

PATHS AND CATEGORIES

by

DAVID MICHAEL PESETSKY

B.A., Yale College
(1977)

SUBMITTED TO THE DEPARTMENT OF
LINGUISTICS AND PHILOSOPHY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 1982

© David Michael Pesetsky

The author hereby grants to M.I.T. permission to reproduce and
distribute copies of this thesis document in whole or in part.

Signature of Author: _____

Certified by: _____

Noam Chomsky
Thesis Supervisor

Accepted by: _____

Samuel Jay Keyser
Departmental Graduate Committee

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
Hum. DEC 2 1982

PATHS AND CATEGORIES

by

DAVID MICHAEL PESETSKY

Submitted to the Department of Linguistics and Philosophy
on 30 August 1982 in partial fulfillment of the requirements for
the Degree of Doctor of Philosophy

ABSTRACT

This study has two parts. In Part I, we consider the role of categorial selection in the theory of grammar. Categorial selection governs the category of a predicate's arguments, determining, for example, whether an argument must be an NP, an S', etc. We study asymmetries between subjects and objects in three quantificational constructions in Russian: the genitive of negation, distributive phrases with po, and numeral phrases. The subject/object asymmetries can be explained by Chomsky's (1981a) Empty Category Principle (ECP), if we separate categorial selection from θ -theory, which governs the assignment of thematic (θ -) roles. We propose that categorial selection applies only at the level of Logical Form (LF), while θ -theory applies at the three levels D-structure, S-structure and LF. This has wide consequences for the categorial status of empty categories, which in turn allows us to explain the subject/object asymmetries. We extend our analysis to Russian infinitival free relatives and secondary predicate constructions. Finally, after separating categorial selection from θ -theory, we argue that categorial selection itself should be subsumed under a more general semantic theory.

In Part II, we turn our attention to the ECP itself. We argue that the ECP, the Subject Condition (Chomsky 1973), and well known constraints on the crossing and nesting of dependencies fall together under a general theory governing the interaction of paths in a syntactic tree. A new subsystem of grammar is proposed, Path Theory, which contains a definition of paths, based on work by Kayne (1981a, 1982), and a Path Containment Condition (PCC). We compare the PCC to other accounts of Crossing effects, and to the ECP, in Chapter Three.

In Chapter Four, we extend Path Theory to constructions with multiple gaps, adapting ideas of Kayne (1982). We deal first with the "parasitic gap" construction of Taraldsen (1981), Engdahl (1981) and Chomsky (1981b). We then turn our attention to coordinate structures, arguing that most of Ross's (1967) Coordinate Structure Constraint derives from the PCC.

Applying the general theory of multiple gaps to coordinate structures, we explain immediately the possibility of "Across the Board" exceptions to Ross's constraint, including certain subject/object asymmetries discovered by Williams (1978). We discuss constraints on the conjunction of tensed and infinitival clauses in the context of Across the Board exceptions to Ross's constraint, and consider in this connection the status of expletive null subjects in so-called "pro-drop" languages.

In Chapter Five, we consider other topics in Path Theory. We present a PCC account of ECP and Superiority Condition effects in WH-in-situ constructions. We discuss some distinctions between indicative and subjunctive clauses. Finally, after considering some problems that arise in Path Theory, we note some desirable implications of Path Theory for the definition of government.

Thesis Supervisor: Noam Chomsky

Title: Institute Professor

PREFACE

It is said that a camel is a horse built by a committee, but this camel would never have made it across the desert, were it not for my thesis committee.

Special gratitude goes to my adviser, Noam Chomsky. The debt that this dissertation owes to his work will be abundantly clear. More specifically, however, he encouraged me in my work when I needed encouragement, went over the chapters with a magnifying glass, and even came up with the title, when I was stuck -- though I won't swear he didn't say "Categories and Paths".

More big thanks to Morris Halle, ДИРЕКТОР. Without his comments and kibbitzing, this dissertation and its author would have been a lot poorer. Big thanks also to Ken Hale, for valuable advice and comments, particularly on Chapter Two. Extreme gratitude, once again, to Jim Higginbotham, who was often able to explain to me what I meant much better than I could.

Two people have also belonged to my own mental thesis committee, even if they did not know it. One is Richard Kayne, whose ideas have inspired almost every page of this thesis, as the reader will soon learn. Without his work, none of this thesis would exist. The other is Jean-Roger Vergnaud, to whom I explained the last three chapters in an hour, and whose comments made me hope that there was something real in what I was doing.

Many other teachers have contributed to this work, by showing me what linguistics is and can be. I thank particularly Paul Kiparsky, Joan Bresnan, and John McCarthy, her T.A. in my first MIT syntax course,

for making linguistics exciting. Jay Keyser has been at MIT as long as I have. He has kept the department not only going, but thriving -- no mean feat.

This dissertation would be nowhere without my classmates and friends Hagit Borer, Ken Safir and Tim Stowell. I don't know why I was so lucky as to be in their class, but I was. Thank you particularly Hagit and Tim for listening to my thesis long distance in California, and Ken, my long-ago collaborator, for joining with me in a mutual aid society in this year of years.

Isabelle Haïk has suffered through this dissertation probably more than I have. Each idea was tried out on her, and what merit there is is due as much to Isabelle as to me. Thank you.

I have been fortunate at MIT to know, like, and talk to Jim Huang, Kyle Johnson, Maria-Luisa Zubizarreta, Dominique Sportiche, Anne Rochette, Lauri Carlson, Carol Neidle, Mario Montalbetti, Mamoru Saito, Rita Manzini, Juliette Levin, Luigi Burzio, Alec Marantz, Ann Farmer, Osvaldo Jaeggli, Shelly Lieber, Mark Baltin, Joel Rotenberg, Robert May, Denis Bouchard, Eric Wehrli, Pino Longobardi... Among the visitors who help to keep the department lively, I have benefited particularly from Eric Reuland, Henk van Riemsdijk, Adriana Belletti, Luigi Rizzi, Tarald Taraldsen, Neil Elliott, Carmen Picallo, Michael Brody and Edwin Williams.

Special kudos to my good friends Donca Steriade and Barry Schein, for services including and beyond linguistics. May they live a thousand years!

Thanks to a grant from the International Research and Exchanges Board, I was able to spend the 1979-1980 academic year in the USSR.

The research done during this year led to a large part of Chapter Two.

Among my friends and colleagues in the USSR, my greatest debt of thanks is to M.G. Sinicyn. Much of the first section of Chapter Two derives from joint research with him, and much more would have been done jointly, were I a better letter writer. It was he who made me realize how interesting Russian syntax is, and who introduced me to the topic of negation in Russian. I also profited enormously from working with O. Vinogradova, whose help, as linguist and informant, was essential. I also benefited from discussion with I. Boguslavskij, A. Kibrik, E. Padu^Vceva, E. Savvina and A. Spencer. Thanks also to my advisor in Leningrad, V. Kolesov, for creating conditions in which I could pursue my research.

Guy Carden taught my introduction to modern linguistics while I was a freshman at Yale, kept me going through four more years, and has remained a friend since. Also at Yale, A. Liberman showed me how phonetics can teach us about the mind. S. Martin and W. Cowgill each expanded my linguistic krugozor in important ways.

Antepenultimately, but crucially, I wish to thank my informants. My Russian informants are too numerous to remember. They include my friends and colleagues in the Soviet Union; a librarian in the New York Public Library's Slavic Collection whose name I never learned; Boris Katz; and especially V. Schiller, who proved herself indefatigable. My French informants have included Isabelle Haïk, Anne Rochette, Dominique Sportiche, and Jean-Roger Vergnaud. For Spanish, Mario Montalbetti, Maria-Luisa Zubizarreta and Osvaldo Jaeggli. For Italian, Rita Manzini and Luigi Burzio. Other informants I will thank in the text.

For completely absolutely and positively indispensable aid in preparing the manuscript in a very, very short time, I thank Isabelle Haïk, Kyle Johnson and Susan Davco. Special thanks also to sometime typist and proofreader Juliette Levin.

Finally, very special and separate thanks go to my parents, who have always been there when I needed them. This thesis is dedicated, with love, to them.

* * * * *

Some notes about the structure of this study. This thesis has two parts. Part I largely concerns Russian, and studies a number of subject/object asymmetries in Russian quantificational constructions. In this section we show that Chomsky's (1981a) Empty Category Principle can explain these asymmetries if certain assumptions are made about the role of categorial selection in the grammar. Part II considers the Empty Category Principle in greater detail, and argues that it can be collapsed with Chomsky's (1973) Subject Condition and some well-known constraints on crossing and nesting dependencies. The connection between the two parts is loose, but present, and some material from Part I is reanalyzed in Part II.

Much of the material in Part I has been presented in talks at the State Universities of Groningen and Utrecht; the University of Amsterdam; Université de Paris VII; USC; UCLA; UC Irvine; UQAM; the Second Soviet-American Conference on Russian Linguistics (College Park, Maryland); NELS XII (MIT) and the 1981 Winter Meeting of the LSA. The research in

Part I was partially supported by an NSF Graduate Fellowship, as well as by a grant from the International Research and Exchanges Board, mentioned above.

Throughout the dissertation I use "prime" notation for denoting levels of X-bar structure, rather than "bar" notation. Thus, we use S' for \bar{S} , etc. This means that we do not use an apostrophe to form the plural of a node: Ss is the plural of S, and S's is the plural of S'. Let no one say that punctuation is not a rich deductive system.

TABLE OF CONTENTS	<u>page</u>
ABSTRACT	2
PREFACE	4
<u>PART ONE: CATEGORIES</u>	
CHAPTER 1: INTRODUCTION	
1. UG and Core Grammar	12
2. θ -theory, the Projection Principle, and Binding Theory	16
3. The Projection Principle, θ -criterion, and Categorical Selection	29
Footnotes	38
CHAPTER 2: RUSSIAN QUANTIFICATION AND CATEGORIAL SELECTION	
1.0 Three Descriptive Conditions	40
1.1 Genitive of Negation	40
1.1.1 The D-structure [XP, VP] Restriction	41
1.1.2 The Non-Obliqueness Restriction	64
1.1.3 Obligatory Quantifier Movement	65
1.2 <u>Po</u> -phrases	69
1.2.1 <u>Po</u> -phrases and the D-structure [XP, VP] Restriction	70
1.2.2 <u>Po</u> -phrases and the Non-Obliqueness Restriction	72
1.2.3 <u>Po</u> -phrases and Obligatory QR	73
1.3 No-agreement Numeral Phrases	74
1.3.1 No-agreement Numeral Phrases and the D-structure [XP, VP] Restriction	77
1.3.2 No-agreement Numeral Phrases and the Non-Obliqueness Restriction	79
1.3.3 No-agreement Numeral Phrases and Obligatory QR	82
2.0 The Non-Obliqueness Restriction and the QP Hypothesis	86
2.1 The QP Hypothesis	
2.2 Oblique vs. Structural Case	
2.3 The Non-Obliqueness Restriction	
3.0 QPs and the Empty Category Principle	
4.0 The ECP, Category Blindness, and Conditions <u>A</u> and <u>C</u>	122
4.1 QP and the ECP	
4.2 Categorical Selection and Condition <u>C</u>	129
4.3 Categorical Selection and Condition <u>A</u>	134

4.4 Two More Cases	149
4.4.1 Infinitival Free Relatives	149
4.4.2 Secondary Predicates	158
4.4.2.1 Secondary Predicates and the D-structure [XP, VP] Restriction	165
4.4.2.2 Secondary Predicates and the Non-Obliqueness Restriction	169
4.4.2.3 Secondary Predicates and Small Clauses	170
5.0 Categorical and Semantic Selection	180
Footnotes	206

PART TWO: PATHS

CHAPTER 3: CROSSING EFFECTS, THE ECP, AND THE SUBJECT CONDITION

0. Introduction	248
1.0 The Structure of S	250
1.1 The Definition of Subject	255
1.2 Contraction	255
2.0 Remarks on the ECP	264
3.0 Crossing and Nesting	267
3.1 Crossing Effects	267
3.2 Complementizer-Trace Phenomena	272
4.0 Crossing Dependencies vs. Paths	277
4.1 Crossing in Two Dimensions	277
4.2 Paths	288
4.3 Paths vs. Linear Crossing	297
4.4 CTP Effects	305
4.5 The Subject Condition	313
4.6 The Subject Condition in Small Clauses and S'-deletion Infinitivals	319
4.7 Conclusions	337
Footnotes	339

CHAPTER FOUR: MULTIPLE GAPS AND PATHS

1.0 Introduction	356
2.0 Paths and Parasitic Gaps	357
2.1 Parasitic Gaps and Subject Condition Violations	357
2.2 Parasitic Gaps and "Connectedness"	368
2.3 Left Branch Extractions	387
2.4 Russian QPs Revisited	406
2.5 Parasitic Gaps and Crossing Effects	420

3.0	Coordination and Multiple Gaps	433
3.1	The Coordinate Structure Constraint	433
3.2	Across the Board	443
3.3	The Path Between INFL and COMP	459
	Appendix to 3.3	484
4.0	Multiple Gaps, the PCC, and Subjacency	488
4.1	A Unified Theory of Multiple Gap Constructions	490
4.2	Chomsky's Analysis of Parasitic Gaps	498
4.3	Subjacency and the PCC	505
4.4	The PCC and LF; Counteranalyses of ATB Phenomena	511
4.5	<u>Move α</u> , Subjacency, and the PCC	530
5.0	Conclusions	551
	Appendix: Slash Categories and Coordination	554
	Footnotes	559
CHAPTER FIVE: FURTHER PCC EFFECTS		
1.0	Introduction	585
2.0	Infinite Paths: WH-in-situ and Gapping	586
2.1	WH-in-situ	586
2.2	ECP, Superiority, and the PCC	595
	2.2.1 Pure ECP Effects	595
	2.2.2 ECP/Superiority Effects	598
	2.2.3 Pure Superiority Effects	602
	2.2.4 Infinite Paths and WH-in-situ	605
2.3	Multiple WH-in-situ and Absorption as Coindexation	613
2.4	WH-in-situ and Logical Form	629
2.5	Focus and Gapping	640
	2.5.1 Focus and the PCC	640
	2.5.2 Gapping	642
3.0	Subjunctives and W-verbs	659
4.0	Problems for the PCC	682
	4.1 The PCC and QR	683
	4.2 The PCC and Preposition Stranding	687
	4.3 The PCC and NP-trace	690
5.0	Path Theory and the PCC	694
	5.1 Path Theory and the Definition of Government	695
	5.2 Conclusions	702
	Footnotes	705
	REFERENCES	726

PART ONE: CATEGORIES

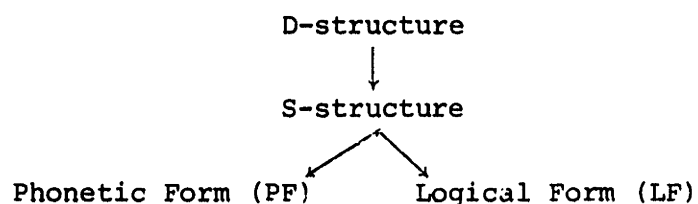
CHAPTER ONE: INTRODUCTION

1. UG and Core Grammar

Our point of departure will be the theory of generative grammar developed in Chomsky (1981a) and much other recent work. This theory characterizes two sorts of objects: a "universal Grammar" (UG) and a set of "core grammars". Universal Grammar is a set of operations and principles. These operations and principles contain certain variables, whose range may also be determined by UG. By replacing these variables with elements of their range, we obtain the set of "core grammars". In other words, UG is parameterized: the different values chosen for the various parameters determine the class of core grammars. To a large extent, the variation among languages may be blamed on the choice of core grammar, although further "peripheral" factors may also contribute.

To give a simple example, we assume that UG contains the following levels of representation:

(1)



To introduce some terminology: we speak of the operations mapping D-structures onto S-structures as applying in the syntax. The rules mapping S-structures onto LF representations apply at LF; and the rules mapping S-structures onto PF representations apply at PF.

As a principle of UG, D-structures are mapped onto S-structures by

the rule Move α . We will have more to say about D-structure, S-structure, and Move α in the following pages. As far as UG is concerned, α ranges freely over a set of features.

Let us consider one feature briefly, the feature [+WH]. We suggest in Chapter Five that UG contains an inviolable principle, which requires [+WH] to be in COMP at LF. How do we get such WH-elements into COMP? There are two places where this might happen: in the syntax, and at LF. WH can move to COMP in the syntax in English:

- (2)a. [_S, [_{COMP}] [_S you saw who]] D-structure
b. [_S, [_{COMP} who_i] [_S did you see e_i]] S-structure and LF

Movement in the syntax is possible because the range of α in Move α includes [+WH] in English. Although in (2)b WH must move in the syntax, for many speakers, for reasons discussed in Chapter Five, there are cases where movement may take place at LF. For example, we will argue that the LF representation of (3)a is (3)b:

- (3)a. [_S, [_{COMP} who_i] [_S e_i saw what]] S-structure
b. [_S, [_{COMP} who_i what_j] [_S e_i saw e_j]] LF

In Chinese, on the other hand, [+WH] is not a possible value for α in the syntax. As Huang (1982, forthcoming) demonstrates, [+WH] moves to COMP at LF. Thus, while (4)a, and not (4)b is a well-formed S-structure, (4)b is the LF representation of (4)a:

- (4)a. [_S, [_{COMP}] [_S ni kanjian-le shei]] S-structure
 you see-ASP who?
b. [_S, [_{COMP} shei_i] [_S ni kanjian-le e_i]] LF

The difference between Chinese and English is readily explained as a difference in the choice of core grammar. English, but not Chinese, includes [+WH] in the range of α in Move α .

This simple example of parametric variation also allows us to demonstrate, not only the distinction between UG and core grammar, but also the deductive structure of a grammar that is a set of principles and operations. A maximally simple theory of core grammars that allows [+WH] to be excluded from syntactic movements should also allow [+WH] to be excluded from movement at LF. This should yield four types of languages:

(5)	[+WH]-movement at:	<u>S-structure</u>	<u>LF</u>
	(i)	+	+
	(ii)	+	-
	(iii)	-	+
	(iv)	-	-

Type (i) is exemplified by English. Type (ii) might be exemplified by Italian, which seems to require all WH-elements to be in COMP at S-structure, and disallows multiple interrogations like (3)a. Type (iii) is exemplified by Chinese. What about (iv)?

Suppose, as seems reasonable, that all languages have [+WH] elements, to form interrogatives and similar operator structures. If these elements are ever found outside of COMP, then this language cannot exemplify (iv). We will soon see that WH-elements that are arguments must be outside of COMP at D-structure. Therefore, as long as a language has [+WH] arguments, it cannot be of type (iv).

While this prediction does not seem to be of great import, it illustrates rather neatly how a basic goal of research in generative grammar can be pursued. In working out the core grammars of actual

languages, we find various restrictions on the class of actual core grammars, and implicational relationships between various properties of the set of core grammars, of the type discovered in the work of Greenberg. If we view UG as a set of principles and parameters, which interact, we can explain, and not merely stipulate, the restrictions and implicational relations found in the set of core grammars. By doing so, we help explain how the child discovers the grammar of his native language during the process of acquisition. In the present work, we will be studying in detail the interactions of principles of UG in particular core grammars.

The principles of UG, and various definitions and assumptions associated with them, group themselves naturally into a number of sub-theories, giving UG a highly "modular" character. For Chomsky (1981a), these subtheories include:

- (6) (i) Bounding Theory
- (ii) Government Theory (Empty Category Principle)
- (iii) Case Theory
- (iv) θ -theory
- (v) Binding Theory
- (vi) Control Theory

Bounding Theory will play a minor role in our theory (except in Chapter Four, section 4.5). It contains the subjacency condition of Chomsky (1973), and a parameterized list of bounding nodes (Rizzi 1978a). The subjacency condition imposes rather severe locality conditions on the rule Move α in syntax, and possibly on other relations as well (as in Chapter Four, loc. cit.). The subjacency condition prevents Move α

from applying across more than one bounding node β , where β ranges over NP, S', S, and possibly other nodes such as PP (cf. Van Riemsdijk 1978a, Baltin 1978), depending on the core grammar.

Government Theory and Case Theory impose other conditions on syntactic categories, which will be of great importance to us. We will introduce both of these categories in Chapter Two, as they become relevant to our discussion. Government Theory will assume particular importance in the chapters of Part Two, where we argue at length that it should be replaced by a new subsystem of grammar: Path Theory. Path Theory concerns the interactions of paths in a phrase marker, and will be introduced in Chapter Three.

θ -theory concerns the syntactic reflexes of the assignment of thematic roles like "agent" and "patient" to arguments. θ -theory is the main topic of Part One. For this reason we will treat it in detail in the next section. Binding Theory and Control Theory deal with the relations between elements that are [+anaphor] or [+pronoun] and their antecedents, as well as with the properties of names and variables, which are [-anaphor, -pronoun]. Since these subtheories use concepts based on θ -theory, we will introduce these theories briefly after discussing θ -theory. Binding and Control theory will be invoked rather frequently in our discussion, but will play peripheral roles.

2. θ -theory, the Projection Principle and Binding Theory

θ -theory is, in a sense, the key to all the subsystems of UG, since it is the theory that concerns one of the most basic relationships in syntactic theory -- the relationship between an argument and the

element that assigns it a thematic role. Following Chomsky, we will use the term θ -role for the thematic roles that are distinguished by θ -theory. Certain lexical items assign θ -roles to, or " θ -mark", other categories, which we may call arguments. This assignment is obligatory and biunique, governed by the θ -criterion, a basic principle of θ -theory:

(7) θ -criterion

Each argument bears one and only one θ -role, and each θ -role is assigned to one and only one argument.¹

The θ -criterion is clearly related to strict subcategorization. We will say more about this relationship shortly.

At what level or levels does the θ -criterion hold? As Chomsky notes, it is at least relevant at LF, where something like the θ -criterion seems a minimal condition for adequacy of representations. Nonetheless, as noted in earlier work by Freidin (1978), May (1977), and Borer (1979), something like the θ -criterion seems a desirable constraint on the other levels of representation as well. Suppose we formulate a Projection Principle to capture this suggestion, as in (8):

(8) Projection Principle

Representations at each syntactic level (LF, D-structure, S-structure) are projected from the lexicon, in that they observe the θ -marking properties of lexical items.

This Projection Principle differs slightly from Chomsky's formulation, in ways we will consider shortly.

Let us give some content to the Projection Principle in (8). First, what positions can be θ -marked? Essentially, they are positions that bear the grammatical relations "subject-of" and "object-of". Thus, in (9), the

object of buy, the [NP, VP] in the notation of Chomsky (1965), bears the θ -role theme or patient; the subject of the sentence, the [NP, S], bears the θ -role agent:

(9) [_S John [_{VP} bought the book]]

θ -role assignment appears to observe a rather strict locality requirement, For example, V may θ -mark NP_{α} in (10)a, but not in (10)b:

(10)a. [_{VP} V NP_{α}]

b. [_{VP} V [_S, ... NP_{α} ...]]

In this case, the condition that θ -marking appears to obey is the condition of government. Let us call θ -marking by a lexical category direct θ -marking, following Chomsky:

(11) α directly θ -marks β if and only if α θ -marks β and $\alpha = X^{\circ}$

(12) If α directly θ -marks β then α governs β .

Let us assume an extremely simple definition of government proposed by Aoun and Sportiche (1982; stated as in Chomsky 1981a, 164):

(13) Government (def)

In the structure:

[β ... γ ... α ... γ ...], where

(i) $\alpha = X^{\circ}$

(ii) where ϕ is a maximal projection, ϕ dominates α if and only if ϕ dominates γ .

α governs γ .

In Chapter Two, where we discuss Government Theory (the Empty Category Principle), we will need to complicate this definition; one of the virtues of Path Theory will be that we can return to (13).

The combined effect of (11) and (12) is to force all θ -marking by a lexical category to take place within the maximal projection of the lexical category. This has the desired results in (10): only in (10)a can V assign a θ -role to NP_α .

What about θ -marking by a non-lexical category? There is some evidence, discussed by Chomsky (1981a), Marantz (1981), and others, that the subject in a sentence like (9) is assigned its θ -role by the VP, rather than by the lexical V. By (11), such θ -marking is not direct, but rather indirect. Here too some locality condition obtains, perhaps requiring sisterhood, as suggested in Chapter Three below.

We thus distinguish direct θ -marking of an object and indirect θ -marking of a subject. Returning to example (9), it follows from our Projection Principle that the subject John and the object book are θ -marked at each level of representation. Now consider an S-structure like (14):

(14) $[_{S_1} \text{John}_i [_{VP_1} \text{seems} [_{S_2} \underline{e}_i [_{VP_2} \text{to have bought the book}]]]]]$

In (14), the θ -role assigned by VP_2 is assigned to the empty category \underline{e}_i , the trace of John. Suppose the Projection Principle is correct. To satisfy the θ -criterion at S-structure, John must bear a θ -role. Clearly, this θ -role does not come from seems. Seems does not assign a θ -role to its subject position, as the possibility of expletive non-arguments in this position shows:

(15) $[_{S} \text{it} [_{VP} \text{seems} [_{S'} \text{that John has bought the book}]]]]]$

To satisfy the θ -criterion, then, John in (14) must be θ -marked by VP_2 . To accomplish this, we may adopt the convention proposed by Chomsky (1981a,

37):

- (16) α θ -marks the category β if α θ -marks the position occupied by β or the trace of β .

Ignoring some imprecision in the distinction between θ -marking a category and θ -marking an argument, made precise in Chomsky's discussion, we may conclude that VP_2 in (14) θ -marks John by θ -marking its trace, satisfying the Projection Principle.

Convention (16) has another, important consequence, as noted by Borer (1979) and Chomsky (1981a). It entails that elements can be moved only into positions that are not assigned a θ -role. Thus consider (17):

- (17) * $[_S \text{ John}_i [_{VP} \text{ helped } \underline{e}_i]]$

The empty category \underline{e}_i cannot be the trace of John. The verb help directly θ -marks the position occupied by \underline{e}_i ; by (17). If \underline{e}_i is the trace of John, then help will also θ -mark John. But the whole VP also indirectly θ -marks John. Thus, John will be θ -marked both by help and by VP, and violate the θ -criterion.

