



National Library of Canada  
Collections Development Branch

Canadian Theses on  
Microfiche Service

Bibliothèque nationale du Canada  
Direction du développement des collections

Service des thèses canadiennes  
sur microfiche

## NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION  
HAS BEEN MICROFILMED  
EXACTLY AS RECEIVED

## AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formulés d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ  
MICROFILMÉE TELLE QUE  
NOUS L'AVONS REÇUE

DISCOURSE AND SENTENCE GRAMMAR

by

Catherine Beauchamp

A thesis submitted to the  
School of Graduate Studies  
and Research in partial fulfillment  
of the requirements for the  
degree of Master of Arts

Department of Linguistics  
University of Ottawa

July 1981

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
CHAPTER I. SENTENCE AND DISCOURSE GRAMMAR	2
I.1. Introduction	2
I.2. Chomsky and Sentence Grammar	3
I.3. Williams and Discourse Grammar	6
I.4. Other Approaches	13
I.5. Summary	20
Footnotes to Chapter I.	21
CHAPTER II. THE CORE AND THE PERIPHERY	22
II.1. Introduction	22
II.2. Approaches to Core and Periphery	22
II.3. Chomsky and Rules of Core Grammar	26
II.4. Koster and the Distinction between Core and Periphery	37
II.5. Summary	50
Footnotes to Chapter II.	51
CHAPTER III. GAPPING	52
III.1. Introduction	52
III.2. Gapping Characteristics	53
III.3. Historical Overview	57
III.4. A Proposal	97
III.5. Summary	103
Footnotes to Chapter III.	104
CONCLUSION	105
BIBLIOGRAPHY	106

## I N T R O D U C T I O N

Recent developments within the extended standard theory have revolved around distinctions between sentence and discourse grammar, and further, between the core and peripheral areas of sentence grammar. This thesis outlines these developments and examines one rule of grammar, Gapping, in the light of the above distinctions. Chapter I sets forth the distinctions between sentence and discourse grammar as reflected in Chomsky (1977) and Williams (1977). Other approaches to these two areas are mentioned. Chapter II examines the distinctions between core and peripheral grammar as seen in Chomsky (1977) and Koster (1978). Chapter III represents an attempt to place the rule of Gapping within the core/peripheral distinctions outlined in this thesis.

## CHAPTER I

### SENTENCE AND DISCOURSE GRAMMAR

#### I.1. Introduction

Within recent years in linguistics a distinction has been made between sentence and discourse grammar. While a definite boundary between the two is difficult to establish, it is possible to distinguish certain characteristics of each in relation to the framework of Chomsky's revised extended standard theory. Chomsky and Lasnik (1977) have presented a model for sentence grammar, and Williams (1976) has attempted to describe a particular rule of discourse grammar. Chomsky's model applies strictly to structures included within the boundaries of the sentence and the rules of syntax which he outlines apply at the level of the sentence. Williams has attempted to make a distinction between rules which apply in sentences and rules which can apply across sentence boundaries within a discourse. It is this distinction which will be dealt with in this chapter, as well as certain other treatments of stylistic rules (Emonds 1976 and Rochemont 1978), which seem to fall between sentence and discourse grammar.

## I.2. Chomsky and Sentence Grammar.

Within Chomsky's theory of sentence grammar is contained the part of language he considers to be innate. According to his rationalist approach, humans are innately equipped for the process of language learning. Linked to the principle of innateness is the idea that all languages possess certain common features known as linguistic universals, which can be considered part of the innate capacity. A theory of grammar must account for universals in language. A more detailed description of the relationship between Chomsky's theory of sentence grammar and innateness and universals will be presented in Chapter II.

Chomsky's model of sentence grammar includes a base with a lexicon and a categorial component, a transformational component, two systems of interpretive rules, and a phonological and semantic component. Base phrase markers derived from the categorial component and the lexicon result in surface structures from application of rules in the transformational component. These in turn result in representations in universal phonetics and logical form (LF) through application of the rules of the phonological and semantic components. A sample derivation will be provided later on. I will concentrate on Chomsky's model as reflected in "On WH-movement" (1977). More recent developments in the model have not significantly changed the areas of interest to this thesis.

The idea of logical form is important for discourse grammar. Chomsky and Lasnik (1977:428) consider it to be a "universal system of representation" which incorporates the meaning resulting from sentence grammar without involving other factors such as speaker's intentions or situational context.

Within sentence grammar Chomsky distinguishes core grammar and peripheral grammar in an attempt to develop a theory of markedness. Chapter II will include a discussion of these two areas of sentence grammar, as well as comments on the theory of markedness in syntax.

The model of sentence grammar presented by Chomsky and Lasnik is as follows:

1. Base
  2. Transformations (movement, adjunction, substitution)
- 
- |                     |  |
|---------------------|--|
| 3a. Deletion        | 3b. Construal                          |
| 4a. Filters         | 4b. Quantifier<br>Interpretation, etc. |
| 5a. Phonology       |  |
| 6a. Stylistic Rules |  |

In this model, the rules of the base generate base structures. The transformations convert the base structures to surface structures.

Stillings (1977) gives the following example of a derivation based on Chomsky:

a. the base component generates the structure in (1):

(1)                    Δ seems [S̄ John to like Bill ]

b. the transformational component moves a noun phrase as in (2):

(2)                    MOVE NP → John seems [S̄ t to like Bill ]

The resulting sentence is John seems to like Bill.

Semantic interpretation is done on surface structures (recently called S-structures) by the rules 3b and 4b. That is to say, surface structures are mapped into logical form. Independently, surface structures are mapped into universal phonetics by the rules 3a, 4a, 5a, and 6a.

Incorporated into this model of sentence grammar is the trace theory of movement rules. From a base-generated structure as in (3):

(3)                    [S̄ [S [NP<sub>e</sub>] [VP<sub>was</sub> [AP<sub>en</sub> [VP [V<sub>hit</sub>] [NP<sub>Bill</sub>] ]]]]]]

we can apply a movement transformation to yield (4):

(4)                    [S̄ [S [NP<sub>Bill</sub>] ] [VP<sub>was</sub> [AP<sub>en</sub> [VP<sub>hit</sub> [NP<sub>e</sub>] ]]]]]]

This corresponds to the sentence Bill was hit. The NP [ Bill ] has been moved, leaving a trace [ <sub>NP<sup>e</sup> ]. A movement rule always leaves a trace which is coindexed with its antecedent.</sub>

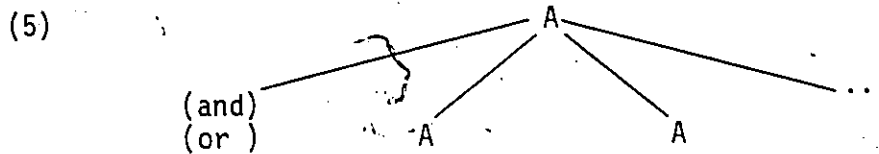
The rules of construal map surface structures to LF. They associate antecedents and anaphors (PRO) or each other by co-indexing. The manner of indexing distinguishes the trace of a movement rule from the PRO of a rule of construal. There has been much discussion about the distinction between trace and PRO, but I will not deal with this topic here.

The above constitutes a very brief outline of what Chomsky considers sentence grammar to be. The rules and conditions will be considered in more detail in Chapter II, where a distinction will be made between the two components of sentence grammar: the core and the periphery. The outline given here is meant to serve as a guide in defining the relation between sentence and discourse grammar.

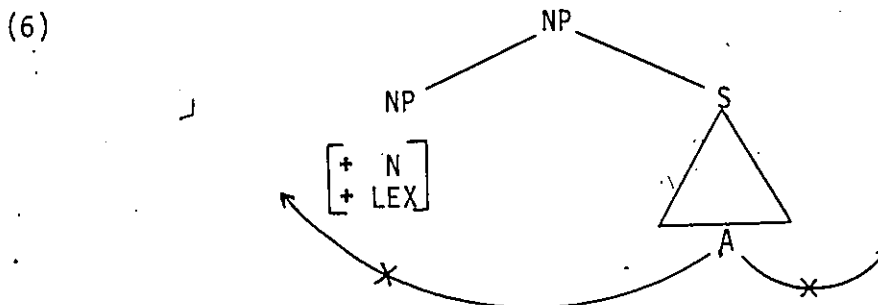
### I.3. Williams and Discourse Grammar.

Williams (1976) makes a distinction between rules that are sentence bound (e.g. Comparative Deletion) and define the form and meaning of sentences, and rules which apply across sentence boundaries in a discourse and do not obey the constraints of sentence grammar outlined in Ross (1967).

Ross' Coordinate Structure Constraint stated that no conjunct, or any element contained in a conjunct, could be moved out of a coordinate structure such as (5):



The Complex Noun Phrase Constraint outlined in Ross stated that a transformation could not move an element out of a sentence dominated by a noun phrase with a lexical head noun, as in (6):



A rule such as Comparative Deletion, which accounts for structures such as (7), obeys both these constraints.

(7) Mary ate more peaches than John ate.

However, Williams points out that a rule such as VP Deletion does not obey Ross' constraints, as illustrated in (8) and (9):

- (8) The man who didn't leave knows the man who did.
- (9) John didn't immediately open the door — first he shut the window, and then he did.

In (8), VP Deletion does not obey the Complex Noun Phrase Constraint, as the missing VP occurs in a complex NP that does not contain the antecedent leave. In (9) the same rule does not obey the Coordinate Structure Constraint. The antecedent of the missing VP (open) is not contained in the same coordinate structure as the missing VP.

Rules such as Comparative Deletion cannot occur across sentence boundaries.

- (10) A. Did John see cows?  
B. Yes, but Sam saw more horses than John saw.

In (10), the missing object must be horses, and not cows, as Comparative Deletion does not allow an antecedent to occur in a sentence different from the one containing the missing object. This is termed, by Williams, a rule of sentence grammar. In contrast to these rules are those which do not obey Ross' constraints, and which may apply across sentences in discourse. I will outline Williams' treatment of one such rule.

Williams deals exclusively with the rule of Verb Phrase Deletion (VPD) as an example of a rule of discourse grammar, but makes some generalizations about rules of this type. All rules of discourse grammar apply after all rules of sentence grammar. The rules of discourse grammar apply to the level of logical form as an output of sentence grammar. The rules of construal and quantifier interpretation map surface structures into LF in sentence grammar. The rules of discourse grammar then apply to the sentences represented in LF. While Chomsky's rules of sentence grammar apply strictly to sentences, a rule of discourse such as VPD operates across sentence or utterance boundaries. This means, for example, that an antecedent for a missing verb phrase (VP) in VPD can be a different sentence from that of the missing VP itself. Williams gives the following example:

(11) A. Did John leave?

(12) B. Yes, he did.

The missing VP in (12) has its antecedent (leave) in (11).

The rules of discourse grammar take into account the linguistic context of a sentence. Although these rules do not obey the formal constraints of sentence grammar, we might say that sentence grammar is considered the primitive element of discourse grammar.

Williams prefers an interpretive approach to the rule of VPD, instead of regarding it as a rule of deletion. To relate the following sentences

(13) a. Who can do it?

(14) b. John can.

and show that the VP do is the same in both, (14) is given a surface structure with a phonologically null anaphoric element in place of the VP, which is anaphorically related to the VP in (13). The terminals of the antecedent VP replace the deltas of the missing VP. If (13) and (14) have these surface structures,

(15) who  $[[t]_{NP} \text{ can } [[do]_V \text{ } [[it]_N \text{ } NP]_{VP}]_S$

(16)  $[_S \text{ John can } [[\Delta]_V \text{ } [[\Delta]_N \text{ } NP]_{VP}]_S$

the application of the VP rule gives (17) and (18) respectively:

(17) who  $[[t]_{NP} \text{ can } [[do]_{VP} \text{ } [[it]_N \text{ } NP]_{VP}]_S$

(18) John can  $[[do]_V \text{ } [[it]_N \text{ } NP]_{VP}$

Williams describes the relationship between discourse and sentence grammar in the following way. Of the different levels which constitute sentence grammar, the only one relevant to discourse is the level of logical form. In VPD, deltas of a VP in logical form have their antecedent in another sentence of logical form.

Williams also says there may be a semantic rule of sentence grammar which can identify deltas needing interpretation in discourse grammar. He suggests a sentence grammar rule of F-subscripting, for example, to identify "free" positions in logical form for rules of sentence grammar. In other words, sentence grammar identifies certain structures that will receive their interpretation in discourse grammar. Williams formalizes this notion with subscripting devices. I will not discuss F-subscripting here.

Williams recognizes three cases of ellipsis in English which need treatment by rules of delta interpretation in discourse grammar. These are:

(19) VP Interpretation: John did  $[[\Delta]_V]_{VP}$

as in: Who ate the cake?

John did.

(20) One's Deletion: mine  $[\Delta]_N$  is fine

as in: Whose book is broken?

Mine is fine.

(21) Sluicing: do you know who \* $[\Delta \dots]_S$

Each of these cases of ellipsis has an antecedent outside of the sentence. These three discourse rules of delta-interpretation apply to the logical forms of sentence grammar.

Williams discusses only one rule of discourse grammar, that of VPD, which follows all rules of sentence grammar by his principle of "strict utterance". (All rules of discourse grammar apply after all rules of sentence grammar):

(22) Who left?

(23) John did.

