill, although John never freaked out on LSD.

(b) Who freaked out on LSD? It happened to Bill, although John never freaked out on LSD.

Definite pronouns seem to be the only anaphors which are subject to BAC even when they have another possible controller in the preceding discourse.

The same in fact seems to be true of the Transitivity Condition. Recall the facts cited by Wasow (following Williams (1971)) to show that VPD obeys the Transitivity Condition:

(4.3.61) Because Mary didn't $\emptyset$, James didn't want to join the party.

(4.3.62) Because nothing happened until some time after Sam joined the party, James didn't want to $\emptyset$.

(4.3.63) *Because Mary didn't $\emptyset$ until sometime after Sam joined the party, James didn't want to $\emptyset$.

The way I see it, the reasoning here is impeccable, but the facts upon which the reasoning is based have not been examined carefully enough. (4.3.62) is indeed grammatical, but only in a discourse where joined the party is a VP in some preceding sentence. Every single one of my informants rejects

(4.3.62) in isolation, or in a context like (4.3.64), but accepts it in a context like (4.3.65).

(4.3.64) *Why did James leave? Because nothing happened until some time after Sam <u>joined the party</u>, James didn't want to <u>ø</u>.

(4.3.65) Why didn't James <u>join the party</u>? Because nothing happened until some time after Sam <u>joined the party</u>, James (just) didn't want to <u>ø</u>.

Notice that the only thing wrong with (4.3.64) (and hence the only thing wrong with (4.3.62) in isolation) is that, for some reason, the embedded VP <u>joined the party</u> is not a possible controller for the null VP anaphor to its right. The fact that the following discourse is well-formed makes this clear:

(4.3.66) Why did James leave? Because nothing happened until some time after Sam <u>joined the party</u>, James didn't want to <u>join the party</u>.

In contexts like (4.3.65), which are the only kind of contexts in which (4.3.62) is well-formed, the null VP anaphor is therefore in an anaphoric relation only with the VP in the preceding sentence.

But if these are the facts, then the entire argument that VPD obeys the Transitivity Condition collapses. The ungrammaticality of (4.3.63) (in isolation) follows from the fact that (4.3.62) is ungrammatical in isolation. The only undeleted VP in (4.3.63) (repeated here for the reader's convenience) can control neither of the null VP anaphors.

(4.3.63) *Because Mary didn't <u>ø</u> until sometime after Sam <u>joined the party</u>, James didn't want to <u>ø</u>.

Thus, without any relevant preceding context, the only way (4.3.63) could arise would be if each of the deleted VP's was the other's controller. But totally unrecoverable deletions of this sort are never allowed. Note that this conclusion is further strengthened by the fact that (4.3.63) is actually not ungrammatical at all, if it occurs in a context like the following one.

(4.3.67) Why didn't James join the party? Because Mary didn't ø until sometime after Sam joined the party, James just didn't want to ø.

The acceptability of (4.3.63) is then a function of the acceptability of (4.3.62), and has nothing to do with the Transitivity Condition.

Note further that all the other anaphoric processes Wasow discusses work essentially like VPD. Consider do it Anaphora, for instance. For some reason that I do not understand, the do it analogue of (4.3.62) above seems perfectly possible, even in the absence of a controller in the preceding discourse:

(4.3.68) Because nothing happened until some time after Sam took off all his clothes, James didn't want to do it.

Predictably, the do it analogue of (4.3.63) is also possible without a controller in the preceding discourse:

(4.3.69) Because Mary didn't do it until some time after Sam took off all his clothes, James didn't want to do it (either).

If (4.3.70) is ungrammatical, that has no effect whatsoever on the possibility of (4.3.69).

(4.3.70) *Mary didn't do it until some time after Sam took off all his clothes.[16]

Thus is would seem that do it Anaphora, like VPD, fails to obey the Transitivity Condition. It will be left as an exercise for the reader to verify similar facts with Sentential-it Anaphora, Sluicing, and the like.

Notice that further facts along these lines can be observed with respect to the facts involved in the last of Wasow's arguments. Thus (4.3.71) is ungrammatical only in isolation. The context in (4.3.72) is sufficient to restore grammaticality.

(4.3.71) *A proof that God exists does ø. [ø = exist]

(4.3.72) What exists? A proof that God exists does ø. [ø = exist]

What all these facts show, first of all, is that it is impossible to explain BAC violations as a single constraint on the application of a single interpretive rule which associates both definite pronouns and null VP anaphors with their antecedents. Whatever the appropriate statement of BAC for definite pronouns is it must have the Transitivity Condition as a consequence. The proper formulation of BAC for VPD must not have the Transitivity Condition as a consequence. The facts, therefore, suggest that the two phenomena are quite distinct.

I do not pretend to have an adequate understanding of BAC. Two hypotheses, however, do come to mind: Perhaps (1) there is some way to reconcile the discrepancies we have observed with the existence of BAC as a unified entity, operating as a very superficial "filter" or as a rule stipulating that two elements in a certain configuration cannot be anaphorically related (the former is suggested by Sag and Hankamer (1976), the latter by Lasnik (1976)).

This would require some way of defining the notion "anaphorically related" in some way so as to include deletion anaphora and interpretive anaphora (see Hankamer (ms.) for one attempt to do this). Alternatively (2) there is no unified BAC. Rather there are general constraints on the application of rules of a certain class (those rules which perform associations of entities of logical form, perhaps).

The choice between the two at this point in time and space is not clear. It is clear, however, that these last three arguments of Wasow's, in light of the wider range of facts we have considered here, do not show at all that we must treat all anaphoric processes interpretively. Moreover, since there are at least two good arguments for distinguishing deletion rules from interpretive rules (the two we gave at the beginning of this subsection), we will continue to do so.