In (14), on the other hand, if we assume that a VP containing seems does not assign a θ -role to the subject, as evidenced by (15), John will bear only the one θ -role it receives from \underline{e}_i , and the θ -criterion is not violated.

We thus distinguish θ -positions, like the subject of help, from non- θ -positions, like the subject of seems. Other non- θ -positions are COMP and positions resulting from adjunction. Suppose now we define D-structure as the level at which (16) does not hold:

- (18) At D-structure, α θ -marks β if and only if α θ -marks the position occupied by β .

It follows that D-structure will be a pure representation of thematic structure, of the pairing of θ -roles with arguments bearing grammatical functions (GFs). In Chomsky's terms, D-structure is a pure representation of GF- θ . The D-structure of (14) thus looks much like (15):

(19) [_{S₁} e [_{VP₁} seems [_{S₂} John [_{VP₂} to have bought the book]]]]

The rule Move α , which maps D-structures onto S-structures, is thus a mapping from a pure representation of GF- θ onto a representation in which some θ -roles are assigned to arguments in non- θ -positions, by virtue of convention (16). We will have more to say about Move α in Chapters Two and Four.

Notice that the passive construction does not differ with respect to its properties from the raising constructions that we have been considering. Consider an S-structure like:

(20) [_S John_i [_{VP} was seen e_i]]

The verb see directly θ -marks e_i, and thus θ -marks John as well. It follows that the subject of a passive VP is in a non- θ -position, and that (20) is derived by Move α from a D-structure like (21):

(21) [_S e [_{VP} was seen John]]

That the subject of a passive verb is not a θ -position can be seen by the possibility of expletive elements in that position:

(22) ?[_S there [_{VP} were seen some men]]

Compare (23), where the subject is a θ -position:

(23) *[_S there [_{VP} helped Mary]]

Chomsky (1981a) develops a slightly different way to satisfy the θ -criterion at S-structure. In an S-structure like (14) or (20), one can construct a sequence out of the argument and its trace, of the form (John_i, e_i) , or of the GFs of the positions these categories occupy: $([\text{NP}, S_1], [\text{NP}, S_2])$ for (14) and $([\text{NP}, S], [\text{NP}, \text{VP}])$ for (20). We can call these sequences chains.

Each element of a chain locally binds the next, where we define "local binding" as follows:

(24) Local Binding (def) (Chomsky 1981a, 59)

α locally binds β if and only if:

- (i) α and β are coindexed
- (ii) α c-commands β
- (iii) There is no γ coindexed with α such that α c-commands γ and γ c-commands β .

For reasons discussed by Chomsky, it is well to consider only a limited type of chain as relevant for the θ -criterion. A natural distinction can be made between positions bearing grammatical functions, to which θ -roles may be assigned at D-structure, and positions that do not bear grammatical functions. We can call the former positions A-positions (to suggest "argument") and the latter positions \bar{A} -positions ("non-argument"). A-positions are subject and complement positions. \bar{A} -positions include COMP and adjoined positions. We can restate the θ -criterion and principle (16) in terms of chains of A-positions, called A-chains:

(25) θ -criterion

- (i) Consider the A-chain $C = (\alpha_1, \dots, \alpha_n)$ and the θ -role R . For all positions $P \in C$, for some $\alpha_i \in C$, if α_i bears R , then P bears R
- (ii) Each argument bears one and only one θ -role, and each θ -role is assigned to one and only one argument.

(25) is our formulation, but is equivalent to Chomsky's (1981a, 335) formulation, for our purposes in this study. Note that (25) allows an argument that is not part of any chain to bear a θ -role. This possibility is realized in the context of Case theory, where we will want to say that S's and certain other arguments do not belong to chains, but bear θ -roles.

Consider the following definition:

- (26) Chain C bears θ -role R if and only if for some $\alpha \in C$, α bears R .

From the θ -criterion, we derive the following theorem:

- (27) Each A-chain bears at most one θ -role, and each θ -role is assigned to at most one chain.

Suppose chain C bears two θ -roles: call them α and β . By (26), some member of C bears α and some member of C bears β . By clause (i) of (25), all members of C bear α and β . Both α and β must be assigned to an argument in C , but this argument will bear two θ -roles by clause (i), violating clause (ii). Hence, the A-chain must bear no more than one θ -role.

The following theorem can also be derived:

- (28) Each A-chain contains at most one argument.

If chain C contains two arguments, they must each bear a θ -role. Suppose they bear the same θ -role. Then (ii) is violated. We know from (i) that

they cannot bear distinct θ -roles. Thus C cannot contain two arguments. By induction C cannot contain more than two arguments.

Notice that the θ -criterion does not require that a chain bear any θ -role, so long as it does not contain an argument. This is desirable, since chains may be formed by raising expletive elements.

A final definition: the member of a chain that c-commands all other members of that chain is called the head of a chain,

We have not so far given any content to the notion argument. Clearly in chains of the form (lexical NP,...), the lexical NP may be the argument, since such chains may be of length 1 in sentences where no NP-movement takes place. Indeed, we may redefine D-structure as the level at which all chains are of length 1. At D-structure, clearly, lexical NPs are arguments.

We may ask whether empty categories can also be arguments. The answer is yes; the conditions under which an empty category may be an argument are determined by Binding Theory and by a typology of empty categories. We end our theoretical overview with a brief discussion of these issues.

Not all A-chains are well-formed. Consider the chain (John, e_i) in (29):

(29) *John_i seems [_S, that e_i came]

An insight of Chomsky (1973) was the relation between (29) and (30):

(30) *John_i said [_S, that himself_i came]

In the framework of Chomsky (1981a), an NP-trace like e_i in (29) and himself in (30) are both anaphors, and are subject to the locality

principles imposed by the Binding Theory:

(31) Binding Theory (Chomsky 1981a)

- A. An anaphor is bound in its governing category.
- B. A pronominal is free in its governing category.
- C. An R-expression is free.

Anaphors include reflexives, reciprocals and NP-traces like e in (29).

Pronominals include lexical pronouns like he, us, etc. R-expressions

include names like John, book, and perhaps, certain empty categories

bound from \bar{A} -positions, called variables.

An anaphor is referentially dependent, lacking any ability to refer independently of a grammatical antecedent. A pronominal may be referentially independent, or may pick an antecedent for its reference.

An R-expression is referentially independent.

The exact definition of the governing category referred to by principles A and B of the Binding Theory is a matter of some controversy.

A number of suggestions are explored by Chomsky (1981a), Huang (1982) and others. Roughly speaking (Chomsky, p. 188):

(32) α is the governing category for β if and only if α is the minimal category containing β and a governor of β , where $\alpha = \text{NP}$ or S .²

Let us return to (29) and (30), and see how the Binding Theory rules them out. Chomsky assumes that the structure of S is as in (33), a structure we will assume in Chapter Two and modify in Chapter Three:

(33) $\text{S} \longrightarrow \text{NP INFL VP}$

INFL (for "inflection") contains the constituents AGR and TNS ("agreement" and "tense") in a tensed sentence. Supposing that the lexical projection

of INFL is AGR and that S is also a projection of INFL, the subject NP is governed by AGR in a tensed S. Thus, (29) and (30) have the structure below:

(29) * $[_{S_1} \text{John}_i [_{VP_1} \text{seems } [_{S'} \text{that } [_{S_2} \text{e}_i [_{INFL_2} \text{TNS AGR}] [_{VP_2} \text{came}]]]]]]]$

(30) * $[_{S_1} \text{John}_i [_{VP_1} \text{said } [_{S'} \text{that } [_{S_2} \text{himself } [_{INFL_2} \text{TNS AGR}] [_{VP_2} \text{came}]]]]]]]$

Since the subject of S_2 is governed by AGR, its governing category is S_2 . Both e_i and himself are anaphors, and are required to be bound in their governing category. Since they are not, the structures are ruled out.

Notice that principle A of the Binding Theory applies to elements that are [+anaphor, -pronominal], principle B to elements that are [-anaphor, +pronominal]. Principle C applies to elements that are [-anaphor, -pronominal]. What about elements that are [+anaphor, +pronominal]? Principle A requires such elements to be bound in their governing category; principle B requires them to be free in their governing category. If they have a governing category, a contradiction is reached. Thus, [+anaphor, +pronominal] elements must lack a governing category. If they are governed by X^0 , they will have a governing category. Thus, they must not be governed by X^0 . The subject position of an S' whose INFL lacks AGR meets this criterion, and essentially no other position, in English. Infinitives lack AGR; thus [+anaphor, +pronominal] elements can only be the subject of an infinitival S at the level at which the Binding Theory applies. Case Theory, which we consider in Chapter Two, prevents a lexical NP from occurring in this position: hence only an

empty category can be [+anaphor,+pronominal]. This empty category is called PRO. Being a pronominal, it has the necessary features to be an argument; hence PRO, the ungoverned empty category, can serve as an argument.

This discussion leads us to the typology of empty categories in the theory we assume. We have seen that PRO, [+anaphor, +pronominal], falls under both principles A and B of the Binding Theory, requiring it to be ungoverned. Since PRO is an argument, it is either free, or has an antecedent which bears an independent θ -role:

(34) John_i hopes [_S, PRO_i to leave on time]

John and PRO bear distinct θ -roles.

We have examined a type of empty category that is [+anaphor, -pronominal]. It falls under principle A of the Binding Theory, and must be bound in its governing category. This empty category is not an argument, and has a local antecedent in an A-position which does not bear an independent θ -role. The empty category and its antecedent belong to the same chain. This empty category is often called "NP-trace", because it is the trace resulting from NP-movement operations in passive and raising constructions.

An empty category whose antecedent is in an \bar{A} -position and does not bear an independent θ -role is [-anaphor, -pronominal] and falls under principle C of the Binding Theory. These empty categories may result from WH-movement to COMP, and may be called "WH-trace".

Finally, we expect there to exist empty categories that are [-anaphoric, +pronominal]. These are discussed by Chomsky (1981b), and

will be considered in Chapter Four, where we discuss the "parasitic gap" construction.

Much attention has been devoted to justifying and explaining the typology of empty categories outlined above. In much of Chomsky (1981a), the various empty categories are taken to intrinsically differ in their feature composition beyond the differences given by [\pm anaphor, \pm pronominal]. Chomsky later notes that the empty categories form a near partition of the structural environments available to them. This "complementary distribution" suggests, here as in phonology, that there is only one empty category -- the result of the optionality of phonetic content in the base -- the properties of whose "allocategories" are determined by the various subsystems of grammar. This is the view we take in this study, but the issue is tangential to our concerns. On the other hand, we will be dealing very closely in Chapter Two with certain of the features of empty categories -- specifically their categorial features for [\pm N] and [\pm V].

The choice of antecedent for NP- and WH-trace is determined by the Binding Theory, Move α , and also Bounding Theory. The choice of antecedent for PRO is determined by Control Theory, which we will discuss only marginally (cf. Manzini 1981 for discussion of Control Theory).

Finally, Chomsky argues that the Binding Theory must apply at least at S-structure, and certainly not at D-structure. It may also apply at LF, as argued by Aoun (1982).

The preceding discussion has, we hope, provided enough background to follow the argument we will present in this study. For a much more complete presentation of the issues and arguments for our various

assumptions, the reader is referred to Chomsky (1981a, 1981b). For a thorough discussion of the theory of chains, see particularly Safir (1982). In the next section, we return to the Projection Principle, and lay some groundwork for our discussion in Chapter Two.

3. The Projection Principle, θ -criterion and Categorial Selection

The θ -criterion and the Projection Principle deal in the most basic sense with what we might call the "syntactic properties" of lexical items. Trivially, certain lexical items in any language are restricted to specific syntactic environments. In English, for example, a verb may "take" a single object, or a pair of objects, or may be intransitive, lacking an object entirely. Some nouns must cooccur with an article; others must not cooccur with an article. A linguistic theory that does not at least characterize properties of this sort is inadequate at the simplest level. The interesting question is which of the many possible ways of characterizing these syntactic properties is correct. θ -theory is a characterization of some of these properties.

This question is interesting because the specification of lexical items' syntactic properties is not only a minimal requirement for descriptive adequacy. It is also a centerpiece of any theory which attempts to deal with the fundamental problem of language acquisition. As Chomsky (1981a) notes, the primitives of any theory that aims to explain how a child acquires his language must meet a strict criterion of "epistemological priority". The properties of these primitives must allow a child, armed only with UG, to develop knowledge of his language-particular grammar from his linguistic and extralinguistic experience.

As Chomsky and others have frequently noted, the poverty of the linguistic stimuli available to the child suggests that the criterion of epistemological priority imposed on the primitives of linguistic theory should be stringent indeed. Notions like "precedes", "agent", "patient", and the like may have the appropriate prelinguistic basis, but notions like "subcategorizes for" or "is bound by" do not. Knowledge of these notions must be part of UG, and the implementation of these notions in the child's particular grammar (e.g. "eat subcategorizes for an NP") must result from the interaction of experience with UG.

Nonetheless, many of the properties of specific lexical items -- in particular their θ -related properties -- must be learned by experience, however much even this knowledge may be guided by innate principles and generalizations about possible worlds. The child can only learn that eat takes two arguments, an agent and a patient, by making some connection with the world around. Given this fact, and the desire to keep the primitives of linguistic theory to a bare, epistemologically prior minimum, it is interesting to investigate the extent to which the surface properties of linguistic systems may be deduced exclusively from these θ -related properties of lexical items, within a rich theory of UG.

Chomsky (1981a), in an important step in this direction, proposes that many of the properties of phrase structure, previously assumed to derive from a set of language-specific base rules, are deducible from the θ -properties of lexical items. We discuss this line of reasoning in greater detail in the first part of Chapter Three. Stowell (1981)

follows this program explicitly, showing how the phrase structure of a large number of constructions follows from the interaction of θ -theory with other principles of grammar.

Results of this kind, in turn, depend on the correct formulation of the basic principles of grammar, and on a correct understanding of the projection of lexical properties onto grammatical representations. As our postulated principles become deep enough to yield results of the sort desired, we expect that slight differences in our formulation of these principles will have immediate and testable consequences for our understanding of linguistic phenomena. In particular, given the central position of the syntactic properties of lexical items in any explanatory linguistic theory, we expect that the way in which these properties are projected onto grammatical representations will be crucial to our view of the rest of the theory. This question has been discussed repeatedly in the literature. We will be referring especially to Chomsky (1965), Grimshaw (1979) and Chomsky (1981a).

In Chapter Two, we will present a proposal about the projection of lexical properties onto syntactic representations which differs from the proposals made in Chomsky (1981a) in a rather minor way. We will show, however, that this issue is so important to the general structure of syntactic theory that our minor change actually has important consequences for seemingly unrelated issues in domains as diverse as the theory of empty categories, LF, Government Theory, and others, as well as for the analysis of a range of constructions.

Our analysis will be motivated mainly by evidence from Russian,

where a group of apparently disparate constructions share an odd collection of surface properties which are not related in any obvious way. We will show that our proposal to modify the relationship between lexical properties and syntax provides the missing link between the Russian constructions and their properties.

Let us now become more concrete and consider in detail which lexical properties are relevant to syntactic representations. Consider a verb like receive. Anyone who knows this verb knows at least two things about its "syntactic environment". First, he knows the positions that this verb requires to exist. In its standard use (apart from receiving lines and American football), receive must have an object. Second, he knows the syntactic categories that can fill the obligatory positions. The object of receive, for example, must be an NP, and not a clause or some other type of category.

In the Standard Theory, as developed in Chomsky (1965), both positional and categorial requirements fall under the theory of strict subcategorization, insofar as they apply to complements and not to subjects. The mechanism of subcategorization features introduced in the Standard Theory does not, in fact, make any distinction between positional and categorial requirements. A verb subcategorizes a position by subcategorizing a particular category in that position.

In the theory of Chomsky (1981a), there is a good deal of overlap between subcategorization and θ -theory. In Chomsky's system, however, certain distinctions remain. First of all, consider the basic configuration in which subcategorization takes place. Consider (35), where α is an immediate constituent of γ :

(35) [γ ... α ... β ...] or [γ ... β ... α ...]

Where $\alpha = X^0$, β must satisfy the subcategorization frame of α . In other words, β is present if and only if it is subcategorized. If we understand Θ -theory to require a position to exist whenever Θ -marking of that position is possible (an assumption Chomsky modifies), then it follows that subcategorization entails Θ -marking, at least when an argument is subcategorized.³

On the other hand, consider the case in which α is a maximal projection. For example, suppose $\alpha = VP$ and $\gamma = S$. Then α might indirectly Θ -mark β , but it does not subcategorize it. This is because only objects are subcategorized, not subjects.

Thus it might seem that we can eliminate entirely the mechanisms of subcategorization in favor of Θ -theory, particularly if Θ -theory, like subcategorization, dictates the syntactic categories that can receive Θ -roles. Chomsky suggests, however, that this is not a desirable step. The argument concerns the obligatoriness of subjects, and can be reconstructed as follows:

Suppose we claim, as we have, that Θ -theory makes certain positions obligatory. To be more precise, suppose we claim that if α assigns Θ -role R to position P in S , then P must be present in S . We derive the obligatoriness of the subject position in a structure like (9), which we repeat:

(9) [_S John [_{VP} bought the book]]

The VP assigns a Θ -role to the subject position; therefore, the subject position must be present. Θ -theory, however, does nothing to make the subject position obligatory in a sentence like (15), and yet it is obligatory:

(15) [_S it [_{VP} seems [_S, that John has bought the book]]]

It thus appears that we need an independent principle like (36):

(36) Clauses have subjects.

Notice that (36) is now redundant with the principle that θ -assignment requires a position to exist, for the case of subject position of a clause. Chomsky goes on to argue that this principle is too strong in any case. Consider the assignment of θ -roles in NP. The complement position must, it seems, exist whenever a θ -role is assigned to it:

(37)a. John's swift [_N, publication of the book]

b.*John's swift [_N, publication]

On the other hand, the subject position, even though there is a θ -role available for it, is optional:

(37)c. the swift [_N, publication of the book]

Chomsky concludes that θ -marking does not require positions to exist. Rather:

(38) If a structural position that can be θ -marked is obligatory, then it is obligatorily θ -marked by an element that can θ -mark it; if such a position is only optionally present, then θ -marking of this position is correspondingly optional. (p. 40)

Two principles are assumed which make positions obligatory. The first is subcategorization, which makes a θ -marked object position obligatory. The second is principle (36), requiring clauses (but not NPs) to have subjects.

Notice that, even so, subcategorization in the sense of Chomsky (1965) -- a theory of positional and categorial selection -- is not really necessary. Its role in the theory is to create a subject/object asymmetry with respect to the obligatoriness of positions. It really is simply a factoring out of a complication that could as well be included in (38), in which case (38) would read like (39):

(39) If an element can directly θ -mark a position, then that position is obligatory.⁴ If an element can indirectly θ -mark a position, then:

- (i) if it is obligatory, then it is obligatorily θ -marked by the element that can θ -mark it;
- (ii) if such a position is only optionally present, then θ -marking of this position is correspondingly optional.

In a sense, then, subcategorization simply factors direct θ -marking out of θ -theory. Given (39), we may still hold that subcategorization is fully subsumed by θ -theory.⁵ Chomsky (1981a) argues that subcategorization falls under the Projection Principle, as well as θ -theory. If subcategorization is fully subsumed by θ -theory, and we assume our version of the Projection Principle in (8), then subcategorization falls under the Projection Principle trivially.

But now let us return to the distinction between positional and categorial requirements discussed above. Recall that subcategorization inextricably mixes these requirements -- subcategorization is thus a theory both of positional selection and of categorial selection. We will abbreviate categorial selection as c-selection. Thus, if a verb has a subcategorization frame +[_NP], this frame positionally selects an object, and c-selects an NP. It is a simple matter to extend θ -theory so that it too includes both positional selection and c-selection, but is such an extension correct?

At first sight, θ -theory seems to be an even better home for c-selection than subcategorization. This is because subcategorization extends only to complements, while c-selection appears to range over both complements and subject (as O. Jaeggli has pointed out to me). For

example, the verb imply, but not kick may take a clausal complement:

(40)a. John implied [that Mary was incompetent]

b.*John kicked [that Mary was incompetent]

This is a simple matter of c-selection. Notice that the same fact holds of the subject positions of these verbs:

(41)a. [that Mary was incompetent] implied the truth of my conjecture

b.*[that Mary was incompetent] kicked John in the shins

These facts too seem to be a matter of c-selection. If c-selection is part of subcategorization, then it can capture (40)a-b, but not (41)a-b. If c-selection is part of θ -theory, however, it can capture both contrasts.

We are not going to deny that there is a connection between c-selection and θ -theory, a connection to which we shall return at the end of the next chapter. Let us ask, however, how close the connection is. In particular, does c-selection fall under the Projection Principle?

If the answer is no, we do not have a disaster on our hands. The significant results of the Projection Principle, discussed above and by Chomsky, derive exclusively from the "positional" aspects of θ -theory, and not from any categorial aspect.

Suppose then that c-selection does not fall under the Projection Principle. C-selection properties would then not be projected to every level of representation. The only level at which satisfaction of c-selection must be relevant, as a minimal criterion of adequacy, is LF. If we were to assume that c-selection applied only at LF, then it would perforce be independent of θ -theory, at least at D-structure and S-structure. The θ -criterion at those levels would then be category-blind

in an important sense: we might expect to find derivations in which θ -roles are assigned in the syntax to constituents that are not of the normally required category -- so long as these derivations somehow satisfy categorial requirements by LF.

In Chapter Two, we shall show that Russian provides derivations of exactly this sort. Hence, we shall have empirical evidence that the satisfaction of thematic requirements is independent from the satisfaction of categorial requirements. We will thus argue that a new subsystem of grammar should be added to the roster: c-selection. On the other hand, in the final section of Chapter Two, we submit the primitives of our theory of c-selection to the test of epistemological priority and find them wanting. We present empirical arguments that the apparent c-selectional properties of predicates actually derive from semantic selection, as developed in Grimshaw (1973), and from Case theory.

CHAPTER ONE: FOOTNOTES

1. The θ -criterion is a strengthened version of principles first developed by Freidin (1979).

2. Chomsky later asks why NP and S should be the governing categories. This leads him to a theory incorporating the following definition (p. 211):

- (i) α is a governing category for β if and only if α is the minimal category containing β , a governor of β , and a SUBJECT accessible to β .

The term SUBJECT is a cover term for the lexical subject position or AGR. Thus, for the subject of a tensed S, that S is a governing category, because it contains a governor (AGR) and a SUBJECT (AGR). The notion "accessible" involves a particular filter on coindexing, and is relevant for some cases we shall discuss only marginally (in Chapter Four, 4.5). See Chomsky (1981a) for discussion; also Huang (1982) and Aoun (1982).

3. We omit discussion of cases in which non-arguments may be sub-categorized: e.g. the verb word, which for some speakers requires a manner adverb:

- (i) John worded the letter $\%$ (carefully)

Verbs like put, which require a directional expression, may actually involve θ -marking of a "small clause" headed by the directional. See discussion in Chapter Two.

4. We may wish to strengthen this to a biconditional, to fully subsume subcategorization: A position that can be directly θ -marked exists in a structure S if and only if an element θ -marks that position in S. Note that this rules out subject-to-object raising, as subcategorization and the Projection Principle does for Chomsky (1981a).

5. Note that subcategorization still exists as a theory: we are claiming that a more perspicuous presentation of that theory subsumes subcategorization under θ -theory. This is particularly true once we bleed the Projection Principle of categorial selection, as discussed below. Then the only content to subcategorization is the subject/object asymmetry in the obligatoriness of positions (pace footnote 4).

CHAPTER TWO: RUSSIAN QUANTIFICATION AND CATEGORIAL SELECTION

1.0 Three Descriptive Conditions

In this section, we will present three seemingly distinct quantificational constructions in Russian. We will demonstrate that these constructions all share a set of three odd properties, which appear to have nothing to do with each other. To begin with, we will provide a precise descriptive characterization of these properties within the general framework that we assume, but we shall wait until section 2 to begin showing how this framework explains the coincidence of properties.

1.1 Genitive of Negation

With preverbal (sentential) negation, certain nominal expressions in Russian may optionally appear in the genitive Case. We will call this construction the "genitive of negation":

- (1)a. ja ne polučal^v pis'ma
 I NEG received letters
 (acc pl)
- b. ja ne polučal^v pisem
 (gen pl)
- (2)a. ni odna gazeta ne byla polučena^v¹
 not one newspaper NEG was received
 (fem mon sg) (fem sg)
- b. ni odnoj gazety no bylo polučeno^v
 (fem gen sg) (neut sg)

Immediately below, we will be concerned with the distribution and interpretation of the genitive of negation. For the moment, we wish to emphasize two facts:

(1) The appearance of a phrase in the genitive Case is optional, as the (a) examples above show.

(2) When a genitive phrase alternates with a nominative subject, as in (2)a-b, the appearance of the genitive variant correlates with the absence of overt subject agreement on the verb. Thus, regardless of the person, gender or number of the genitive phrase in such examples, the verb will appear in the third person, neuter singular form. We take this form to be unmarked.²

Finally, a note about word order. The order of major constituents in a Russian clause is quite free. To emphasize the surface parallel between examples like (2)a, with a nominative subject, and (2)b, with a genitive phrase corresponding to the nominative subject, we have placed the genitive phrase before the verb. While (2)a represents an unmarked order for Russian, (2)b does not. In fact, the unmarked order for sentences like (2)b appears to be the opposite: such sentences are more common and more neutral with the genitive phrase following the verb (as noted, for example, by Peškovskij 1956, 367), as in (3) (cf. (2)b):

(3) ne bylo polučeno ni odnoj gazety

This fact will be very relevant later. For the time being, we will give examples of this type with the order genitive phrase + verb, as in (2)b.

We now turn to the important restrictions on the occurrence of the genitive of negation.

1.1.1 The D-structure [XP, VP] Restriction

We have already seen that genitive phrases under negation may

correspond to accusative direct objects, as in (1), or to nominative subjects, as in (2). With a few odd exceptions (cf. Ravič^V 1971), any accusative direct object may be "replaced" by a genitive phrase under negation. In contrast, "replacement" of a nominative subject by a genitive phrase under negation is quite restricted.