Example (23) will have the following surface structure in sentence grammar:

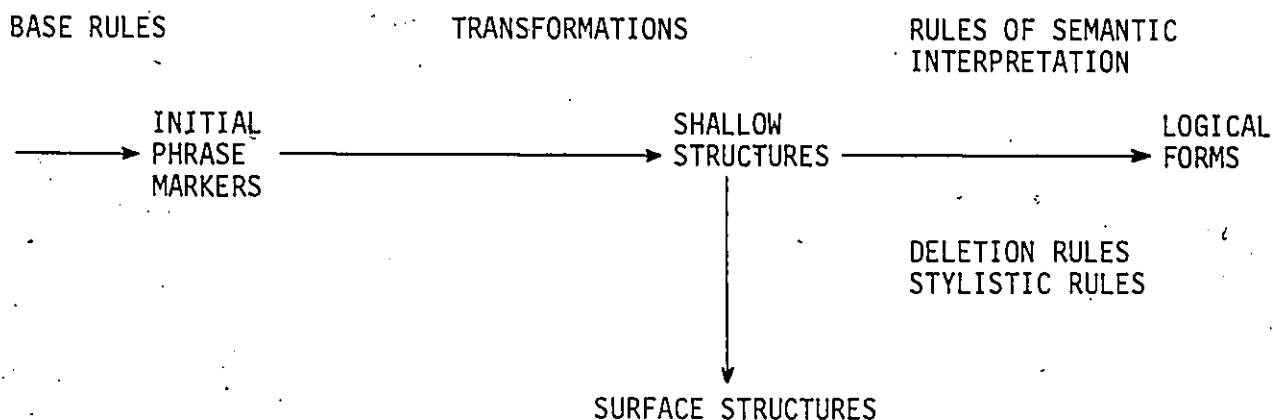
(24) John did  $[\Delta]_{VP}$

This is the input to discourse grammar. The VPD rule of discourse grammar interprets the delta as an anaphor of the VP in (22).

The trend in recent approaches (Chomsky and Lasnik 1977) is to eliminate deletion, including identity deletions such as VPD, from sentence grammar.

#### I.4. Other Approaches

Sag (1977) discusses VPD as a rule which can apply in sentence grammar or in discourse grammar. Sag proposes to include a level he terms "Shallow structure" in his model for sentence grammar, which can be contrasted with the model proposed by Chomsky and Lasnik. Their model refers to surface structure only.



Sag formulates an optional rule of VPD:

(25) Verb Phrase Deletion (optional)

	X	AUX	VP	Y
S.D.	1	2	3	4
S.C.	1	2	∅	4

He cites examples of VPD both in sentences and at the level of discourse. He does not classify the rule as belonging to sentence or discourse grammar.

(26) VPD in the sentence:

John loves Mary and Peter does  $\emptyset$  too.

(27) VPD in discourse:

- a. Who hit the home run?
- b. Betsy did  $\emptyset$ .

Sag claims that his rule accounts for both cases of VPD, as the deletion trigger is not mentioned in the structural description.

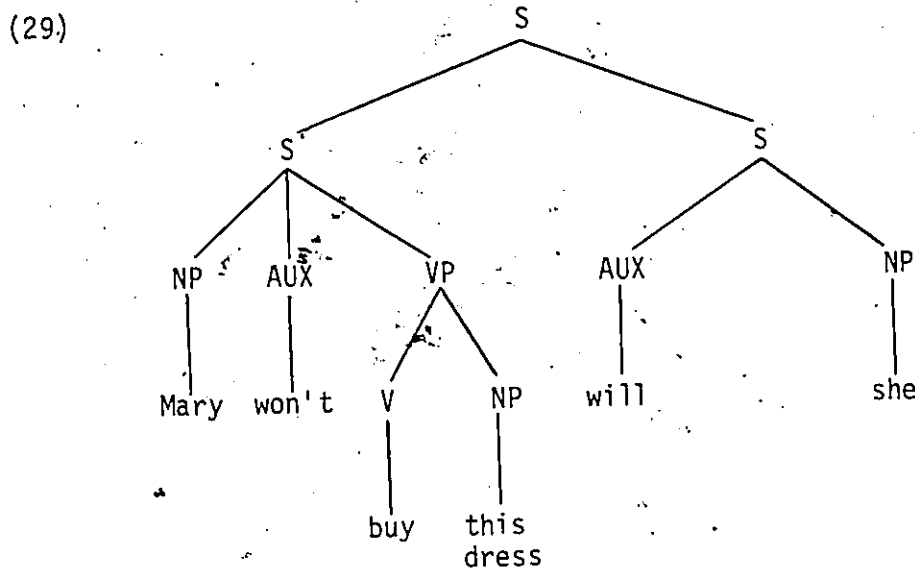
There have been other approaches to the problem of rules which do not fit into the framework of what we now consider sentence grammar. One of these approaches is seen in Emonds (1976), in a discussion of rules which precedes the sentence/discourse distinction.

Emonds (1976) has discussed stylistic rules, based on Banfield (1973). Emonds distinguishes between root transformations and structure preserving transformations. He defines root transformations as those which, "move, copy, or insert a node C into a position in which C is immediately dominated by a root S in derived structure". A root S is an S that is not dominated by a node other than S. In most cases, a root S is similar to an independent clause.

A structure-preserving transformation "introduces or substitutes a constituent C into a position in a phrase marker held by a node C." An example of a root transformation in Emonds' work is the transformation which produces a tag question.

(28) Mary won't buy this dress, will she?

The tag formation rule gives the following surface structure, in which the right-hand S is a root S:



An example of a structure-preserving transformation is Subject-Raising, transforming (30) into (31):

(30) It appears that John is sick.

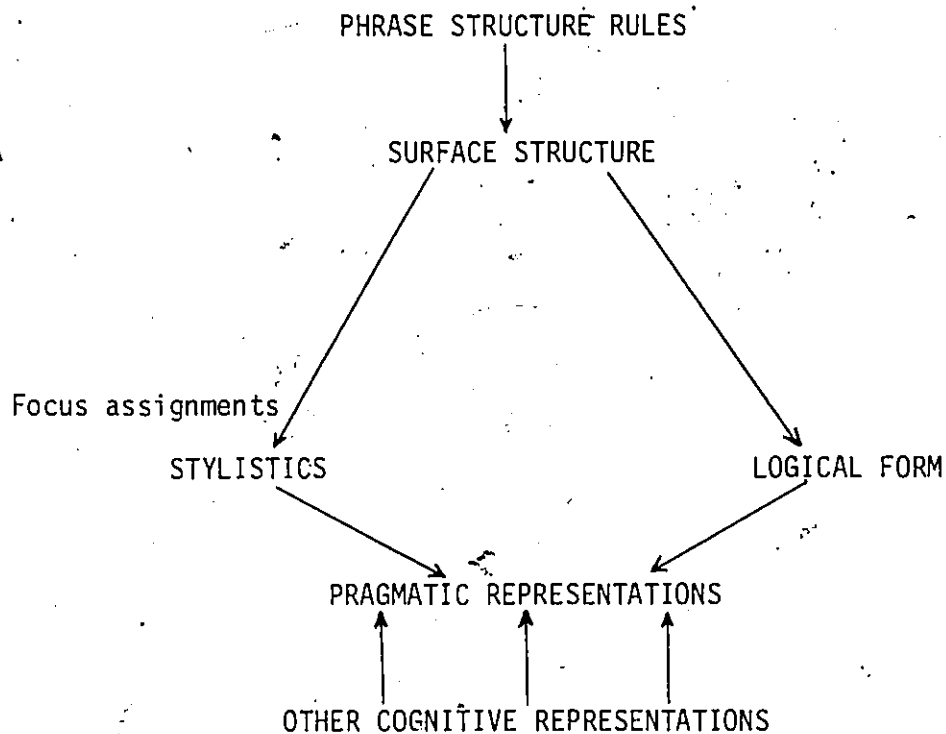
(31) John appears to be sick.

Emonds also states the structure-preserving constraint, which says that major transformations are either root or structure-preserving.

In his book, Emonds considers that stylistic transformations may violate the structure-preserving constraint, as they allow certain constituents to move freely within a sentence. These rules do have certain formal properties, which have been outlined in Banfield (1973). These properties distinguish them from grammatical transformations. Emonds points out that rules of modern English which could be considered stylistic, such as Topicalization, do obey the structure-preserving constraint. Those which do not obey the constraint apply in literary or poetic language. He suggests that in these cases the input to the transformation is more grammatical than the output. Emonds sees stylistic rules applying after all grammatical transformations. They are optional rules.

As Emonds' work comes before the Chomsky/Williams distinction between sentence and discourse grammar, it is difficult to put it into the context of this distinction. Emonds himself recognized the problem of deciding exactly where these rules should be included in a grammar. It appears that these rules could apply as in Chomsky and Lasnik's model of sentence grammar. Chomsky and Lasnik have included stylistic rules within the model, but they occur after all other transformations and are included with deletion and filters which map surface structures into universal phonetics. Emonds' stylistic rules, occurring optionally after all other rules, appear to do the same thing.

There are more attempts to characterize the relationship between sentence grammar and rules affected by discourse. Rochemont (1978) discusses stylistic rules which are supposed to account for the effect of social discourse and utterer's meaning on linguistic processes. While these stylistic rules are discourse bound, as sentences resulting from their application must be appropriate to the context of the discourse, Rochemont incorporates them into sentence grammar. He assumes the model presented by Chomsky and Lasnik (1977), but adds to the model, discussing aspects not treated by Chomsky and Lasnik.



In relation to what Chomsky has in mind for grammar, Rochemont's model appears less restricted. He views certain rules as being less formal than other rules of sentence grammar, which can still not be considered real rules of discourse.

There is a gradience seen from strict rules of sentence grammar, towards rules that might be considered "intermediary" between sentence and discourse grammar.

Rochemont's stylistic rules have the properties of transformations, but are not considered syntactic. They do not influence interpretation. For example, Rochemont discusses the Inverted Directional Construction (IDC)<sup>1</sup>.

(32) Into the house ran John.

(33) Out walked Mary.

(34) Down the hill rolled the carriage.

He claims two processes are involved here: PP Preposing, a syntactic operation, and Subject Postposing, a stylistic operation. These are independent operations as seen in (35), (36), and (37):

- (35) Into the house John ran.
- (36) Down the hill the carriage rolled.
- (37) Out Mary walked.

In all of the above, PP Preposing applies alone. Rochemont analyzes IDC sentences in terms of Focus, which is a factor of discourse. Stylistic rules such as Subject Postposing apply to surface structures. Rules of focus assignment apply before these stylistic rules. The rule of Subject Postposing formulated by Rochemont is as follows:

- (38) Subject Postposing

(	X	NP	(AUX)	(V)	Y)	1	∅	3	4	2	5
	S			VP							
	1	2	3	4	5						
	Focus										

An IDC sentence such as (39) is derived in the following way:

- (39) IDC: Into the house (there) ran a man in a blue suit.

Application of syntactic rules gives the surface structure (40):

- (40) [[Into the house] [a man in a blue suit] [ran]]

Subject Postposing applies to this:

(41) Subject Postposing

[[Into the house] [ran] [a man in a blue suit]]

Another stylistic rule of 'there' insertion gives the final result:<sup>2</sup>

(42) [[Into the house] there [ran] [a man in a blue suit]]

Rochemont says his stylistic rules operate independently of formal syntactic and semantic processes. The stylistic rules apply to the output of syntactic rules.

#### I.5. Summary

In summary, it has been shown that the Chomsky/Lasnik model for sentence grammar and the discourse rule of VP Deletion presented by Williams constitute a basis for the distinction between sentence and discourse grammar. Other attempts have been made by Emonds and Rochemont to clarify the differences seen in the status of certain rule types.

## CHAPTER I

FOOTNOTES

1. Emonds (1976) discussed the rule of Directional Adverb Preposing as a root transformation. Two processes are involved in deriving sentences such as:

Down the street rolled the baby carriage.

In came John.

The first is adverbial PP preposing; the second is subject-simple verb inversion. Rochemont is changing the status of the rule as seen by Emonds.

2. Emonds (1976) considers 'there' - insertion to be a structure - preserving (non-stylistic) transformation as in:

A hatless stranger appeared.

There appeared a hatless stranger.

Rochemont changes the status of this rule and treats it as stylistic.

## CHAPTER II

### THE CORE AND THE PERIPHERY

#### II.1. Introduction

In this chapter I propose to outline the distinctions between core and peripheral grammar as seen by Chomsky and Koster. Chomsky's theory of sentence grammar will be examined, with special reference to elements contained in the core. Koster's views on the core and the periphery will be presented, with his criteria for distinguishing between the two. His approach will be contrasted with that of Chomsky.

#### II.2. Approaches to core and periphery

Chomsky has distinguished between two different aspects of sentence grammar: the core and the periphery. Core grammar represents the part of sentence grammar that is related to innate principles. As mentioned in Chapter I, within Chomsky's rationalist framework humans are considered to possess an innate capacity for language learning. Core grammar, as part of sentence grammar, represents part of this innate capacity, or is determined by it. In other words, this language faculty common to humans is biologically based. A child at birth is equipped for language learning.

Part of this innate capacity is a restriction on the form of grammars acceptable for language learning. A child possesses, in the innate capacity for language learning, an evaluation procedure which enables him to choose the optimal grammar available to him.

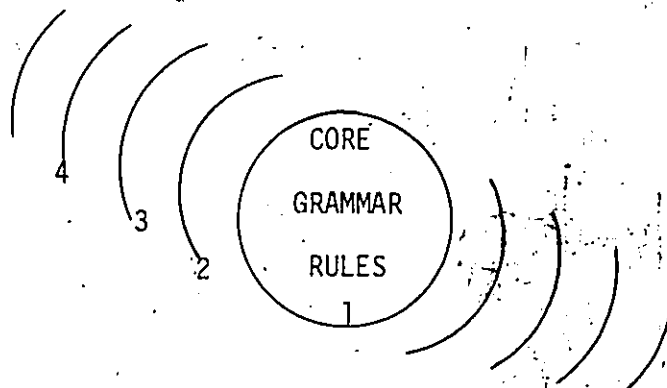
Closely related to the innate capacity are universals of language. These are the rules and principles which all human languages possess. Although languages may differ in other properties, all conform to these universal principles of sound, meaning, and structure, termed universal grammar (UG). A child has tacit knowledge of these universals. Chomsky maintains that a human language learner would be unable to learn a language constructed in violation of UG.