4.4 Summary of Chapter Four

Chapter Four has dealt with a variety of anaphoric processes which have been seen to exhibit certain distinguishable properties. We have examined Wasow's (1972) arguments that a single mechanism (or type of mechanism) is involved in the interpretation of all anaphors, and found those arguments to be unconvincing. Similarly, we have considered all of Wasow's arguments against a deletion theory of such phenomena as VPD and Sluicing and have shown that they are not arguments against the theory of deletion we have developed in the preceding chapters. In the process of doing this, we have uncovered various facts which serve to divide English anaphoric processes into two coherent classes. Wasow's Empty Structures Hypothesis, as we have seen, provides no basis for accommodating this fundamental dichotomy. In order to obtain an adequate account of the processes studied in this chapter, we must assume a "mixed" theory of anaphora. Certain anaphors recieve their interpretation via rules of semantic interpretation. Other anaphors arise by deletion.

Footnotes to Chapter Four

1. This, of course, renders the notion "syntactically anomolous" completely vacuous, for if anything is deviant for syntactic reasons, the examples in (4.1.1) certainly are. I can't figure out, however, if this objection, which some people feel in their gut is decisive, is very serious.

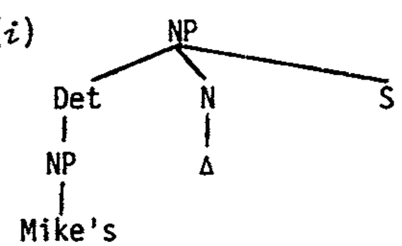
2. See Williams (1975, 1976).

3. See also the remarks in Section 3.4 of Chapter Three.

4. This is as good a place as any to register my general skepticism for the "Strict Cycle Principle" in the domain of syntax. As far as I know, no one has ever presented a single convincing case of a fact it would explain. Therefore these last two arguments of Wasow's are, in my opinion, even more dubious than is indicated in the text.

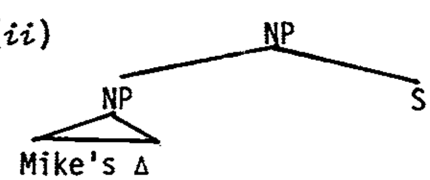
5. As Jorge Hankamer points out to me, however, this argument only goes through under the assumption that the constituent structure of the relevant NP's in (4.1.27) is as in (*i*).

(*i*)



If instead, the structure in (*ii*) is the correct one, then a proponent of Wasow's theory can argue that there still exists a partial lexical head.

(*ii*)



While the structure in (*ii*) strikes me as counter-intuitive, I have no good arguments against it at the moment.

Some people, by the way, find sentences similar to (4.1.27), but where no extraction has taken place, to be of reduced acceptability:

(*iii*) (?)Joan believed Bill's claim that he liked Sue, but she denied Sam's Δ that he liked Mary.

No speakers, however, fail to find a contrast between (*iii*) and (4.1.27), which leaves the essence of the argument unimpaired.

6. Sag and Hankamer define anaphoric process as "any grammatical device which allows the interpretation of an element to be chosen from an infinite number of potential values, the choice in a particular instance being determined by context" (nt. 3).

This conception of anaphora, is an explicit statement of what has come to be a standard (though etymologically untrue) conception of that notion in the linguistic literature. On this view, anaphora subsumes what are sometimes distinguished as anaphora vs. cataphora (see Halliday (1967)) or anaphora vs. epiphora (see Hiż (1969)). The term anaphor (due to Edes (1968)) is roughly equivalent to Hiż's referential.

7. The prose here echoes the lament of Bach, Bresnan and Wasow (1974) concerning judgements of sentences involving "sloppy identity".

8. Possible counterarguments to the validity of this example are discussed in Sag and Hankamer (1976).

9. As a quick comparison of (4.2.21) with (*i*) will show, the proper formulation of the rule of Stripping is by no means obvious.

(*i*) *Alan took his hat off, but not Betsy took her hat off.

10. See also the discussion in Section 2.2 of Chapter Two and the references cited there.

11. I note in passing that a very similar process (see (*i*)) referred to by Jackendoff (1971) as $\bar{N}$-Gapping seems to be a case of Surface Anaphora.

(*i*) Peter's pictures of Nixon are terrible, but Betsy's ø of Agnew aren't bad at all. [ø = pictures]

As (*ii*) shows, $\bar{N}$-Gapping cannot be pragmatically controlled.

(*ii*) [We walk into a room full of Peter's pictures, each of which is a portrait of Richard M. Nixon, whereupon I say...]
*Betsy's ø of Agnew must be in the next room.

12. If I read Williams (1976) correctly, he also suggests this.

13. I have taken the liberty of rescuing poor Mary from the ordeal Wasow's examples put her through.

14. I have taken the liberty of reformulating Wasow's statement of BAC (p. 52) in light of his subsequent comments (p. 88) and in terms of "precedes" rather than "to the left of".

15. This point came up in viva voce discussion with Bob Fiengo.

16. Actually, I think this example is not well-formed without a controller for do it in the previous discourse.

Inconclusion

As is only appropriate for an essay on the topic of deletion, there is much that has been left unsaid. You may have wondered, for instance, just what the rules of SI-1 referred to in Chapter Two looked like. Well, the rule that interprets quantifiers would like something like this (where i is an index on a quantifier word):

$$(1) \quad [_S\, W_1 - Q_i - W_2] \Rightarrow (Q\, v_i)\, [_S\, W_1 - v_i - W_2]$$

This might be modified in the following way to account for relative clauses which modify quantifier words.

$$(2) \quad [_S\, W_1 - [Q_i - \bar{S}] - W_2\,] \Rightarrow (Qv_i\!: \bar{S}\,)[_S\, W_1 - v_i - W_2\,]$$

A proposal along these lines, taken together with the Derived Verb Phrase Rule (modified to allow double λ-extraction in some cases) and whatever principles guarantee that certain quantifiers must have certain scopes, will generate most of the formulas discussed in Chapter Two. There is clearly much more to be said about this than I am able to make precise here.

It is also of interest to ponder the general relation of sentence grammar and deletion rules. Are there two kinds of deletion rules, as Williams (1976) suggests, those that are part of sentence grammar and those that belong to "discourse grammar" (however we are to construe that term)?

Given the relation we have established between the recoverability of deletion and logical form, it may well be the case that no such distinction is necessary. Williams argues that Gapping, for instance, is a rule of Sentence Grammar and that VPD is not. The possibility of (3) however, casts doubt on this claim about Gapping (see also the remarks in Section 3.1 of Chapter Three).