A variety of suggestions have appeared in the literature to account for the restrictions in question (see Babby 1980 for a survey). We will suggest that the restriction is to be described rather simply: genitive phrases under negation may correspond to S-structure nominative subjects only if those nominative subjects actually derive from D-structure objects.³ In this section, we present evidence for this statement of the restriction.

First, the occurrence of genitive phrases in passive constructions is as unrestricted as the occurrence of genitive phrases corresponding to accusative direct objects. Every passive verb, as far as I can tell, admits a genitive phrase under negation. Examples (4)a-b are from Chvany (1975):

- (4)a. ni odin gorod ne byl vzjat vragom
 not one city NEG was taken enemy
 (masc nom sg) (masc sg) (instr)
- 'not one city was taken by the enemy'
- b. ni odnogo goroda ne bylo vzjato vragom
 (masc gen sg) (neut sg)

Additionally, certain other monadic⁴ predicates admit the genitive of negation. This set is familiar: it is roughly the set of verbs that participate in impersonal constructions in languages like English (presentational there constructions) or French (impersonal il constructions),

and more roughly the set of verbs which take the auxiliary essere in Italian (cf. Perlmutter, to appear; Rosen 1981; Burzio 1981). These are mostly, but not exclusively, verbs of existence and of appearance:

(5)a. zdes' xorošie ljudi ne sušcestvujut
 here good people NEG exist
 (nom pl) (3 pl)

b. zdes' xorošix ljudej ne sušcestvuet
 (gen pl) (3 sg)

(6)a. griby zdes' ne rastut
 mushrooms here NEG grow
 (nom pl) (3 pl)

b. gribov zdes' ne rastet
 (gen pl) (3 sg)

(7)a. otvet iz polka ne prišel
 answer from regiment NEG arrived
 (masc nom sg) (masc sg)

b. otveta iz polka ne prišlo
 (masc nom sg) (neut sg)

(8)a. nikakie dokladčiki ne pojavilis'
 no speakers NEG showed up
 (masc nom pl) (pl)

b. nikakix dokladčikov ne pojavilos'
 (masc gen pl) (neut sg)

With examples (5)-(8), compare the following:

(9)a. v pivbarax kul'turnye ljudi ne p'jut
 in beerhalls refined people NEG drink
 (masc nom pl) (3 pl)

b.*v pivbarax kul'turnyx ljudej ne p'et
 (masc gen pl) (3 sg)

- (10)a. ni odin rebenok ne prygnul
 not one child NEG jumped
 (masc nom sg) (masc sg)
- b. *ni odnogo rebenka ne prygnulo
 (masc gen sg) (neut sg)
- (11)a. na zavode nikakie žensčiny ne rabotajut
 at factory no women NEG work
 (fem nom pl) (3 pl)
- b. *na zavode nikakix žensčin ne rabotaet
 (fem gen pl) (3 sg)

- (12)a. takie sobaki ne kusajutsja
 such dogs NEG bite
 (fem nom pl) (3 pl)
- b. *takix sobak ne kusaetsja
 (fem gen pl) (3 sg)

The verbs in (9)-(12), which do not allow the genitive of negation, contrast with the verbs in (5)-(8) in being obligatorily agentive. This generalization appears to hold over the entire range of monadic verbs in Russian, defining a necessary (though perhaps not sufficient⁵) condition for the genitive of negation construction. This fact has not been clearly noted by earlier investigators of the genitive of negation, mostly because of the absence of sure criteria to distinguish agentive from non-agentive arguments (cf. Babby 1980, 65). This is not unexpected: the semantic notion of agentivity clearly relies on inferences about the real-world status of predicates, and these are extralinguistic judgments.

Nonetheless, this generalization is easily confirmed by examining verbs that are ambiguous between agentive and non-agentive uses. A nice example is provided by the verb plavat'. This verb may be translated

into English as either 'swim' or 'float'. Suppose that these two English verbs have an identical core of meaning, having something to do with flotation. They differ only in that the notion of "swimming" normally involves the active participation of an animate body, while "floating" does not. In other words, plavat' as 'swim' assigns the θ -role of agent to its single argument, while plavat' as 'float' will assign some θ -role like theme.

If the genitive of negation construction is limited to arguments that are not assigned the role of agent, then we expect that this construction will disambiguate plavat'. This is exactly what we find:⁶

- (13)a. v bassejne nikakoj rebenok ne plavaet
 in pool no child NEG floats/swims
 (masc nom sg) (3 sg)
- b. v bassejne nikakogo rebenka ne plavaet
 floats/*swims
 (masc gen sg) (3 sg)

Also, compare the example given in Babby (1980, 18):

- (14) v supe ne plavalo nikakogo mjaso
 in soup NEG floated no meat
 (neut sg)(masc gen sg)

(14) shows a typical usage of the genitive of negation with this verb, to denote completely inactive flotation, characteristic of an inanimate object.

Examples like (13)-(14) thus justify our claim that the genitive of negation is possible with monadic verbs only if the genitive phrase is a non-agent. Nonetheless, the impossibility of the genitive of negation in (9)-(12) cannot be directly related to the semantic notion of agentivity. This can be seen by turning to transitive constructions.

Let us call verbs that take an accusative direct object and a nominative subject transitive polyadic verbs.⁷ We do this to reserve the term transitive for a more general case discussed below. Crucially, transitive polyadic verbs, regardless of the semantic status of their subjects, never allow genitive phrases to replace nominative subjects under negation. This is a truly exceptionless generalization, as noted first (perhaps) by Peškovskij (1956, 367):

(15)a. ni odna gazeta ne pečataet takuji erundu
 not one newspaper NEG prints such nonsense
 (fem nom sg) (3 sg) (fem acc sg)

b. *ni odnoj gazety ne pečataet takuju erundu
 (fem gen sg)

(16)a. studenty ne smotrjat televizor
 students NEG watch television
 (masc nom pl) (3 pl) (masc acc sg)

b. *studentov ne smotrit televizor
 (masc gen pl) (3 sg)

(17)a. ni odna devuška ne polučala naše pis'mo
 not one girl NEG received our letter
 (fem nom sg) (fem sg) (neut acc sg)

b. *ni odnoj devuški ne polučalo naše pis'mo
 (fem gen sg) (neut sg)

(18)a. takie mašiny ne proizvodjat vpečatlenie
 such cars NEG produce impression
 (fem nom pl) (3 pl) (neut acc sg)

b. *takix mašin ne proizvodit vpečatlenie
 (fem gen pl) (3 sg) (neut acc sg)

Semantically, neither the subject of polučat' 'receive' in (17)a nor the subject of the idiom proizvodit' vpečatlenie in (18)a is an agent. None-

theless, these nominative subjects may not be replaced by a genitive phrase under negation.

We are thus left with a curious distribution of genitive phrases under negation. Genitive phrases under negation may correspond to:

- (19) a. any accusative direct object
 b. any nominative subject of a passive verb
 c. a non-agent subject of a monadic verb
 d. no nominative subject of a transitive polyadic verb

What generalization can we make about this set of positions?

Excluding for the moment (19)c -- monadic verbs -- we ask what criterion distinguishes direct objects and subjects of passives, on the one hand, from subjects of transitive verbs, on the other. The answer is obvious. Both surface direct objects and surface subjects of passives are objects at D-structure. The surface subjects of transitive polyadic verbs are subjects at D-structure.

This analysis is forced on us by the Θ -criterion applying at the level of D-structure, as required by the Projection Principle. Let us see why this is so. To do so, we will need to introduce some of the principles of Case theory.

First let us consider a transitive polyadic verb and its passive counterpart. It is an elementary observation that the surface object of the transitive polyadic verb bears the same Θ -role as the surface subject of its passive counterpart. Thus, city bears the same Θ -role in (20)a and (20)b, as does gorod in the Russian translation in (21)a and (21)b:

- (20)a. the enemy took the city
 b. the city was taken

- (21)a. vrag vzjal gorod
 enemy took city
 (nom) (acc)
 b. gorod byl vzjat
 city was taken
 (nom)

We can explain this easily if (20)a has the D-structure in (22)a, while (20)b has the D-structure in (22)b:

- (22)a. [_S the enemy INFL [_{VP} took the city]]
 b. [_S e INFL [_{VP} was taken the city]]

In each case take directly θ -marks city. By the Projection Principle, the S-structures are as in (23):

- (23)a. [_S the enemy INFL [_{VP} took the city]]
 b. [_S the city_i INFL [_{VP} was taken e_i]]

In (23)b, the city is directly θ -marked by take by membership in the chain (the city, e_i).

Descriptively, the passive b sentences differ from the active a sentences in two respects:

- (24) I. The VP in the active structure assigns a θ -role to the subject position. The VP in the passive structure does not.
 II. A D-structure NP object of a passive verb must move to subject position in the syntax.

As Chomsky (1981a) notes, II implies I. If the D-structure object of the passive verb were required to move into subject position, and if the subject position were a θ -position, the Projection Principle would be violated. Recall from Chapter One that no movement is possible into a θ -criterion.

Property II derives from Case Theory, a subsystem of grammar that we have not discussed as yet. In its simplest form, Case Theory consists of a set of rules for the assignment of Case features and a Case Filter. Case features are assigned under government. In English, the following conventions may be assumed:

(25) Case Assignment (English)

- i. If NP is nominative, then it is governed by AGR.⁸
- ii. If NP is objective, then it is governed by [-N].⁹
- iii. If NP is neither nominative nor objective, it lacks Case.

(26) Case Filter

*NP, if NP has phonetic content and has no Case.

(25)-(26) account in large part for the distribution of lexical NPs in English, as noted first by Vergnaud (1982).¹⁰ Thus, a lexical NP may be the subject of a tensed S with agreement, or the complement of certain verbs and prepositions. The NP to which the reader's attention is drawn is underlined:

- (27) a. [_S John [_{INFL} TNS AGR] [_{VP} ate meat]]
 b. it is difficult [_S, PRO to [_{VP} suggest an answer to our question]]
 c. John is [_{PP} in the room]

A lexical NP cannot be the ungoverned subject of an infinitival S' nor can it be the complement of a noun or adjective, which are [+N].

- (28)a. *[_S, [_S John [_{INFL} to] [_{VP} eat meat]]] would surprise me
 b. *this observation is [_{AP} suggestive an answer to our question]]
 c. *John's [_N, suggestion an answer to our question]

In structures like those of (28)b-c, a dummy [-N] Case marker must be inserted, so that the lexical NP in question will not violate the Case Filter:

- (29)a. this observation is [_{AP} suggestive [of an answer to our question]]
 b. John's [_N, suggestion [of an answer to our question]]

Positions in which no Case is assigned are non-Case-marked positions. The insertion of a dummy Case marker in (29) transforms a non-Case-marked position into a Case-marked position, and allows the underlined lexical NP to satisfy the Case Filter.

An NP in a non-Case-marked position may satisfy the Case Filter in another way, however. It may move to a Case-marked position, if such movement is not ruled out by other principles like the Binding Theory or Projection Principle. Such is the case in the passive construction.

We have phrased our Case Assignment conventions as conditionals, not bi-conditionals, allowing some [-N] elements not to assign Case. Suppose in particular that passive verbs differ from active verbs in not assigning Case to the object position that they govern. In that Case, a D-structure like (30) cannot yield a well-formed S-structure, without something happening that gives the underlined NP Case:

(30) [_S e INFL [_{VP} was taken the city]]

In (30), one thing that can happen is movement to the position of e, which is governed by AGR and marked with nominative Case:

(31) [_S the city_i INFL [_{VP} was taken e_i]]

While it remains unexplained why no dummy Case marker can save (30) without movement in English, Case theory provides a generally elegant explanation of property (24)II of passive verbs: the D-structure object of the passive verb must move to subject position in order to receive Case. If we assume that the Case Filter applies at least at S-structure, we explain why this movement must apply in the syntax.

Another factor leads us to believe that Case Theory is behind the obligatory movement of an NP object of a passive verb. Notice that the Case Filter does not apply to non-NPs, a fact that will be extremely important to us in this chapter. This is a stipulation, of course, but it is a stipulation of some explanatory depth. (We will revise the stipulation somewhat later, but it will remain unexplained.)

Notice that an S' has no problems as the complement of a [+N] category:¹¹

(32)a. John is [_{AP} happy [_S, that we have arrived on time]]

b. John's [_N, suggestion [_S, that bear-baiting is immoral]]

No dummy Case marker is necessary, suggesting that S' does not fall under the Case filter. Now notice that S' need not move in passive constructions:

(33) it was [_{VP} suggested [_S, that bear-baiting is immoral]]

The only rule specific to (33) is a general rule inserting expletive it in the non- θ -subject position, necessary in English for independent reasons, to which we return later. E. Williams has observed (Chomsky 1981a, 149 note 121) that some verbs do not even allow movement of S':

- (34)a. it was reasoned (held, rumored...) that John arrived
 b. *that John arrived was reasoned (held, rumored...)

This rules out an analysis in which movement to subject is followed by extraposition.

The Case Assignment conventions and the Case Filter of (25)-(26) come close to being adequate for Russian as well, except that [+N] elements and some [-N] elements can assign oblique Cases to their NP:¹²

(35) Case Assignment (Russian)

- i. If NP is nominative, then it is governed by AGR.
- ii. If NP is accusative, then it is governed by [-N].
- iii. If NP is oblique, then it is governed.
- iv. If NP is neither nominative, accusative or oblique, then it lacks Case.

The oblique Cases are genitive, dative, instrumental and prepositional (also known as locative). In certain instances, accusative might also function as an oblique Case, when assigned by a preposition, as we shall see later. We assume, of course, the same Case Filter for Russian as for English.

Returning to our main topic, we see that we can explain the surface syntax of passive constructions with nominative subjects if we assume that their S-structure subjects are D-structure direct objects. Passive

verbs, like their active counterparts, may thus be transitive. Russian also allows impersonal passives to some degree, which take no direct object at any level:

(36) bylo napisano [ob ètom] [v gazete]
 was written about this in newspaper
 (prep) (prep)

These constructions are not relevant to our discussion, and will be ignored.

We thus justify grouping together surface direct objects with subjects of passive verbs as D-structure objects. Both surface direct objects and subjects of passive verbs may be replaced by the genitive under negation. By contrast, the transitive polyadic verbs like polučat'^V 'receive' assign a θ -role to their surface subject, indicating that their surface subject is also a D-structure subject. These subjects may not be replaced by the genitive under negation.¹³

We can make our analysis of the D-structure of transitive polyadic verbs tighter by referring to a generalization discovered by Burzio (1981). Burzio observes that, as a general rule, verbs which assign Case to their objects generally form VPs which indirectly θ -mark their subject position. All the transitive polyadic verbs whose subject cannot be replaced by the genitive of negation are of this type. Thus, the subjects of all these verbs that take accusative direct objects receive a θ -role at D-structure. The S-structure occupant of such a subject position cannot have moved into that position, by the Projection Principle, and thus cannot be a D-structure object.

Thus, the Θ -criterion and the Projection Principle, aided by Case Theory and Burzio's generalization (which we discuss in greater detail below), allow us unify (19)a, b and d, repeated below, under the generalization in (37):

(19) Genitive phrases under negation may correspond to:

- a. any accusative direct object
- b. any nominative subject of a passive verb
- d. no nominative subject of a transitive polyadic verb

(37) A genitive phrase under negation is a D-structure [XP, VP], where XP is some (unknown) category.

Let us now return to (19)c:

(19) c. a non-agent subject of a monadic verb

How can we assimilate (19)c to our generalization in (37)? We wish to capture the fact that a genitive phrase under negation corresponds to the argument of a monadic verb if and only if that argument is not an agent. (19)c falls under generalization (37) if the following statement is true:

(38) If the surface subject of a monadic verb is an [XP, VP] at D-structure, then it is not an agent.

Since a monadic verb has only one argument, (38) may be stated more generally as (39):

(39) If the argument of a monadic verb is an [XP, VP] at D-structure, then it is not an agent.

In other words, in order to extend our D-structure generalization about passives and polyadic transitives to monadic verbs like those in (5)-(12), we must assume that the S-structure nominative subjects of some, but not all, monadic verbs are actually D-structure objects. Only these nominative subjects may correspond to genitive phrases under negation.

Thus, the D-structure of (7)a, with a nominative subject for prišel^V 'arrived', differs from the D-structure of (11)a, with the verb rabotajut 'work', in that the former, but not the latter, may be transitive in D-structure:

- (40)a. [_S e [_{VP} ne prišel^V [otvet iz polka]]] (D-structure for (7)a)
 NEG arrived answer from regiment
- b. [_S nikakie žensčiny^V [_{VP} ne rabotajut^{VV}]] (D-structure for (11)a)
 no women NEG work

We may look at (39) in another way, since D-structure is a pure representation of the assignment of Θ -roles to GFs. What (39) says is that the Θ -role of agent is always indirectly assigned by a phrasal node, while non-agent roles may be directly assigned by a lexical node, where direct and indirect Θ -marking are defined as in Chapter One. Monadic verbs thus fall into two classes. Some take an agent argument, in which case the Θ -role is assigned by the VP containing the monadic V. Others take a non-agent argument, in which case the Θ -role may be assigned by the V itself.

This distinction between two types of monadic verbs is not novel. Its consequences have been worked out in great explanatory detail for Italian by Burzio (1981) within the Extended Standard Theory. As the

"Unaccusative Hypothesis", the idea was earlier explored in the literature on Relational Grammar (Perlmutter, forthcoming; Rosen 1981). With reference to the genitive of negation in Russian, the distinction dates at least to Ružicka (1963), and is adopted by Chvany (1975).

We will use Burzio's terminology, and call verbs like prisel^V in (40)a, which are underlyingly transitive, ergative verbs. An ergative verb is a monadic transitive verb. Verbs like rabotajut in (40)b, which are underlyingly intransitive, we may call simply intransitive or agentive intransitive verbs.¹⁴

What principle requires D-structure objects of ergative verbs like prisel^V to move to subject position in syntax? That they are in subject position at S-structure is clear from their nominative Case marking, as well as from other aspects of their behavior which we shall discuss later. Following Burzio, we take the principle in question to be the Case Filter. Burzio argues that ergative verbs, like passives, do not assign Case to the objects they govern. Hence, the object of an ergative verb must move to a Case-marked position by S-structure. In this respect, the derivation of sentences with ergative verbs is identical to that of passive sentences. In each case, a non-Case-marked NP must move to subject position to receive Case. In each case, also, the subject position must be a non- θ -position, or else the θ -criterion and the Projection Principle will be violated.

Returning to the generalization in (39), we might ask why there should be a connection between the position to which a θ -role is assigned and the essentially semantic notion of agentivity. This question relates in turn to the question of how a child acquires knowledge of the syntactic

properties of lexical items. Somehow, the child progresses from knowledge, perhaps situationally based, of the meaning of a predicate to knowledge of the Θ -properties of that predicate. Ideally, this "grammaticalization" of meaning should be guided by UG, making the construction of a lexical entry trivially simple. We therefore expect UG to contain principles relating semantic properties to Θ -properties of lexical items. In turn, UG may contain other principles relating Θ -properties of lexical items to even more specifically linguistic properties, like Case assignment. For example, we have already seen a generalization of this second type, suggested by Burzio:

(41) Burzio's Generalization

For a verb V_i , if V_i assigns Case to an NP it directly Θ -marks, then VP_i indirectly Θ -marks its subject.

This principle, almost but not quite exceptionless,¹⁵ told us that the surface subjects of transitive verbs are also D-structure subjects, and are not candidates for the genitive of negation.

As for generalizations of the first type, relating semantic properties of lexical items to Θ -properties, (39) is a candidate for such a generalization. Probably, (39) is a special case of a more general Agent Rule, which we may state as a descendent of Anderson's (1979) rule of the same name:

(42) Agent Rule

If α assigns the Θ -role agent to β , then α indirectly Θ -marks β .¹⁶

This tells us, not only that a verb like work cannot be ergative, but also

that of the two arguments associated with the active verb kick, the kicker will be the subject of the sentence, and, by default, the kicked will be the object. (cf. Marantz 1981).

Notice that the Agent Rule allows monadic verbs whose one argument is non-agentive to θ -mark this argument either directly or indirectly. Thus, a verb like prisel^V 'arrived' might take either a D-structure subject or a D-structure object, being either intransitive or ergative. Is this freedom correct? We could force prisel^V to be exclusively ergative with a Theme Rule:

(43) Theme Rule

If α assigns the θ -role theme to β , then α directly θ -marks β .

An advantage of having both a theme rule and an agent rule is that we can predict the argument structure of a verb like receive, which takes a goal and a theme, but (presumably) no agent. By (43), we know that the theme will be the object; by default, the goal will be the subject.

For our purposes, it does not really matter whether or not UG or the core grammar of Russian contains a Theme Rule, so long as it contains an Agent Rule. Russian provides tests that show whether a predicate may be ergative, it provides no tests that show whether a predicate must be ergative. We are free to assume that every ergative verb has an intransitive counterpart, if we wish.

Without deviating too much from our main argument, we should add one remark about the status of the Agent Rule (and the Theme Rule, if it exists), which is relevant to anyone working on the genitive of negation construction in Russian. Rules like the Agent Rule may be interpreted in two ways. They may be seen as generalizations over lexical entries for

Θ -assigners, which allow redundant information to be omitted from the lexicon. On this view, the Agent Rule renders otiose a lexical entry for a verb like rabotat' 'work' that says both that rabotat' indirectly Θ -marks its argument and that it assigns the role agent. By the Agent Rule, the lexicon need only say that the role assigned is that of agent, and the fact that it is indirectly assigned follows independently.

Alternatively, rules like the Agent Rule may be seen as principles of language acquisition, helping the child learn a lexicon quickly, and not as principles operative in the grammar acquired. On this view, the lexicon will state both that rabotat' indirectly Θ -marks its argument and that it assigns the Θ -role of agent.

Comparison of Russian with a language like Italian suggests that a principle like the Agent Rule is always a principle of language acquisition, but that its status as a principle of the grammar acquired, simplifying lexical entries, varies across languages. In a language like Russian, the distinction between ergative and intransitive monadic verbs, as seen through the genitive of negation, is fuzzy, and heavily dependent on context and the speakers perception of the degree of agentivity involved in the action denoted by the verb. Although almost all speakers agree on the extreme cases -- that verbs meaning 'arrive' accept the genitive of negation, and that verbs meaning things like 'bite' do not -- there is much disagreement on the cases in between. This is expected if the Agent Rule is a principle of grammar, relating semantics directly to GF- Θ in the core grammar of Russian. As a speaker's perceptions of the predicate will vary accordingly, and be reflected in the possibility of the genitive of negation.

In a language like Italian, on the other hand, where the ergative/intransitive distinction is well-motivated by a number of tests, speakers' judgments about individual predicates are not at all fuzzy. While some predicates function both as ergatives and intransitives, the assignment of a predicate to one or the other or both classes is fully lexicalized and reflected in the selection of verbal auxiliary (essere for ergatives, avere for intransitives). While in Italian, as in Russian, something like the Agent Rule probably governs the acquisition of lexical entries, the Agent Rule may not be part of the grammar acquired. Judgments about ergativity/intransitivity in Italian do not seem dependent on semantic criteria and on context, but rather on a fully specified lexical entry for each verb. Hence the sharpness of the judgments in Italian.

This difference between Italian and Russian is not surprising. Suppose, with Chomsky and Lasnik (1977), that a child cannot learn from negative evidence -- i.e. from the fact that he does not hear a sentence. In Russian, ergative verbs are distinguished from intransitives only by the absence of sentences with constructions like the genitive of negation. This absence does not help the child at all. Rather, only lexical semantics guides the classification of monadic verbs into ergative and intransitive categories. In Italian, on the other hand, monadic verbs are readily distinguishable on the basis of the auxiliary that they take, a pervasive and easily observable phenomenon in the language. The Italian child can (and must) grammaticalize the ergative/intransitive distinction, removing it from the domain of lexical semantics, in order to choose correctly an auxiliary verb.

Returning to our main discussion, we can now make all the statements

about the distribution of the genitive of negation in (19) fall under generalization (37), which we will restate and call condition A:

A: XP must be an [XP,VP] at D-structure, where XP = genitive phrase under negation.

Before discussing some other properties of the genitive of negation construction, let us extend this generalization to three other environments.

First, recall that the Agent Rule in (42) does not prevent non-agents from being indirectly Θ -marked, in addition to agents. Consider not the following principle:

(44) Attributive and Identificational Rule

If α assigns the Θ -role of attribute or identificand to β , then α indirectly Θ -marks β .¹⁷

We suppose that the role of attribute is the Θ -role that most adjectives assign to their surface subjects, as in (45)a. The role of identificand is the role of the subject in identificational sentences like (45)b:

- (45)a. Boston is clean (attribute)
 b. these young people are our freshman (identificand)

(44) tells us that the S-structure subjects in (45)a-b are also D-structure subjects. We thus expect, correctly, that attributes and identificands may not correspond to genitive phrases under negation in Russian:

- (46)a. ni odin gorod ne byl $\overset{V}{\text{čist}}$
 not one city NEG was clean
 (masc nom sg) (masc sg)
- b. *ni odnogo goroda ne bylo $\overset{V}{\text{čisto}}$
 (masc gen sg) (neut sg)

- (47)a. studenty zdes' — ne duraki
 students here NEG fools
 (masc nom pl) (masc nom pl)
- b. *studentov zdes' — ne duraki
 (masc gen pl)
- (48)a. `eti molodye ljudi — ne naš^vi pervokursniki
 these young people NEG our freshmen
 (masc nom pl) (masc nom pl)
- b. *etix molodyx ljudej — ne naš^vi pervokursniki
 (masc gen plur)

The (b) sentences are excluded by condition A and by the Attributive and Identificational Rule. The relevant condition cannot be morphological. The Russian passives we have been considering are formed with the copula (null in the present tense) and a passive participle. These constructions are indistinguishable on the surface from copular constructions with an adjectival predicate. Nonetheless, as we have seen, the subjects of passives, but not of adjectives, may be replaced by a genitive phrase under negation. This is because passive participles and adjectives, though morphologically indistinguishable, have radically different θ -properties.