The theory of UG incorporates what is known as the theory of markedness. It is within this theory of markedness that it is possible to differentiate between the core of sentence grammar and the periphery. Riemsdijk (1978) considers the theory of markedness to be the part of linguistic theory concerned with the evaluation procedure possessed by the child for choosing the optimal available grammar. Grammatical properties which rank high in the evaluation procedure are unmarked, and those obtaining a low ranking are marked.

Core grammar is then a subset of universal grammar. The core grammar for one language selects among the universal devices, which are innate.

Koster (1978) calls the unmarked structures the core grammar, and terms the marked structures of a language the periphery. This means that language learning involves the establishment of the limits of core grammar, and the addition of marked structures up to the periphery. Transformations which violate core grammar specifications are marked. They may have structural descriptions more complex than those allowed for in core grammar or violate some other condition of core grammar.

The implications of this theory of markedness for the language learner are clear. Structures or transformations which are highly marked are more difficult for the language learner to master. They will presumably be learned much later than the unmarked rules of core grammar. Stillings (1977) uses the following diagram to indicate the difference between core and peripheral grammar:



The core grammar is innate for the language learner. It contains the universal properties of natural language which are the generalizations that can be made across language boundaries. The processes in the core are the most easily learned grammatical structures.

The outer circles 2, 3, 4 represent the more highly marked rules and transformations. These rules are more complex and less general, and are learned later than the rules in the core. This is the periphery of sentence grammar.

White (1980) has approached the problem of the relationship between grammatical theory and language acquisition. She makes some comments about the link between markedness and learnability, or the facility with which a structure may be learned. She says it seems reasonable to assume there is a certain connection between grammatical complexity, as defined by markedness, and acquisitional complexity. White is careful to note that the grammar should not account for perceptual difficulties, these being outside its scope. The grammar accounts for grammatical complexity only. This restriction will be important in considering the work of Kuno in Chapter III.

A further approach to the distinction within sentence grammar between the core and the periphery is outlined in Hirschbühler and Rivero (1980). They propose an interpretation of the distinction which considers a rule of grammar to be an "aggregate of dimensions",

some of which are marked and belong to the periphery, and some of which are unmarked and belong to the core. In this interpretation, the differing elements of the core and the periphery are embodied in a single rule. This interpretation will be considered more extensively in Chapter III.

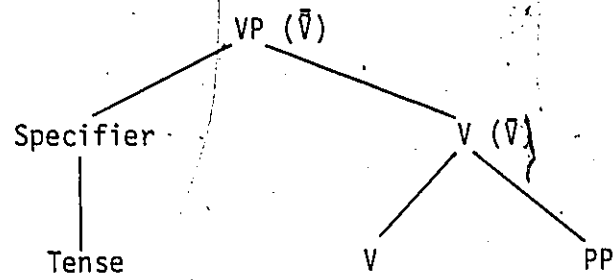
### II.3. Chomsky and Rules of Core Grammar.

In Chapter I, I outlined briefly the Chomsky-Lasnik model of sentence grammar. The base rules generate base structures and the transformations convert the base structures to surface structures. In addition, there are certain conditions which must be met for a structure generated by the base rules and the transformations to be considered grammatical. The conditions in this way act as filters. Chomsky sees core grammar as a restricted number of general rules with conditions governing their operation.

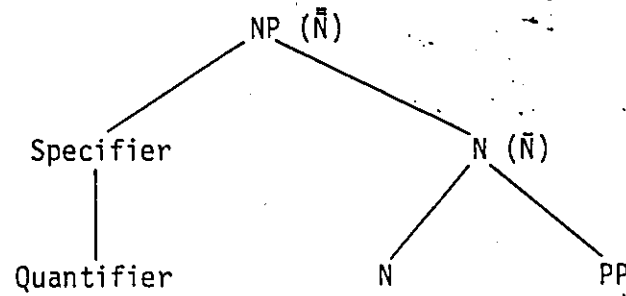
I will examine more closely the different areas of this grammar, with an example of their operation. I will base my presentation on the material found in Chomsky's (1977) "On WH-Movement". The rules of the base follow Chomsky's X-bar notation, a theory intended to reflect the parallel behavior of major categories in syntax: noun phrases (NP), verb phrases (VP), and adjective

phrases (AP). The X-bar notation generalizes these structures by giving them levels of "embedding".<sup>1</sup>

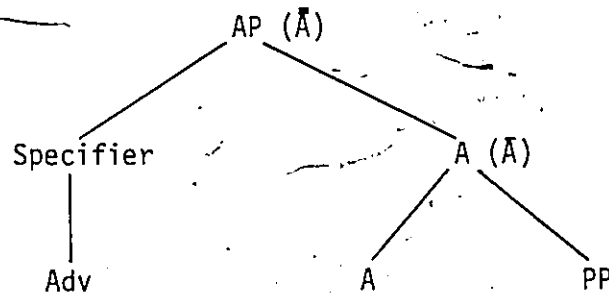
(1)



(2)



(3)



Chomsky now recognizes one transformation in the transformational component of core grammar: Move  $\alpha$ . Transformations in core grammar are restricted in form and function. The formal representation of Chomsky's WH-Movement is shown in (4):

(4) (COMP, vbl, wh-phrase, vbl)

Move wh-phrase into COMP

The structural description (SD) of all transformations must be  $(a_1 \dots a_n)$  with each  $a$  as a variable, X-bar category, or terminal string.

All major movement rules are cyclic, and NP Movement obeys the structure-preserving hypothesis put forth by Emonds. This means that an NP can only move into the position of a base-generated NP. The transformations are unordered and optional.

Included in core grammar are interpretive rules and rules of construal which convert the output of transformational rules into logical form. Semantic rules act on LF. Chomsky has three construal rules in core grammar; (in "WH-Movement", for example). Construal rules assign antecedents to anaphors. These are:

- a. The Reciprocal Rule: This rule assigns each other the feature [+anaphoric to  $_i$ ] in a structure containing  $NP_i$ .

- b. The Bound Anaphora Rule assigns the feature [+anaphoric to  $i$ ] to a pronoun in a structure containing  $NP_i$  in the context  $[_{NP} - Possessive - N_X]$ .
- c. The Disjoint Reference Rule functions to assign nonanaphoric reference to a pronoun in relation to some  $NP_i$ .

The transformations and interpretive rules of core grammar must meet three conditions: the Subjacency condition, the Propositional Island Condition (PIC), and the Specified Subject Condition (SSC). These conditions can be language specific. They have been modified and renamed in recent work by Chomsky, but I will not be concerned with problems relating to the modifications.

The condition of Subjacency states that a transformation cannot apply across more than one cyclic node at a time. This condition applies to transformations only, and not to interpretive rules. Chomsky's (1977) statement of the condition is:

"No rule can involve X, Y, X superior to Y,  
if Y is not subjacent to X."

(5)  $[[who]_{COMP_1} \text{ did Mary claim } [ [ ]_{\bar{S}} \text{ PRO to have seen } t ]_{COMP_2}]$

who moves by WH-Movement to  $COMP_2$  from  $t$  and then to  $COMP_1$ . The movement occurs across one cyclic node,  $\bar{S}$ . (The cyclic nodes are  $\bar{S}$  and NP).

Chomsky handles rules that are apparent counterexamples to Subjacency, such as WH-Movement, by showing that the WH element moves from COMP to COMP, one cycle at a time, thus avoiding any violation of Subjacency. The derivation for "whom do you think they say that they saw?" is as follows:

(6) [<sub>S</sub>COMP [<sub>S</sub> you think [<sub>S</sub>COMP [<sub>S</sub> they say [<sub>S</sub>Comp [they saw whom]]]]]

In the example, whom moves from COMP to COMP, one cyclic node at a time. Subjacency explains the Complex NP and WH Island Constraints of Ross (1967).

The Propositional Island Condition makes extraction from a tensed sentence impossible. It applies to transformations and interpretive rules and in later work it is considered as a filter on LF, and not as a condition that blocks the application of the rules of sentence grammar. Chomsky states:

"No rule can involve X, Y in the structure

...X...[<sub>S</sub> ...Y...]...

where <sub>S</sub> is a tensed sentence.

Subject Raising, for example, extracts from an infinitive structure where there is no tense.

(7) John seems to be sick.

In French, in the L-Tous transformation, the PIC prohibits

(8) \* J'ai tout voulu que Marie mange.

from

(9) J'ai voulu que Marie mange tout.

The PIC prohibits the movement of tout from the tensed sentence in (9), blocking the derivation of (8).

The Specified Subject Condition prohibits extraction from a sentence with a specified subject. This refers to a lexical subject or a non-anaphoric pronoun and also trace. This condition, like the PIC, applies to both transformations and rules of interpretation, and has been reinterpreted as a filter on LF in recent work.

"No rule can involve X, Y in the structure

...X...[ $\mathcal{C}$ ...Z...-WYZ...]...

where Z is the specified subject of WYZ in  $\mathcal{C}$ ."

(10) COMP you saw [<sub>NP</sub>John's picture of who]

(11) \*Who did you see John's pictures of?

The rule applies in the case of  $\mathcal{C}$  as S or NP.

Stillings clarifies the operation of core grammar with two examples, showing how the various components produce sentences.

EXAMPLE 1: 1. The base generates a structure:

$\Delta$  seems [<sub>S</sub> John to like Bill ].

2. A movement transformation operates on the output of the base.

Move NP:  $\longrightarrow$  John seems [<sub>S</sub> t to like Bill ].

3. The sentence is grammatical as none of the three conditions is violated, and the rules of interpretation do not apply. The resulting surface structure is:

John seems to like Bill.

EXAMPLE 2: 1. The base generates this structure:

They expect [<sub>S</sub>(someone) to read what books  
to each other (by  $\Delta$ )].

2. The following transformations operate on the output of the base component:

a. Move NP<sub>1</sub>

$\longrightarrow$  They expect [<sub>S</sub> $\Delta$  to be read what books to each other (by someone)].

b. Move  $NP_2$

→ They expect [ $\xi$  what books to be read to each other (by someone)].

c. Move WH

→ What books do they expect [ $\xi$  to be read to each other...].

3. The interpretive reciprocal rule must apply to assign to each other anaphoric reference to they.

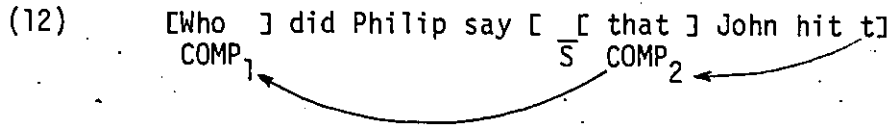
What books do they  $_i$  expect [ $\xi$   $\bar{t}$  to be read to each other  $_i$ ].

4. The SSC is violated as they and each other are separated by a specified subject. The structure is then ungrammatical.

\*What books do they expect to be read to each other?

Chomsky's rule of WH-Movement is a good example of a core transformation in English. It handles processes that were once considered to be handled by many other rules. For example, under WH-Movement Chomsky places topicalization and infinitival relatives. WH-Movement is optional and cyclic, and obeys Subjacency, the PIC and the SSC. It always leaves a trace  $t$  in the position from which the WH-phrase moved, to be later interpreted by the rules of construal.

An example of WH-Movement is the following:



On the first cycle, the transformation moves who from t to COMP<sub>2</sub>.

On the second cycle it is moved from COMP<sub>2</sub> to COMP<sub>1</sub>. This is WH-Movement out of a clause. WH-Movement within a clause deals with structures such as:



There are three further aspects of Chomsky's analysis of WH-Movement which must be considered. The first of these concerns a prohibition in core grammar which states that a COMP node cannot be filled by more than one WH-phrase at a time. The following is ungrammatical..

(14) \*Which problems were you wondering who solved?

Only one element can be questioned at a time.

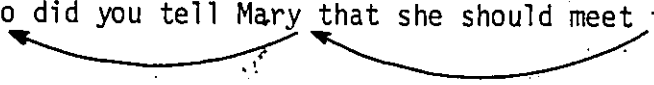
(15) [ $\bar{S}$ COMP<sub>1</sub> [ $\bar{S}$ you were wondering [ $\bar{S}$ COMP<sub>2</sub> [ $\bar{S}$ who solved which problem]]]]

WH-Movement applies to who. Then which cannot be moved to COMP<sub>2</sub> because it is filled, nor can it be moved to COMP<sub>1</sub> because this would violate the PIC. The sentence is therefore ungrammatical.

The Double COMP Prohibition in this way accounts for WH-Island Constraints (Ross 1967).

The second aspect to be considered is that of bridges<sup>2</sup>. WH-Movement is freely permitted in sentences like (13), which are examples of WH-Movement within a clause. However, in English the following, an apparent violation of the PIC, is also permitted:

(16) Who did you tell Mary that she should meet t ?



Who appears to be moved out of a tensed sentence. Chomsky has suggested two answers to the problem. The first is that certain words can act as bridges, allowing WH words to be moved out of tensed sentences. For example, in sentence (16), tell acts as a bridge for the extraction of who from a tensed sentence. The second solution proposed by Chomsky is that the conditions on core grammar operations can be language specific. In this way it is possible to make a proviso on the PIC and SSC in English allowing extraction over bridge words if the WH word has been moved to the COMP of an embedded S. The PIC would state that in the structure

...X...[ < ...Y... ]...

nothing can involve X and Y unless Y is in COMP.

This would account for the grammaticality of (16). This solution is also known as the COMP Escape Hatch.