(3) Peter says one thing. Betsy ∅ another.

Of the deletion rules we have been discussing (I mean to exclude from this EQUI and the like, which may simply be pronoun-deletion rules), the deletion phenomenon which is the best candidate for a purely intra-sentential process is the ellipsis that goes on in the comparative clause:

(4)(a) I think you draw much more attention to yourself with them on than ∅ off. [∅ = you draw Δ much attention to yourself with them] (Dennis Foley to Mary Hartmann)

(b) I gave more people books on Tuesday than you ∅ on Wednesday. [∅ = gave Δ many people books]

Notice, however, that the deleted string in all such examples corresponds to a logical entity containing a variable which must be bound by an operator (the comparative operator) in the same sentence. Such variables can never be bound by an operator elsewhere in the discourse.

Given this, however we write this deletion rule, the theory of recoverability of deletion developed in Chapter Two will have as a consequence the fact that ellipsis in the comparative clause functions only intra-sententially. The structural description of the rule(s) need only mention <u>than</u> for this to be the case.

But if this is true, then perhaps all deletion rules are in some sense part of "discourse grammar", ellipsis in the comparative clause, say, being the limiting case: a rule of discourse grammar which, because of its nature and because of the nature of the recoverability of deletion, applies only within one-sentence discourses.

There is much that could be said about how to constrain the class of discourse grammars. The nature of the rules examined here suggests that the class of possible structural descriptions of deletion rules, for instance, may be very restricted. Moreover, if we restrict sentence grammar so as to disallow deletion rules of the sort we have examined here (allowing, say, only deletion of specified elements or pronominal entities), then we have drastically reduced the class of possible (sentence) grammars, and perhaps also solved the well-known problems which (as Peters and Ritchie have observed) deletion rules pose for the theory of transformational grammar. We'll see.

Bibliography

Akmajian, A. (1970) *Aspects of the Grammar of Focus in English*, unpublished doctoral dissertation, MIT.

________ (1973) "The Role of Focus in the Interpretation of Anaphoric Expressions", in Anderson and Kiparsky, eds. (1973), p. 215-226.

Akmajian, A.,and F. Heny (1975) *An Introduction to the Principles of Transformational Syntax*, Cambridge, Mass.: MIT Press.

Akmajian, A., and T. Wasow (1975) "The Constituent Structure of VP and AUX and the Position of the Verb Be", *Linguistic Analysis*, 1.3.

Anderson, S., and P. Kiparsky, eds. (1973) *A Festschrift for Morris Halle*, Holt, Rinehart, and Winston, New York.

Andrews, A. (1975) *Studies in the Syntax of Relative and Comparative Clauses*, unpublished MIT doctoral dissertation.

Åqvist, L. (1965) *A New Approach to the Logical Theory of Interrogatives Part I*, Uppsala.

Aronoff, M. (1974) *Word-Structure*, unpublished doctoral dissertation, MIT.

Bach, E. (1965) "On Some Recurrent Types of Transformations" in C.W. Kreidler, ed., *Sixteenth Annual Round Table Meeting On Linguistics and Language Studies*, Georgetown University Monograph series on Languages and Linguistics 18.

________ (1971) "Questions", *Linguistic Inquiry*, 2: 153-166.

Bach, E., J. Bresnan, and T. Wasow (1974) "'Sloppy Identity': an Unnecessary and Insufficient Criterion for Deletion Rules", *Linguistic Inquiry*, 5.4: 609-614.

Baker, C.L. (1968) *Indirect Questions in English*, unpublished doctoral dissertation, University of Illinois.

________ (1971) "Stress Level and Auxiliary Behavior in English", *Linguistic Inquiry*, 2.2: 167-182.

Baltin, M. (forthcoming) "On the Proper Formulation of Q-Float", unpublished paper, MIT.

Bartsch, R., and T. Vennemann (1972) *Semantic Structures. A Study in the Relation Between Semantics and Syntax*, Frankfurt a.M.: Athenäum.

Belnap, N. (1963) *An Analysis of Questions: Preliminary Report*, System Development Corp., Santa Monica, Cal.

Bennett, M. (1974) *Some Extensions of a Montague Fragment of English*, unpublished doctoral dissertation, U.C.L.A.

Bierwisch, M., and K. Heidolph, eds. (1970) *Progress in Linguistics*, Mouton: The Hague.

Binnick, R.I., A. Davidson, G. Green, and J. Morgan, eds. (1969), *Papers From the Fifth Regional Meeting of The Chicago Linguistic Society*, University of Chicago, Chicago, Ill.

Bolinger, D. (1958) "A Theory of Pitch Accent in English" *Word*, 14: 109-140.

________ (1965) *Forms of English: Accent, Morpheme, Order* (edited by I. Abe and T. Kanekiyo), Harvard University Press, Cambridge, Mass.

________ (1971a) "The Nominal in the Progressive" *Linguistic Inquiry* 2.2: 246-250.

________ (1971b) "A Further Note on the Nominal in the Progressive" *Linguistic Inquiry*, 2.4: 584-586.

Bouton, L. (1970) "Antecedent-contained Pro-forms" in *Papers From the Sixth Regional Meeting of the Chicago Linguistic Society*, University of Chicago, Chicago, Ill.

Bowers, J. (1969) "Some Adjectival Nominalizations in English", unpublished paper, MIT.

Bresnan, J. (1970) "On Complementizers: Toward a Syntactic Theory of Complement Types", *Foundations of Language* 6: 297-321.

________ (1971) "A Note on the Notion 'Identity of Sense Anaphora'", *Linguistic Inquiry*, 2.4: 589-597.

________ (1972) *Theory of Complementation in English Syntax*, unpublished doctoral dissertation, MIT.

________ (1973) "Syntax of the Comparative Clause Construction in English" *Linguistic Inquiry*, 4: 275-343.

________ (1975) "Comparative Deletion and Constraints on Transformations" *Linguistic Analysis*, 1.1: 25-74.

________ (1976a) "On the Form and Functioning of Transformations" *Linguistic Inquiry*, 7.1.

________ (1976b) "Toward a Realistic Model of Transformational Grammar", paper presented at MIT- AT&T Convocation on Communications, MIT.