Interestingly, passive participles in Russian, as in English, can be "adjectivized", often with a slight shift in meaning (see Kalakuckaja 1971, for a thorough study). Predictably, adjectivization of a passive participle produces a predicate which is subject to the Attribute and Identificational Rule and which does not allow the genitive of negation. For example, the passive participle prinjat 'accepted' may be used as an adjective, with roughly the meaning it acquires in similar circumstances in English -- 'acceptable', 'comme il faut' (as in the accepted norms). The two usages contrast with respect to the genitive of negation, as expected:

- (49)a. takie studenty nikogda ne prinjaty v universitet
 such studente never NEG accepted in university
 (masc nom pl) (pl)
 'such students are never accepted to the university'
- b. takix studentov nikogda ne prinjato v universitet
 (masc gen pl) (neut sg)
- (50)a. takie manery nikogda ne prinjaty v xorošix klubax
 such manners never NEG acceptable in good clubs
 (fem nom pl) (pl)
- b. *takix maner nikogda ne prinjato v xorošix klubax
 (fem gen pl) (neut sg)

The second environment that we will consider is the object position of a PP. Here, the object is not an [XP, VP], and we predict that the genitive of negation should be impossible. This prediction is correct:

- (51)a. ja nikogda ne namekal ni na kakie plany po ètoj linii
 I never NEG hinted not at any plans along this line
 Masc nom pl)
- b. *ja nikogda ne namekal ni na kakix planov po ètoj linii
 (masc gen pl)

Although this correct prediction is made by condition A, we will be accounting for these facts somewhat differently in later discussion.

Finally, condition A predicts that if a non-argument should also be found in VP, as in [XP, VP], then this non-argument too should be able to participate in the genitive of negation construction, all things being equal. In Russian, expressions of duration, which are not θ -marked or selected by V, are assigned accusative Case, presumably because they are governed by V. Such expressions of duration may be marked genitive under negation, exactly as expected (sentences due to M.G. Sinicyn):

- (52)a. ja ni odnu minutu ne spal
 I not one minute NEG slept
 (fem acc sg)
- b. ja ni odnoj minuty ne spal
 (fem gen sg)

We will provide some evidence that these expressions of duration are indeed [XP, VP] later. In footnote 21, we also discuss a semantic difference between (52)a and (52)b. For now, note simply that under the assumption that these phrases are in VP, we can account for the possibility of the genitive of negation, given condition A. Notice that these phrases show that the condition on the genitive of negation is purely structural, and not relational in nature, confirming our statement in condition A.

1.1.2 The Non-Obliqueness Restriction

We now turn to another restriction on the genitive of negation. As we have shown, condition A allows direct objects to occur freely in the genitive of negation construction. In Russian, as in many languages with overt Case marking, some predicates require a particular oblique Case of nouns they θ -mark.

As Babby (1980) has observed, such oblique objects never participate in the genitive of negation construction. This Non-Obliqueness Restriction (Babby's "Direct Case Condition") is exceptionless, and we may state it as condition B:

- B: XP may not occur in positions where oblique Case is required, where XP = genitive phrase under negation.

Thus, for example, the verb pomogat' 'help' requires a dative object,

and the verb upravljat' 'manage', 'control' requires an instrumental object. Neither of these verbs allows a genitive phrase as an object under negation:

(53)a. ja ne pomagaju nikakim devuškam^V
 I NEG help no girls
 (fem dat pl)

b. *ja ne pomagaju nikakix devušek^V
 (fem gen pl)

(54)a. ja ne upravljaju ni odnim zavodom
 I NEG manage not one factory
 (masc instr sg)

b. *Ja ne upravljaju ni odnogo zavoda
 (masc gen sg)

Condition B will play an essential role in the discussion in section 2 of this chapter.

1.1.3 Obligatory Quantifier Movement

Up to this point, we have been deliberately vague about the interpretation of sentences with the genitive of negation. In fact, the interpretation of genitive phrases under negation is more restricted than that of their non-genitive counterparts.

In the traditional literature about this construction, it is usually said that genitive phrases under negation, unlike their non-genitive counterparts, are obligatorily indefinite. (Note that Russian has no articles to signal definiteness.) Thus, consider the following examples (with inverted order in both, which makes the indefinite reading more

apparent)¹⁹:

- (55)a. ne pojavilis' studenty = (56)a, (56)b
 NEG showed up students
 (pl) (masc nom pl)
- b. ne pojavilos' studentov = *(56)a, (56)b
 (neut sg) (masc gen pl)

(55)a may have as a logical paraphrase either (56)a, or (56)b below.

(55)b, however, with the genitive of negation, may be paraphrased only by (56)b:

- (56)a. \neg the students showed up (cf. 'the students didn't show up')
- b. $\neg \exists x, x$ students (x showed up) (cf. 'no students showed up')

Essentially the same data hold when a genitive phrase under negation replaces an accusative object:

- (57)a. ja ne polučal^V pis'ma = (58)a, (58)b
 I NEG received letters
 (neut acc pl)
- b. ja ne polučal^V pisem = *(58)a, (58)b
 I NEG received letters
 (neut gen pl)

For reasons we will not be able to explain, the reading of (58)b below is much less accessible for (57)a than was the corresponding reading (56)b for (55)a.²⁰ Nonetheless, (58)b is a possible reading for (57)a, while (58)a is completely excluded for (57)b:²¹

- (58)a. \neg I received the letters (cf. 'I didn't receive the letters')
- b. $\neg \exists x, x$ letters (I received x) (cf. 'I didn't receive any letters')

The (b) readings in (56) and (58) differ from the (a) readings in several important ways. We will focus on only one of them, for the most part. The crucial difference between the (b) readings and the (a) readings is the following: the (b) representations are quantifier-variable structures, while the (a) representations are not. Also important is the fact that the quantifier appears to be existential and the fact that negation always has scope over the quantifier (as shown by Babby 1980). We will have more to say about both these facts later, but they will not be central to our discussion.

We will want a principle that will tell us that the genitive of negation obligatorily produces a quantifier-variable structure, while corresponding sentences without the genitive of negation produce such structures only optionally. Following proposals of Chomsky (1976), and particularly of May (1977), we will assume that quantificational structures like those in (56)b and (58)b actually correspond closely to representations at the level of LF. In particular, quantifier-variable structures are derived at LF from S-structures with the quantifying expression in situ by means of an adjunction rule: May's Quantifier Rule (QR). We will assume that it applies roughly as in (59):

(59) QR: Adjoin a quantifying expression to S.

Very possibly, as argued in May (forthcoming), QR is a special case of Move α at LF, but we will refer to it as if it were a special rule by itself. The trace left by this movement is interpreted as the variable bound by the moved quantifier.

We will have much more to say about QR quite soon, and we shall be

more specific about the LF representations we are assuming for the genitive of negation. For now, let us concentrate on how we can state the difference in interpretation between sentences with and without the genitive of negation.

We may capture our generalization about the LF interpretation of the genitive of negation construction in the following way: QR applies optionally to non-genitive phrases under negation, but obligatorily to genitive phrases in the genitive of negation construction.

Thus, QR applies optionally in (55)a, yielding either a representation corresponding to (56)a or a representation corresponding to (56)b. In (55)b, QR applies obligatorily, yielding only a representation like (56)b. A parallel story is told for (57)-(58). Since the optionality of QR in the normal case follows from the general optionality of movement rules, we need only state a generalization about the genitive phrases under negation, which we may call condition C:

C: XP obligatorily undergoes QR, where XP = genitive phrase under negation.

We have now examined briefly three conditions that govern the occurrence of the genitive of negation. The most elementary considerations of learnability tell us that these conditions must result from deeper principles. Why should these conditions exist? If a child simply "learns" a construction like the genitive of negation, why should he not generalize from sentences with an ergative verb to sentences with a non-ergative intransitive verb, eliminating condition A? How does the child know that the construction does not extend to environments in which an oblique Case is required? How can he learn that the genitive of negation requires the

creation of a quantifier-variable structure in LF? These questions become still more pressing when we note that a number of other constructions show the same coincidence of properties, albeit somewhat more weakly. We will discuss two such constructions in the next sections.

1.2 Po-phrases

Russian expresses distributive quantification by means of a "particle"

po:

(60)a. ja dal mal'^vcikam jabloko
 I gave boys apple
 (masc dat pl) (neut acc sg)

'I gave the boys (collectively) an apple'

b. ja dal mal'cikam [po jabloku]
 po apple
 (masc dat sg)
 'I gave the boys an apple each'

Po is a preposition in other usages, and may be a preposition in the construction we will consider here, an issue we will discuss later. Like po in its clearly prepositional usages, po assigns dative Case in sentences like (60)b. Po has some peculiar properties with respect to Case marking, discussed by Mel'^vcuk (1981), and by Babby (1980). Although these properties are clearly of interest, we will ignore them here, since they are rather wild in ways that are probably tangential to our discussion. Instead, we will be mainly interested in the external distribution of distributive po-phrases. We will show that they are subject to conditions A, B and C, which we have seen to govern the genitive of negation. We present the evidence briefly.

1.2.1 Po-Phrases and the D-structure [XP, VP] Restriction

Distributive po-phrases, like genitive phrases under negation, are limited to D-structure positions in VP, although the relevant contrasts are less clear here for some speakers. For all speakers, po-phrases freely correspond to direct objects of polyadic transitive verbs, to subjects of passive and ergative verbs, and to accusative expressions of duration:

- (61)a. ja polučal [po pis'mu] v den'
 I received po letter in day
 (nom (masc sg) (neut dat sg))
 'I received a letter each day'
- b. každyj den', [po gorodu] bylo vzjato vragom
 each day po city was taken enemy
 (masc dat sg) (neut sg) (masc instr sg)
 'each day, a (different) city was taken by the enemy'
- c. [po jabloku] upalo s každogo dereva
po apple fell from each tree
 (neut dat sg) (neut sg)
 'a (different) apple fell from each tree'
- d. ja spal [po času] v den'
 I slept po hour in day
 (nom) (masc dat sg)
 'I slept an hour per day'

(Example (61)c is from Babby 1980, 45.)

Note that, like genitive phrases under negation, distributive po-phrases do not trigger agreement on the verb when they correspond to nominative subjects.²² In addition, the unmarked positions for a po-phrase appears to be post-verbal -- again recalling the genitive of negation.

For almost all speakers, po-phrases are markedly worse as subjects

of polyadic transitive verbs, agentive intransitives, subjects of attributive and identificational predicates, or as objects of prepositions. In the first two cases, judgments vary from a single question mark to a star; we compromise below:²³

- (62)a. ??[po studentu] ubilo košku v každoj gruppe
po student killed cat in each group
 (masc dat sg) (neut sg) (fem acc sg)
 'a (different) student killed a cat in each group'
- b. ??[po sobake] kusaetsja v každoj kletke
po dog bites in each cage
 (fem dat sg) (3 sg)
 'a (different) dog bites in each cage'
- c. *[po gorodu] bylo čisto v každoj oblasti
po city was clean in each district
 (masc dat sg) (neut sg) (neut sg)
 'a (different) city was clean in each district'
- d. *[po studentu] — pervokursnik v každom otrjade
po student (is) freshman in each detachment
 (masc dat sg) (masc nom sg)
 'a (different) student is a freshman in each detachment'
- e. *ja namekal na [po planu] v den'
 I hinted at po plan in dat
 (masc dat sg)
 'I hinted at a plan per day'

In Chapter Four, we present an elaboration of the theory we will develop in this chapter, which accounts for the relative weakness of the judgments in (62)a-b. In this chapter, we will treat them as ungrammatical.

The contrasts between (61)a-d and (62)a-e suggest that we should extend condition A to include distributive po-phrases. Po-phrases must thus be included with genitive phrases under the rubric "XP":

A: XP must be an [XP, VP] at D-structure, where XP = genitive phrase under negation or a distributive po-phrase.

1.2.2 Po-phrases and the Non-Obliqueness Restriction

In section 1.1.2, we saw that genitive phrases under negation can never occur in direct object position where an oblique Case is required by the governing verb. This absolute prohibition extends to po-phrases. Recall that pomogat 'help' requires a dative object (cf. (53)), and that upravljat 'manage' requires an instrumental object (cf. (54)). Distributive po-phrases, like genitive phrases under negation, cannot correspond to the object of either of these verbs (despite the fact that po itself assigns dative Case):

(63)a. *ja pomogal [po devuške] v den'
 I helped po girl in dat
 (nom) (fem dat sg)
 'I helped a girl a day'

b. *ja upravljaju [po zavodu] v každy gorode
 I manage po factory in each city
 (nom) (masc dat sg)
 'I manage a (different) factory in each city'

We should thus generalize condition B as we generalized condition

A:

B: XP may not occur in positions where oblique Case is required, where XP = genitive phrase under negation or a distributive po-phrase.

1.2.3 Po-phrases and Obligatory QR

It should be clear that the LF representation of a sentence with distributive po must be a quantifier-variable structure. An accurate account of the logical form of po-sentences goes somewhat beyond the scope of this study (cf. Crockett 1976a for some discussion). One possibility, probably incorrect, is represented in (64)b:

(64)a. ja dal mal'^Vcikam [po jabloku]
 I gave boys po jabloku
 (nom) (masc dat pl) (neut dat sg)

'I gave the boys an apple each'

b. $(\forall x: x \text{ a boy}) (\exists! y: y \text{ an apple})$ I gave y to x

(64)b does not imply a one-to-one correspondence between boys and apples. If such a correspondence is implied by po-phrases, as it might be, we might represent po as a two-place quantifier resulting from the operation of "absorption" discussed by Higginbotham and May (1981):²⁴

(65) PO(x, y): x a boy, y an apple (I gave x to y)

Po defines a bijective function from values of x to values of y, and entails (correctly, it seems) that the LF of sentences like (64)a contains at least two variables in its LF, accounting for ambiguities of the sort discussed by Crockett (op. cit., (4)).

What is clear, and important for our purposes, is that QR is obligatory for distributive po-phrases. In the case of po-phrases, unlike the genitive of negation, this fact might be trivial, since po is probably uninterpretable without QR. Nonetheless, as we shall see in section 1.3.3, other quantifiers can receive an interpretation as

part of a name, in the absence of QR -- an interpretation apparently lacking in the case of po-phrases. Thus, we take it as a fact to be accounted for that po-phrases must undergo QR, and we extend condition C, in turn, to include po-phrases:

C: XP obligatorily undergoes QR, where XP = genitive phrase under negation or a distributive po-phrase.

We have seen that po-phrases should be included in the class of XPs subject to conditions A, B, and C. In the next section, we turn to another construction that also obeys these three descriptive conditions. Here, the evidence will be somewhat less direct.

1.3 No-agreement Numeral Phrases

Numeral phrases in Russian (translations of six men, two cars, etc.)²⁵ have an involved internal syntax, which we shall avoid discussing as much as possible. Many of the complications are apparently artifacts of the prescriptive-grammatical tradition, and do not seem to represent "core" facts about Russian -- confusing native speakers almost as much as foreigners. For a useful discussion of some of these peculiarities, particularly those about which native speakers do have intuitions, see Crockett (1976)b. In the following sections, we shall examine certain properties of simple numeral phrases, which consist of a numeral followed by an unmodified noun.

First, we declare our intention to ignore the numeral odin '1', and all numerals ending in odin. Numeral phrases with odin have all the properties of ordinary NPs, in which odin functions as an adjectival modifier. None of the following discussion of numeral phrases applies

to phrases with odin. In the discussion that follows, we will use the term "numeral phrase" with the understanding that phrases with odin are excluded.

When (other) numeral phrases correspond to nominative or accusative NPs, the numeral appears in its citation form. We shall be cautious about calling this form either nominative or accusative, for reasons that shall become apparent in section 2. A noun following the numeral bears genitive Case. This genitive noun will be singular or plural, depending on the numeral preceding it: 2 through 4 and numerals ending in 2-4 require the genitive singular, while 5-20, numerals ending in 5-20, multiples of 10 and numerals ending in multiples of 10 require the genitive plural:

- (66)a. ja polučil^v [tri priemnika]
 I received three radio
 (masc gen sg)
- b. ja polučil^v [šest' priemnikov]
 I received six radios
 (masc gen pl)

The facts that we shall discuss are clearest with respect to those numerals which require the plural of the following genitive. We shall accordingly choose our examples from that class.

Recall that when either a genitive phrase under negation or a distributive po-phrase corresponds to a nominative subject, the verb of its clause does not show any sort of agreement; rather, it is found in an unmarked, third person neuter singular form. When a numeral phrase corresponds to a nominative subject, the verb may either show normal

agreement of the sort expected with a plural subject, or may lack agreement:

- (67)a. [^Všest' studentov] prišlo
 six students arrived
 (masc gen pl) (neut sg)
- b. [^Všest' studentov] prišli
 (pl)

We will suggest that numeral phrases in sentences like (67)a -- which correspond to nominative subjects but do not trigger agreement -- are XPs. We will demonstrate that these numeral phrases, which we shall call no-agreement numeral phrases, obey conditions A, B and C. By contrast, numeral phrases that do trigger agreement, as in (67)b, do not obey conditions A, B and C. These numeral phrases, which we may call agreement numeral phrases, thus give no sign of being XPs.

Independently of conditions A, B, and C, we may note a first piece of circumstantial evidence linking no-agreement numeral phrases to XPs. Recall that genitive phrases under negation and po-phrases are most natural when postverbal. As noted by Svedova (1970, 554), Revzin (1978, 270), Corbett (1979, 81), and others, the same is true of no-agreement numeral phrases. Agreement numeral phrases, on the other hand, are most natural preverbally. Thus, (68) represents the natural orders for (67)a-b:

- (68)a. prišlo [^Všest' studentov]
 (neut sg)
- b. [^Všest' studentov] prišli
 (pl)

In the following sections, we show that no-agreement numeral phrases resemble genitive phrases under negation and po-phrases in other ways as well.

1.3.1 No-agreement Numeral Phrases and the D-structure [XP, VP] Restriction

As an initial hypothesis, let us assume that agreement numeral phrases differ from no-agreement numeral phrases in their internal structure -- a hypothesis which we shall make specific in section 2. Let us suppose that either of these types of numeral phrases occurs in a variety of syntactic positions. Superficially, the two types of numeral phrases are identical. Thus, we can tell them apart only when they correspond to nominative subjects. This is because the only difference between them that is visible is their differing behavior with respect to verbal agreement, and Russian verbs agree only with subjects. Given a numeral phrase corresponding to an accusative object, for example, we are unable to tell whether it is a no-agreement or an agreement numeral phrase, because objects do not trigger agreement in any case.

This problem must be borne in mind during our discussion of conditions A and B. Let us consider condition A:

A: XP must be an [XP, VP] at D-structure ...

We will show that no-agreement numeral phrases are subject to A, while agreement numeral phrases are not. Recall, however, the class of [XP, VP]s that we were able to examine in earlier sections. These included surface objects (direct objects, expressions of duration) as well as surface subjects (of passive and ergative verbs). With respect to numeral phrases, we will have to limit our attention to the second class: D-structure objects which, in the case of Case-marked NPs, we can identify as S-structure subjects.

Within these inherent limitations, we appear to find what we are looking for. Once again, the effects of condition A are weaker for many speakers in the domain of numeral phrases than they were for the genitive of negation. Once again, we will offer an explanation of this fact in Chapter Four.

In the case of D-structure [XP, VP]s, both agreement and no-agreement numeral phrases are possible:

(69)a. [^Vsest' gorodov] bylo vzjato vragom
 six cities was taken enemy
 (masc gen pl) (neut sg) (masc instr sg)

b. [^Vsest' gorodov] byli vzjaty vragom
 six cities were taken enemy
 (masc gen pl) (pl)

(70)a. [^Vsest' studentov] prišlo
 six students arrived
 (masc gen pl) (neut sg)

b. [Sest' studentov] prišli
 (pl)

In the case of D-structure [XP, S]s, only the agreement numeral phrase is fully acceptable:²⁶

(71)a. ??[^Vsest' studentov] ubilo košku
 six students killed cat
 (masc gen pl) (neut sg) (fem acc sg)

b. [^Vsest' studentov] ubili košku
 (pl)

(72)a. ??[^Vsest' sobak] kusaetsja v ètoj kletke
 six dogs bites in this cage
 (fem gen pl) (3 sg)

b. [^Všest' sobak] kusajutsja v `etoj kletke
(3 pl)

(73)a. * [^Všest' gorodov] bylo čisto v našej oblasti
six cities was clean in our district
(masc gen pl) (neut sg)

b. [^Všest' gorodov] byli čisty v našej oblasti
(plur)

These facts will fall under condition A if no-agreement numeral phrases are XPs:

A: XP must be an [XP, VP] at D-structure, where XP = genitive phrase under negation, a distributive no-phrase, or a no-agreement numeral phrase.

1.3.2 No-agreement Numeral Phrases and the Non-Obliqueness Restriction

The canonical form of numeral phrases that we have considered, both agreement and no-agreement, contains a numeral phrase in its "citation form" followed by a noun in the genitive Case. This canonical form is found in positions that allow the non-oblique nominative and accusative Cases, as the preceding examples and (74)a below show. Numeral phrases in this canonical form never occur in positions where oblique Case is required, as (74)b-c show. Recall that pomogat' 'help' requires a dative object, and that upravljat' 'manage' requires an instrumental object:

(74)a. ja polučil [^Všest' priemnikov]
I received six radios
(nom) (masc gen pl)

b. *ja pomogaju [^Všest' devušek]
I help six girls
(nom) (fem gen pl)

- c. *ja upravljaju [^Všest' zavodov]
 I manage six factories
 (nom) (masc gen pl)

Suppose that all no-agreement numeral phrases, which are indistinguishable from agreement phrases in the sentences we have seen so far, show this "canonical form". (74) may be taken to show that condition B holds for at least the no-agreement numeral phrases:

B: XP may not occur in positions where oblique Case is required, where BP = genitive phrase under negation, a distributive po-phrase, or a no-agreement numeral phrase.

What about agreement numeral phrases? Since they are not subject to condition A, we expect that they should not be subject to B either. Now bear in mind that agreement numeral phrases are identical to no-agreement numeral phrases in non-oblique positions. In oblique positions, numeral phrases do occur, but they do not have the canonical form. Instead, the numeral and following noun behave like a normal constituent containing an adjective and a noun. The noun bears the oblique Case required in its position, and the numeral agrees in Case. This non-canonical form is impossible in a non-oblique position:

- //
- (75)a. *ja polučil [^Všest' priemniki]
 I received six radios
 (nom) (masc acc pl)
- b. ja pomogaju [^Všesti devuškam]
 I help six girls
 (nom) (fem dat pl)
- c. ja upravljaju [^Všest'ju zavodami]
 I manage six factories
 (nom) (masc instr pl)

We have claimed that no-agreement numeral phrases obey conditions A and B. We have seen that agreement numeral phrases do not obey condition A. If, as we shall show, the cluster of conditions A, B and C have a unified explanation, we expect that agreement numeral phrases, which do not obey A, should not obey B either. This immediately suggests that the "adjective-noun" numeral phrases of (75)b-c belong to the class of agreement numeral phrases, and that agreement phrases idiosyncratically do not show the "adjective-noun" form in non-oblique positions.

If this situation is not merely an artifact of our analysis, we would expect that the "idiosyncrasy" of agreement numeral phrases in non-oblique positions that leads them to look like no-agreement numeral phrases is in some way marked. This is not too surprising, given the other marked features of the Russian numeral system, to which we have alluded earlier.²⁷ What is interesting is that this marked property is absent from the numeral systems of at least one closely related language. In Polish, the phenomenon of "agreement" and "no-agreement" numeral phrases has a somewhat different character. The "paucal" numerals 2 through 4 are the only numerals that trigger agreement in Polish. When they do, the numeral phrase always shows "adjective-noun" behavior, regardless of its Case:²⁸

(76)a. [dwaj panowie] przyszli
 two men arrived
 (masc nom pl human) (plur)

b. ja pomagam [dwóm panom]
 I help two men
 (nom) (masc dat pl)

We will return to the "adjective-noun" numeral phrases of Russian, which only occur in oblique positions, in section 2. For now, our conclusions in this section are merely an unconfirmed hypothesis: we suppose that the numeral phrases that turn up in oblique positions are agreement numeral phrases, and that no-agreement phrases are, as we wish, subject to condition B. We will be more specific about the ad hoc Russian-specific rule we propose for agreement numeral phrases later. We assure the reader, however, that this will be the only such ad hoc proposal we will make in the course of our analysis, and that if a Slavic language contains ad hoc rules, the numeral system is certainly the place one would most expect to find them.

1.3.3 No-agreement Numeral Phrases and Obligatory QR

In this section, we show that no-agreement numeral phrases, which fall under conditions A and B, also fall under C, and that agreement numeral phrases do not. Recall the content of C:

C: XP obligatorily undergoes QR ...

Obviously, C is an interesting generalization only if we assume that QR, like movement rules in the syntax, is formally optional. We exploited this optionally in 1.1.3, where we claimed that an ambiguity in non-genitive NPs under negation resulted from the optionality of QR, while genitive NPs under negation lacked this ambiguity due to C. The proposal that QR is optional is argued for most convincingly by Higginbotham (forthcoming). We rely on his proposals in this section.

Higginbotham notes that if QR is optional we may characterize a classic ambiguity of numeral phrases in a simple way (cf. Lakoff 1970;

also Paduceva^V 1974, 111-124 for Russian examples and enlightening discussion). For the moment, we draw our examples from English.

Consider a sentence like (77):

(77) six mathematicians proved the theorem

This sentence has two readings. On one, if the six mathematicians have names A through F, the following claim is true:

(78) A proved the theorem; B proved the theorem; ...; and F proved the theorem too.

In other words, each mathematician individually proved the theorem. This interpretation naturally corresponds to an LF representation which is a quantifier-variable structure:

(79) for $6x$: x a mathematician (x proved the theorem)²⁹

Let us call this the individual reading.