The third consideration is that of a change in interpretation of the conditions in core grammar. Chomsky originally postulated the A-over-A condition, which stated that no transformation could affect any category except the dominating one:



Chomsky (1968) stated:

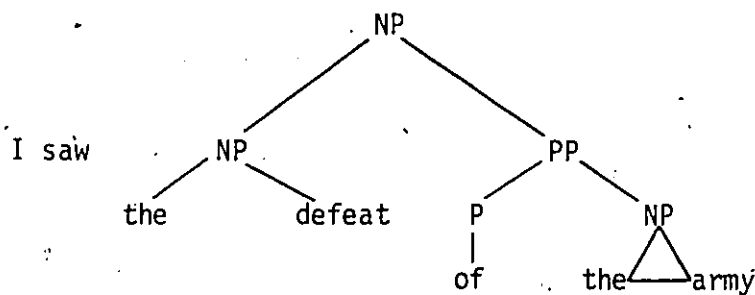
If a transformation applies to a structure of the form

$$[ \dots [ A \dots ]_A \dots ]_S$$

for any category A, then it must be so interpreted as to apply to the maximal phrase of the type A.

For example, to obtain the passive from a sentence such as I saw the defeat of the army with the structure in (17), the A-over-A principle allows for the movement of the dominating NP only, to derive The defeat of the army was seen by me.

(17)



This absolute interpretation of the A-over-A principle represents the unmarked case in core grammar. Chomsky now also accepts a non-absolute interpretation which allows for the marked case. In certain marked cases, a transformation may affect the lower category. WH-Movement in the following can move WHO/WHOM in isolation:

- (18) They took a picture of whom?
- (19) Whom did they take a picture of?
- (20) You would approve of my seeing whom?
- (21) Whom would you approve of my seeing?

#### II.4. Koster and the Distinction between Core and Periphery

Koster (1978) has contributed to the discussion of what is involved in core grammar, particularly in drawing a distinction between the core and the periphery. His view is that the theory of core grammar will lead to the elaboration of a theory of markedness in syntax. The unmarked rules and conditions of the core are more easily learned, and include operations that are common to all languages. The periphery, on the other hand, includes marked structures which are less readily accessible to the language learner.

Koster deals with two examples of the difference between the core and the periphery. The first of these involves islands. He states that it is possible for a phrase to be an island of core grammar, but not to be an island in the peripheral sense. The examples he uses are those seen above (13 and 16) of WH-Movement, within a clause and out of a clause.

(22) Who did Mary meet t ?

(23) Who did you tell Mary that she should meet t ?

Koster concludes that clause bound WH-Movement is more universal than WH extraction from a clause, as some languages do not even allow the latter. In addition, extraction from a clause is dependent on the presence of a bridge, as seen earlier, or a particular lexical structure, making it more restricted. The hypothesis suggested by Koster is that a full clause is an island of core grammar, while extraction belongs in the periphery. English has the marked alternative in allowing extraction out of a clause. In such a way rules reveal their markedness by their cost. Those which are less general and which depend on such things as lexical structure have a higher cost in terms of acquisition. They fall outside of core grammar.

The second example of the difference between the core and the periphery suggested by Koster concerns root and non-root sentences.

Emonds (1976) defines a root sentence as an S that is not dominated by a node other than S. This is related to the notion of independent clause in traditional grammar, but a root S in some cases may be from a dependent clause.

(24) Had John arrived earlier, everyone would have been upset.

(25) Will we be welcome, do you think?

The notion of root transformation as discussed in Emonds was presented in Chapter I. Chomsky terms a root transformation one that applies only to the full sentence structure, not to an embedded sentence. Emonds gives several examples of root transformations, one of them being Subject-Auxiliary Inversion. This is the inversion of Subject NP and auxiliary in direct questions.

(26) Is Harry coming?

This rule is formalized in Emonds as:

(27) COMP - NP - Aux - X  $\longrightarrow$  1-3-2-4

where 1 dominates WH or NEG. It applies only in the highest S or root S. This also covers cases of conjoined sentences immediately dominated by the highest S:

(28) She didn't do the dishes, and why should she?

A nonroot transformation may apply to something other than the matrix sentence. Chomsky points out that Extraposition applies to the underlying subject NP of the following:

(29) [<sub>NP</sub>the only one that I like of Tolstoy's novels] is out of print.

The result is:

(30) The only one of Tolstoy's novels that I like is out of print.

Hooper and Thompson (1973) have disputed the adequacy of Emonds' views on root transformations. They suggest that root phenomena also occur in embedded clauses under certain conditions. Their view is that, with the semantic notion of assertion, it can be shown that root transformations have applicability restricted to asserted clauses. Because most embedded clauses cannot be asserted or made emphatic, root transformations do not apply to them. However, Hooper and Thompson point out that in some cases, under certain conditions, root transformations may apply to embedded clauses. Hooper and Thompson give examples such as (31) and (32), in which a root transformation is acceptable in an embedded sentence following a verb of saying. The embedded sentence occurs as reported discourse, and is an example of assertion.

(31) Sally plans for Gary to marry her, and he vows that marry her he will:

(32) I exclaimed that never in my life had I seen such a crowd.

Koster places the root/nonroot controversy within the distinction between core and peripheral grammar. Embedded root phenomena are peripheral because of their cost. Root phenomena occur in embedded clauses in a restricted way, depending upon certain conditions which are lexically governed. In certain verb complements, root phenomena can occur. These complements contain the same verbs that allow extraction.

(33) He explained that each part he has examined carefully.

Root phenomena are not allowed in NP or PP complements.

(34) \*I resent that each part he has examined carefully.

The fact that embedded root phenomena are peripheral could also explain the variation in judgment in many cases, and the fact that some languages make a very strict distinction between root and nonroot phenomena. In languages such as Dutch and German, root phenomena never occur in embedded clauses.

Koster then makes the assumption that extraction out of a clause and embedded root phenomena fall outside of core grammar in the periphery.

To understand Koster's position, it is necessary to outline his model in a general way. Koster argues against the need for major transformations. He favours a phrase structure grammar with moved phrases (WH words and NPs) generated directly in the base. He eliminates movement transformations, replacing them with a rule of coindexing. His argument is that there is no distinction between movement rules and rules of construal or coindexing.

I have already outlined the Chomsky/Lasnik model for sentence grammar and explained the operation of rules with the formulation Move  $\alpha$  (Move NP; Move WH). The rules transform the deep structures of the base into surface structures containing traces left by Move  $\alpha$ , unexpanded nodes, and material which is to be deleted. Surface structure is also subject to rules of construal which co-index anaphors and antecedents.

In Chomsky's model, the Raising transformation operates as in (35).

(35)  $\Delta$  appears [  $\xi$  Mary to be sick ].

The transformation Move  $\alpha$  operates to raise the NP (Mary), to yield (36).

(36) Mary<sub>i</sub> appears t<sub>i</sub> to be sick.

Mary is connected to the empty  $\Delta$  subject through the trace left by the transformation.

Koster finds that Move  $\alpha$  is unnecessary. He says that movement and construal can be distinguished by the property of Subjacency only. He reduces Subjacency to two principles - the Bounding Condition and the Locality Principle. His model eliminates cyclic transformations, but maintains minor movements, deletions, and stylistic rules.

1. Base rules

2. Coindexing

---

3a. Deletion rules, etc.

3b. Interpretive rules, etc.

The base generates WH-phrases and NP subjects. Unexpanded nodes (e) appear where movement originates. An antecedent node is coindexed with a consequent, or a verb is coindexed with an unexpanded verbal node. Koster's treatment of the operation in (37) is shown in (38).

(37) Mary appears [<sub>S</sub> e to be sick].

Unexpanded nodes  $e$  appear in those positions from which a phrase was usually said to have been extracted.

- (38) Mary appears to be sick.  
 Coindex  $\longrightarrow$   
 Mary<sub>i</sub> appears e<sub>i</sub> to be sick.

The coindexing can also apply to verbs, as in (39).

- (39) John saw<sub>i</sub> Susan and Mary e<sub>i</sub> Bob.

The coindexing rule is constrained in various ways. For example, it cannot apply across a tensed S boundary, as in Chomsky's Propositional Island Constraint.

- (40) \*Mary<sub>i</sub> appears [<sub>S</sub> that e<sub>i</sub> is sick].

The Bounding Condition and the Locality Principle constrain the rule of coindexing. According to the Bounding Condition an empty node must have an antecedent within the same phrase. This means that phrases are islands, in the unmarked case, in core grammar. The Bounding Condition accounts for some of the phenomena also accounted for by Subjacency. In (41), Subjacency and the Bounding Condition predict the same thing.

- (41) Who<sub>i</sub> do you know [<sub>NP</sub> the boy<sub>j</sub> [<sub>S</sub> e<sub>j</sub> that [<sub>S</sub> e<sub>j</sub> saw e<sub>i</sub> ]]].

The sentence is prohibited by Subjacency because who<sub>i</sub> and  $e_i$  are separated by more than one cyclic node: NP and  $\bar{S}$  or S. The Bounding Condition rules out the sentence because the antecedent of  $e_i$  is not in  $\bar{S}$ .

The Locality Principle handles rules involving an antecedent and a consequent. The antecedent for a given consequent is the closest one in terms of phrase structure. The Locality Principle permits (42).

(42) Mary<sub>i</sub> appears NP<sub>i</sub> to like Dave.

It prohibits (43) because Mary<sub>i</sub> and NP<sub>i</sub> are separated by Dave, a possible antecedent.

(43) \*Mary<sub>i</sub> appears Dave to like NP<sub>i</sub>.

This principle partly replaces Chomsky's Specified Subject Condition. The Cojacency Principle, one instance of the Locality Principle, will be discussed in CHAPTER III.

Koster discusses the distinction between core and periphery, including his rules of construal, with the conditions on them, in the core. He proposes auxiliary hypotheses to account for language specific modifications of core grammar. At the core is a

minimum formulation of a rule, the essential idea, with no exceptions. Because the rules of this core are too strict for natural language, he provides auxiliary hypotheses which list the exceptions in as general terms as possible. For example, the following sentence is unmarked. It reflects a universal rule of core grammar.

(44) What did you see e ?

As shown above, WH is generated in the base and coindexed with e. The next two sentences are marked, both falling outside the ideals of core grammar. The first is acceptable. The second is unacceptable and reflects the cost of the first. The first is derived by a rule specific to English.

(45) Who did you see a picture of e ?

(46) ?? Who did you give a picture of e to Bill?

Koster's approach in some of these cases is similar to Chomsky's treatment of WH-Movement over a bridge, which has been discussed in this paper. For English, he makes a special proviso to the PIC which is the equivalent of Chomsky's COMP Escape Hatch, permitting an apparent violation of the PIC. Koster's auxiliary hypotheses handle such cases by allowing for an additional hypothesis following the core grammar rule. The additional hypothesis handles problems specific to a language.

Koster therefore makes a distinction between core grammar and language specific modifications that may be required to enable core grammar rules to work for natural language. The specifications of these additional hypotheses are possible only at a cost. The auxiliary hypotheses constitute generalizations of core grammar specifications. Universal grammar establishes the principles of core grammar, but also points to the range of language-specific alterations that may fall outside of the essential core.

In addition to the preceding contributions to core grammar, Koster lists three characteristics of noncore, or marked, structure: variation across languages, variation in judgment, and susceptibility to lexical and structural complexity. The acceptability of noncore structure may be affected by complexity factors which are part of other cognitive systems which interact with grammar. Examples of these complexity factors are depth of embedding, nesting and double binding. Koster feels that marked structures have a greater sensitivity to phenomena from other cognitive systems which might influence them.

The first complexity factor, depth of embedding, reflects the interaction of lexical factors. Koster notes the difference in acceptability between the following two sentences.

- (47) Who did you see [NP a picture of e]
- (48) \*Who did you destroy [NP a picture of e]

These correspond to Who did you see a picture of ? and Who did you destroy a picture of ? respectively. He also notes that many examples of embedding such as this one cannot occur in Dutch, "indicating a variation across languages. Some examples show a variation in acceptability.

- (49) Which country is there money for a trip to?

Within the framework outlined by Koster, these sentences fall outside the limitation of core grammar, but are still not at the extreme periphery.

Nesting is the second complexity factor mentioned in Koster. A sentence containing a nested construction may be unacceptable due to processing or perceptual difficulties. The structure (50) is an example of an unacceptable nested construction:

- (50) \*the man a brother of e I know.

The Nesting Hypothesis predicts that (51), the underlying structure for (52), is unacceptable.

(51) Who do you believe [<sub>S</sub> to [e] [<sub>S</sub> Mary thinks Joan talked]]

(52) \*Who/whom do you believe to Mary thinks Joan talked?

The third complexity factor is double binding. The Double Binding Constraint states:

No rule refers to X, Y in:

...X...[ $\alpha$  ...Y...]...Z...

where Y is empty (e), and Z is linked to  $\alpha$ .

This constraint excludes e's in COMP-internal categories.

(53) \*the man [ $\alpha$  a brother of e] I know e

X

Y

Z

(54) \*who do you believe [ $\alpha$  to e] Mary thinks Joan talked e

X

Y

Z

This rules out cases of dangling prepositions in the COMP.

No rule can operate to derive sentences in (53) and (54), as the Y of both structures contains an empty node e. The same examples have been shown to be ruled out by the Nesting Hypothesis.

#### II.5. Summary

In summary, it has been shown that Chomsky's theory of core grammar allows for a rule of MOVE  $\alpha$  and interpretive rules, all governed by Subjacency, the PIC and the SSC. Koster's core rule of coindexing replaces movement rules and is subject to the Bounding Condition and Locality Principle. Koster has provided a view of core and peripheral grammar within a theory of markedness.

## CHAPTER II

FOOTNOTES

1. The  $\bar{X}$ -notation has been explored in detail in Jackendoff's  $\bar{X}$ -Syntax: A Study of Phrase Structure. 1977. Cambridge, Mass.: MIT Press.
2. Bridges have been discussed in Erteschik, Nomi. "On the Nature of Island Constraints". Unpublished doctoral dissertation. 1973. MIT. Distributed by Indiana University Linguistics Club.