Burt, M.K. (1971) *From Deep to Surface Structure*, Harper and Row, New York.

Carden, G. (1970) "A Note on Conflicting Idiolects", *Linguistic Inquiry*, 1.3: 281-290.

Carrier, J. (1974) "Stress and Anaphora", unpublished paper, MIT.

Catlin, J. and Catlin, J.-C. (1972) "Intentionality: A Source of Ambiguity in English?" *Linguistic Inquiry*, 3.4: 504-508.

Channon, R. (1974) "The NAC and How Not to Get It" in Kaisse and Hankamer, eds. (1974), pp. 276-298.

________ (1975) "Acceptable Ambiguity" in Cogen et al., eds. (1975), pp. 37-46.

Chomsky, N. (1955) *The Logical Structure of Linguistic Theory*, MIT microfilm [published in part by Plenum, New York - 1975].

________ (1957) *Syntactic Structures*, Mouton: The Hague.

________ (1964) "Current Issues in Linguistic Theory" in Fodor and Katz, eds. (1964).

________ (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Mass.

________ (1968) *Language and Mind* (enlarged edition), Harcourt Brace Jovanovich.

________ (1970) "Remarks on Nominalization", in Jacobs and Rosenbaum (1970) [also reprinted in Chomsky (1972c)].

________ (1972a) "Deep Structure, Surface Structure, and Semantic Interpretation", in Chomsky (1972c).

________ (1972b) "Some Empirical Issues in the Theory of Transformational Grammar", in Chomsky (1972c) [also in Peters, ed. (1972)].

________ (1972c) *Studies on Semantics in Generative Grammar*, Mouton: The Hague.

________ (1973) "Conditions on Transformations" in Anderson, S. and P. Kiparsky, eds. (1973).

________ (1975a) *Reflections on Language*, Pantheon, New York.

________ (1975b) "Conditions on Rules of Grammar", unpublished paper, MIT.

________ (1975c) "Questions of Form and Interpretation" *Linguistic Analysis*, 1.1: 75-109.

Church, A. (1941) *The Calculi of Lambda-Conversion*, Princeton University Press, Princeton, N.J. [reprinted by Kraus Reprint Corp., N.Y. - 1965].

Cogen, C., H. Thompson, G. Thurgood, K. Whistler, and J. Wright, eds. (1975) *Proceedings of the First Annual Meeting of the Berkeley Linguistic Society*, University of California, Berkeley.

Corum, C., T.C. Smith-Stark, and A. Weiser, eds. (1973) *Papers From the Ninth Regional Meeting of the Chicago Linguistic Society*, Chicago, Ill.

Corum, C. (1973) "Anaphoric Peninsulas", in Corum et al., eds. (1973), pp. 89-97.

Curry, H.B., J.P. Seldin, and J.R. Hindley (1971) *Combinatory Logic Vol. 2*, North Holland Company.

Curry, H.B., R. Seys, and W. Craig (1958) *Combinatory Logic Vol. 1*, North Holland Company.

Dahl, O. (1972) "On So-Called 'Sloppy Identity'", *Gothenburg Papers in Theoretical Linguistics* 11, University of Göteborg.

________ (1974) "How to Open a Sentence: Abstraction in Natural Language", *Logical Grammar Reports* 12, University of Göteborg.

Dougherty, R. (1969) "An Interpretive Theory of Pronominal Reference", *Foundations of Language*, 5.4: 488-519.

________ (1970) "A Grammar of Coordinate Conjoined Structures: I" *Language*, 46.4: 850-898.

________ (1971) "A Grammar of Coordinate Conjoined Structures: II" *Language*, 47.2: 298-339.

Dretske, F.I. (1972) "Contrastive Statements", *Philosophical Review*, 81: 411-437.

Dummett, M. (1973) *Frege: Philosophy of Language*, Duckworth.

Edes, E. (1968) "Output Conditions on Anaphoric Expressions with Split Antecedents", unpublished paper, Harvard University.

Emonds, J. (1970) *Root and Structure Preserving Transformations*, unpublished MIT doctoral dissertation [available from Indiana University Linguistics Club]

Fedorowicz-Bacz, B. (1973) "Gapping in Polish and the No-Ambiguity Condition" in University of Michigan Papers in Linguistics, 1.2: 43-53.

Fiengo, R. (1974) *Semantic Conditions On Surface Structure*, unpublished doctoral dissertation, MIT.

Fiengo, R., and H. Lasnik (1976) "Some Issues in the Theory of Transformations", *Linguistic Inquiry*, 7.1.

Firbas, J. (1964) "On Defining the Theme in Functional Sentence Analysis", *Travaux Linguistiques de Prague*, 1: 267-280.

Fodor, J.A. (1975) *The Language of Thought*, Crowell, New York.

Fodor, J.A. and J. Katz, eds. (1964) *The Structure of Language - Readings in the Philosophy of Language,* Prentice Hall, Englewood Cliffs, N.J.

Garvin, P. (1964) *A Prague School Reader on Esthetics, Literary Structure, and Style,* Georgetown University Press, Washington, D.C.

Grice, H.P. (1975) "Logic and Conversation" in Cole, P., and J. Morgan, eds. (1975), *Syntax and Semantics Vol. 3: Speech Acts,* Academic Press, New York [extracted from Grice's 1967 William James Lectures, Harvard University].

Grinder, J., and P. Postal (1971) "Missing Antecedents", *Linguistic Inquiry,* 2.3.

Grosu, A. (1973) "On the Non-unitary Nature of the Coordinate Structure Constraint", *Linguistic Inquiry,* 4.1: 82-92.

________ (1975) "A Plea for Greater Caution in Proposing Functional Explanations in Linguistics", in Grossman, R., L.J. San, and T. Vance, eds., *Papers From the Parasession on Functionalism,* Chicago Linguistic Society, University of Chicago, Chicago, Ill.

________ (to appear) "On the Status of Positionally-defined Constraints in Syntax", to appear in *Theoretical Linguistics.*

Halitsky, D. (1975) "Left Branch S's and NP's in English: A Bar Notation Analysis", *Linguistic Analysis,* 1.3: 279-296.