There is a second reading, however, on which (78) is not true. Rather, a group consisting of six mathematicians is collectively responsible for the theorem's solution. No one mathematician solved the theorem alone. Higginbotham suggests that this reading corresponds to an LF close to or identical with the S-structure (77): QR has not applied. Let us call this reading the group reading. Thus, the distinction between group and individual readings of numeral phrases may be traced back to the optionality of QR, following Higginbotham.

It follows that if QR is obligatory for a numeral phrase -- if, for example, a no-agreement numeral phrase were subject to C -- then this numeral phrase should exclude the group reading and have only the

individual reading. At the level of intuition, this appears to be the case. Insofar as informants can distinguish group from individual readings with predicates that must be ergative or passive by condition A, they report that no-agreement numeral phrases generally have an "existential" or "individuated" flavor which is not obligatory for agreement numeral phrases. Somewhat similar intuitions have been reported elsewhere in the literature, for example by Crockett (1976b, 349). Revzin (1978, 262-272), in an extremely interesting discussion of the question, claims that no-agreement numeral phrases are always "indefinite" (neopredelennyj), while agreement numeral phrases are ambiguously definite or indefinite. They argue that the same distinction held in earlier stages of Russian -- the so-called "Potebnja's law".

Fortunately, a more precise test shows that condition C applies to no-agreement numeral phrases. As is well-known, certain verbs, like gather, disperse, meet and their synonyms cross-linguistically, require that one of their arguments be a group:

- (80)a. the group dispersed quietly
 b. #the mathematician dispersed quietly

(The sign # indicates a violation of selectional restrictions.)

Since (81) below is acceptable, logic dictates that the numeral phrase is interpreted as a group, and hence, by Higginbotham's hypothesis, is not undergoing the optional rule of QR:

- (81) six mathematicians dispersed quietly

Since a verb like disperse requires a numeral-phrase argument not

to undergo QR, a numeral phrase that obeys condition C will not be able to be this argument of such a verb, since C requires the application of QR that the verb would like to prevent. We thus make the following prediction: with a Russian verb of the disperse type, a sentence like (81) should be acceptable with an agreement numeral phrase, which does not have to undergo QR, but should be unacceptable with a no-agreement numeral phrase, which is subject to C and must undergo QR -- if our hypothesis is correct.

The Russian verb razlučit'sja^V 'part company', 'go one's separate ways' is a verb of this type, and appears to be ergative.³⁰ It shows the paradigm of (80)a-b in (82)a-b below, and -- as we hope -- allows only agreement numeral phrases, as we see in (82)c-d:

- (82)a. gruppa razlučilas'^V na mostu
group parted-company on bridge
- b. #matemat'ik razlučilsja^V na mostu
mathematician parted-company on bridge
- c. [šest' matematikov]^V razlučilis' na mostu
six mathematicians parted-company on bridge
(pl)
- d. [šest' matematikov]^V razlučilos' na mostu 31
(neut sg)

The paradigm of (82) shows clearly that C applies to no-agreement numeral phrases and does not apply to agreement numeral phrases:

C: XP obligatorily undergoes QR, where XP = genitive phrase under negator, a distributive po-phrase or a no-agreement numeral phrase.

This completes our exposition of the basic data with which we will be dealing.

2. The Non-Obliqueness Restriction and the QP Hypothesis

In this and subsequent sections, we shall derive generalizations A, B, and C from independent principles of grammar. In this section, we shall account for condition B: XP may not occur in positions where oblique Case is required. We shall argue that B is an artifact of a general theory of Case, the Θ -criterion and a hypothesis about the structure of phrases showing conditions A, B, and C. In section 3, we show that the predictions made by this account of B are valid, and that our explanation of B accounts for otherwise mysterious facts about complementizer-trace phenomena in Russian.

2.1 The QP Hypothesis

What is the structure of phrases that we have called XPs? Let us first consider po-phrases and no-agreement numeral phrases. Both of these XPs appear to have the general form:

(83) [_{XP} Q N]

We leave open, as irrelevant to our discussion, the number of bars on Q and on N. Thus:

- (84)a. ja dal mal'^Včikam [_{XP} [_Q po] [_N jabloku]]
 I gave boys po apple
 (masc dat pl) (neut dat sg)
- b. [_{XP} [_Q ^Všest'] [_N studentov]] prišlo^V
 six students arrived
 (masc gen pl) (neut sg)

Quantifiers like po and šest' each "take" a particular Case in their noun. Some other Russian quantifiers, like vse 'all' are adjectival.

Among the non-adjectival quantifiers, po is the only one we will examine that does not take genitive Case in its noun -- a legacy from its prepositional past (but see section 3 below). If an unmarked non-adjectival quantifier takes the genitive Case, as seems likely, then we can assimilate the genitive of negation to the structure in (83)-(84) by blaming the genitive Case on a null quantifier:

(85) ja ne polučal [XP [Q e] [N pisem]]
 I NEG received letters
 (neut gen pl)

The null [Qe] has the properties of lexical quantifiers like Vest'.

This analysis of the genitive of negation construction, so far, derives from Kayne's (1975, 1981a) analysis of the French pas-de construction, which strongly resembles the Russian genitive of negation (cf. fn. 20). Relating the genitive of negation to the genitive with quantifiers dates at least to Jakobson (1936-1971). In Kayne's analysis, XP = NP, which will not be our hypothesis, as we shall soon see.

Notice that this analysis so far makes no connection between the genitive of negation construction and the presence of negation. Despite the familiar name of the construction, this may be very nearly the right move, if our goal is to explain properties A through C. The reason negation licenses the construction will be considered in section 3. In giving the negation a secondary role, we differ from most analyses of the construction (e.g. Chvany 1975; Babby 1980).

Note a further feature of this analysis. If null Qs -- that is, Qs to which lexical insertion has optionally failed to apply -- can assign genitive Case, then it follows that genitive Case after quantifiers is

not the result of idiosyncratic lexical properties of certain lexical quantifiers, but is rather a general property of the node Q. Q does not fall in the two-feature inventory of categorial features introduced in Chomsky (1970), although Jackendoff (1977) does give features for a category Q. We will return to the extremely interesting question "what is Q?", once more, in section 3.

Given the X-bar convention for phrase structure, our next task is to identify the "X" in "XP" as either Q or N. At this point, we make a hypothesis which will be absolutely crucial to the analysis. We will derive properties A, B, and C in pages to come from the assumption that, in phrases which show these conditions, the head of our [Q N] constituents is Q. In other words, XPs, including genitive phrases under negation, distributive po-phrases and no-agreement numeral phrases, are all QPs. We call this the QP hypothesis:

- (86)a. [_{QP} [_Q po] [_N jabloku]]
 b. [_{QP} [_Q ^v sest'] [_N studentov]]
 c. [_{QP} [_Q e] [_N pisem]]

Evidence for the QP hypothesis will be our derivation of generalizations A, B, and C. Nonetheless, we can give a quick suggestive argument in favor of this hypothesis here. As shown in section 1.3, numeral phrases come in two varieties: agreement and no-agreement. We claimed that the no-agreement numeral phrases obey A, B, and C, while agreement numeral phrases do not. If we will derive A-C from the hypothesis that constituents like no-agreement numeral phrases are QPs, what are agreement numeral phrases?

Under the X-bar convention, in a phrase of the form [_{XP} Q N], X must be either Q or N. If X is not Q in an agreement numeral phrase, then it must be N. We thus suggest that what distinguishes agreement from no-agreement numeral phrases is the choice of head. A no-agreement numeral phrase is a QP; an agreement numeral phrase is an NP:

- (87)a. [_{QP} [_Q ^vsest'] [_N studentov]] no-agreement numeral phrase
 b. [_{NP} [_Q ^vsest'] [_N studentov]] agreement numeral phrase

If we assume that a verb agrees with an NP, but not with a QP, we account for the agreement facts; we will, however, have a more sophisticated suggestion later.

It may be objected that (87)b represents a rather odd NP, in which a non-head assigns Case to the head noun (though similar NPs have been proposed for English by Selkirk 1977; cf. references therein). We wish to agree with this objection. Recall our discussion of the non-obliqueness restriction in numeral phrases (1.3.2). We were forced to say that no-agreement numeral phrases could not occur where oblique Case was required, if we wanted to claim that they are subject to conditions A, B, and C. Translated into the terms of the QP hypothesis, this means that QPs cannot occur in oblique positions. Numeral phrases of some sort do occur in oblique positions, however, and we claimed, without argument, that these were actually agreement numeral phrases -- now called NPs.

Recall now what these oblique numeral phrases looked like. As we showed in (75)b-c, they looked like normal adjective-noun pairs, in which the numeral agreed with the noun, which showed the required oblique Case. Under our theory, these numeral phrases are NPs, and these numeral phrases

certainly look like normal NPs. Thus, while we still assume, as an ad hoc and language specific rule, that NPs containing numerals irregularly show genitive marking on N in non-oblique positions, this locus of irregularity now seems more reasonable, in the context of our structural hypothesis about numeral phrases. Finally, we may say that Polish (cf. (76)a-b) simply lacks this ad hoc feature: its NP numeral phrases (only the paucal numerals) act like true NPs in all positions.

One other point follows from this discussion. It is often noted in the literature (Švedova 1970, 544; Revzin 1978, 269; among others) that when a numeral phrase is preceded by an adnominal modifier -- an adjective or adjectival quantifier -- verbal agreement is obligatory:

- (88) kakie-to / vse / sčastlivye [šest' studentov] prišli/*prišlo
 some / all / (the happy six students arrived
 (masc gen pl) (pl) / (neut sg)

It is reasonable to suppose that such adnominal modifiers, preceding the core of the numeral phrase, can only modify an N with some number of bars, and not a Q. Hence, these modifiers will be able to occur in the position they do only if the numeral phrase is an NP. If only NP numeral phrases trigger agreement, we account for the facts of (88).³²

2.2 Oblique vs. Structural Case

We now develop a theory of oblique and structural Case. In 2.3, we will combine this with our QP hypothesis to explain condition B.

One of the facts a speaker of Russian knows is that the deep object of ponogat' 'help' must be dative, and that the deep object of upravljat' 'manage' must be instrumental. It may be the case that universal semantic principles of some sort provide a markedness theory of oblique Case

assignment: surely it is not accidental, for example, that the object of the verb meaning 'help' bears the same Case as the indirect object of 'give' (i.e. dative) in language after language. Nonetheless, it seems that knowledge of oblique Case requirements is to a great degree language-specific, and must be learned through experience. Russian children do seem to learn these requirements gradually, much as they learn irregularities of conjugation or declension.

The phenomenon of "quirky" oblique Case requirements is common in languages with overt Case marking. Following work on oblique Case in Icelandic by Levin (1981) (cf. also Levin and Simpson 1981), we suggest that oblique Case requirements imposed by certain verbs are intimately linked to Θ -role assignment. Specifically, certain verbs will Θ -mark a constituent only if it bears a particular Case.³³ By the Projection Principle, this Θ -marking is obligatory at all levels of representation. Hence, if a verb requires oblique Case in order to assign a Θ -role, then (1) this Case must be marked on the appropriate constituent, and (2) this Case must be present at D-structure, S-structure and LF. Consequence (1) captures the obligatoriness of "quirky" oblique Case requirements. Consequence (2) is independently argued for by Freidin and Babby (forthcoming), who call this the "Principle of Lexical Satisfaction".

Thus, the lexical entries of pomogat' 'help' and upravljat' 'manage' will contain statements like the following:

- (89)a. pomogat' directly Θ -marks a [+dative] constituent
- b. upravljat' directly Θ -marks a [+instrumental] constituent

While oblique Case appears to be linked to the assignment of a

particular Θ -role by a particular predicate, non-oblique Case appears to be purely structural. All things being equal, any NP subject of a tensed sentence may be marked with nominative Case when governed by AGR, and any NP governed by V may be marked with accusative Case. The assignment of non-oblique Case appears to be independent of the assignment of a Θ -role, and is independent of the Projection Principle.³⁴ This leads us to expect the following three consequences.

1. Constituents not Θ -marked by a verb may bear non-oblique Case, if they meet the structural requirements of our Case conventions in (35).

2. Non-oblique Case marking is optional, except when required by the Case Filter in (26).³⁵

3. Non-oblique Case marking need not be present at all levels of representation, since it is unconnected to the Projection Principle. It need only be present at the level at which the Case Filter applies.

We shall argue that these three consequences are all true. Consequence (1) can be shown to hold by looking at the expressions of duration considered in section 1.1.1:

- (90)a. ja ni odnu minutu ne spal
 I not one minute NEG slept
 (fem acc sg)
- b. ja ni odnoj minuty ne spal
 (fem gen sg)

The genitive of negation found in (90)b tells us (via condition A) that the expression of duration is governed by V. (90)a shows us that these expressions receive accusative Case, although they are not Θ -marked by the verb spat' 'sleep'.³⁶

Consequence (3) is shown to hold by the phenomenon of passive and

ergative verbs. These verbs have the property of forbidding normal assignment of Case to a Θ -marked constituent in VP.³⁷ To satisfy the Case Filter, as we have seen, an NP object of such a verb must move to a position in which it can receive Case. Since such Case marking is not available to it until after the application of Move α , it is clear that this Case marking does not have to hold at all levels, and one might ask whether it applies at D-structure at all.

As for consequence (2), this consequence will play a role in our derivation of condition B for QPs, which follows immediately.

If non-oblique Case works as we have claimed, then verbs like polučat'^V 'receive', which "take" accusative objects, do not refer to the Case of their objects at all in the lexicon. Case marking of their objects will follow from purely syntactic considerations. The lexical entry of polučat'^V will thus contrast with the lexical entries for pomogat' and upravljat' in (89) by stating simply:

(91) polučat'^V indirectly Θ -marks a constituent.³⁸

A technical problem now arises. We must prevent the random generation of oblique constituents as the complements to verbs like polučat'^V that do not care about the Case of their object. The necessary condition is simple and natural:

(92) Oblique Biconditional

α bears an oblique Case C if and only if for β the Θ -marker of α , β requires C as a condition on Θ -marking α .

Actually, (92) follows from right to left from the Θ -criterion: if Θ -marking is contingent on α bearing Case C, then the Θ -criterion requires α to

bear C. Only the conditional from left to right is actually an independent principle: that α bears C only if C is required by its Θ -marker. Nonetheless, we will leave (92) as a biconditional, for clarity.

The Oblique Biconditional is useful to us because it excludes sentences like (93), where a "wild" dative NP has been generated:

(93) *ja polučal^V pis'mam
 I received letters
 (neut dat pl)

(92) is, in effect, a simple statement about markedness. It says that oblique Case is found when lexical entries are more complex. As such, it is not ad hoc, but merely captures a basic intuition about oblique Case requirements: that they represent an irregularity.³⁹

2.3 The Non-Obliqueness Restriction

In section 1.1.1, we briefly outlined a Case-theoretic approach to the obligatoriness of NP-movement with passive and ergative verbs. This approach singles out two properties of these verbs, which may be connected via Burzio's generalization: (1) they do not allow their Θ -marked object to bear Case in VP; (2) they do not indirectly Θ -mark their subject.

Recall that property (2) allows the D-structure object of such verbs to move to the subject position, without violating the Θ -criterion or Projection Principle. Property (1), in tensed sentences, requires a D-structure object NP to move to subject position, to escape the effects of the Case Filter. Thus, in the D-structure representation (94)a, the subject position is a non- Θ -position, and hence not filled by an argument. The object argument NP is in a position where it cannot receive Case. If it does not move to the Case-marked subject position, it will remain

Caseless, and violate the Case Filter. Hence it must move, yielding the S-structure (94)b:

- (93)a. [S e [INFL TNS AGR] [VP was received [NP the letter]]]
 b. [S [NP the letter]_i [INFL TNS AGR] [VP was received e_i]]

Recall also that S's do not fall under the Case Filter. Hence a D-structure like (94) may yield a surface structure that is identical -- the only condition being the insertion of the lexical expletive it in subject position, for reasons we consider later in this chapter:

- (94)a. [S it [INFL TNS AGR] [VP was claimed [S, that John came]]]

Let us now consider the status of our newly postulated QPs with respect to Case Theory and the Case Filter. There are three possibilities. (i) QPs may be assigned to Case features and are subject to the Case Filter. (ii) QPs may be assigned Case features, but are not subject to the Case Filter. (iii) QPs may not be assigned Case features, and are not subject to the Case Filter.

By "subject to the Case Filter" we mean, of course, subject to a filter that stars Caseless QPs as well as NPs. Our case Filter in (26) applies, by stipulation, only to NP. Notice that a fourth logical possibility for the status of QPs under Case Theory is excluded as a contradiction: (iv) QPs may not be assigned Case features, but are subject to the Case Filter. Possibility (iv), if true, implies that lexical QPs cannot exist, since they would all fail the Case Filter.

We will argue that possibility (iii) is true for QPs -- as, in fact, its counterpart is for S's. This claim, obviously, makes certain immediate predictions about QPs in passive and ergative constructions, which we

shall argue to be true. Before examining these predictions, let us see how possibility (iii) derives condition B. We will also examine the status of NPs and QPs under the Case Filter if the Case Filter is reinterpreted as part of θ -theory, a move suggested by Chomsky (1981a).

As illustrated in (91), a verb like polu^Vcat' 'receive' does not care about the Case features of its object. Let us consider what sort of object this verb might take. Suppose we were to give polu^Vcat' an NP object, but not assign it any Case features. Polu^Vcat' could assign it a θ -role (under our present assumptions; cf. infra), but the Case Filter would rule it out. This situation is illustrated in (94)a below. Suppose we now give polu^Vcat' an NP object and assign some oblique Case features to this NP -- e.g. dative. Polu^Vcat' once more can assign the NP a θ -role, and the Case Filter will not rule the structure out. In this Case, however, the Oblique Biconditional (92) will intervene, as it does in (93), ruling out (94)b below. Suppose we give polu^Vcat' an NP object and mark it accusative; as we expect, nothing will rule this structure out. This is seen in (94)c.

Finally, suppose we give polu^Vcat' a QP object. If QP neither bears nor needs Case features, as follows from possibility (iii) above, the verb may assign its θ -role to the Caseless QP, the Case Filter will not affect the QP, and the Oblique Biconditional will not be relevant. The structure will be acceptable, as we see in (94)d:

- (94)a. * $\text{polu}\overset{v}{\text{cat}}'$ + [$\overset{\text{NP}}{\text{-Case}}$] (Case Filter)
- $\left. \begin{array}{l} | \\ \text{---} \\ | \end{array} \right\} \Theta \left. \begin{array}{l} \text{---} \\ | \\ \uparrow \end{array} \right\}$
- b. * $\text{polu}\overset{v}{\text{cat}}'$ + [$\overset{\text{NP}}{\text{+dative}}$] (Oblique Biconditional)
- $\left. \begin{array}{l} | \\ \text{---} \\ | \end{array} \right\} \Theta \left. \begin{array}{l} \text{---} \\ | \\ \uparrow \end{array} \right\}$
- c. $\text{polu}\overset{v}{\text{cat}}'$ + [$\overset{\text{NP}}{\text{+accusative}}$]
- $\left. \begin{array}{l} | \\ \text{---} \\ | \end{array} \right\} \Theta \left. \begin{array}{l} \text{---} \\ | \\ \uparrow \end{array} \right\}$
- d. $\text{polu}\overset{v}{\text{cat}}'$ + [$\overset{\text{QP}}{\text{-Case}}$]
- $\left. \begin{array}{l} | \\ \text{---} \\ | \end{array} \right\} \Theta \left. \begin{array}{l} \text{---} \\ | \\ \uparrow \end{array} \right\}$

The story is rather different with verbs like pomogat' 'help'. As we saw in (89)a, pomogat' Θ -marks a constituent which bears dative Case. Suppose we give pomogat' an NP object, but do not give it any Case features. Not only will the Case Filter rule it out, as in (94)a, but it will not be Θ -marked by pomogat' and the Θ -criterion will be violated. This situation is demonstrated in (95)a below. Suppose now we give the verb an NP object and assign it dative Case. The NP will satisfy the Case Filter, pomogat' will Θ -mark it successfully, and the structure will be grammatical, as seen in (95)b. Now suppose we give pomogat' an accusative NP object. The NP will satisfy the Case Filter, but pomogat' will not Θ -mark it, violating the Θ -criterion. This ungrammatical structure is seen in (95)c.

Finally, and most crucially, suppose we generate a QP as the object of pomogat'. If QPs neither need nor receive Case features, they will, in particular, not receive dative Case features. As a result, pomogat' will be unable to assign its Θ -role to the QP object, and the structure

will be ruled out. We see this in (95)d:

- (95) a. *pomogat' + [^{NP}
-Case] (θ-criterion, Case Filter)
- | ↑
_____ e _____
- b. pomogat' + [^{NP}
+dative]
- | ↑
_____ e _____
- c. *pomogat' + [^{NP}
+accusative] (θ-criterion)
- | ↑
_____ e _____
- d. *pomogat' + [^{QP}
-Case] (θ-criterion)
- | ↑
_____ e _____

This result is quite general. If the QP hypothesis is true, and if the Case Filter stands as we have given it, and if QPs do not receive Case features or fall under the Case Filter, then nothing we have called "XP" can occur where an oblique object is required. Extend the notion of "oblique object" to a few cases where an "inherent θ-role" may be assigned to an oblique NP -- e.g. to instrumental NPs functioning like English by-phrases -- and we have derived condition B: XPs cannot occur where oblique Case is required.

Condition B thus reduces to the QP hypothesis:

- (96) B: XPs are QPs

Our theory of Case and our interpretation of the Case Filter seem to be extremely simple, and do capture the facts of Russian. In following sections, we will assume that they are correct. Nonetheless, the skeptical

reader may object that we have merely substituted for the ad hoc condition B the ad hoc assumption that QPs do not bear Case features. In section 3, we will show that this assumption is not ad hoc; in particular it has further consequences that follow without further stipulation. These consequences appear correct, and advance us towards an explanation of conditions A and C.

First, however, let us briefly consider a reinterpretation of the Case Filter offered by Chomsky (1981a, Chapter 6), following suggestions of Aoun. Chomsky suggests that the obligatoriness of Case features on lexical NPs follows from the θ -criterion, in particular from the θ -criterion as a condition of A-chains, a notion we have discussed in Chapter One.

Recall we assumed the following convention in Chapter One (example (26)):

(97) Chain C bears θ -role R if and only if for some $\alpha \in C$, α bears R.

We may assume a parallel convention for Case (as in Chomsky, p. 334, (16)):

(98) Chain C bears Case K if and only if for some $\alpha \in C$, α occupies a position assigned K by β .

Thus, by conventions (97)-(98), the chain (John, e_i) in (99) bears the θ -role assigned by seen by virtue of e_i , and the nominative Case assigned by AGR by virtue of John:

(99) [_S John_i [_{INFL} TNS AGR] [_{VP} was seen e_i]]

We may then state the Case Filter as in (100) (= Chomsky, p. 334, (17)):

(100) Every lexical NP is an element of a chain with Case.

A generalization of (100) subsumes (100), and also captures the fact that PRO and NP-trace are not subject to the Case Filter:

(101) Case Filter (Chomsky, p. 334, (18))

Suppose that the position P is marked with the θ -role R and $C = (\alpha_1, \dots, \alpha_n)$ is a chain. Then P assigns R to C if and only if for some i , α_i is in position P and C has Case or is headed by PRO.

NP-trace is locally A-bound -- i.e. it does not head its chain. A chain containing an NP-trace will satisfy (101) if it has Case or is headed by PRO. There is no requirement that NP-trace itself be in a Case-marked position.

(101) has an empirical advantage over our earlier formulation of the Case Filter. If we restrict our attention to A-chains, as we have, it follows from (101) that chains headed by WH-trace will have to bear Case, even though no element in such chains is lexical. We thus unify (102)a-b with (103)a-b:⁴⁰

(102)a. *John, who_i it is impossible [_S, e_i to come]

b. *I wonder what_i it was seen e_i

(103)a. *it is impossible [_S, Bill to come]

b. *I know that it was seen Bill

In (102)a-b, the singleton chain (e_i) neither has Case nor is headed by PRO. (Since e_i is bound by who_i, it neither is free, nor has an antecedent bearing an independent θ -role, as required by the contextual definition

of PRO given in Chapter One.) The same applies to the singleton chain (Bill) in (103)a-b. By the Case Filter in (101), neither chain may bear its θ -role, and the θ -criterion is violated. By contrast, our previous Case Filter, applying only to lexical NPs, rules out only (103)a-b.

How can the Case Filter in (101) capture the properties of lexical categories that do not need Case, but which nonetheless seem to bear θ -roles? S' belongs to this category, and -- we have claimed -- QP as well. The answer lies in a stipulation equivalent in this theory to our stipulation that the Case Filter of this chapter's (26) applies only to NPs. The form of this stipulation is ours; Chomsky takes a slightly different approach:

(104) If an NP bears a θ -role, it is a member of some chain.

S' or QP, on the other hand, may bear a θ -role without being a member of a chain, and thus may bear a θ -role without being subject to (101).

We may even want to strengthen this conclusion, on Chomsky's theory, to say that QP (and S'?) is never a member of a chain. This would follow from (101) if we have the following additional convention:

(105) Consider the A-chain $C = (\alpha_1, \dots, \alpha_n)$ and Case K. For all positions' $P \in C$, for some $\alpha_i \in C$, if α_i bears K, then P bears K.

(105) says that a Case marked position in a chain transmits its Case to all the other members of the chain, and, more importantly, entails that no member of a chain may bear Case if some member of that chain cannot bear Case. Thus, a chain containing QP could never contain a Case-marked position; by (98), such a chain would not bear Case; by (101) such a chain could not bear a θ -role. We leave open whether this strengthening

of (101) is justified; in particular we will give another reason why QP cannot be a member of a chain somewhat later, and show that this result may be correct.

In any case, we have shown that our hypothesis about the status of QP under Case Theory may be placed in the framework of Chomsky's reduction of the Case Filter to the Θ -criterion, via (101), as well as in the simpler framework of our Case Filter (26).