CHAPTER IIIGAPPING

## III.I. Introduction

In this chapter I will deal with the rule of Gapping, seen in the light of the sentence/discourse, core/peripheral distinctions. I will present a historical overview of treatments of Gapping from Ross (1967) to Koster (1978). While most of these works were not produced with the sentence/discourse distinction in mind, it is possible in some cases to identify the approach, in retrospect, within recent distinctions. Some of the treatments of Gapping do not lend themselves so easily to such classification. Where possible, I will attempt to place an approach to Gapping within sentence or discourse grammar.

I will also suggest that Gapping is a rule of sentence grammar, but does not fit neatly into either the core or the periphery. Rather, it appears to conform to an interpretation of the core/peripheral distinction as seen by Hirschbühler and Rivero (1980). The rule, according to their approach, could be considered an "aggregate of dimensions", some of which belong to the core and some to the periphery.

The historical treatments of Gapping range from those which place the rule outside sentence grammar, as part of discourse, to those which place Gapping within the strict core area of sentence grammar.

### III.2. Gapping Characteristics

The rule of Gapping is first discussed in Ross (1967) and several times after that in transformational literature. Sag (1977) summarizes Gapping and the related constraints.

Gapping deletes simple verbs and auxiliary verbs in coordinate structures as in (1) and (2):

- (1) John ate peaches, and Susan, pears.
- (2) John has spoken to his mother, and Susan, to her father.

If there is a negation in the auxiliary, Gapping cannot occur.

- (3) \*I didn't play hockey, and Bill, football.

However, Gapping can occur if the negation is in the conjunction.

- (4) I didn't play hockey, nor Bill football.

In coordinate structure, only the highest S can undergo Gapping.

(5) \*John said yes, and I know that Melanie no.

There can be no deletion of said after Melanie in (5) because it is inside an embedded sentence.

Coordinating conjunctions and, or and nor are acceptable in sentences which undergo Gapping. But is less acceptable as a coordinate conjunction in a gapped sentence.

(6) ?Mary played the piano, but Anne the guitar.

In English, backward Gapping does not occur.

(7) \*John novels, and Mary read short stories.

In a sentence undergoing Gapping, identical pre-verbal adverbs can be gapped. Non-identical adverbs cannot be gapped.

(8) is acceptable because the adverb that is gapped in the right conjunct is identical to the adverb of the left conjunct. In (9) the adverbs are not identical and Gapping is not permitted.

(8) Pam sometimes plays tennis, and Lucy, squash.

(9) \*Pam sometimes plays tennis, and Lucy frequently squash.

Gapping remnants are discussed in detail by Sag. While most treatments of Gapping deal with gapped clauses containing two remnants, as in the above examples, Sag gives acceptable gapped sentences with three remnants. The following is Sag's example.

- (10) Peter talked to his boss on Tuesday, and Betsy to her supervisor on Wednesday.

There is much controversy over the acceptability of some gapped clauses containing three remnants, as in (11), an example from Jackendoff (1971).

- (11) ? Charlie entered the bedroom at 5:30, and Vera the kitchen at 6:00.

The second remnant in a gapped clause (in which the verb only has been gapped) can be an NP, an adjective phrase (AP), a that-clause, a for-to clause, or a prepositional phrase (PP).

- (12) Sally saw Anne, and John, Mark. (NP)
- (13) Mary looked upset, and Alice calm. (AP)
- (14) Sam said he was late, and Phil that he was early.  
(that-clause)

(15) She prefers for her mother to go with her, and Tom for his father to go with him. (for-to clause)

(16) Anne waited at the bus-stop, and Susan on the steps. (PP)

In cases in which the second gapping remnant is an S or  $\bar{S}$ , a complementizer must be present, even if the complementizer could otherwise be deleted.

(17) His mother said he was polite, and his father that he was rude.

Another constraint on Gapping is that the NP immediately following the verb cannot be gapped.

(18) \*Susan gave a book to her mother, and Lawrence to his father.

Ross (1967b) has discussed the case of no NP being gapped, and the numerous possibilities which result in a sentence like (19):

(19) John tried to begin to write a play, and Harry

(a) to begin to write a novel

(b) to write a novel

(c) a novel

In addition, parts of NP's cannot be gapped by themselves.

- (20) \*John washed his brother's sleek, low racing car, and Peter
- (a) brother's battered old jalopy
  - (b) battered old jalopy
  - (c) old jalopy
  - (d) jalopy

### III.3. Historical Overview

In order to present the major treatments of Gapping found in the literature, I have divided the relevant material into three groups. The first five are treatments of Gapping from 1967 to 1975. As the core/peripheral distinction did not exist at the time these were written, I will attempt to reinterpret them within the distinction outlined in the first two chapters of this thesis. The five treatments of Gapping are those of Ross (1967a and b), Koutsoudas (1971), Jackendoff (1971), Langendoen (1974), and Stillings (1975). A sixth approach, that of Hankamer (1973), is impossible to place within the present framework as it was conceived within the framework of generative semantics. I will include it with the above five approaches for chronological reasons.

The second group will deal with Gapping from the point of view of linguists who regard it as a rule which lies outside of sentence grammar. Kuno (1976) sees the rule as subject to perceptual

factors and he places it in a non-formal component, outside the framework of sentence grammar as presented by Chomsky. A second approach, that of Sag (1977), is more formal than that of Kuno, but Sag accepts Kuno's perceptual principles as they regard Gapping.

### III.3A Ross

Ross (1967a and b) gave the first detailed account of Gapping. He saw it as a rule deleting non-constituents, converting the underlying structure of (21) into the underlying structure of (22):

(21) I ate fish, Bill ate rice, and Harry ate roast beef.

(22) I ate fish, Bill rice, and Harry roast beef.

Ross' rule of Gapping in English affects only clause-internal and right-peripheral ellipsis. One application of Gapping does both deletions in (23).

(23) Betsy talked to Peter on Sunday, and Alan  $\emptyset_1$  to Sandy  $\emptyset_2$ .

$\emptyset_1$  = talked

$\emptyset_2$  = on Sunday

Left-peripheral ellipsis results from some other rule such as Conjunction Reduction.

(24) John talked to Bill and  $\emptyset$  seemed upset.

$\emptyset$  = he = John

According to Ross (1967b) two types of Gapping exist in natural language. If identical elements are on left branches of a structure, forward gapping occurs. If identical elements are on right branches of a structure, backward Gapping occurs.<sup>1</sup> (The latter was shown to be unacceptable in English in an earlier example.)

Ross considered Gapping a universal rule. He said that it could be stated in a general form in the theory of language. However, he considered that it could be used in any grammar in some form. Ross said that particular languages might impose their own conditions on the rule of Gapping. He stated that English has only forward Gapping. Therefore, in English, the general, universal rule of Gapping has a condition prohibiting backward Gapping, (25) is the same as an earlier example (7).

(25) \*John novels, and Mary read short stories.

If we look at Ross' approach in terms of the later distinctions between core and periphery, we can relate it to the ideas of both Chomsky and Koster that a rule can be stated in a general way in the core and have a particular application in a language in the periphery. The idea of auxiliary hypotheses presented by

Koster and seen in Chapter II could be applied here. The universal, core rule of Gapping includes two kinds of Gapping: forward and backward; or alternatively, Gapping is formulated in a very general way in the core without restriction as to its directionality. This is comparable to Move  $\alpha$  in Chomsky's approach to rules of movement. An auxiliary hypothesis for English makes the condition that only forward Gapping can occur in English. This is a device that restricts the core but does not contradict it.

### III.3B. Jackendoff

Jackendoff (1971) discusses his rule of Gapping and some related rules. He presents his view of Gapping within the  $\bar{X}$ -notation. Jackendoff's rule is different from the rule of Conjunction Reduction. It deletes the verb of one or more clauses conjoined to the right of a clause containing the same verb.

(26) Max ate the apple and Sally the hamburgers.

Jackendoff recognizes two types of conditions on Gapping. There are conditions concerned with the connection between the clauses and conditions concerned with the structure and deletions in the second clause. The conditions dealing with the connection between the clauses cover restrictions already discussed. Among these are the restrictions on but as a coordinating conjunction in a gapped sentence, and the restrictions on subordinate conjunctions.

The conditions on deletion restrict Gapping with unlike auxiliaries.

- (27) \*John has written the words, and Paul  $\left\{ \begin{array}{l} \text{will} \\ \text{is} \end{array} \right\}$  the music.

Gapping does not occur with unlike adverbs preceding the verb, as seen previously and in (28).

- (28) \*Max sometimes beats his wife and Ted frequently his dog.

As mentioned earlier, Gapping cannot occur when there is a negation in the auxiliary.

Jackendoff maintains that Gapping requires an unlike constituent in the VP. It is not simply a deletion of a variable. He gives cases of Gapping with one unlike constituent in the VP. Identical NPs or PPs occurring next to the verb can gap if they are followed by another constituent.

- (29) John writes poetry in the garden and Max in the bathroom.

Gapping occurs if the complement is an infinitive only.

- (30) Bob tried to wash himself and Mary to read the funnies.

Gapping deletes identical NPs, PPs, clauses and infinitives that are not adjacent to the verb, if there is contrastive stress in the first clause.

(31) Jack begged Elsie to get married and Wilfrid, Phoebe.

As Ross also pointed out, any number of embedded infinitives can delete if one unlike constituent remains.

(32) Max seemed to be trying to begin to love Harriet, and Fred (((to be trying) to begin) to love) Sue.

In the case of Gapping with two unlike constituents, acceptability of the output varies. With two NPs the results are unacceptable.

(33) \*Millie will send the president an obscene telegram, and Paul, the Queen, a pregnant duck.

The results are also unacceptable when the constituents are an NP and a clause or an infinitive.

(34) \*Frank forced Tom to shave himself, and Sam, Harry, to watch.

With a complement NP-PP, the results are questionable.

(35) ?Willie put the flowers in the vase, and Charlie the book on the table.

(36) ?Charlie entered the bedroom at 5:30 and Vera the kitchen at 6:00.

Sentences such as (35) and (36) were mentioned previously as being controversial due to the three remnants in the gapped clause. Jackendoff prefers (36) in which the PP is not strictly subcategorized by the verb.

Jackendoff presents a hypothesis regarding the use of syntactic features to express generalizations in Gapping and Deletion rules within the  $\bar{X}$ -theory. He says these generalizations are not easily expressed by traditional syntactic formalism. He relates Gapping to  $\bar{N}$ -Gapping, and VP-Deletion to  $\bar{N}$ -Deletion. The first two collapse to  $\bar{X}$ -Gapping, the second to  $\bar{X}$ -Deletion.

It is clear that Jackendoff's approach to the rule of Gapping is within sentence grammar. He applies the X-bar theory in an attempt to express generalizations of the rule. It is difficult to classify his approach within sentence grammar as being in favour of Gapping as a core rule or a peripheral rule, but it seems to me that his attempt to collapse the rule into  $\bar{X}$ -gapping could be interpreted as identifying the core-aspect of the rules, with the restrictions being the periphery for English.

## III.3C Koutsoudas

Koutsoudas (1971) argues against an independent rule of Gapping. He collapses the rules of Gapping and Conjunction Reduction into one rule, Coordinate Deletion. He says both Gapping and Conjunction Reduction perform the same operations and meet the same conditions. They delete identical elements, have the same directionality, are both last cyclic, and have the same place in the rule sequence. The only difference between the two seen by Koutsoudas is that Gapping deletes only verbs with no regrouping of constituents, and Conjunction Reduction deletes any grammatical category with a regrouping of constituents.

Koutsoudas explains the collapsing of the two rules as follows. Gapping is collapsed with the rule for object-reduced coordinations. The latter has the same directionality constraint as Gapping regarding identical elements on the right (forward Gapping) and identical elements on the left (backward Gapping). In object-reduced coordinations the constraint explains how (37) can be converted to (37a) but not (37b).

(37) John chased Mary and Pete caught Mary.

(37a) John chased and Pete caught Mary.

(37b) \*John chased Mary and Pete caught.

Koutsoudas collapses Gapping and the rule for object-reduced coordinations into the rule of VP Reduction. This optional rule states:

- (38) Given a coordination in which each conjunct includes a verb or object which is identical to the verb or object of each other conjunct, all but one of these identical constituents may be deleted, the undeleted constituent being that of the first conjunct if it is a left-branching constituent, and that of the last conjunct if it is a right-branching constituent.

The result is that Gapping is no longer an independent rule in any grammar. Koutsoudas provides three arguments against an independent rule of Gapping in favour of VP-Reduction. The directionality constraint operates in all languages in which objects can be reduced. In a well-defined position it can be shown that one language deletes a verb, another an object, and vice versa. For example, in English middle position, a verb is deleted; in Japanese middle position, an object is deleted. Rule ordering does not separate the two rules. Furthermore, Koutsoudas claims that from the evidence of thirty-two languages it can be shown that Gapping and Object-Reduction are either both present in a language, or neither is present. Therefore, Gapping is not an independent rule.

Koutsoudas lets VP-Reduction operate independently of grammatical categories, as it is subject to heavy restrictions. The result is the Coordinate Deletion rule.

## (39) Coordinate Deletion

$$[X_1 W [AX_2] X_3] [X_4 W [AX_5] X_6]$$

$$1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8$$

$$1 \quad 2 \quad 3 \quad 4 \quad \cancel{5} \quad \emptyset \quad 7 \quad 8$$


Conditions: 1. A is an immediate constituent of W.

2.  $X_2 \neq \emptyset$  and  $X_5 \neq \emptyset$

It operates to derive (41) from (40)

(40) John saw the dog and saw the cat.

(41) John saw the dog and the cat.

Koutsoudas criticizes Ross' use of the rules of Gapping and Conjunction Reduction as separate rules in a grammar. He says Conjunction Reduction must account for object-reduced coordinations as in (42) and (42a).

(42) John washed the dishes and Mary dried the dishes.

(42a) John washed and Mary dried the dishes.