Halle, M. (1973) "Stress Rules in English: A New Version", *Linguistic Inquiry,* 7.4: 451-464.

Halliday, M.A.K. (1967) "Notes on Transitivity and Theme Part 2", *Journal of Linguistics* 3: 199-244.

Halliday, M.A.K., and R. Hasan (1973) *Cohesion in Spoken and Written English,* London: Longmans.

Hankamer, J. (1971) *Constraints on Deletion in Syntax,* unpublished doctoral dissertation, Yale University.

________ (1972a) "On the Nonexistence of Mirror Image Rules in Syntax", in Kimball, J., ed. (1972), *Syntax and Semantics Vol 1,* Seminar Press.

________ (1972b) "Pied Wiping", paper presented at meeting of the Northeastern Linguistic Society.

________ (1973a) "Unacceptable Ambiguity", *Linguistic Inquiry,* 4: 17-68.

________ (1973b) "Why There are Two Thans in English", in Corum et al., eds. (1973).

________ (ms.) "The Semantic Interpretation of Anaphoric Expressions", unpublished manuscript.

Hankamer, J., and I.A. Sag (1975) "Result Clauses in English: So-Flip, or So What?" paper presented to the Fourth Conference on New Ways of Analyzing Variation, Etc., Georgetown University, Washington, D.C.

Harris, Z.S. (1946) "From Morpheme to Utterance", *Language,* 22: 161-183 [also reprinted in Joos, ed. (1957)].

________ (1951) *Methods in Structural Linguistics,* Chicago: University of Chicago Press.

________ (1957) "Co-occurrence and Transformation in Linguistic Structure", *Language,* 33: 283-340 [also reprinted in Fodor and Katz, eds. (1964)].

________ (1964) "The Elementary Transformations", Transformations and Discourse Analysis Paper No. 54, University of Pennsylvania, Philadelphia.

________ (1965) "Transformational Theory", *Language,* 41: 363-401.

________ (1968) *Mathematical Structures of Language,* Interscience Publishers, New York.

Higgens, F. (1974) *The Pseudo-Cleft Construction in English,* unpublished doctoral dissertation, MIT.

Hindle, D., and I.A. Sag (1975) "Some More on Anymore", in Fasold, R. and R. Shuy, eds. (1975), *Analyzing Variation in Language,* Georgetown University Press, Washington, D.C.

Hindly, J.R., B. Lercher, and J.P. Seldin (1972) *Introduction to Combinatory Logic,* London Mathematics Society Lecture Note Series 7, Cambridge University Press.

Hiž, H. (1962) "Questions and Answers", *Journal of Philosophy,* 59.

________ (1969) "Referentials", *Semiotica* 1: 136-166.

Horn, L. (1972) *On the Semantic Properties of Logical Operators in English,* unpublished doctoral dissertation, U.C.L.A.

Hornstein, N. (1975) "S andthe $\overline{X}$ Convention", *Montreal Working Papers in Linguistics,* 4: 35-71.

Hughes, G., and M. Cresswell (1968) *An Introduction to Modal Logic,* Methuen, London.

Hull, R. (1972) *Towards a Logical Analysis of Questions and Answers,* unpublished dissertation for Diploma of Linguistics, Cambridge University.

Ioup, G. (1976) "Some Universals for Quantifier Scope" in Kimball, J., ed. (1976), *Syntax and Semantics Volume 4,* Academic Press, New York.

Jackendoff, R.S. (1968) "Quantifiers in English", *Foundations of Language*, 4.4: 422-442.

________ (1971) "Gapping and Related Rules", *Linguistic Inquiry*, 2.1: 21-36.

________ (1972) *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, Mass.

________ (1974) "Introduction to the $\bar{X}$ Convention", Indiana University Linguistics Club.

________ (1975) "Tough and the Trace Theory of Movement Rules", *Linguistic Inquiry*, 6.3: 437-446.

________ (1976) *Lexicalist Syntax: A Study of Phrase Structure*, unpublished manuscript.

Jacobs, T., and P. Rosenbaum, eds. (1970) *Readings in Transformational Grammar*, Ginn and Co., Waltham, Mass.

Jacobson, P., and P. Neubauer (forthcoming) "Rule Cyclicity: Evidence from the Intervention Constraint", to appear in *Linguistic Inquiry*, 7.

Joos, M., ed. (1957) *Readings in Linguistics I*, University of Chicago Press, Chicago, Ill.

Kaisse, E., and J. Hankamer, eds. (1974) *Papers From the Fifth Annual Meeting of the Northeastern Linguistic Society*, Harvard University, Cambridge, Mass.

Kalish, D., and R. Montague (1964) *Logic: Techniques of Formal Reasoning*, New York: Harcourt, Brace and World.

Kamp, H. (1973) "Free Choice Permission", *Proceedings of the Aristotelian Society*.

Karttunen, L. (1974) "Until", in La Galy et al., eds. (1974), pp. 284-297.

Katz, J. (1972) *Semantic Theory*, Harper and Row, New York.

Katz, J., and P.M. Postal (1964) *An Integrated Theory of Linguistic Descriptions*, MIT Press, Cambridge, Mass.

Kayne, R. (1975) *French Syntax: The Transformational Cycle*, MIT Press, Cambridge, Mass.

Kean, M.-L. (1974) "The Strict Cycle In Phonology", *Linguistic Inquiry*, 5.2: 179-204.

Keenan, E. (1970) "A Logical Base for a Transformational Grammar of English", *Transformations and Discourse Analysis Papers*, 85, University of Pennsylvania, Philadelphia.

_________ (1971a) "Names, Quantifiers, and Sloppy Identity Problem", *Papers in Linguistics,* 4.2.

_________ (1971b) "Two Kinds of Presupposition in Natural Language", in C. Fillmore and D.T. Langendoen, eds. (1971), *Studies in Linguistic Semantics,* Holt, Rinehart, and Winston, New York.