3. QPs and the Empty Category Principle

As we have seen, a passive or ergative verb Θ -marks an object argument, but does not allow Case to be assigned to it. If the object is an NP, the interaction of this property with the Case Filter will force NP to move to a Case-marked position -- in a tensed sentence, the subject position of that sentence. We show this schematically in (106), where the D-structure (106)a must yield an S-structure (106)b, with the object moved to the subject position. The feature [-Case] on V indicates that it is passive or ergative:

- (106)a. [_S e INFL [_{VP} [^V [-Case] NP]]] D-structure
 b. [_S NP_i INFL [_{VP} [^V [-Case] e_i]]] S-structure

What is important is that the movement to subject position is formally completely optional. Only the independent Case Filter forces NP to move by S-structure (assuming the Case Filter applies at S-structure). If the NP does not move, it will violate the Case Filter.

We thus take a "modular" approach to movement: surface complexities of rules like passive really result from the interaction of various

autonomous subsystems, each of which is simple and general, and plays a wider role in the system (see Jaeggli 1981 for further discussion).

We have claimed that QPs neither need nor receive Case. Suppose we replace NP with QP in (106)a. No principle whatsoever will require this QP to move to subject position, given our treatment of passive and ergative constructions. In other words, while D-structures like (106)a must always yield an S-structure like (106)b in a language like Russian or English, a corresponding D-structure like (107) below may yield an S-structure that is essentially identical to the D-structure:

(107) [_S e INFL [_{VP} [_V -Case] QP]] D-structure and S-structure

Clearly, if we can find evidence that this result is correct, we support both our argument in the preceding section and the modular approach to passive and similar processes discussed above.

If our QPs were subject to the Case Filter,⁴¹ and if they could receive Case, D-structures like (107) would yield S-structures in which QP was a subject. On our assumptions, they yield S-structures in which QP is an object. We need a test that can distinguish obligatory S-structure subjects from S-structure objects. An obvious candidate is the so-called "complementizer-trace" or "that-trace" phenomenon. We will use the term CTP effects (Complementizer-Trace-Phenomena), following Pesetsky (1982). This is the well-known contrast in English between "long movement" of a subject, on the one hand, and "long movement" of an object or any kind of "short movement", on the other hand:

- (108)a. [_S, what book_i [_S did Mary INFL [_{VP} read e_i]]]
 b. [_S, what book_i [_S e_i INFL [_{VP} charmed Mary]]]
 c. [_S₁, what book_i [_S₁ did John claim [_S₂, that Mary INFL [_{VP} read e_i]]]]]
 d. * [_S₁, what book_i [_S₁ did John claim [_S₂, that e_i INFL [_{VP} charmed Mary]]]]]

As noted in Pesetsky (1981) (and by Chvany 1975; Zaliznjak and Paduceva 1979, among others), this paradigm is found in Russian as well, although the contrast is "fuzzier", as Sinicyn (1982b) notes:

- (109)a. kakuju knigu_i Maša^v pročla e_i^v
 what book Maša^v read
 (fem acc sg) (fem nom sg) (fem sg)
- b. kakaja kniga_i e_i^v očarovala Mašu^v
 what book charmed Maša^v
 (fem nom sg) (fem acc sg)
- c. ?kakuju knigu_i Ivan xotel [čtoby Maša^v pročla e_i^v]
 what book Ivan wanted that Maša^v read
 (fem acc sg) (nom) (subjunc) (nom)
- d. *kakaja kniga_i Ivan xotel [čtoby e_i^v očarovala Mašu^v] 41 42
 what book Ivan wanted that charm Maša^v
 (fem nom sg) (subjunc) (acc)

Before considering QPs in this paradigm, we must digress for a while about the account given for CTP effects in the framework of Chomsky (1981a).

In Chomsky (1981a), CTP effects fall under Government Theory, a theory which controls the distribution of empty categories. Government Theory has as its core a principle and a definition:

(110) Empty Category Principle (ECP)

An empty category [_αe] must be properly governed.

(111) α properly governs β if and only if α governs β [and $\alpha \neq \text{AGR}$]

In Chapter Three, we will suggest that CTP effects more properly fall under Path Theory, which we develop there. For the moment, Chomsky's Government Theory will suffice to explain conditions A and C in the coming sections. For other reformulations of Government Theory, cf. Kayne (1981a, 1982), which we discuss in detail in Chapter Four (also Brody (1982) and Aoun (1982) for proposed reductions of Government Theory to the Binding Theory).

Recall the definition of government we assumed in Chapter One, from Aoun and Sportiche (1982):

(112) Government (def)

In the structure:

$[\beta \dots \gamma \dots \alpha \dots \gamma \dots]$, where

(i) $\alpha = X^0$

(ii) where ϕ is a maximal projection, ϕ dominates α if and only if ϕ dominates γ .

α governs γ .

With this definition of government, the ECP accounts straightforwardly for the (a), (c) and (d) sentences of (108) and (109), but not for the (b) sentences.

In the (a) structures, where an object has been short-moved within its clause, the empty category left behind is governed by V. By (111), it is properly governed, and the ECP is satisfied.

The same applies to the (c) structures, in which an object has been long-moved outside its clause. The empty category is once more governed by V, properly governed, and satisfies the ECP.

In the (d) structures, however, a subject has been long-moved outside its clause. The empty category left behind is governed, by AGR in INFL. Government by AGR, however, does not count as proper government, according to (111). The empty category in subject position is thus not properly governed, and the ECP is violated. The structure is correctly ruled out.

What about the (b) structures? According to the definition of government in (112), these structures should also violate the ECP. In the (b) sentences a subject has been short-moved within its clause. The empty category left behind is governed only by AGR in INFL, according to (112), and should violate the ECP. Obviously, the definition of government must be changed to allow this case under the ECP. The change adopted by Chomsky (1981a) is to allow, not only "structural" government, as in (112), but also government by coindexation, so that a WH-phrase in COMP can properly govern a subject trace. Let us state this revised definition of government as in (113):

(113) Government - Revised (def)

In the structure:

[$\beta \dots \gamma \dots \alpha \dots \gamma \dots$], where

- (i) $\alpha = X^0$ or properly binds γ
 - (ii) where ϕ is a maximal projection, ϕ dominates α if and only if ϕ dominates γ .
- α governs γ .

(114) Proper Binding

α properly binds β if and only if

- (i) α and β are coindexed
- (ii) α c-commands β
- (iii) α is a possible antecedent for β

The definition in (113) is close to Chomsky's (1981a, 250) definition, but not identical. Chomsky alters clause (ii) to incorporate some facts about Italian that we consider below. Our definition of Proper Binding (after Fiengo 1974) anticipates our discussion in the next section, particularly its reference to "possible antecedent" -- a notion whose relevance will become clear shortly.

If we extend government as in (113), to include government by coindexation, we account for the full paradigm in (108) and (109). In particular, the problem with the (b) structures disappears. In (108)b and (109)b, where a subject has been short-moved to COMP, the empty category left behind is not properly governed by AGR in INFL, since AGR is not a proper governor. It is, however, properly governed by its binder in the nearest COMP, since no maximal projection intervenes between the two.

In (108)d and (109)d, however, where the subject has been long-moved to COMP of a higher clause, the ECP still assigns a star, as desired. Although the trace left by this movement is properly bound, its proper binder is not a governor, since maximal projections intervene between the trace and its binder.⁴³

Chomsky suggests that the ECP also governs the distribution of NP-trace. Consider the contrast between (115)a-b:

- (115)a. John_i is [_{AP} likely [_α e_i to win]
 b. *John_i is [_{AP} probable [_α e_i to win]

In each instance, John_i is in a non-θ-position at S-structure, as can be

seen from (116):

- (116)a. it is [_{AP} likely [_S, that John will win]]
 b. it is [_{AP} probable [_S, that John will win]]

It follows that (John, e_i) must form a chain in (115)a-b, in order for John not to violate the θ -criterion. e_i is thus an NP-trace.

If α is maximal in (115), e_i will violate the ECP, since it will be un-governed. If α is non-maximal, e_i will be properly governed by A, and will satisfy the ECP. To relate (115) to the ECP, then, it is necessary to take α to be non-maximal in (115)a -- i.e. as S -- and to take α to be maximal in (115)b -- i.e. as S':

- (117)a. John_i is [_{AP} likely [_S e_i to win]
 b. *John_i is [_{AP} probable [_S, e_i to win]

(117)b is now ruled out by the ECP.

To account for this distinction, Chomsky (1981a) (following a proposal of Rouveret and Vergnaud 1980) proposes that the clausal complement of predicates like likely is subject to a lexically governed, language-specific rule of S'-deletion, which allows the subject of the complement to be governed by the higher predicate. S'-deletion applies to the complement of likely, but not to the complement of possible, accounting for the contrast under NP-movement.

One problem arises with the ECP, that must be noted. The ECP must clearly not apply to empty categories that are completely un-governed. We saw in Chapter One that empty categories that are [+pronominal, +anaphor] are forced by the Binding Theory to be un-governed. Since these un-governed categories exist, they are obviously exempt from the ECP:

(118) it is unclear [_S, where [_S PRO to go]]

Chomsky speculates that PRO may bear features that make it not "empty" for the ECP; the point remains that PRO must somehow be excluded from the ECP.

Recall now that we are going to use the ECP as a test for whether QPs that are the D-structure object of a passive or ergative verb must move to subject position, i.e. whether a D-structure like (119) can also be an S-structure:

(119) [_S e INFL [_{VP} [_V ^V-Case] QP]]

Before making this test, however, we need to modify the ECP slightly. If (119) is an acceptable S-structure, why doesn't e in subject position violate the ECP? The answer lies in the suggestion of Safir (1981, 1982) that empty categories that are expletive -- that do not belong to a θ -marked chain -- are exempt from the ECP. Putting this claim together with the fact that PRO is not subject to the ECP, we have a revised ECP, as in (120):

(120) Revised Empty Category Principle

An empty category [e] must be properly governed unless

(a) it is PRO or

(b) it is expletive (does not belong to a θ -marked chain)

In Chapter Four, we discuss some ways in which Path Theory, which subsumes the ECP and Government Theory, can derive the stipulated exceptions for PRO and expletive empty categories. In this Chapter, (120) is the ECP we will assume.

Safir (1982) presents a great many arguments in favor of the exclusion of expletive empty categories from the ECP, most of them embedded in his analysis of the so-called "pro-drop" phenomenon. Before returning to our discussion of Russian QPs, let us briefly consider the role of Safir's hypothesis in his analysis of pro-drop phenomena, as well as the status of Russian with respect to this phenomenon.

It was observed first by Perlmutter (1971), that the paradigm of CTP effects demonstrated in (108)-(109) does not hold on the surface in all languages. For example, it does not hold in Italian:

- (121)a. mi domando [che libro Maria ha letto]
I wonder what book Maria read
- b. mi domando [che libro [ha affascinato Maria]]
I wonder what book charmed Maria
- c. mi domando [che libro [Paolo ha detto [che Maria ha letto]]]
I wonder what book Paolo said that Maria read
- d. mi domando [che libro [Paolo ha detto [che ha affascinato Maria]]]
I wonder what book Paolo said that charmed Mary

As noted by Kayne (1980) and by Rizzi (1982), Italian differs in another way from English: Italian allows the subject of a tensed sentence to appear post-verbally:

- (122)a. Giovanni mangia la pasta
Giovanni eats the pasta
- b. mangia la pasta Giovanni
- (123)a. Giovanni ha telefonato
Giovanni has telephoned
- b. ha telefonato Giovanni

- (124)a. Giovanni è arrivato
Giovanni has arrived
- b. è arrivato Giovanni

Rizzi argues that the inverted subject with polyadic transitive verbs and agentive intransitive verbs, as in (122)-(123), is adjoined to VP:

- (125)a. [_S e INFL [_{VP} [_{VP} mangia la pasta] Giovanni]]]
- b. [_S e INFL [_{VP} [_{VP} ha telefonato] Giovanni]]]

With ergative verbs like arrivare in (124), as well as passives, an inverted subject remains in its D-structure object position:

- (126) [_S e INFL [_{VP} è arrivato Giovanni]]

The inverted subject with regative verbs, when it is a bare quantifier, is like a direct object in allowing and requiring cliticization of ne 'of (pronoun)':

- (127)a. [Maria INFL [*(ne) ha letto molti]]
Maria of-them has read many
- b. [e INFL [*(ne) sono arrivati molti]]
of-them have arrived many

The inverted subjects in (125) do not allow cliticization of ne, and, in fact, may not be bare quantifiers:

- (128)a. *[e INFL [_{VP} [_{VP} (ne) mangiano la pasta] molti]]
of-them eat the pasta many
- b. *[e INFL [_{VP} [_{VP} (ne) telefonano] molti]]
of-them telephone many

Returning to the absence of CTP effects in Italian, notice that we have suppressed the traces in (121). This is because, given the possibility of free inversion of the subject, we do not yet know where the trace of WH is in (121)b and d. Concentrating on (121)d, suppose the trace is in the preverbal subject position, as in (129):

(129) [che libro_i [Paolo ha detto [_S, che [_S e_i INFL [_{VP} ha affascinato Maria]]]]]

The subject trace in (129) violates the ECP, just as its counterparts in (108)d and (109)d do. If (129) is an acceptable structure, we must either claim that the ECP does not apply to \bar{A} -bound traces in Italian, or else that the subject of a tensed sentence is properly governed by Italian AGR.

Alternately, suppose the trace is in the postverbal subject position, as in (130):

(130) [che libro_i [Paolo ha detto [_S, che [_S e INFL [_{VP} [_{VP} ha affascinato M.] e_i]]]]]

Under the definition of government in (113), e_i in the adjoined position is not governed by the verb in VP, since an instance of the maximal projection VP intervenes. Following Chomsky, however, we may alter the definition of government so as to allow V to govern NP in the configuration:

(131) [_{VP} [_{VP} V ...] NP]

Chomsky makes the definition of government refer to a particular definition of c-command that allows this case:⁴⁴

(132) Government (based on Chomsky 1981a, 250)

In the structure:

$[\beta \dots \gamma \dots \alpha \dots \gamma \dots]$, where

- (i) $\alpha = X^0$ or properly binds γ
 - (ii) where ϕ is a maximal projection, if ϕ dominates γ then ϕ dominates α
 - (iii) α c-commands γ
- α governs γ .

(133) C-command (Chomsky 1981a, 166)

α c-commands β if and only if

- (i) α does not contain β
- (ii) Suppose that $\gamma_1, \dots, \gamma_n$ is the maximal sequence such that
 - (a) $\gamma_n = \alpha$
 - (b) $\gamma_i = \alpha^j$
 - (c) γ_i immediately dominates γ_{i+1}

Then if δ dominates α , then either

- (I) δ dominates β , or
- (II) $\delta = \gamma_i$ and γ_1 dominates β

Rizzi argues that, given the possibility of free subject inversion, the ECP functions identically in Italian and in English, applying to WH-trace and not counting AGR as a proper governor. The trace of WH in sentence (121)d will satisfy the ECP only if it is located postverbally as in (130). Similarly, grammatical long movement of the subject of an ergative (or passive) verb will satisfy the ECP only if the trace of WH is located in the inverted, object position. That is, (134) may not have the structure in (135)a, but may only have the structure in (135)b:

(134) che libro Paolo ha detto che è arrivato
'what book P. said that has come'

- (135)a. *[che libro_i [Paolo ha detto [_S che [_S e_i INFL [_{VP} è arrivato
e_i]]]]]]
b. [che libro_i [Paolo ha detto [_S che [_S e INFL [_{VP} è arrivato
e_i]]]]]]

(The postverbal trace is an NP-trace in (135)a, but a WH-trace in (135)b.)

He presents two important arguments in favor of this hypothesis, the first involving the possibility of wide-scope quantification, an argument we take up briefly in the next section, the second involving the clitic ne that we examined at (127)-(128). We saw that ne may and must be extracted from a bare quantifier in direct object position. By contrast, ne may and must not be extracted from a bare quantifier in preverbal position, thus:

- (136)a. [_S e INFL [_{VP} *(ne) sono arrivati molti]]
 of-them have arrived many
b. [_S molti_i INFL [_{VP} *(ne) sono arrivati e_i]]

Now consider long movement of a bare quantifier. If long movement is possible from preverbal subject position, as in (135)a, then we expect ne to be missing. If long movement is only possible from postverbal object position (of an ergative or passive verb), then we expect ne to be obligatorily present. The second possibility is correct:

(137) quanti_i [Paolo ha detto [che *(ne) sono arrivati]]

(137) shows that the ECP applies to WH-trace in Italian, just as it does

in English, and also that AGR is not a proper governor in Italian, just as it is not in English.⁴⁵

But an obvious problem remains. Why doesn't the ECP rule out the e in subject position of (125)a-b, (126), (127)b, (130), (135)b, and (136)a? Italian appears to show null subjects freely in inversion constructions:

- (138)a. [_S e INFL [_{VP} è arrivato Giovanni]]
 b. [_S e INFL [_{VP} [_{VP} ha telefonato] Giovanni]]

-- when the null subject is interpreted as a definite pronoun:

- (139) [_S e INFL [_{VP} ha telefonato]]
 has telephoned
 (3 sg)
 'he telephoned'

-- and when the null subject is expletive, as in (140)a, or the impersonal passive (140)b:

- (140)a. [_S e INFL [_{VP} sembra [_S, che Luigi è venuto]]
 seems that Luigi has come
 'it seems that Luigi has come'
 b. [_S e INFL [_{VP} gli fu detto del pericolo]
 to-them was said about the danger
 'they were told about the danger'

A number of theories have tried to account for why the subject e in all these constructions is not subject to the ECP. Chomsky (1981a) proposes, adapting ideas of Jaeggli (1980b), that e in these sentences is PRO, ungoverned as a result of a rule cliticizing AGR to V applying in the

syntax. Chomsky has criticized this proposal in more recent work (Chomsky 1981b), noting among other problems the fact that e in these constructions does not have the semantic or contextual properties of the [+pronominal, +anaphor] empty category PRO.

An approach we find convincing, which we will presuppose here, is that taken by Safir (1982). He suggests, in effect, that the key to the proper treatment of e in these constructions is provided by (140)a-b, where e is clearly expletive. Suggesting that expletive e is not subject to the ECP, (140) no longer poses problems for the ECP. Similarly, in the ergative inversion construction of (138)a, there is no problem if we suppose that expletive e is not subject to the ECP. For the intransitive inversion construction of (138)b, and for polyadic transitive ergative constructions, Safir proposes that the inverted subject receives its indirect θ -role in situ, as a sister of the smaller VP, leaving the preverbal subject position expletive. In the "null pronoun" construction of (139), Safir proposes that the θ -role is borne by a subject clitic, phonologically null in literary Italian, but present in many Italian dialects. Assignment of the indirect θ -role to the subject clitic, in turn, leaves the subject e without a θ -role, hence expletive and immune from the effects of the ECP.

Clearly, Safir's proposal carries over at least in part to Russian. Russian clearly has null expletive subjects, in Russian equivalents of Italian (140)a-b:

- (141)a. [_S e INFL [_{VP} kažetsja [_S, čto Viktor prišel]]]
 seems that Viktor has come
- b. [_S e INFL [_{VP} bylo napisano ob opasnosti]]
 was written about danger

Russian has null subjects interpreted as definite pronouns:

- (142)a. [_S e INFL [_{VP} pojdu v Inotdel]]
 will go to foreign students office
 (1 sg)
 'I'm going to the foreign students office'
- b. [_S e INFL [_{VP} ne ponimaet po-russki]]
 NEG understands Russian
 (3 sg)
 'he doesn't understand Russian'

-- although there are certain differences between Russian and Italian in this respect, discussed in Pesetsky (1982).

On the other hand, does Russian have inverted subjects like Italian? Clearly our answer must be no, since Russian, unlike Italian, shows CTP effects, albeit somewhat "fuzzily" for some speakers. Russian does, of course, freely invert the subject:

- (143) ^Vžret soljanku Ira
 gobbles soljanka Ira
 (3 sg) (fem acc sg) (fem nom sg)
 'Ira is gobbling soljanka (a soup)'

But this inversion is part of a much larger phenomenon of scrambling:

- (144) soljanku tam rybnuju ^Vžret Ira vkusnen'kuju
 soljanka there fish gobbles Ira tasty
 (fem acc sg) (fem acc sg) (fem nom sg) (fem acc sg)
 'Ira is gobbling the tasty fish soljanka there'

We have no new suggestions to offer for handling scrambling phenomena here (cf. Hale 1979; Farmer 1980; Chomsky 1981a, 128-135; Stowell 1981;

among others). What is likely, however, is that the scrambled structures exhibited by languages like Russian are not the result of simple substitutions and adjunctions of the sort that normally distinguish S-structures from D-structures, but result from some more radical mapping from S-structures⁴⁶ onto surface structures, one which does not seem to obey the Projection Principle. For our purposes, we may simply assume a "scrambling rule" in PF,⁴⁷ although we could equally well assume that scrambling structures are co-represented along with "virtual", unscrambled structures at S-structure, as suggested by Vergnaud and Zubizarreta (1982). What the CTP effects appear to indicate is that sentences like (143) cannot have a structure like (145) at the level at which the ECP applies:

(145) * $[_S \underline{e} \text{ INFL } [_{VP} [_{VP} \overset{V}{Zret} \text{ soljanku}] \text{ Ira}]]$

How can we prevent structures like (145) in languages like Russian, or, for that matter, English?

(146) *eats the soup Ira

The answer appears to lie in Case Theory: Italian, unlike Russian or English,⁴⁸ has a way of assigning nominative Case to a postverbal subject. The exact mechanisms allowing this do not concern us (cf. Safir 1982 for discussion); what is important is that Italian has such a mechanism, while Russian and English do not.

Finally, let us consider one area in which Russian is closer to Italian than to English. If the ECP does not apply to expletive empty categories, we correctly predict the existence of empty expletive subjects in Russian and Italian. What, however, rules out empty expletive subjects in English?

(147) *(it) seems that John is here

Safir suggests that the answer lies in another parameter of variation. We have seen that Italian, but not Russian and English, allows Case assignment to a postverbal subject. As a second parameter, Safir suggests that English, but not Italian or Russian, requires AGR to assign nominative Case to some lexical subject. It is this parameter, which Safir calls "NOM-drop", that forces the existence of the lexical expletive in English, but allows the null expletive in Italian and Russian.

At this point, armed with the Revised ECP of (120), which excludes expletive e , we are ready to return to our main line of argument. Let us recall what we are looking for. In section 2, we argued that we could derive the Non-Obliqueness Condition (condition B) if our XPs were QPs, and if QPs did not bear Case or need Case features. This immediately predicts that QPs, unlike NPs, are not forced by the Case Filter to move from the object position of an ergative or passive verb. In other words, we predict that, while (148)a must yield an S-structure like (148)b, (149)a may yield an S-structure like (149)b:

- | | | |
|----------|---|--------------------|
| (148) a. | $[_S \underline{e} \text{ INFL } [_{VP} \begin{matrix} V \\ \text{-Case} \end{matrix}] \text{ NP}]]$ | <u>D-structure</u> |
| b. | $[_S \text{ NP}_i \text{ INFL } [_{VP} \begin{matrix} V \\ \text{-Case} \end{matrix}] \underline{e}_i]]$ | <u>S-structure</u> |
| (149) a. | $[_S \underline{e} \text{ INFL } [_{VP} \begin{matrix} V \\ \text{-Case} \end{matrix}] \text{ QP}]]$ | <u>D-structure</u> |
| b. | $[_S \underline{e} \text{ INFL } [_{VP} \begin{matrix} V \\ \text{-Case} \end{matrix}] \text{ QP}]]$ | <u>S-structure</u> |

First, let us note that movement from the subject position of a passive or ergative sentence, like movement from any subject position, shows CTP effects, derivable from the ECP. Here, for some speakers, the effects are fuzzier still (recalling Portuguese; Zubizarreta,

forthcoming). Still, the effects are present for most speakers:

- (150) ?*[kakaja kniga_i [Ivan xotel [_S ^Vctoby [_S e_i¹ INFL [_{VP} byla pročitana e_i²]]]
 what book Ivan wanted that by read
 'what book did Ivan want that be read'

The object trace e_i² is an NP-trace; it is the subject trace e_i¹, locally \bar{A} -bound from COMP, that violates the ECP. Note that e_i¹, although in a non- θ -position, is not expletive, since it is part of the chain (e_i¹, e_i²), which is directly θ -marked by the verb pročitana.

If we are correct about QP not needing Case, we predict that the equivalent of (150), with extraction of a QP instead of an NP, should not violate the ECP. In (150), the trace of NP must be in a Case-marked position; the trace of QP, on the other hand, should not need Case. This should allow WH-movement from the postverbal object position of a passive or ergative verb, despite the fact that Russian, unlike Italian, does not assign Case to this position. As we noted in Pesetsky (1982), this prediction is fulfilled:⁴⁹

- (151) ?[[_{QP} [_Q e] kakix knig] [Ivan xotel [_S ^Vctoby [_S e INFL
 what books Ivan wanted that
 (fem gen pl)
 [_{VP} bylo procitano e_i]]]]]
 be read
 (neut sg)

(151) forms a minimal pair with (150). Despite the fuzziness of (150) and the awkwardness of long movements in general (giving the "?" to (151)), the contrast between the two is apparent to most informants.⁵⁰

We take this contrast as good evidence for our conjecture that XPs

are categories that do not need Case features. In turn, if XPs do not need Case features, then they are not NPs. If they are not NPs, then by X-bar theory they must be QPs. Thus, we take our QP hypothesis, and the derivation of condition B from this hypothesis to be well-founded.

The results of this section will also form the basis for our derivations of principles A and C in the next section, and will bring us back to the basic questions about c-selection discussed in Chapter One. Before leaving this section, however, we should note that the data just discussed show that QPs in passive and ergative constructions are not obliged to move to subject position. Hence the lack of CTP effects. We have not shown that they cannot move to subject position. Should this also be the case, it might follow from Chomsky's reduction of the Case Filter to principles of θ -theory, discussed above. In the next section, we show, inter alia, that there is some reason to believe that QPs in fact cannot move to subject position, but our derivation of conditions A and C will provide another explanation for this fact.⁵¹

4.0 The ECP, Category Blindness, and Conditions A and C

4.1 QP and the ECP

In this section, we will assume the QP hypothesis to be correct, and we will take for granted the modification of the ECP discussed in the preceding section. Let us now restate conditions A and C in light of the QP hypothesis:

(152) A: QP must be a [QP, VP] at D-structure

B: QP obligatorily undergoes QR.