If Gapping accounts for verb-reduced coordinations and Conjunction Reduction for object-reduced coordinations, with two rules we cannot account for the fact that in any language either both objects and verbs reduce or neither does. The two rules are formally and substantively related, with the same constraint on directionality. It is unnecessary, according to Koutsoudas, to have two rules obeying the same constraint which both specify identity deletion.

Notice that this approach is similar to Chomsky's attempt to reduce many movement rules to Move  $\alpha$  in the core, identifying the general symptomatic properties. Presumably there is only one rule in the core, and peripheral conditions in relation to some structures create an appearance of two or more rules.

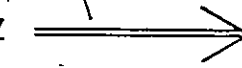
From the standpoint of present approaches to sentence and discourse grammar, it could be said that Koutsoudas, like Jackendoff, would place the Gapping rule within sentence grammar. Again, it is difficult to make a further classification as to core or periphery. The idea of collapsing rules to allow for one rather than two would be consistent with recent trends in sentence grammar. (The question of whether the analysis provided by Koutsoudas is correct is not a concern of this paper.)

## III.3D Hankamer

Hankamer (1973) presents an approach to Gapping consistent with the generative semantics framework. He establishes a transderivational constraint to generalize the restrictions on Gapping.

Hankamer says Gapping deletes rightmost occurrences of identical verbs in conjoined sentences, but it may delete more than the verb. He considers left-peripheral ellipsis to be part of Gapping, but does not provide any mechanism for collapsing the two rules. In Hankamer's view of Gapping, the deleted material may be on both sides of the non-deleted NP. His formulation of the rule showing the Gapping operation as deletion of a variable is as follows:

(43) Gapping

NP X A Z and NP X B Z 

NP X A Z and NP B

where A and B are non-identical major constituents.

Hankamer suggests a constraint to generalize the restrictions on Gapping. This No Ambiguity Condition (NAC) is a transderivational constraint which accounts for all the conditions which

operate to prevent Gapping from producing ambiguity. It is stated in (44).

- (44) 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 NAC makes the following predictions.

1. Any application of Gapping resulting in configuration-NP NP, or NP X on the right is blocked when there is a like configuration in the VP of the left conjunct.

- (45)a. Jack calls Joe Mike and Sam Harry.  
 b. Jack calls Joe Mike and [Jack calls] Sam Harry.  
 c. Jack calls Joe Mike and Sam \*[calls Joe] Harry.

2. Gapping in (46b) is possible but in (46a) the internal Gapping is prohibited. The output of (a) is the same as the output of peripheral Gapping.

- (46)a. Max gave Sally a nickel yesterday, and\*[Max gave] Susan [a nickel] last week.  
 b. Max gave Sally a nickel yesterday and [Max gave Sally] a dime last week.

There are several exceptions to the NAC which are discussed by Hankamer. (47) is from Jackendoff.

- (47) Massachusetts elected McCormack Congressman, and Pennsylvania, Schweiker.

Hankamer considers this to be a counterexample only if "elected McCormack Congressman" is VP NP NP. He considers "Congressman" a titular NP and argues that, as such, it is not an NP.

A second counterexample requires a modification of the NAC, according to Hankamer. It involves (48), (49), (50), which vary in acceptability.

- (48) ?Max writes plays in the bedroom, and Harvey in the basement.
- (49) ??Max is writing a play in the bedroom, and Harvey in the basement.
- (50) ???Max is memorizing the play in the bedroom, and Harvey in the basement.

The NAC must be modified to distinguish constituents required by the verb, and constituents which are optional adjuncts to the VP. In (48), Hankamer says grammaticality depends on the fact that plays and writes plays is generic. The result in (49) is worse due to the presence of a definite direct object.

A third counterexample is seen in cases like (51).

- (51) Paul Schachter has informed me that the basic order in Tagalog and related languages is VOS; Ives Goddard that the unmarked order in Algonkian is OVS; and Guy Carden that the basic order in Aleut is OSV.

As far as the NAC is concerned, a pronoun contiguous to a verb does not count as an NP. The exception would cease to be a problem if English had a rule for cliticization of pronouns in immediate post-verbal position. Hankamer gives evidence in favour of such a rule. He maintains that Gapping must distinguish between (51) and (52).

- (52) Paul Schachter has informed Haj Ross that the basic order in Tagalog and related languages is VOS; Ives Goddard that the unmarked order in Algonkian is OVS; and Guy Carden that the basic order in Aleut is OSV.

Hankamer further states that the exceptions noted do not destroy the value of the NAC. It captures a generalization which a multicondition approach would not. The NAC on Gapping is intended as a universal constraint to reflect a general restriction on deletions.

In addition to the NAC, Hankamer posits the Structural Recoverability Hypothesis.

- (53) Deletion rules involving variables are univ<sup>er</sup>sally subject to a transderivational condition which prevents them from applying in such a way as to introduce structural ambiguity.

This is a universal constraint of which the NAC is a particular case. The Peripheral Gap Principle, a further hypothesis of Hankamer, is stated in (54).

- (54) If any interpretation is possible for an unacceptably ambiguous structure, it will be that interpretation under which the location of the deletion site is peripheral rather than internal.

Hankamer discusses numerous examples and counterexamples to these principles, most of which do not involve the rule of Gapping with which I am concerned. I will not include them here.

I will not try to classify Hankamer's work on Gapping along the lines of the distinctions now recognized between sentence and discourse grammar. His approach does not fit the formal grammar as seen by Chomsky. Within the framework of generative semantics, it is difficult to separate the different levels of grammar. If we were to translate the approach of Hankamer into the revised extended standard theory, it would be completely outside the realm of sentence grammar. A transderivational constraint such as the one suggested by Hankamer would be inconsistent with present theories.

A rule does not have the power to look ahead to see which derivation will produce a non-ambiguous output. Sag (1977) points out that many of the examples provided by Hankamer are very controversial. He cites (55) as an acceptable sentence in Quirk et al. (1972) which violates the NAC.

- (55) Peter is playing football for his school and Paul for his club.

Hankamer does, however, raise many problems for which other linguists have attempted to provide answers in their treatments of Gapping.

### III. 3E Langendoen

Langendoen (1974),<sup>2</sup> working within the standard theory, maintains that the constraint on Gapping is grammatical, and not transderivational, as suggested by Hankamer. His constraint is termed the Non-Left Peripheral Noun Phrase Constraint (NLPNPC). This is an attempt to "translate" a condition which is outside the standard theory into a formal constraint within the limits of the model by eliminating its transderivational nature.

- (56) No NP can be deleted by Gapping unless it is either left-peripheral or involved in a continuous left-peripheral gap.

In Langendoen's view, the NAC does not account for (57) and (58).

(57) Max sent Sally the messenger last week, and Susan yesterday.

(58) Max wanted Ted to persuade Alex to see Mary, and Walt, Ira.

The NLPNPC is intended to account for the above, as well as for the data outlined in Hankamer. He cites (59) as an example that the NLPNPC is independently motivated.

(59) \*Mary sang the Bach cantata, and Sam played  $\emptyset$ .

Langendoen collapses Gapping and Conjunction Reduction with the formulation in (60).

(60) Conjunction Reduction

$X_1 - Z_1 - X_2 - Z_2 - X_3 - (\text{and} - X_1 - Z_1^1 - X_2 - Z_2^1 - X_3)^*$

1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 - 11

opt.  $\longrightarrow$

(A) 1 - 2 - 3 - 4 - 5 - 6 -  $\emptyset$  - 8 -  $\emptyset$  - 10 -  $\emptyset$

(B) 1 - ~~2~~~~6~~~~8~~ - 3 - 4  $\emptyset$  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; and 4  $\neq$  10

2. A is precluded if:

(a) 4, 5 =  $\emptyset$  and 2 is not VP, or

(b) 3 or 5 contain an NP other than a clitic pronoun adjoined to its V.

Langendoen's remarks indicate that Gapping would be a rule of sentence grammar, but it is a rule with a large number of specific conditions that make it look suspicious and highly marked. His approach has certain problems and Sag points out that the NLPNPC leaves some of the data unaccounted for. In this way, his approach is not necessarily superior to the others which have been discussed here. Sag cites the following examples as unexplained by Langendoen's proposal. Many of these examples are problematic for most approaches, and some have been mentioned previously in this paper.

(61) Max writes plays in the bedroom, and Harvey in the basement.

(62) Joan will cook the meals today, and Barbara tomorrow.

(63) Peter is playing football for his school and Paul for his club.

(64) At Betsy's house we play bridge, and at our house poker.

(65) John took Harry to the movies, and Bill, Mike.

### III.3F Stillings

Stillings' (1975) work on Gapping represents another attempt to characterize the constraints on Gapping. She formulates a rule to accompany her conclusion that direct object nouns are not deleted along with the verbs that precede them. Her rule restricts the deletion target of Gapping.

Prior to a formulation of her rule, Stillings outlines the characteristics of Gapping as she sees them.

a. Gapping derives (66) from (67).

(66) Ornette plays the plastic alto and Roscoe, the harmonica.

(67) Ornette plays the plastic alto and Roscoe plays the harmonica.

b. Gapping does not delete noncontiguous verbs.

(68) \*Oscar clearly intended to tell Sally to sing vibrato, and Sasha, Wilfred.

c. Gapping does not delete over a variable in the right conjunct.

(69) \*Gesualdo wrote madrigals to assuage his guilty conscience, and Ives, wrote symphonies, his hatred of Central Park.

d. The variable to the right of the deletion site must represent a single constituent.

(70) \*Terry thinks poorly of Elliott Carter and Emmon, well of Flatt and Scruggs. (\*ADV PP)

To account for the above, Stillings includes a variable type C which ranges over single constituents. C is any single non-terminal node NP, PP, ADV, VP, S.

e. The rightmost variables in the left and right conjuncts must represent the same constituent type.

(71) \*Fred eats at Arby's and my brother-in-law, pickled beets.

f. Gapping can delete a string of verbs.

(72) Kipnis tends to overemphasize ornamentation and Leonardt, heavy-handedness.

Example (73) is Stillings' formulation of Gapping.

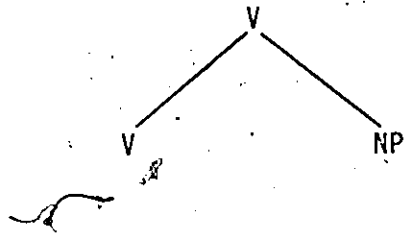


Here Stillings refers to Hankamer as saying that preverbal adverbs and postverbal pronouns are cliticized to the verb, so this does not dispute her view that only verbs are gapped.

Sentence (77) is a second counterexample.

- (77) The FBI kept tabs on Calvin Coolidge and the IRS, on Hamilton.

Stillings claims this is not a counterexample to her view that direct object NPs cannot be gapped. She regards keep tabs as a complex verb, with the NP falling within the range of V.



Stillings briefly mentions other factors which may interact with Gapping. Bresnan's rule of Truncation may sometimes apply with Gapping to account for two gaps. A semantic restriction may apply when two deleted variables are so closely related semantically that one cannot be deleted without the other. Finally, there may be a focus stress condition necessary for double Gapping.

Stillings' approach is interesting when we consider the core/peripheral distinctions made recently. The brief remarks she

makes at the end of her article would suggest that there are perhaps factors lying outside the normal domain of sentence grammar which may interact with the rule of Gapping. She considers semantic factors and focus stress, both of which lie outside the core and possibly outside the periphery of sentence grammar. While the bulk of her article deals with the rule of Gapping that we could place within sentence grammar, the extra factors she mentions make her approach difficult to classify. Her consideration of these other factors is interesting to keep in mind in reviewing the work of Kuno, in the next section.

### III.36 Kuno

Kuno (1976) has presented a functional approach to the problems of Gapping. His discourse-oriented approach is the most extreme one among those who regard Gapping as a phenomenon lying outside the formal component of sentence grammar. Kuno sees the Gapping operation as conditioned by the interaction of external factors — perceptual, discourse-based, and semantic — with syntax.

Kuno lists counterexamples to the NAC proposed by Hankamer and to the NLPNPC suggested by Langendoen. All these examples have something other than proper names as the remnants in the gapped clauses.

- (78) Some people live in this city because they like living here, and others because they don't have the means to move to the suburbs.
- (79) One of the muggers hit Mary with a baseball bat, and another with a bicycle chain.
- (80) Some people go to Europe every year, and others every other year.
- (81) 50% of the constituents asked the Senator to vote for the bill, and 25% to vote against it.
- (82) Two days ago John took Mary out to dinner, and this afternoon to the movies.

Kuno explains that the examples considered in previous works contained proper names and resulted in incorrect proposals. Proper names are "old information". In a gapped sentence, the remnants should be "new information". He presents two examples, within discourse, in which the NAC and NLPNPC are violated even when proper names appear as new information.

(83) Q.- Who persuaded who to examine Mary?

A.- John persuaded Dr. Thomas to examine Mary, and Bill

$\emptyset_1$  Dr. Jones  $\emptyset_2$ .

\*  $\emptyset_1$  = persuaded

$\emptyset_2$  = to examine Mary

(84) Q. With what did John and Bill hit Mary?

A. John hit Mary with a stick, and Bill  $\emptyset$  with a belt.

$\emptyset$  = hit Mary

Kuno argues in favour of a perceptual principle, his Minimal Distance Principle (MDP), as an explanation for some of the Gapping phenomena.

(85) Minimal Distance Principle

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.

There is a tendency in Gapping to couple the two constituents left as remnants with the rightmost constituents of the first conjunct. The principle, in interaction with three others, is intended to account for Gapping phenomena and the great variation in acceptability accorded to gapped sentences. The other three principles are:

a. The Functional Sentence Perspective Principle (FSPP).

This states that the closer a given constituent is to sentence-final position, the newer the information it presents.

Constituents left by Gapping are new information and are paired with constituents of the first conjunct that are also new information. The following sentence can then be regarded as unambiguous.