_________ (1972) "On Semantically Based Grammar", *Linguistic Inquiry,* 3.4: 413-461.

Keenan, E., and R. Hull (1972) "The Logical Presuppositions of Questions and Answers" in Petöfi, J., and D. Franck, eds. (1973), *Präsuppositionen in Philosophie und Linguistik,* Athenäum, Frankfurt.

Keyser, S.J. (1968) Review of Sven Jacobson, *Adverbial Positions in English, Language,* 44: 357-374.

Kleene, S. (1952) *Mathematical Logic,* John Wiley and Sons, New York.

Klima, E.S. (1964) "Negation in English", in Fodor and Katz, eds. (1964), pp. 246-323.

Koutsoudas, A. (1971) "Gapping, Conjunction Reduction, and Coordinate Deletion", *Foundations of Language,* 7: 337-386.

Kroch, A.S. (1974) *The Semantics of Scope in English,* unpublished doctoral dissertation, MIT [available from Indiana University Linguistics Club].

Kuno, S. (1972) "Functional Sentence Perspective: A Case Study from Japanese and English", *Linguistic Inquiry,* 3.3: 269-320.

_________ (1974) "Lexical and Contextual Meaning", *Linguistic Inquiry,* 5.3.

_________ (1975a) "Conditions for Verb Phrase Deletion", *Foundations of Language,* 13: 161-175.

_________ (1975b) "Three Perspectives in the Functional Approach to Syntax" in Grossman, R., L. San, and T. Vance, eds., *Papers From the Parasession on Functionalism,* Chicago, Ill.

_________ (1976) "Gapping: A Functional Analysis", *Linguistic Inquiry,* 7.2.

Labov, W. (1972) "For an End to the Uncontrolled Use of Intuition in Linguistic Analysis", paper presented to winter meeting of the Linguistic Society of America, Atlanta, Georgia.

La Galy, M., R. Fox, and A. Bruck, eds. (1974) *Papers From the Tenth Regional Meeting of the Chicago Linguistic Society,* Chicago, Ill.

Lakoff, G. (1965) *On the Nature of Syntactic Irregularity,* Report NSF-16, Harvard University Computation Laboratory [reprinted as *Irregularity In Syntax,* Holt, Rinehart and Winston, New York (1970)].

________ (1968) "Deep and Surface Grammar" [available from Indiana University Linguistics Club].

________ (1969) "On Derivational Constraints" in Binnick et al., eds. (1969), pp. 117-139.

________ (1970) "A Note on Vagueness and Ambiguity", *Linguistic Inquiry,* 1.

________ (1971) "On Generative Semantics" in Jakobovits, L., and D. Steinberg, eds., *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology,* Cambridge University Press.

Lakoff, G., and J.R. Ross (1966) "A Criterion for Verb Phrase Constituency" in *Computation Laboratory of Harvard University Report No. NSF-17 to the National Science Foundation.*

________ (1972) "A Note on Anaphoric Islands and Causatives", *Linguistic Inquiry,* 3.1.

Lakoff, R. (1969) "Some Reasons Why There Can't Be Any Some-Any Rule", *Language,* 45: 608-615.

Langacker, R. (1969) "On Pronominalization and the Chain of Command", in Reibel and Schane, eds. (1969), pp. 160-186.

Langendoen, D.T. (1974) "Acceptable Conclusions from Unacceptable Ambiguity" paper presented at the conference "Testing Linguistic Hypotheses", University of Wisconsin, Milwaukee - May, 1974.

Lasnik, H. (1972) *Analyses of Negation in English,* unpublished doctoral dissertation, MIT.

________ (1976) "Remarks on Coreference", to appear in *Linguistic Analysis.*

Lasnik, H., and R. Fiengo (1974) "Complement Object Deletion", *Linguistic Inquiry,* 5.4: 535-572.

Lees, R. (1960) *The Grammar of English Nominalizations,* MIT doctoral dissertation [published by Indiana University - Bloomington, Mouton: The Hague, 1968].

Lees, R., and E. Klima (1963) "Rules for English Pronominalization", *Language,* 39: 17-28 [also reprinted in Reibel and Schane, eds. (1969)].

Liberman, M. (1974) "On Conditioning the Rule of Subject-Auxiliary Inversion", in Kaisse and Hankamer, eds. (1974).

Liddell, S. (1975) "What About the Fact That 'On Certain Ambiguities' Says What It Says?", *Linguistic Inquiry,* 6.4: 568-578.

Lust, B. (ms.) "On Gapping", unpublished manuscript, MIT.

Lyons, J. (1968) *An Introduction to Theoretical Linguistics*, Cambridge University Press.

Maling, J. (1972) "On 'Gapping and the Order of Constituents'", *Linguistic Inquiry*, 3.1.

McCawley, J. (1967) 'Meaning and the Description of Languages', *Kotoba no uchu* 2 [also in McCawley, J. (1973), *Grammar and Meaning*, Tokyo: Taishukan Publishing Company].

________ (1970) "English as a VSO Language", *Language*, 46.2.

________ (1971) "Where do Noun Phrases Come From?" in Steinberg, D., and L. Jakobovits, eds., *Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, Cambridge University Press [also in Jacobs and Rosenbaum, eds. (1970)].

________ (1973) "Syntactic and Logical Arguments for Semantic Structures" in O. Fujimura, ed., *Three Dimensions of Linguistic Theory*, Tokyo: The TEC Company.

________ (1976) "How to Get an Interpretive Theory of Anaphora to Work", to appear in *Linguistic Inquiry*, 7.

Montague, R. (1973) "The Proper Treatment of Quantification in Ordinary English", in Montague (1975), pp. 247-270.

________ (1975) *Formal Philosophy* (R.Thomason, ed.), Yale University Press, New Haven, Conn. [see especially Montague (1973)].

Moravcsik, J., and D. Gabbay (forthcoming), untitled paper on the interpretation of comparative sentences.

Morgan, J. (1972) "Some Aspects of Relative Clauses in English and Albanian", in *The Chicago Which Hunt*, The Chicago Linguistics Society.

________ (1973) "Sentence Fragments and the Notion 'Sentence'", in Kachru, B., R. Lees, Y. Malkiel, A. Pietrangeli, and S. Saporta, eds. (1973), *Issues in Linguistics: Papers in Honor of Henry and Renee Kahane*, University of Illinois Press, Urbana.