Notice now that the revised version of A above may itself be revised. If the QP hypothesis is correct, D-structure [QP, VP]s always have at least the option of remaining [QP, VP]s at S-structure, since no principle like the Case Filter forces them to move elsewhere. It is thus equally possible that condition A holds at S-structure. From the facts alone, we cannot tell.

Suppose we conjecture that condition A should be revised once more:

(153) A: QP must be a [QP, VP] at S-structure.

This newest version of A is actually a tighter condition than the earlier versions, since some D-structure objects are S-structure subjects, but no S-structure objects are non-objects at D-structure, by the Projection Principle. Thus, if (153) is true, then (152) is true. It will turn out, in fact, that we can maintain (153), and that it is (153), and not (152)'s version of condition A, that will allow us to derive A from something deeper.

There is another way to think about conditions C and A. We can consider the range of possible configurations involving QPs at LF, and

note that only one of them corresponds to a grammatical structure:

- (154)a. * $[_S \text{ QP INFL VP}]$
 b. * $[_S \text{ subject INFL } [_{VP} \text{ V...QP...}]]$
 c. * $[_S \text{ QP}_i [_S \text{ e}_i \text{ INFL VP}]]$
 d. $[_S \text{ QP}_i [_S \text{ subject INFL } [_{VP} \text{ V...e}_i \text{...}]]]$

(154)a-b both show structures in which QR has not applied, and QP is in an A-position. The stars on these structures restate condition C: QPs must undergo the movement rule QR at LF.

(154)c-d show structures in which QR has applied. In (154)c, the trace of QR is in subject position of a tensed sentence. The star on this structure restates condition A. In (154)d, the trace of QR is in the object position. The lack of a star again restates condition A.

In a sense, (154) is a restatement of conditions A and C as conditions on representations at LF. As such, (154) illustrates an important point. Although condition C, requiring QR, can be described as a generalization about derivations that map S-structures onto LF representations, and although condition A is most easily described as a condition on representations at S- or D-structure, the trace theory of movement rules allows us to suppose that the true source of these generalizations might lie in conditions on representations at the level of LF alone. From a descriptive point of view, these restatements of conditions A and C are all more or less equivalent. On the other hand, these restatements differ in their ability to lead us towards a theory that deduces A and C from deeper principles. As a purely tactical decision, we will suppose that treating A and C as conditions on

representations at LF is the right approach. We will try to show that this tactical decision has interesting consequences.

As an initial step, let us assume condition C to hold by stipulation. (We shall, of course, provide an explanation of C later.) We thus take for granted the fact that all QPs undergo the rule QR in the mapping from S-structure to LF, or, alternatively, that no QPs occupy A-positions at LF. We concentrate our attention on condition A -- the difference between (154)c and (154)d.

Condition A tells us that movement of a QP (assuming C) must take place from a position governed by V. Movement, for example, from the position governed by AGR -- the subject position of a tensed S -- is not allowed. A subject/object asymmetry of this sort immediately suggests an effect of the ECP, which we discussed in the previous section. If the ECP were involved in this contrast, however, it would have to be the case that the ECP applies at LF, so that it could analyze traces resulting from QR. Fortunately, there is evidence which supports the view that the ECP applies at LF, first presented by Kayne (1981e).⁵²

Kayne observed that the scope of the quantifier personne 'nobody' is determined by the negative particle ne. If scope is assigned by the adjunction rule QR, as argued by May (1977), then the particle ne determines the level at which personne adjoins. As a consequence, in a multiclausal structure, placing ne in a clause higher than the minimal clause containing personne has the effect of forcing "long movement" of personne. This is exhibited in (155)-(156). The (b) examples are LF representations for the (a) sentences:

- (155)a. j'ai exigé [_S, qu'ils n'arrêtent personne]
I required that they ne arrest nobody
- b. [_S₁ [_S₁ j'ai exigé [_S₂ que [_S₂ personne_i [_S₂ ils n'arrêtent e_i]]]]]
- (156)a. ?je n'ai exigé [_S, qu'ils arrêtent personne]
I ne required that they arrest nobody
- b. [_S₁ [_S₁ personne_i [_S₁ j'ai exigé [_S₂ que [_S₂ ils arrêtent e_i]]]]]

(155) shows a short movement of a direct object to an \bar{A} -position. (156) shows long movement of the same object. Both are acceptable for most speakers, although (156) may require some extra stress on personne.

On the other hand, the situation changes radically when personne is a subject. In this case, Kayne discovered, ne cannot appear in a higher clause. It must appear in the same clause as personne. Placing ne in a higher clause forces long movement of the subject personne:

- (157)a. j'ai exigé [_S, que personne ne soit arrêté]
I required that nobody ne be arrested
- b. [_S₁ [_S₁ j'ai exigé [_S₂ que [_S₂ personne_i [_S₂ e_i INFL ne soit arrêté]]]]]

- (158)a. *je n'ai exigé [_S₂ que personne soit arrêté]
I ne required that nobody be arrested
- b. [_S₁ [_S₁ personne_i [_S₂ j'ai exigé [_S₂ e_i INFL soit arrêté]]]]]

As Kayne observes, the LF contrast between (156)a, with long movement of an object, and (158)b, with long movement of a subject, is obviously related to the parallel contrast with WH-movement -- to CTP effects. In the present framework, the contrast between (156)b and (158)b follows

immediately from the ECP -- if the ECP applies at LF. The trace in (156)b is properly governed by V, but the trace in (158)b is not properly governed.^{53 54}

Even if the ECP applies at LF, however, it is not immediately clear how we can use the ECP to derive condition A. In the cases we have examined -- CTP effects with WH-movement and Kayne's personne facts -- the ECP produces a subject/object asymmetry only in the event of long movement outside S'. This is because a subject empty category can be properly governed by a binder inside S', and an object empty category is always properly governed by V. Looking at the consequences of condition A as displayed in (154)c-d, it seems as if the trace of QP should display no subject-object asymmetry, because the movement of QP in our examples is always "short" -- that is, S'-internal. In (154)d, where e_i is properly governed by V, it is clear why ECP does not intervene; in (154)c, however, e_i should also be properly governed -- by QP. We repeat the structures below:

(154)c. $*[{}_S QP_i [{}_S e_i INFL VP]]$

d. $[{}_S QP_i [{}_S \text{subject INFL } [{}_{VP} V \dots e_i \dots]]]$

Now suppose we were to find some reason why QP_i , contrary to appearances, does not properly bind its trace in (154)c-d. The ECP would still not rule out (154)d, since an object trace does not satisfy the ECP because of proper binding. On the other hand, (154)c would be correctly ruled out by the ECP, since it is proper binding that allows the subject trace to satisfy the ECP. In the discussion that follows, we will justify exactly this assumption. We will show that QP does not properly bind e_i in (154)c-d, and will derive condition A from the ECP

at LF. At first, our proposal will not look very appealing. In the next section, however, we will show that the proposal actually reduces immediately to the more general issue of the place of categorial selection in the grammar, which we raised in Chapter One.

In our statement of a revised definition of government in (132), we have altered Chomsky's statement "where $\alpha = x^0$ or is coindexed with γ " to "...or properly binds γ ". In (114), we gave a definition of Proper Binding, which we repeat:

(159) Proper Binding⁵⁵

α properly binds β if and only if

- (i) α and β are coindexed
- (ii) α c-commands β
- (iii) α is a possible antecedent for β

What does "possible antecedent" mean? Of course, we have placed this term in (159) with malice aforethought, but some such requirement seems natural and uncontroversial. Suppose we assume the following definition:

(160) Possible Antecedent (def)

α is a possible antecedent for β if and only if α and β bear non-distinct values for number, gender and categorial features.

For number and gender, this condition is straightforward, particularly when WH is a possible antecedent for its trace, where the number and gender of the trace are in any case determined by the range assigned to it by the WH-operator.⁵⁶ The part of this definition that we will make heavy use of involves categorial features -- the features that distinguish N from V, V from Q, etc. This part of the definition has

a crucial consequence: if α properly governs β by coindexation, they must be of the same category.

What syntactic category does the trace of QP in (154)c-d belong to? In earlier work on the "trace theory of movement rules" (e.g. Fiengo 1977, and cf. the resumé in Chomsky, 1977b, 14; 1981a, 85-89), a "trace" is viewed as the former position of a moved element, which is not deleted after movement, but remains, with null lexical content. On this view, as a null hypothesis, the trace of a moved QP will be a QP. Consequently, we would expect a fuller representation of (154)c-d to be as in (161)a-b:

- (161)a. $*[{}_S QP_i [{}_S [{}_{QP} e]_i INFL VP]]$
 b. $[{}_S QP_i [{}_S \text{subject} INFL [{}_{VP} V...[{}_{QP} e]_i...]]]$

As noted above, (161)a in no way violates the ECP.

Suppose, however, that this view of trace theory were slightly mistaken. Suppose one could view trace theory somewhat differently, so that in certain cases the trace of a moved constituent might be of a syntactic category different from that of the moved constituent. In other words, suppose there were some reason to say that the trace of a moved αP , where α is some bundle of syntactic features, could be a βP , $\alpha \neq \beta$. In particular, suppose that the following descriptive statement were to be true:

- (162) Movement of a QP leaves an NP trace.

If (162) were true, then we would be able to derive the contrast between (154)c and (154)d from the ECP. (154)c-d would be more accurately represented, not as (161)a-b, but as (163)a-b below:

- (163)a. * $[_S QP_i [_S [_{NP} e]_i INFL VP]]$
 b. $[_S QP_i [_S subject INFL [_{VP} V...[_{NP} e]_i...]]]$

(163)b does not violate the ECP, since the object trace is properly governed by V, whatever its categorial specification may be. (163)a, on the other hand, does violate the ECP, since QP_i does not properly bind the NP trace. QP is coindexed with the NP trace, and does c-command it, but it is not a possible antecedent, since they are categorially distinct. Since neither QP nor AGR in INFL properly govern the subject trace in (163)a, it violates the ECP.

Thus, if (162) were true, we could derive condition A from the ECP applying at LF, if we assume condition C as a stipulation. This is a small advance, since we have factored from our surface generalization A the effects of a universal condition, the ECP. Given the ECP, we may say that condition A reduces to the proposition in (162):

(164) A: Movement of a QP leaves an NP trace.

Nonetheless, this whole discussion is so far rather fantastical, since the proposition in (162) is entirely unmotivated and at present rather bizarre. We will answer this objection in the next two subsections. In the next subsection, we will derive condition C.

4.2 Categorial Selection and Condition C

In Chapter One, we promised that our analysis of Russian quantificational constructions would provide evidence about the status of c-selection (categorial selection). Recall that we considered two possible views.

One view, implicit but not crucial in Chomsky (1981a), holds that c-selection, like the Θ -criterion, falls under the Projection Principle. This means that a verb that c-selects an NP at one level c-selects an NP at all levels: D-structure, S-structure and LF. The other view, which we argue for here, holds that only the Θ -criterion in its strictest sense, and not c-selection, falls under the Projection Principle. If this view is correct, a verb which Θ -marks a position and c-selects an NP in that position will have to find that position filled by some category at every level, by the Projection Principle, but will not have to find that position filled by an NP at each of these levels. Only at one level will this position have to be filled by an NP -- the level at which c-selection applies. We suggested that this level is LF.

Crucially, we are not rejecting the Projection Principle. Rather, we are suggesting that the Projection Principle is category-blind. Categorical requirements are not projected by it to all levels. Only Θ -requirements are.

The reader may have noticed that we have actually been assuming this proposition in advancing the QP hypothesis. In section 2, we proposed that various problems, particularly the Non-Obliqueness Restriction (condition B), can be solved if we allow Russian to have representations at D-structure and at S-structure of the form:

- (165)a. ja ne polučal^V QP
 I NEG received
- b. e prišlo^V QP
 arrived
 (neut sg)

What are the categorial requirements of the verbs in (165)? Intuition and the general facts of the matter tell us that verbs meaning 'receive' or 'arrive' select an NP -- a semantic "term". It is clear, for example, that they do not c-select a clause:

(166)a. *John received [that Bill argued with Mary]

b. *[for Bill to leave] arrived yesterday

-- nor do they c-select an AP:

(167)a. *John received [pleased with Mary]

b. *[hot enough for June] arrived yesterday

One may well presume that they also do not select a quantifier. A quantifier is a logical operator: operators do not arrive, nor are they received, nor do they do any of the things we have let our QPs do in the previous examples. When six men arrive, it is not the operator six that arrives, but a set of men.

Supposing it to be true that verbs like receive and arrive do not select QPs, but do select NPs, it is clear that representations like (165)a-b will violate any theory of c-selection that applies to them. If (165)a-b are actual D-structures and S-structures, as we have argued, then c-selection must not apply at either of these levels. In advancing the QP hypothesis, we have thus committed ourselves to the position that c-selection does not apply at D-structure or at S-structure. Of course, the fact that we have committed ourselves to a position does not argue for it. Rather, we must show that this position explains phenomena other than those that motivated it in the first place.

We do have two stray phenomena which are not yet explained. The first

is condition C, which makes QR obligatory for QPs. The second is condition A, revised as in (164). Let us look first at condition C.

We have found that the QP hypothesis entails that c-selection does not apply at D-structure or at S-structure. It must apply at some level, however. In Chapter One, we suggested that this level is LF. The problem, of course, is that representations like (165)a-b do not satisfy c-selection at LF any more than they do at D-structure or at S-structure. The consequence of this is that some rule must "change" representations like (165)a-b in the mapping from S-structure to LF, so that they can meet the requirements of c-selection. What should these sentences look like, in order to meet the requirements of c-selection? Wherever (165)a-b show a QP, the corresponding representations at LF should show an NP. What rule can yield an NP in these positions? The answer, of course, is QR -- assuming the revision of condition A in (164): Movement of a QP leaves an NP trace. Given this revision of condition A, it follows that QR must apply to QPs in positions where an NP is c-selected, simply in order to satisfy c-selection at LF. With respect to all the examples we have considered so far, this is condition C.⁵⁷

To recapitulate: in the previous section we showed that our previous formulations of condition A can be replaced by (164) and the ECP, if we stipulate condition C. Now we see that, given the hypothesis that c-selection applies only at LF, condition C itself can be derived from (164). Representations like (165)a-b, which would not satisfy c-selection at LF, can be mapped onto representations that do satisfy c-selection at LF by the rule QR, yielding (168), schematically:

- (168)a. [_S QP_i [_S ja ne polučal [_{NP} e]_i]]
- b. [_S QP_i [_S e prišlo [_{NP} e]_i]]

The next step, obviously, is to derive the odd assumption that QP leaves an NP trace from some more general principles.

Before considering how to derive (164), let us note one respect in which we have not derived condition C in full, and ask whether we are correct not to do so. From the discussion just ended it follows that QP will have to undergo QR whenever it finds itself in a position where an NP is s-selected. On the other hand, if QP is in a position where QP is in fact s-selected, nothing will require it to move. The ECP will then not intervene, and we should not find any of the subject/object asymmetries that motivated condition A in the first place.

Predicates that s-select a QP are naturally rather rare. Nonetheless, some may exist: predicates that take logical operators as arguments. Recall, for example, that adjectives always indirectly θ -mark the argument to which they assign the role attribute. From this it followed that attributes are always D-structure subjects, from which it follows that a QP may not occupy a position in which the θ -role attribute is assigned and an NP is c-selected:

- (169) * [_{QP} šest' gorodov] bylo čisto v našej oblasti
 six cities was clean in our district
 (neut sg)

The "arithmetic" adjective ravno 'equal', however, may be an adjective that assigns the role of attribute to a QP (as well as an NP). This is semantically plausible, since 'equal' is a predicate that may range over quantities. We may thus explain the fact that ravno is one of the very

few adjectives that may have a QP as its subject:

(170) [_{QP} ^V šest' gorodov] bylo ravno trem oblastjam
 six cities was equal three districts
 (neut sg) (fem dat pl)
 'six cities are equivalent to three districts'

Various constructions involving obligatory genitives may also be examples of predicates selecting QP -- in this case a QP with a null Q, as in the genitive of negation (cf. also 5.5 below, on partitive genitives):

(171) [_{QP} [_Q e] sosisok] bylo navalom
 hot dogs was galore
 (fem gen pl) (neut sg)
 'there were hot dogs galore'

If more evidence can be brought to bear on the analysis of these constructions, and they should prove to be cases of predicates selecting QPs, which violate our earlier formulation of C, and of A, then we will have some additional confirmation of our derivation of C and the earlier formulations of A from (164), the ECP, and the assumption that c-selection only applies at LF. In fact, a solid argument of just this kind will be provided in section 5, where we discuss infinitival free relatives.

4.3 Categorial Selection and Condition A

The pieces of our story now include the QP hypothesis, Case Theory, the ECP, the theory that c-selection applies only at LF, and condition A rewritten as (164). We are moving towards an explanatory account of our original three descriptive conditions -- explanatory because these three conditions are unlearnable, odd restrictions on surface phenomena as they stand. Case Theory, the ECP and our theory about c-selection, which now

interact to yield most of A, B and C, are all plausible candidates for innate principles of Universal Grammar. The knowledge that our structures are QPs is derivable from very obvious facts about Russian (many quantifiers are followed by genitive Case) and the unmarked principle that heads of phrases assign Case to their complements. The one implausible component of our story so far is (164). We therefore would like to derive the stipulation that QPs leave NP traces from some other principle.

We shall show that the solution to this problem is already in hand. We have motivated the obligatory application of QR to QPs by the theory of c-selection applying at LF, on the assumption that the trace of QP in an NP. We can demonstrate, however, that it is the same theory of c-selection that allows and forces the trace of QP to be an NP in the first place.

Let us have another look at the Projection Principle, which, we are claiming, projects the θ -criterion, but not c-selection, from the lexicon to D-structure and S-structure, as well as to LF. Chomsky (1981a) notes a happy consequence of the Projection Principle: it implies the existence of traces in a large number of cases. Speaking crudely, suppose Move α applies to a constituent in a θ -position, but no empty category were left in that position in the derived structure. Clearly, the θ -criterion would be violated. The θ -role assigned to that position before movement would not be unassigned, and, by the Projection Principle, θ -roles must be biuniquely assigned to constituents at all levels. Hence, an empty category must be present in a θ -position after movement, because of the θ -criterion and Projection Principle. In the case of NP-movement to a non- θ A-position, the coindexation of the trace with its binder is forced again by the θ -criterion: a moved argument would otherwise lack a θ -role. In the case of movement to an \bar{A} -position, other principles

force the coindexation -- for example, Chomsky's (1981b) principle that operators must bind a variable, in the case of WH-movement.

Interestingly, other principles of the grammar derive trace theory for movement from a non- θ -position. For example, the Case Filter requires \underline{e}_i^1 to exist in (172):

(172) [_S, who_i [_S \underline{e}_i^1 INFL [_{VP} was seen \underline{e}_i^2]]]

Unless \underline{e}_i^2 is part of the chain (e_i^1, e_i^2), it will lack Case, and will not be able to bear its θ -role, given Chomsky's derivation of the Case Filter.

In (173), \underline{e}_i^3 is motivated by the Projection Principle; \underline{e}_i^2 and \underline{e}_i^1 are required by principle A of the Binding Theory:

(172) John_i seems [_S \underline{e}_i^1 to be likely [_S \underline{e}_i^2 to be invited \underline{e}_i^3]]]

The Projection Principle requires at least the chain (John_i, \underline{e}_i^3). But \underline{e}_i^3 is an anaphor, and must, by the Binding Theory, be bound within its S: hence it is bound by \underline{e}_i^2 , establishing the chain (John_i, \underline{e}_i^2 , \underline{e}_i^3). But now \underline{e}_i^2 is an anaphor, and must be bound within the next highest S, establishing the chain (John_i, \underline{e}_i^1 , \underline{e}_i^2 , \underline{e}_i^3).⁵⁸

The only empty categories traditionally posited under the "trace theory of movement rules" that do not seem to be motivated by independent subtheories of grammar are certain traces in COMP, like \underline{e}_i^1 in (173):

(173) [_S, who_i [_S does John think [_S, [_{COMP} \underline{e}_i^1] [_S Bill likes \underline{e}_i^2]]]]

I know of no principle that will force \underline{e}_i^1 to exist,⁵⁹ and it is perhaps safe to assume that such traces in COMP do not have to exist.

A trace in COMP is obligatory, however, when the ECP forces it to exist in order to bind the subject position (cf. Pesetsky 1982; Chapter

Three):

(174) [_S, who_i [_S does John think [_S, [_{COMP} e_i¹] [_S e_i² likes Bill]]]]

Given the almost complete redundancy of the "trace theory of movement rules" with other subsystems of grammar, it is tempting to derive trace theory in toto from the other subsystems (see Bouchard 1982 for a development of this program). Instead, we might suppose that empty categories are freely assumed at any level of representation, with free coindexation, subject to independent principles like those of Binding Theory, Case Theory, the ECP, etc.

Let us now combine this treatment of trace theory with our proposal about the status of c-selection in the grammar. All of the QPs we have considered so far (with the exception, possibly, of the expressions of duration; cf. note 57) occupy θ -positions at D-structure and S-structure, since Case Theory never forces them to move in the syntax.⁶⁰ When these QPs undergo QR, therefore, the existence of traces of these QPs will be motivated by the θ -criterion and Projection Principle, applying at LF. We have seen that the θ -criterion and Projection Principle require the presence of some empty category occupying a vacated θ -position. Do these principles require anything more? In particular, do these principles say anything at all about the categorial identity of the empty category assumed after movement of a QP?

As noted in Chapter One, if the theory of c-selection is not subject to the Projection Principle, then the Projection Principle will be category-blind. The θ -criterion at D-structure and at S-structure will not distinguish among syntactic categories. Insofar as these principles are responsible for the existence of empty categories in θ -positions

vacated by Move α -- i.e. "traces" -- these principles will say nothing at all about the categorial identity of these empty categories. In other words, if trace theory (in the cases we are considering) derives from the Θ -criterion and Projection Principle, and if these principles are category-blind, then trace theory will be category-blind. From this it follows that the trace of a moved constituent may be of any category at all, as far as these principles are concerned. This is an immediate consequence of the proposal that c-selection be separated from the Projection Principle.

What does determine the category of a trace? If trace theory is not an independent subtheory relating traces to Move α , the category of a trace will not have any inherent relation to the category of a moved constituent. Only the interaction of independent principles will determine the category of a trace.

One such independent principle is, of course, the theory of c-selection, applying at LF. Let us return to our final revision of condition A in (164), which represented the residue of condition A once the ECP was factored out. We have already noted that (164) is not a plausible candidate for a principle of UG. What is more, it makes the sort of stipulation about a trace left by movement that we might want to eliminate from the grammar. We can now derive (164) in its essentials from our independently motivated ideas about c-selection.

Suppose (164) did not hold. Suppose a QP were to be moved by QR, and its vacated position were filled by a QP trace. As we saw in (161), this would exclude an ECP account of the subject/object asymmetries that QPs exhibit. Why should such a derivation be excluded? By now the

answer is obvious: if a QP trace is left in a position where an NP is c-selected, the structure will be ruled out at LF. Predicates like polučat^V 'arrive' c-select an NP. If movement of a QP left a QP trace, these selectional requirements would be violated.

If movement of a QP left an NP trace, however, c-selectional requirements would be met. No principle prevents QP from leaving a trace that is an NP. Hence, c-selection dictates that the only possible output of QR applying to a QP (where an NP is required) involves a trace of category NP. Hence, stipulation (164) derives simply from our general proposal about the status of c-selection in the grammar.

Thus, from the QP hypothesis, Case Theory, the ECP, and the assumption that c-selection applies only at LF, we derive all the relevant cases of our original three descriptive conditions. The QP hypothesis is inextricably linked to the proposal that c-selection applies only at LF. From the QP hypothesis we derive the Non-Obliqueness Restriction, condition B. From the QP hypothesis and c-selection it follows that the position occupied by a QP at S-structure in our examples must be occupied by an NP at LF. From c-selection and the elimination of an independent trace theory it follows that submitting QP to QR can leave an NP trace that will satisfy c-selection at LF. From c-selection it follows that this NP trace must be left; hence QR is obligatory; hence condition C. But a QP that leaves an NP trace will not properly bind this trace, hence cannot properly govern it by coindexation. The ECP then dictates that the NP trace cannot be the subject of a tensed S; hence condition A.

Some loose ends need to be tied. First of all, we claim that QP does not properly bind an NP trace. Nonetheless, it must be able to bind it in some sense, or else the NP trace would be taken to be free, e.g. in the following configuration:

(175) [_S QP_i [_S subject INFL [_{VP} V [_{NP} e]_i]]]

If [_{NP} e] is free, then by the contextual definitions of empty categories given by Chomsky (1981a, 1981b) and sketched in Chapter One, [_{NP} e]_i should be PRO, and ungoverned. Since [_{NP} e]_i is governed, the contextual definitions and the Binding Theory should rule (175) out. Clearly what we must say is that the contextual definitions refer, not to proper binding, but just to binding -- i.e. to coindexation and c-command, without the requirement of possible antecedenthood. Thus, while a QP does not properly bind an NP trace, it does bind it, and in particular \bar{A} -bind it, in the sense relevant for the definitions of empty categories.

With this in mind, however, let us consider what might happen to a QP if it underwent movement from a Θ -position where an NP was c-selected to a non- Θ A-position. Recall that we have shown that QPs generated in D-structure Caseless objects positions are not forced by any principle to move to a Case-marked A-position. We noted there that we did not necessarily have any principle that prevents QP from moving to a Case-marked A-position. Possibly, Chomsky's revision of the Case Filter has that effect, as we pointed out, since it might entail that QP cannot belong to a chain. When we temporarily restated condition A in 4.1 as "QP must be a [QP, VP] at S-structure", we were actually claiming that this stronger condition, preventing movement to an A-position, was true, without evidence.