- (86) John persuaded Bill to examine Jane, and Tom, Martha.  
Tom and Martha, as new information, are paired with Jane and Bill, the two constituents on the right that represent new information.

b. The Tendency for Subject-Predicate Interpretation.

The FSPP does not explain the unacceptability of (87), (88), and (89).

- (87) \*Jack asked Mike to wash himself, and Sue to shave himself.

- (88) \*Jack asked Mike to wash himself and [Jack asked] Sue to shave himself.

- (89) Jack asked Mike to wash himself and Sue \*[asked Mike] to shave himself.

The Tendency for Subject-Predicate Interpretation is a semantic condition that says when Gapping leaves an NP and a VP, these constituents are interpreted as a sentence pattern. The NP is considered the subject of the VP, as in (90).

- (90) John entered Harvard to study linguistics, and Bill to study psychology.

According to the principle, Bill and to study psychology are interpreted as a sentence unit.

C. The Requirement for a Simplex-Sentential Relationship.

Kuno's third principle states that the two constituents left by Gapping have a simplex-sentential relationship. Without such a relationship, gapped sentences are much less clear.

(93) is preferable to (91) or (92), according to Kuno.

(91) Who persuaded who to examine Mary?

John persuaded Dr. Thomas to examine Mary, and Bill,  
Dr. Jones.

(92) Who persuaded Dr. Thomas to examine who?

John persuaded Dr. Thomas to examine Jane, and Bill,  
Martha.

(93) Who promised Bill to examine who?

Dr. Thomas promised Bill to examine Jane, and Dr. Jones,  
Martha.

Kuno's examples here are unclear. It is questionable to me whether (93) is to be preferred over (91) and (92).

Finally, Kuno argues in favour of the fact that the acceptability of gapped strings may depend on semantic unity. There is a difference in the acceptability of (94) and (95).

(94) Max writes plays in the bedroom, and ? Harvey in the  
basement.

- (95) Max wanted to put eggplant on the table, and Harvey  
 \*[wanted to put eggplant] in the sink.

Kuno suggests that there is semantic unity in writes plays, making it more acceptable than wanted to put eggplant, which is not a semantic unit. This example is discussed by almost everyone who deals with Gapping. Jackendoff has claimed that identical NPs or PPs next to the verb can gap if they are followed by another constituent. For him, both (94) and (95) would be acceptable:

Hankamer claims that writes plays is generic, and therefore has a greater degree of acceptability than something which is not generic. This approaches the idea of Kuno regarding semantic unity. Stillings is probably suggesting the same thing when she mentions that some deleted variables are so closely related that one cannot be deleted without the other.

Kuno has presented five nonsyntactic factors affecting the acceptability of gapped sentences. The Minimal Distance Principle represents perceptual factors, and the Functional Sentence Perspective is discourse-based. Two factors, the Tendency for Subject-Predicate Interpretation and the Requirement for a Simplex Sentential Relationship are both semantic and perceptual in nature. The requirement for semantic unity of gapped strings is a semantic factor. According to Kuno, all of these support the view that a functional approach to Gapping is necessary. It is difficult to assess some of the claims made by Kuno because of the informal

nature of the discussion of the data. Also, many of his examples are very controversial, and the acceptability judgments vary so much that his claims are weakened. His approach, however, definitely places Gapping outside the domain of sentence grammar as seen by Chomsky and Koster in recent works. His functional approach is inconsistent with the framework of the extended standard theory I have adopted in this paper.

### III.3H Sag

I will briefly outline the proposals made by Sag (1977) for the rule of Gapping. His general approach to grammar was discussed in Chapter I, and his rule formulation will be presented later. Sag's treatment of Gapping is somewhat unclear; however, his proposal can be reinterpreted along the lines of recent proposals in the extended standard theory. Rules are simplified, and they overgenerate. An example of such rule simplification is the rule to move a category  $\alpha$ , that is part of Chomsky's theory of sentence grammar. The overgeneration of rules is corrected by independent constraints. Subjacency is one such constraint.

Sag suggests that Gapping is a rule of both sentence and discourse grammar. His hypothesis regarding Gapping and logical form takes into account the ideas of Kuno about the effect of perceptual factors

on the Gapping rule. He says his rule will overgenerate because the perceptual factors proposed by Kuno will affect the acceptability of Gapped clauses. So in this case, non-formal factors compensate for the overgeneration of the rule.

Sag presents a formulation of the rule of Gapping which will be discussed later in the chapter. The overgeneration of the rule is limited by a surface constraint, which will also be presented later on.

The problem with such a proposal is to determine whether the principles that correct the effects of massive Gapping are conditions of sentence grammar or not. Looking at Sag's treatment within the sentence/discourse distinction, it would seem that the rule is inside sentence grammar, while the principles correcting its overgeneration are outside sentence grammar. It then becomes almost impossible to separate sentence grammar from non-sentence grammar, because every contradiction to sentence grammar could be corrected by an external principle.

Sag appears to fall somewhere between the extreme informalism of Kuno and the strict formalism of Koster, in terms of his approach to Gapping.

## III.3I Williams

Among the last group of linguists, those who place Gapping within the bounds of sentence grammar, is Williams (1977). His distinction between discourse and sentence grammar has already been discussed. According to Williams, Gapping differs from VP Deletion. He points out that VP Deletion does not obey the Coordinate Structure Constraint and the Complex Noun Phrase Constraint of Ross (1967a). In this way it differs from Gapping, which does obey these constraints. He therefore considers VP Deletion to be a rule of discourse grammar and Gapping to be a rule of sentence grammar. The two rules also differ in the fact that VP Deletion can apply across sentence boundaries, while Gapping cannot. In VP Deletion, the antecedent of a deleted verb can be in a sentence other than the one containing the missing VP. In an utterance in which Gapping has applied, this is not possible. (96) shows an application of VPD across sentence boundaries. (97) shows that this is impossible in Gapping.

(96) A. Did Mary sing?

B. Yes, she did.

(97) A. Did Mary sing in the house?

B. \*No, Susan in the garage.

Because Gapping obeys the constraints mentioned above, and because it must apply within the boundaries of the sentence, Williams places it within the domain of sentence grammar.

### III.3J Koster

Koster (1978) represents the extreme formalist point of view with regard to Gapping. His proposals are in sharp contrast to those of Kuno, the most extreme of the discourse-oriented proponents. Koster considers Gapping to be strictly a core rule of sentence grammar. It is not a deletion rule, but a rule of interpreting  $\sqrt{[e]}$ . It is another instance of Koster's coindexing rule of core grammar presented in Chapter II, because it obeys the Bounding Condition and the Locality Principle.

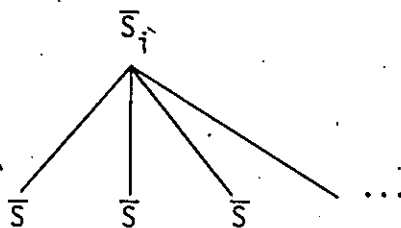
In Chapter II, I outlined Koster's Bounding Condition, which states that an empty node must have an antecedent within the phrase in which it is a constituent. Gapping is constrained by the Bounding Condition. To explain this, Koster reinterprets the Bounding Condition to state:

(98) The Bounding Condition  
 $\gamma$  cannot be free in  $\beta$  in:

[  $\beta$  ... [ $\gamma$   $\underline{e}$  ] ... ]  
 where  $\beta$  is a top node.

A top node, in Koster's reinterpretation, cannot be dominated by a node of the same type with the same number of bars.

(99)



$\bar{S}_i$  is a top node in (99).

The Bounding Condition now explains the contrast between (100) and (101).

(100)  $[\bar{S}_i [ \bar{S} \text{ John hit Mary} ] \text{ and } [ \bar{S} \text{ Bill } [_{Ve} ] \text{ Sue } ] ]$ .

John hit Mary and Bill Sue.

(101)  $*[ \bar{S}_i [ \bar{S} \text{ John hit Mary} ] \text{ and } [ \bar{S} \text{ I don't believe } [ \bar{S} \text{ Bill } [_{Ve} ] \text{ Sue} ] ] ]$ .

\*John hit Mary and I don't believe Bill Sue.

In (100) the Bounding Condition does not apply. There is an empty V in an  $\bar{S}$ , but the  $\bar{S}$  is not a top node as it is immediately dominated by  $\bar{S}_i$ . The antecedent of  $[_{Ve}]$  is hit. Example (101) is ungrammatical because  $[_{Ve}]$  is free in an  $\bar{S}$  that is a top node. Thus the

Bounding Condition predicts that Gapping does not occur in embedded VPs.

Koster's Locality Principle states that a given consequent selects the closest antecedent in terms of phrase structure. (This is related to the Minimal Distance Principle<sup>3</sup>).

The Cojacency Principle outlined by Koster is one instance of the Locality Principle. The following example of Gapping is subject to the Cojacency Principle.

(102) Peter read novels, Mary wrote books, and John  $\emptyset$  papers.

In (102) the gap must be interpreted as wrote, not read, according to the Cojacency Principle or the Locality Principle.

Koster also presents arguments to distinguish between bounded Gapping as a rule of interpretation of [Ve] and an instance of the core rule of coindexing, and free deletion, a peripheral process. Gapping, unlike processes of free deletion, such as VP Deletion, is governed by the Bounding Condition. The following contrasts are presented by Koster as evidence that Gapping is different from VP Deletion, because it obeys the Bounding Condition.

(103) \*John hit Mary after Bill Sue.

(104) John hit Mary after Bill did.

Example (103) is ungrammatical because Gapping does not go into PPs. The Bounding Condition prevents this, because Gapping in this case would leave a free  $[_{V_e}]$  within the PP. The same does not apply to (104), an example of VP Deletion. In (106), VPD goes into complex NPs or islands, but Gapping (105) does not as it would leave a free  $[_{V_e}]$  under a top node.

(105) \*John didn't hit Mary but I know somebody who Sue.

(106) John didn't hit Mary but I know somebody who did.

Koster concludes that two classes of rules exist. One class consists of interpretive rules, like Gapping, which obey the Bounding Condition and link antecedents to anaphors, including null anaphors such as  $[_{V_e}]$ . The second class is that of free deletions, which are not subject to the Bounding Condition. In other terms, Gapping cannot delete as much as VPD. In Koster's terms, the Locality Principle governs both classes, and he considers bound-  
edness a core grammar property, and locality a more general property.

The work of Koster stands in sharp contrast to that of Kuno. Koster considers Gapping to be a rule of core grammar, within the domain of sentence grammar. Kuno sees it as a rule lying outside sentence grammar and affected by perceptual factors.

From Koster's work it is unclear whether all aspects of the Gapping rule meet his constraints. The problematic examples of Gapping which appear throughout the literature are not mentioned in Koster. He has included an idea regarding auxiliary hypotheses for language - specific cases of a rule, but makes no mention of their possible use for Gapping.

One sentence which appears several times in the literature is (105). The problem of its acceptability was discussed earlier.

(105) Max writes plays in the bedroom and Harvey in the basement.

Hankamer considers the sentence to be of questionable acceptability, but better than (106) or (107).

(106) ??Max is writing a play in the bedroom, and Harvey in the basement,

(107) ???Max is memorizing a play in the bedroom, and Harvey in the basement.

He considers the differences to be due to the fact that writes plays is generic.

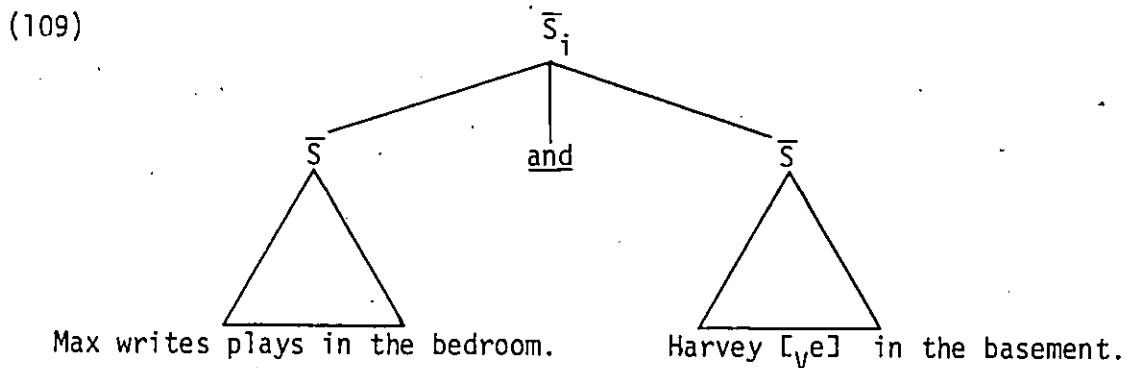
Kuno considers (105) more acceptable than (108).

- (108) Max wanted to put eggplant on the table, and Harvey  
\*[ wanted to put eggplant ] in the sink.

He claims that there is semantic unity in writes plays, while there is none in wanted to put eggplant. Stillings' idea that some deleted variables are so closely related that they must be deleted together seems to be the same as Kuno's approach to (105).

Jackendoff's claim that identical NPs or PPs next to the verb can gap, indicates that he would perhaps find (105) and (108) acceptable.

Koster does not make any statement about such a sentence. It is possible to attempt an application of the Bounding Condition on sentence (105).



In (109),  $S_i$  is the top node, and  $V$  is not free in  $S_i$  because it has the antecedent writes plays in the other  $S$ .

The problem with such an analysis is that writes plays is not an independent verb in isolation (a small V), but a  $\bar{V}$  or  $\bar{\bar{V}}$  (a bigger V). Nothing in Koster's discussion clarifies how high up the verbal structure the Gapping rule can apply. If V is a verb with no complements, the Bounding Condition is no longer applicable. At the same time if Gapping and VP Deletion are to be kept separate, it cannot be the case that gapping can delete the maximal projection of V (that is, the VP).