Neeld, R. (1973) "On the Variable Strength of Island Constraints" in Corum et al. (1973), pp. 450-461.

Neubauer, P. (1970) "On the Notion of Chopping Rule" in *Papers From the Sixth Regional Meeting of the Chicago Linguistic Society*.

Newmeyer, F. (1975) "The Position of Incorporation Transformations in the Grammar" in Cogen et al., eds. (1975).

Nunberg, G. (forthcoming) *Pragmatics and Semantic Interpretation,* doctoral dissertation, C.U.N.Y.

Nunberg, G., and C. Pan (1975) "Inferring Quantification In Generic Sentences", in Grossman, T., L.J. San, and T. Vance, eds., *Papers From the Eleventh Regional Meeting of the Chicago Linguistic Society,* Chicago, Ill.

Parret, H., ed. (1974) *Discussing Language,* Mouton: The Hague.

Partee, B.H. (1972) "Some Transformational Extensions of Montague Grammar", in Rodman, R., ed. (1972), *Papers in Montague Grammar,* Occasional Papers in Linguistics No. 2, UCLA [also appears in *Journal of Philosophical Logic* 2 (1973), pp. 509-534].

_________ (1975) "Montague Grammar and Transformational Grammar", *Linguistic Inquiry,* 6.2: 203-300.

Perlmutter, D. (1970) "On the Article in English", in Bierwisch, M., and K. Heidolph, eds. (1970), pp. 233-248.

_________ (1971) *Deep and Surface Structure Constraints in Syntax,* Holt, Rinehart, and Winston, New York.

Perlmutter, D. and P.M. Postal (forthcoming) *Relational Grammar.*

Peters, P.S., ed. (1972) *Goals of Linguistic Theory,* Prentice Hall, Englewood Cliffs, N.J.

Peters, P.S. and R. Ritchie (1973) "On the Generative Power of Transformational Grammars", *Information Sciences*, 6: 49-83.

Pike, K. (1945) *The Intonation of American English,* Ann Arbor: The University of Michigan Press.

Pope, E. (1972) *Questions and Answers in English,* unpublished doctoral dissertation, MIT.

Postal, P.M. (1969) "Anaphoric Islands", in Binnick et al., eds. (1969).

_________ (1971) *Crossover Phenomena,* Holt, Rinehart and Winston, New York.

_________ (1972) "The Best Theory", in Peters, ed. (1972).

_________ (1974a) *On Raising,* MIT Press, Cambridge, Mass.

_________ (1974b) "On Certain Ambiguities", *Linguistic Inquiry,* 5.3: 367-424.

_________ (1976) "On Avoiding Reference to Subject", *Linguistic Inquiry,* 7.1.

Quine, W.V.O. (1960) *Word and Object,* MIT Press, Cambridge, Mass.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik (1972) *A Grammar of Contemporary English,* Seminar Press, London.

Reibel, D., and S. Schane, eds. (1969) *Modern Studies in English - Readings in Transformational Grammar,* Prentice Hall, Englewood Cliffs, N.J.

Reinhart, T. (1974) "Syntax and Coreference" in Kaisse and Hankamer, eds, (1974).

________ (forthcoming) doctoral dissertation, MIT.

Rodman, R. (1972) "On the Stranding of Prepositions in English", in G. Bedell, ed., *Explorations in Syntactic Theory* [= *UCLA Papers in Syntax,* 2].

Ross, J.R. (1967a) *Constraints on Variables in Syntax,* unpublished MIT doctoral dissertation [available from Indiana University Linguistics Club].

________ (1967b) "Gapping and the Order of Constituents", Indiana University Linguistics Club [also reprinted in Bierwisch and Heidolph, eds. (1970)].

________ (1967c) "Auxiliaries as Main Verbs", unpublished paper, MIT [in Todd, W., ed. (1969), *Studies in Philosophical Linguistics - Series One,* Evanston, Ill., pp. 77-102].

________ (1969a) "On the Cyclic Nature of English Pronominalization", in *To Honor Roman Jakobson,* Mouton: The Hague [also reprinted in Reibel and Schane, eds. (1969)].

________ (1969b) "Guess Who" in Binnick, R.I. et al. (1969).

________ (1971) "The Superficial Nature of the Anaphoric Island Constraint", *Linguistic Inquiry,* 2.4.

________ (1972) "Doubl-ing", *Linguistic Inquiry,* 3.1: 61-86.

________ (1974) "Three Batons for Cognitive Psychologists", in Weimer, W., and D. Palermo, eds., *Cognition and the Symbolic Processes,* Lawrence Erlbaum Associates.

________ (1976) "Clausematiness", in Keenan, E., ed. (1976), *Formal Semantics of Natural Language,* Cambridge University Press, Cambridge.

Russell, B. (1905) "On Denoting", *Mind,* 14 [reprinted in Russell's *Logic and Knowledge,* London: Allen & Unwin - 1958].

Sadock, J., and A. Zwicky (1973) "Ambiguity Tests and How to Fail Them", *Ohio State Working Papers in Linguistics,* No. 16.

Sag, I.A., and J. Hankamer (1976) "Deep and Surface Anaphora", *Linguistic Inquiry,* 7.3.

Savin, H. (1971) "Every Any Means Every", in *Proceedings of the Conference on Current Problems in Psycholoinguistics*, edited by J. Mehler and A. Bresson, Paris: Centre Nationale de la Recherche Scientifique.

Schiebe, T. (1973) "Zum Problem der grammatisch relevanten Identität", in Kiefer, F., and N. Ruwet, eds., *Generative Grammar in Europe*, Reidel.

Selkirk, E. (1970) "On the Determiner Systems of Noun Phrase and Adjective Phrase", unpublished paper, MIT.

Shopen, T. (1972) *A Generative Theory of Ellipsis*, unpublished doctoral dissertation, U.C.L.A. [available from Indiana University Linguistics Club].

Siegel, D. (1974) *Topics in English Morphology*, unpublished doctoral dissertation, MIT.

Sjoblom, T. (1975a) *Derived Conjunction with And, Not, But and Too*, unpublished B. Sc. thesis, MIT.

________ (1975b) "Some Singular Answers to Objections Against Derived Conjunction", unpublished paper, MIT.

Soames, S. (1976) *A Critical Examination of Frege's Theory of Presupposition and Contemporary Alternatives*, unpublished doctoral dissertation, MIT.

Stillings, J. (1975) "The Formulation of Gapping in English as Evidence for Variable Types", *Linguistic Analysis*, 1.3: 247-274.

Stockwell, R., P. Schachter, and B. Partee (1972) *The Major Syntactic Structures of English*, Holt, Rinehart, and Winston, New York.

Tai, J. (1969) *Coordination Reduction*, unpublished doctoral dissertation, Indiana University.

Thomason, R. (1974) "Some Complement Constructions in Montague Grammar", in M. La Galy et al., eds. (1974).

Thrasher, R. (1974) *Shouldn't Ignore These Strings: A Study of Conversational Deletion*, unpublished doctoral dissertation, University of Michigan.

Vachek, J., ed. (1964) *A Prague School Reader in Linguistics*, Indiana University Press, Bloomington.

van Fraassen, B.C. (1971) *Formal Semantics and Logic*, The Macmillan Company, New York.

Vendler, Z. (1967) "Each and Every, Any and All", in Vendler, *Linguistics in Philosophy*, Cornell University Press, Ithaca, N.Y.

Vergnaud, J.-R., M. Halle, and A. Prince (forthcoming) "On the Formalism of Phonological Rules", unpublished paper.

Wasow, T. (1972) *Anaphoric Relations in English*, unpublished doctoral dissertation, MIT.

Wells, R. (1947) "Immediate Constituents", *Language*, 23: 81-117 [also reprinted in Joos, ed. (1957)].

Williams, E. (1974) *Rule Ordering In Syntax*, unpublished MIT doctoral dissertation.

________ (1975) "Free Deletion", unpublished paper, University of Massachusetts.

________ (1976) "Discourse Grammar", unpublished paper, University of Massachusetts.

Williams, G. (1971) *Networks of Anaphora: An Essay in the Syntax of Pronominalization*, unpublished doctoral dissertation, MIT.

Wilson, J.C. (1926) *Statement and Inference*, Vol. 1, Clarendon Press, Oxford.

Witten, E. (1972) "Pronominalization and Sloppy Identity", in *Aiken Computation Laboratory Report No. NSF-28 to the National Science Foundation*, Harvard University, Cambridge, Mass.

Woisetschlaeger, E. (1975) "Conditions On Mixed Terms", unpublished paper, MIT.

________ (1976) "Degree Phrase Shift", unpublished paper, MIT.

Zaenen, A., and J. Pinkham (forthcoming) "The Discovery of Another Island", to appear in *Linguistic Inquiry*, 7.4.

Biographical Note

The author was born on November 9, 1949 in a dreary Ohio town called Alliance, whose sole claim to fame (aside from being the world's largest manufacturer of automatic electric garage door openers and TV antenna rotators) is a mention in Ripley's Believe It or Not which remarks that it is the only town in the world whose main street dead-ends at both ends. He attended the Mercersburg Academy in Mercersburg, Pa., only to be expelled in the last half of his senior year, amidst a flurry of allegations surrounding a field trip whereupon more than a score of adolescents somehow wound up in a local tavern instead of attending what in fact turned out to be a rather mediocre rendition of Molière's *Les Femmes Savantes*. He nevertheless attended college (at the University of Rochester), wherefrom he dropped out at the beginning of the fall term of his senior year to pursue a dubious career as an organist in The Red, White and Blues Band, a cacophonous but inspired organization which he also managed. Two months later, seeing the folly of his ways, he reenrolled at the U.of R. where he somehow managed to earn his B.A. on schedule, but only after a six-month fling as a not-so-always-shrewd rock concert entrepreneur. A financially disastrous experience in the music business, in fact, is what motivated the author to return to academia. He earned an M.A. from the University of Pennsylvania, and has spent his last three happy years as a graduate student at MIT. His publications include:

"On the State of Progress On Progressives and Statives". In Bailey, C.-J.N., and R. Shuy, eds., *New Ways of Analyzing Variation in English*. Georgetown University Press, Washington, D.C. (1973)

"Variable Analysis of Variable Data" (with Peg Griffin and Gregory Guy). *University of Michigan Working Papers in Linguistics*, Vol. 1 No. 2, pp. 54-68. Ann Arbor, Michigan. (1973)

"Prosodic Form and Discourse Function" (with Mark Liberman). In La Galy et al., eds., *Papers From The Tenth Regional Meeting of The Chicago Linguistic Society*, Chicago, Illinois. (1974)

"The Grassmann's Law Ordering Pseudoparadox". *Linguistic Inquiry* Vol. V No. 4, pp 591-607. (1974)

"Some More On Anymore" (with Donald Hindle). In Fasold, R.W., and R. Shuy, eds., Analyzing Variation In Language, pp. 89-110. Georgetown University Press. (1975)

"Syntactically vs. Pragmatically Controlled Anaphora" (with Jorge Hankamer). In Fasold, R.W. and R. Shuy, eds., Proceedings of The Third Conference on New Ways of Analyzing Variation, Etc. Georgetown University Press. (to appear)

"The Intonational Disambiguation of Indirect Speech Acts" (with Mark Liberman). In Grossman, R., et al., eds., Papers From The Eleventh Regional Meeting of the Chicago Linguistic Society, Chicago, Illinois. (1975)

"Deep and Surface Anaphora" (with Jorge Hankamer). Linguistic Inquiry Vol. VII No. 3. (to appear)

"Pseudosolutions to the Pseudoparadox: Sanskrit Diaspirates Revisited". Linguistic Inquiry Vol. VII No. 4. (to appear)

"A Logical Theory of Verb Phrase Deletion". In Mufwene, S., C. Walker, and S. Steever, eds., Papers From The Twelfth Regional Meeting of The Chicago Linguistic Society, Chicago, Ill. (to appear)

"A Note on Verb Phrase Deletion". Linguistic Inquiry Vol. VII No. 4. (to appear)