Another controversial example is sentence (110), which has been mentioned before.

(110) Max wanted Ted to persuade Alex to see Mary, and Walt, Ira.

Example (110) is a difficult sentence to deal with. The most probable interpretation is one which gives the reading in (111).

(111) Max wanted Ted to persuade Alex to see Mary, and Max wanted Ted to persuade Walt to see Ira.

However, if (112) and (113) are considered as possible readings of (110), a problem is created for Koster's analysis, as longer strings are gapped.

(112) Max wanted Ted to persuade Alex to see Mary, and Walt wanted Ted to persuade Alex to see Ira.

(113) Max wanted Ted to persuade Alex to see Mary, and Max wanted Walt to persuade Alex to see Ira.

It is possible most speakers would rule out (112) and (113) as too difficult to process. Kuno presents a similar sentence, (114), which he claims is accounted for by a perceptual principle involving functional sentence perspective, a discourse-based, perceptual factor.

(114) John persuaded Bill to examine Jane, and Tom, Martha.

Example (114) becomes problematic for Koster's analysis if both (115) and (116) are acceptable readings.

(115) John persuaded Bill to examine Jane, and John persuaded Tom to examine Martha.

(116) John persuaded Bill to examine Jane and Tom persuaded Bill to examine Martha.

Example (116), if acceptable, appears to be an example of VP Deletion. In terms of Koster's analysis, the problems of application of the Bounding Condition are compounded because  $v[E]$  has, as antecedent, longer strings.

Koster's treatment of Gapping does not include a discussion of previous proposals. In addition, his definition of Gapping is narrow, as it is a rule applying only to simple verbs. He does not include any approach to the gapping of larger Vs.

### III.4 A Proposal

From the range of approaches to Gapping that have been presented, it is obviously very difficult to select one that can handle all the controversial elements of the rule that have been noted here. The treatments reviewed in this paper cover every possibility, from Koster's core rule of sentence grammar, to Kuno's perceptually governed rule which lies outside of sentence grammar.

I would like to present a view of Gapping which, while retaining it within sentence grammar, allows it to fit into both the core and the periphery. The idea that perceptual factors influence Gapping may be relevant at another level outside sentence grammar.

Kuno has suggested that Gapping is a rule lying outside of sentence grammar, which would place it, in the framework used in this thesis, within discourse grammar. However, Williams' test for discourse can be applied to Kuno's sentences, giving an opposite result.

Williams uses the following example as an indication that Gapping is not the same as VPD and cannot operate in discourse.

(117) Q.- Did Sam go to the store?

A.- \*No, Bill to the supermarket.

The same test can be applied to Kuno's data, giving a similar result. Kuno says (118) is an example of Gapping governed by the Functional Sentence Perspective Principle, a discourse-based principle.

(118) John persuaded Bill to examine Jane, and Tom, Martha.

Williams' test gives the results in (119).

(119) Q.- Did John persuade Bill to examine Jane?

A.- \*No, Tom, Martha.

From this it is clear that Kuno's example cannot be part of discourse. The operation produces an ungrammatical result when applied across sentence boundaries. Other examples taken from Kuno's data produce the same result. Gapping, then, is a sentence grammar rule. This does not exclude the possibility that Gapping, as a rule of sentence grammar, can be influenced by external factors that are formal or perceptual in nature.

A second distinction that can be made between Gapping and rules of discourse such as VPD has been made by Koster. He has shown that Gapping is governed by the Bounding Condition and is

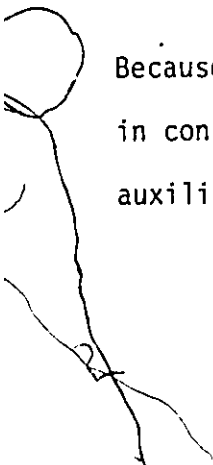
therefore a rule of sentence grammar. VPD is not governed by the Bounding Condition. The two rules, therefore, cannot be collapsed.

A serious problem in dealing with Gapping is to decide what is the projection of V that can be gapped. Sag and Koster both maintain that the rule deletes only simple verbs and their auxiliaries. There are many controversial sentences, some of which are mentioned above, which have varying degrees of acceptability, and which delete more than a simple verb. The degrees of acceptability in such examples make it difficult to make any statement concerning what Gapping can delete in contradiction to Sag and Koster. For example, (120), in which an NP is deleted with the verb, may be acceptable to some people, but not to others.

(120) ?John ate peaches with a fork, and Susan with a spoon.

A similar sentence, (121), seems to me to be less acceptable, although again, some speakers might consider it acceptable.

(121) ??John ate three peaches with a fork, and Susan with a spoon.



Because of examples such as these, I will follow Sag and Koster in considering Gapping as a rule deleting only simple verbs and auxiliaries.

The approach to Gapping presented by Sag is an attempt to take into consideration the great variation in judgment accorded to gapped sentences. His rule generates sentences whose acceptability differs among speakers. The variations are handled by surface constraints which differ slightly.

Sag's approach is interesting when viewed in conjunction with recent approaches to the core and the periphery of sentence grammar. Sag provides a formulation of the Gapping rule as in (122).

(122) Gapping

$$\begin{array}{cccccccccccc}
 W_3 - [S X^2 - W_1 - [X^2]^* - W_2] - \left\{ \begin{array}{l} \text{and} \\ \text{or} \end{array} \right\} - [S X^2 - W_1 - [X^2]^* - W_2] - W_4 \\
 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8' \quad 9 \quad 10 \quad 11 \\
 \implies 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad \emptyset \quad 9 \quad \emptyset \quad 11
 \end{array}$$

The rule overgenerates, but much of the overgeneration is ruled out by a surface constraint of the type in (123).

(123)  $*[S X^2 [V] X^2 - C * ]]$

C\* is any non-null sequence of constituents (within  $V^1$  (VP)).

Sag's approach to Gapping can be seen within the perspective of present approaches to core and periphery. Hirschbühler and Rivero have suggested an interpretation of the core/periphery distinction which could apply to Gapping. They see a rule of sentence grammar as being an "aggregate of dimensions", some of which fall among the unmarked phenomena of core grammar, some of which fall among the marked phenomena of the periphery. They see the peripheral part of a rule as restricting, but not contradicting, the core aspects of a rule. The peripheral dimensions of a rule are restrictive in the sense that they allow only a subset of the dimensions which the core can generate. This approach differs from Koster's view of the core/periphery distinction in that Koster sees some of the peripheral dimensions in contradiction to the core dimensions of a rule. Hirschbühler and Rivero suggest that it is preferable to modify the core and thus avoid any contradiction in the periphery.

Adapting this approach to the distinction between core and periphery, the rule of Gapping can be regarded as being composed of some unmarked aspects which fit into the core, and some marked phenomena, which fall outside the core in the periphery. The core aspects could, for example, include the rule in a form such as Sag has formulated, with an overgeneration. The noncore, or peripheral aspects, would include the necessary constraints needed to restrict the output of the rule.

Such a proposal can be implemented if the core is seen to overgenerate as follows. The core gaps massively; that is, it gaps  $v$  [e], or any of its projections:  $\bar{v}$  [e],  $\bar{\bar{v}}$  [e], for example. The restrictions to this are then built into the periphery including perhaps the restriction of forward Gapping mentioned by Ross. Suppose Gapping is a rule in Universal Grammar. The core grammar of English selects this rule in its most general formulation much as it selects Move  $\alpha$ . The periphery restricts the core by imposing specific conditions such as: only forward Gapping can occur.

As a result, there is a rule of Gapping which is part of sentence grammar. It has two types of elements: a core rule, which overgenerates, as does Sag's rule, gapping  $v$  [e] and its projections but not the maximal projection, that is, the VP because otherwise VP Deletion would be identical to Gapping, which is not correct; the peripheral aspects of the rule are constraints which limit the overgeneration of the core to produce acceptable sentences. Any non-formal, perceptual constraints which might affect the rule would fall within a domain lying outside of sentence grammar. The peripheral constraints of the rule would be the subject of a further study, but their nature can be gathered from the restrictions mentioned by different authors. Among these might be included Ross' restriction on forward Gapping, or conditions such as those mentioned by Jackendoff regarding unlike constituents in the VP.

### III.5. Summary

In this chapter I have presented and commented on the various proposals that have been made concerning the rule of Gapping. These range from the early proposal of Ross, to recent attempts by Koester to justify the consideration of Gapping as a strictly core rule of sentence grammar. The proposals include such differing approaches as the generative semantics framework of Hankamer, and the perceptually based suggestions made by Kunö.

The chapter also includes a suggestion that Gapping can be considered a rule made up of some elements which belong in the core, and others which belong in the periphery. The rule over-generates in the same way that Sag's rule of Gapping does, and has elements at the two levels of sentence grammar, following the proposals of Hirschbühler and Rivero for rules of grammar in general. An enumeration and description of the aspects of the rule which lie in the periphery and constitute constraints limiting the overgeneration of the rule in the core would fall within the scope of a further study of Gapping; but these would probably include such things as restrictions mentioned in the literature regarding forward Gapping and conditions on constituents in the gapped clause.

## CHAPTER III

FOOTNOTES

1. Maling (1972) disagrees with Ross. She states that Gapping operates forward only. She also includes a separate rule of node-raising which always occurs with Gapping.
2. I am presenting Langendoen's treatment of Gapping as reflected in Sag (1977).
3. It is not to be confused with the MDP used by Kuno. The Minimal Distance Principle originated with Rosenbaum, P. (1967). The Grammar of English Predicate Complement Constructions. Cambridge, Mass: MIT Press. Kuno's principle, unlike Rosenbaum's, is not structural.

## C O N C L U S I O N

It has been shown that a distinction exists between sentence and discourse grammar. Certain rules of grammar which operate across the boundaries of the sentence are considered rules of discourse, while others operate within sentences and belong to sentence grammar. Furthermore, sentence grammar can be separated into the two areas of core and periphery. The distinctions that have been made between the core and the periphery represent attempts to outline a theory of markedness for syntax.

Of the several interpretations accorded to the core/peripheral distinction is one which regards a rule of grammar as a composite of both core and peripheral dimensions. The core aspects of a rule generate structures which are restricted by the peripheral aspects of the rule.

An examination of the rule of Gapping shows that this interpretation of the core/peripheral distinction explains how the rule can overgenerate in the core and encounter restrictions in the periphery. An outline of specific peripheral restrictions provides scope for further study of Gapping. In addition, the interpretation of the core and periphery discussed here should be of interest in the examination of other rules of grammar.

B I B L I O G R A P H Y

- Banfield, A. 1973. "Narrative Style and the Grammar of Direct and Indirect Speech". Foundations of Language. 10, 1-39.
- Chomsky, N. 1968. Language and Mind. New York: Harcourt, Brace and World.
- Chomsky, N. 1975. Reflections on Language. New York: Pantheon.
- Chomsky, N. 1977. "On WH-Movement". In P.W. Culicover, T. Wasow, A. Akmajian (eds.), Formal Syntax, New York: Academic Press.
- Chomsky, N. 1980. "On Binding". Linguistic Inquiry. 11, 1-46.
- Chomsky, N. and H. Lasnik: 1977. "Filters and Control". Linguistic Inquiry. 8, 425-504.
- Emonds, J. 1976. A Transformational Approach to English Syntax: Root, Structure-Preserving and Local Transformations. New York: Academic Press.
- Erteschik, N. 1973. "On the Nature of Island Constraints". MIT Doctoral Dissertation.
- Hankamer, J. 1973. "Unacceptable Ambiguity". Linguistic Inquiry. 4, 17-68.
- Hirschbühler, P. and M.L. Rivero. 1980. "Catalan Restrictive Relatives: Core and Periphery". To appear in Language.

- Hooper, J. and S. Thompson. 1973. "On the Applicability of Root Transformations". Linguistic Inquiry. 4, 465-497.
- Jackendoff, R. 1971. "Gapping and Related Rules". Linguistic Inquiry. 2, 21-35.
- Jackendoff, R. 1977. X-Syntax: A Study of Phrase Structure. Cambridge: MIT Press.
- Koster, J. 1978. Locality Principles in Syntax. Dordrecht: Foris Publications.
- Koutsoudas, A. 1971. "Gapping, Conjunction Reduction, and Coordinate Deletion". Foundations of Language. 7, 337-386.
- Kuno, S. 1976. "Gapping: A Functional Analysis". Linguistic Inquiry. 7, 300-319.
- Maling, J. 1972. On "Gapping and the Order of Constituents". Linguistic Inquiry. 3, 101-108.
- Quirk, R., S. Greenbaum, G. Leech, J. Svartvik. 1972. A Grammar of Contemporary English. London: Seminar Press.
- Riemsdijk, H. van. 1978. A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases. Lisse: The Peter de Ridder Press.
- Rivero, M.L. 1980. "Review of J. Koster. Locality Principles in Syntax". Canadian Journal of Linguistics. 25, 97-99.
- Rochemont, M. 1978. "A Theory of Stylistic Rules in English." University of Massachusetts Doctoral Dissertation.

- Ross, J. 1967a. "Constraints on Variables in Syntax". MIT Doctoral Dissertation.
- Ross, J. 1967b. "Gapping and the Order of Constituents". Reprinted in M. Bierwisch and K. Heidolph, (eds.) 1970. Progress in Linguistics. The Hague: Mouton.
- Sag, I. 1977. "Deletion and Logical Form". University of Pennsylvania Doctoral Dissertation.
- Stillings, J. 1975. "The Formulation of Gapping in English as Evidence for Variable Types in Syntactic Transformations". Linguistic Inquiry. 1, 247-273.
- Stillings, J. 1977. "Remarks on Core Grammar". University of Massachusetts Doctoral Dissertation.
- White, L. 1980. "Grammatical Theory and Language Acquisition". McGill University Doctoral Dissertation.
- Williams, E. 1977. "Discourse and Logical Form". Linguistic Inquiry. 8, 101-139.