We simply pointed out that this restatement did not conflict with anything we knew. It might seem, however, that our derivation of condition A from the ECP and our theory of c-selection loses this stronger claim. Consider an S-structure representation in which a D-structure QP in an object position where an NP is c-selected has moved to subject position. By c-selection, its trace must be an NP:

(176) $[_S \text{QP}_i \text{ INFL } [_{VP} \text{V } [_{NP} \text{e}]_i]]$

An LF representation derived from (176) would not violate the ECP. The trace of QP in subject position could be a QP, since no c-selectional requirements are made of this position, and this trace could be assumed to be expletive. The variable would be the NP-trace in object position (as suggested to me by M. Brody):

(177) $[_S \text{QP}_i [_S [_{QP} \text{e}] \text{ INFL } [_{VP} \text{V } [_{NP} \text{e}]_i]]]$

In point of fact, it is not clear that anything in (176) would even force QR in the first place, thus undermining condition C. This suggests that we want to rule structures like (176) out.

As noted, one way to rule (176) out is by appealing to a convention that QP cannot belong to a chain. Another way might concern the definition of chains: reasonably, since elements of a chain seem to share all their features, the members of a chain cannot contribute conflicting categorial features to the chain. This would have as a result that we could not form the chain $(\text{QP}_i, [_{NP} \text{e}]_i)$ in (176), and the Θ -criterion would be violated. Alternatively, we might suppose that there is no direct restriction on a chain of this sort, but that the Binding Theory requires

Proper Binding, like the ECP and unlike the contextual definitions of empty categories. Thus, by the contextual definitions, the NP trace has an antecedent in an A-position that lacks an independent θ -role, and thus is an anaphor; according to the Binding Theory, this anaphor is A-free, and violates principle A.

Is there independent evidence that we wish to rule out structures like (176)? Are QPs never subjects at any level? There is some evidence that bears on this question, although it is rather weak.

First, the reflexive pronoun sebja generally must take an S-structure subject as its antecedent:

- (178) $\overset{V}{\text{Maša}}$ ubil Volodju u sebja
 $\overset{V}{\text{Maša}}$ killed Volodja chez(reflexive)
 (nom) (acc)
 'Maša killed Volodja at her house'
not ' $\overset{V}{\text{Maša}}$ killed Volodja at his house'

Now compare:

- (179)a. [_{NP} ni odin mal' \overset{V} cik] ne byl ubit u sebja
 not one boy NEG was killed chez (reflexive)
- b.??[_{QP} ni odnogo mal' \overset{V} cika] ne bylo ubito u sebja
 (neut sg)
- (180)a. [_{NP} šest' studentov \overset{V}] byli ubity u sebja
 six students were killed chez (reflexive)
 (pl)
- b. ?[_{QP} šest' studentov \overset{V}] bylo ubito u sebja
 (neut sg)

(The data in (179) are alluded to by Babby (1980, 39).)

We might conclude from these data that QPs are always S-structure

objects, and thus cannot be antecedents for reflexives. Of course, there are other explanations available. For example, if the Binding Theory does require proper binding, and not just binding, the categorial mismatch between a QP and the NP reflexive sebja might violate the Binding Theory, no matter where QP was at S-structure.

A similar argument is suggested by C. Neidle (personal communication), who notes that the understood subject of a verbal adverb (often called a "gerund" in Russian grammar) is controlled by an S-structure subject, in literary Russian. As expected, there is a contrast in control between an NP and a QP:

- (181)a. [PRO_i vozvrasčajas' domoj]^{VV}, [NP ni odin mal'cik]_i ne byl ubit
while-returning home not one boy NEG was killed
- b. *[PRO_i vozvrasčajas' domoj]^{VV}, [QP ni odnogo mal'cika]_i ne bylo ubito
(masc gen sg) (neut sg)

A second piece of evidence that QPs may not move to an A-position in syntax comes from the unmarked word order of passive and ergative constructions containing QPs. As we remarked in section 1, the unmarked position of such QPs is after the verb. In view of the clear intuitions of native speakers with respect to word order, it seems likely that some level of representation has this order as a property, and that some function maps this level onto a level with fewer constraints on order (cf. footnote 46). Since the "unmarked order" appears to yield non-discontinuous constituents, where governed elements are adjacent to governors and θ -roles are assigned locally, it is likely that this level is in fact S-structure, since S-structure maps onto LF, where such properties, once more, seem minimal criteria for adequacy. If this is the case, then if the unmarked position

of QPs is postverbal, QPs must be postverbal at S-structure. Since subjects are preverbal in Russian (in the unmarked case), it follows that QPs are never subjects at S-structure⁶¹

Finally, we may return to the fact, discussed throughout section 1, that verbs do not agree with QPs. In section 1, we suggested that this might derive from a simple stipulation that only NPs trigger agreement. This may be the case, but we may now derive the same fact from the simple absence of a subject at S-structure. Russian verbs do not agree with their objects.

Another loose end we should tie concerns the level at which the Case Filter applies. Recall that the Case Filter applies to NPs, but not to QPs, whether the Case Filter is a primitive or derives from the θ -criterion, as argued by Chomsky. This claim allowed us to distinguish the S-structure behavior and distribution of NPs from QPs in a way that was vital to our derivation of the three descriptive conditions of section 1. Thus, for example, (182)a was taken to violate the Case Filter, while (182)b was taken not to violate it:

- (182) a. $[_S \underline{e} \text{ INFL } [_{VP} \overset{V}{-}\text{Case}] \text{ NP }]]$
 b. $[_S \underline{e} \text{ INFL } [_{VP} \overset{V}{-}\text{Case}] \text{ QP }]]$

We now argue, however, that (182)b has the following structure at LF, if the ergative or passive verb in question c-selects an NP:

- (183) $[_S \text{ QP } _i [_S \underline{e} \text{ INFL } [_{VP} \overset{V}{-}\text{Case}] [_{NP} \text{ e }] _i]]]$

The NP trace heads its chain and lacks Case. If (183) is an acceptable LF, we must say one of three things:

(1) If we assume a simple Case Filter like (26), we may simply stipulate that (26) applies only at S-structure, and not at LF. Thus, only S-structure NPs require Case. Similarly, if the Case Filter is reduced to the Θ -criterion and a principle like Chomsky's (101), we may stipulate that (101) is only valid at S-structure, and not at LF. On this hypothesis, we may say that the NP trace does form a (singleton) chain at LF, but no Case requirements are placed on this chain.

Alternatively, (2): Suppose we adopt Chomsky's reduction of the Case Filter to the Θ -criterion and (101). (101) requires a chain that bears a Θ -role to have Case or be headed by PRO. To make (101) work, we needed an auxiliary assumption, which we stated in (104): "If an NP bears a Θ -role, it is a member of some chain". We might assume that (101) applies at all levels, but that (104) is a principle only of S-structure. Thus, the NP trace in (183) need not belong to a chain at LF, and thus need not have Case.

Finally, we might assume (3): Ergative and passive verbs may be taken not to assign Case because of Burzio's generalization, that verbs that do not indirectly Θ -mark their subjects also do not assign Case to their objects. Suppose this generalization, or whatever it may reduce to, applies only at S-structure. Then the NP trace in (183) will in fact receive Case at LF, and no problems arise.

We suspect that the correct approach to (183) limits the Case Filter to S-structure, as in alternative (1), but it is clear in any case that there are a number of ways to resolve the problem. Further research may indicate which of these approaches, if any, is correct.

Another loose end that we might try to tie concerns the categorial status of the elements we have been calling QPs. What is Q? As we noted earlier, Q does not fall naturally into the feature system of Chomsky (1970), which uses the two features [N] and [V] to distinguish the categories N, V, A and P. One possibility worth considering, which was suggested by Ian Roberts (personal communication), is that our Qs are actually Ps, with the additional feature [+operator]. Many Ps do assign genitive Case, particularly compound prepositions, suggesting that genitive is the unmarked Case for prepositions to assign, as we have suggested for Q. We do know that the feature [+operator] also coexists with adjectives (like vse 'all') and nouns (like bol'sinstvo^V 'majority'; cf. Crockett 1976b). We thus actually expect to find prepositions that are also operators:

(184) [_{PP} [_P šest' [_N studentov]]]
 six students

Babby (1982) notes that modern Russian quantifiers like šest'^V were actually nouns in earlier stages of Russian. As such, they did not observe the Non-Obliqueness Condition. (185) was well-formed in Old Russian:

(185) s toju pjat'ju staryx ^V ^{VV} zenscin
 with that five old women
 (fem instr sing) (fem gen pl)

If modern Russian šest'^V 'six', pjat' 'five', etc. are in fact prepositions, then the change from Old Russian to Modern Russian involves one feature: Old Russian [+N, -V, +operator] became [-N, -V, +operator].

In fact, Babby (1982) notes constructions in modern Russian which appear to be quantificational PPs (which he calls QPs, but with essentially

the same analysis). These phrases assign genitive Case:

(186) každyj^V lingvist znaet [[_{pp}ot dvux do pjati] jazykov]
 each linguist knows from two to five languages
 (gen) (gen) (masc gen pl)

(Cf. Babby 1982, (17))

As Babby observes, these quantificational phrases obey the Non-Obliqueness Restriction, condition B. The verb vladet' is essentially synonymous in (187) with znat' 'know' in (186), but requires instrumental Case:

(187) *každyj^V lingvist vladeet [[_{pp}ot dvux do pjati] jazykov/jazykami]
 each linguist knows from two to five languages
 (masc gen/instr pl)

We could explain this contrast if these phrases were themselves PPs, in which the smaller PP functions as the head. Various problems arise, of course, and we have no definitive answer to offer. Nonetheless, the existence of the construction provides a tantalizing suggestion that Roberts' conjecture that our Qs are Ps may be correct.

A final loose end, which we will not tie very adequately, has to do with the relation of negation to the genitive of negation construction. As we noted earlier, we have suggested that conditions A, B, and C have nothing directly to do with negation. First, they show up in constructions other than the genitive of negation construction. Second, a coherent explanation of all these properties is possible without reference to the negation. Nonetheless, there is a relation between negation and the genitive of negation construction, which merits investigation. Certainly QPs with the structure of those found in the genitive of negation are not found in simple positive sentences:

(188) ja *(ne) ljublju [_{QP} [e] kartin]
 I NEG love pictures
 (fem gen pl)

We speculate that the negation serves to "identify" the empty Q and supply it with the features necessary to act as a quantifier at LF. In this case, negation seems to indicate that the empty quantifier is existential. If c-command is a precondition for identification, we can explain Babby's observation that the genitive of negation is always interpreted as in the scope of negation at LF: that is, the negative version of (188) can correspond to a logical representation like (189)a, but not (189)b:

(189)a. $\neg \exists x: x \text{ picture } (I \text{ love } x)$
 b. $\exists x: x \text{ picture } \neg (I \text{ love } x)$

In the LF corresponding to (189)a, negation c-commands the quantifier and can identify it, while in (189)b, the quantifier remains unidentified.

In this connection, it is worthwhile noting that the "genitive of negation" does not only occur with negation, but also with adverbial quantifiers like mnogo 'many' in (190) (cf. Haik 1981 for discussion of the corresponding construction in French):

(190) ja mnogo ljublju [_{QP} [e] kartin]
 I many love pictures
 (fem gen pl)

'I love many pictures'

Here too, if we have analyzed the sentence correctly, the empty Q is identified. In (190), it is the adverbial mnogo that identifies it as the quantifier many.

It is worth noting that locality conditions of a familiar sort appear to govern this "identification" relation: a negation (or adverbial) in a higher clause may identify a null Q in a lower clause only if the lower clause, and all clauses intervening between it and the higher clause, is an infinitive:

(191)a. ja ne xoču^V [_S, PRO pisat' [_{QP}[_Q^e] stixov]]
 I NEG want write verses
 (infin) (masc gen pl)

b. *ja ne skazal [_S, čto ty pišeš^{V V} [_{QP}[_Q^e] stixov]]
 I NEG said that you write verses
 (indic) (masc gen pl)

Obviously, more study is needed to explore the relation between negation, adverbials, and null Qs. We discuss "identification" in another context in Chapter Four, 4.4 (and cf. Jaeggli 1980b; Chomsky 1981b).

Having examined some loose ends, of which we have tied a few, we consider in the next section two other domains of Russian grammar which invite an analysis similar to our analysis of Russian QPs. We also discuss in the next section some of the properties of the negative existential net, which seems to support the QP hypothesis. In the final section of this chapter, we will question the very existence of c-selection, which we have already extracted from θ -theory, and we discuss some issues in English grammar, as well as the Russian partitive construction.⁶²

4.4 Two More Cases

4.4.1 Infinitival Free Relatives

Our explanation of the behavior of Russian QPs has centered on a very small change in the framework we began with -- the relation of

c-selection to LF. This change, we argued, had radical consequences for the nature of the empty categories called "traces". We have suggested that traces may in principle belong to any syntactic category, that their categorial status is independent of the status of constituents moved from their position. If this is true, and our explanations are correct, can we prevent the massive overgeneration that our theory seems to allow? Can independent principles restrict the categorial nature of traces? For example, we do not expect to be able to place an AP in a position where NP is required in constructions like (192)a and legitimate it by QR, yielding (192)b:

(192)a. * [_S John INFL [_{VP} bought [_{AP} drunk]]] (S-structure)

b. [_S [_{AP} drunk]_i [_S John INFL [_{VP} bought [_{NP} e]_i]]] (LF)

What rules out (192) is probably the fact that the NP trace in (192)b cannot be interpreted as a variable: the adjective drunk cannot assign it a range. Adjectives can, of course, participate in quantificational expressions, when they help to assign the range of an appropriate variable:

(193) [how drunk]_i do you think John was [_{AP} e]_i

On the other hand, we do expect to find consequences of our proposal in domains other than QPs. Any phrase that contains an operator that ranges over categories other than the one dominating it should be able to show the same phenomena we have found Russian QPs to exhibit. If an operator ranges over X, but is dominated by Y (as our Qs range over N -- terms -- but are dominated by QP), Y should be able to fill a position where X is c-selected, as long as it vacates that position before LF.

Our familiar effects should then be found.

In this context, consider an infinitival S' whose COMP is filled by a WH-operator:

- (194) a. [_S, [_{COMP} [_{NP} what]_i] [_S PRO to [_{VP} read [_{NP} e]_i]]
 b. [_S, [_{COMP} [_{NP}^Vcto]_i] [_S PRO INFL [_{VP}^Vcitat' [_{NP} e]_i]]]

Such an S' may fill a position where an S' is c-selected. In such a case, the S' is interpreted as an indirect question:

- (195) a. I asked [_S, what_i PRO to read e_i]
 b. ja sprosil [_S,^Vcto_i PRO^V citat' e_i]

Suppose just such an S' is placed in a position where an NP is c-selected. Contrary to prediction, and for unknown reasons, the result is bad in English:

- (196) *I bought [_S, what_i PRO to read e_i]

The result, however, is good in colloquial Russian, and in a number of other languages (including Spanish, which we discuss below):

- (197) ja kupil [_S,^Vcto_i PRO^V citat' e_i]

I bought what read
 (infin)

'I bought something to read'

We can explain this immediately, if the S' undergoes QR in the mapping to LF. The WH-word ^Vcto is an operator, and can logically bind the NP trace that can be left behind by QR:

(198) [_S [_S ^Vcto_i PRO citat' e_i]_i [_S ja kupil [_{NP}e]_i]]⁶³

Thus, the S' in (197) is interpreted as an infinitival free relative.

We do not intend to enter the general debate over the constituency of free relatives here (cf. Bresnan and Grimshaw 1978; Groos and Van Riemsdijk 1981; Levin 1982; among others). As far as these infinitival free relatives are concerned, we may demonstrate that the WH-word is not the head of an NP. It may bear any Case allowed in situ inside the free relative, even where the matrix verb would normally disallow a non-structural, oblique Case because of the Oblique Case Biconditional (92).⁶⁴ Thus, compare (199)a-b:

(199)a. ja kupil [_S ^Vcem_i [_S PRO pisat' e_i]]

I bought what write
 (instr) (infin)

b. ja kupil karandas^V/*karandasom^V

I bought pencil
 (masc acc/instr sing)

With respect to the Oblique Case Biconditional, cem^V does not act like the head of an NP, but rather like a WH-phrase in COMP in an indirect question.⁶⁵

(200) ja sprosil [_S ^Vcem [_S PRO pisat' e_i]]

I asked what write
 (instr) (infin)

'I asked what to write with'

The simplest hypothesis, then, is that both infinitival indirect questions and infinitival free relatives are S's, and do not differ in internal structure.

If this analysis is right, a number of expectations are raised. We

expect the ECP to intervene to block infinitival free relatives from S-structure subject position. The trace left after movement of an infinitival free relative will be an NP bound by an S'. If this trace is the subject of a tensed sentence, it will not be properly governed. This expectation is confirmed:

(201) * $[_S, ^V_{\text{cem}}]_i$ $[_S \text{ PRO pisat' } e_i]$ INFL $[_{VP} \text{ dostavilo udovol'stvie}]$
 what write brought satisfaction
 (instr) (infin)

(201) has an LF representation like:

(202) $[_S [_S, ^V_{\text{cem}}]_i [\text{PRO pisat' } e_i]]_i$ $[_S [_{NP} e_i] \text{ INFL } [_{VP} \text{ dostavilo udovol'stvie}]]]$

The subject trace is not properly bound by S', because of the categorial mismatch. As a consequence, the ECP rules the structure out.⁶⁶

Furthermore, if infinitival free relatives are S', and not NP, we may ask whether they bear Case features, like NP, or whether they may not, like QP. The answer seems to be no; infinitival free relatives obey the Non-Obliqueness Restriction. Thus, the verbs zaxvatit' and ovladet' are near-synonyms, meaning 'seize'. Zaxvatit' takes an accusative object, while ovladet' requires an instrumental. Only the former allows an infinitival free relative:

(203)a. spekuljant zaxvatil $[_S, ^V_{\text{cto}}]_i$ $[\text{PRO prodavat' } e_i]$
 speculator seized what sell
 (acc) (infin)
 'the speculator seized something to sell'

b. *spekuljant ovladel $[_S, ^V_{\text{cto}}]_i$ $[\text{PRO prodavat' } e_i]$

The verb ovladet' does not allow an infinitival free relative even if the WH-word in COMP is instrumental (although some speakers may find a slight improvement):

(204) *spekuljant ovladel [_S,^Vcem_i PRO prodavat' e_i]

If S' does not bear Case features, it, like QP, does not fall under the Case Filter -- a fact that we have already demonstrated for ordinary clausal complements. It follows that an S' placed in a Caseless position will not have to move to a position where it can receive Case. An S' generated, for example, as the object of a passive verb will not have to become a subject. When it undergoes movement in LF, the variable it binds will be an object, and hence will satisfy the ECP regardless of the categorial mismatch forced by c-selection. Like QPs, infinitival free relatives should be possible as arguments of passive and ergative verbs, and, as expected, they are:

(205) a. bylo kupleno [_S,^Vcem_i PRO pisat' e_i]

was bought what write
(neut sing) (instr) (infin)

b. pojavilos' [_S,^Vcem_i PRO pisat' e_i]

appeared what write
(neut sing) (instr) (infin)

(for some speakers, infinitival free relatives are more difficult with ergatives than with passives.)

Thus, infinitival free relatives in Russian seem to be S's, with all the properties of the QPs we considered earlier. This is not surprising, since our analysis of Russian QPs hinged minimally on characteristics

peculiar to QPs, but rather on very general properties of grammar.

A number of problems with our analysis of infinitival free relatives deserve mentioning. After mentioning these problems, we end the section with one problem of Russian grammar that appears to be solved by our analysis.

The first problem is the fact that tensed free relatives, in Russian or in English, do not show ECP effects of the sort described above. It may be that tensed free relatives are COMP-headed: "non-matching" free relatives corresponding to (199) above are ungrammatical for most speakers if tensed:

(206) *ja kupil [_S^V cem_i [_S ty INFL [_{VP} pisal e_i]]]

I bought what you wrote
(instr)

'I bought with what you were writing'

I do not know why there should be a difference of this sort.⁶⁷

The second problem involves Spanish. Spanish has infinitival free relatives, discussed most recently by Plann (1980). These, like the Russian free relatives, show subject/object asymmetries (Plann, p. 126):

(207)a. Julia no encontró [_S [a quien]_i [_S PRO dirigirse e_i]]

Julia NEG found to whom turn-reflex.
(infin)

'Julia didn't find anyone to turn to'

b. * [_S [_S [a quien]_i [_S PRO dirigirse e_i]] INFL [_{VP} no fue encontrado]]

'someone to turn to wasn't found'

If this contrast derives from the ECP, we expect that (207)b will improve if the infinitival free relative is postverbal, properly governed by V.

Unfortunately, such a sentence is no better than (207)b, for all my Spanish informants. Spanish, of course, allows postverbal subjects just as Italian does. We have no solution for this problem.

On the other hand, our analysis of Russian QPs and infinitival free relatives has an unexpected dividend: it helps us explain some of the properties of the negative existential net. Recall that we claimed at the very beginning of this chapter that the genitive of negation was completely optional. This statement is true in most environments, but it is a well-known fact that the negative existential net requires the genitive when it takes a nominal:

(208)a. net knig
there are not books
(gen pl)

b. *net knigi
(nom/acc pl)

Suppose we claim that this genitive is in fact part of a familiar-looking QP:

(209) [_Se INFL [_{VP} net [_{QP} [_Qe] knig]]]

If this analysis is correct, the obligatoriness of the genitive after net is due to the impossibility of an NP. The obvious question: why is an NP impossible?

The next point to notice is that net takes one other sort of complement besides a QP. It takes an infinitival free relative. An irrelevant phonological process cliticizes the free relative's WH-phrase to net, which takes stress and loses its t in the process: 68

(210)a. [_Se INFL [_{VP}né- [_S^včem_i [_SPRO INFL [_{VP}pisat' e_i]]]]]]

there isn't what write
(instr) (infin)

'there isn't anything to write with'

b. [_Se INFL [_{VP}né- [_Skogo_i [_SPRO INFL [_{VP}arrestovat' e]]]]]]

there isn't whom arrest
(acc) (infin)

'there isn't anyone to arrest'

We may thus rephrase our question about net: why can net take a QP or an S' as a complement, but not an NP? The obvious answer lies in the domain of Case Theory. Crucial to our account of the Non-Obliqueness Restriction on both QPs and S's was the claim that these categories do not need Case. Suppose that the negative existential θ -marks a complement, but is not a Case assigner. If the negative existential were like any other verb that does not assign Case to its object or a θ -role to its subject, it would allow an NP complement to move to subject position and receive nominative Case; but the negative existential is not inflected for agreement, preventing nominative Case assignment. Hence there is no way for an NP to be the complement of the negative existential, since it will always violate the Case Filter. On the other hand, QP and S' do not need Case, and thus may be complements of the negative existential. Thus, Case Theory, the QP hypothesis and our analysis of infinitival free relatives as bare S's all work together to explain the odd properties of net. Net is an ordinary ergative verb whose one peculiarity is that it does not cooccur with agreement; everything follows from this one assumption.⁶⁹

4.4.2 Secondary Predicates

In this section, we will sketch an analysis of Russian constructions with "secondary predicates" (the traditional term) that, if correct, supports our analysis of QPs and infinitival free relatives. We will suggest that some secondary predicate constructions, though they do not appear to have anything in common with the quantificational structures examined in previous discussion, should be analyzed in a very similar way. As evidence, we will show that they obey a form of conditions A and B, and we will make a weak case for condition C.

First, we distinguish two types of secondary predicates, which have obviously different thematic properties:

- (211)a. ja scitaju Mašu p'janoj
 I consider Maša drunk
 (fem acc sing) (fem instr sing)
- b. ja vstretil Mašu p'janoj
 I met Maša drunk
 (fem acc sing) (fem instr sing)

In each sentence, Maša bears accusative Case, like any normal object. The secondary predicate p'janoj agrees with its understood subject, Maša, in gender and number, but shows instrumental Case. We will not have much to say about the surface Case marking of secondary predicates in these constructions, although certain aspects of it will be important. See Schein (1982; forthcoming) for some suggestions. For our purposes we will simply suppose that [+N] categories bear instrumental Case when they are secondary predicates.

We say [+N], because the adjectives in constructions like (211) may be replaced by predicative NPs:

(212)a. ja ^Vscitaju Puškina [liričeskim poètom]
 I consider Puškin lyric poet
 (masc acc sing) (masc instr sing)

b. ja uvažaju Puškina [liričeskim poètom]
 I respect Puškin lyric poet
 (masc acc sing) (masc instr sing)

Secondary predicates may also be attributive PPs, in which case they do not, of course, show any Case marking:

(213)a. ja ^Vscitaju takie knigi [v xorošem sostojanii]
 I consider such books in good condition
 (fem acc pl) (neut prep sing)

b. ja pokupaju takie knigi [v xorošem sostojanii]
 I buy such books in good condition
 (fem acc pl)

The thematic difference between the (a) and (b) sentences lies in the question of what the main verb c-selects. The verb we have used in the (a) sentences, scitat' 'consider', appears to directly θ -mark a clause. In (211)a, for example, what "I" have an opinion about is the proposition "that Masa is drunk". Similarly, in (212)a, "I" believe that Puškin is a lyric poet, and in (213)a, that books are in good condition. In fact, the verb scitat' does allow a tensed sentential complement, just like its English counterpart:

(214) ja ^Vscitaju, [_S ^Vcto Puškin byl liričeskim poètom]
 I consider that Puškin was lyric poet

In the (b) sentences, however, the main verb appears to c-select an NP. In particular, the verbs meaning 'meet', 'respect', and 'buy' in (211)-(213) are interpreted in such a way as to suggest that they are θ -marking

the NP that bears accusative Case. This is clear from the entailment relations: if I met Maša^V drunk, then I met Maša^V; if I respect Puškin^V the lyric poet, then I respect Puškin^V; etc. Also, these verbs do not take overtly clausal complements:

- (215) *ja vstretil, [_S ^Vcto Maša byla p'janoj]
 I met that Maša^V was drunk

How can we account for the apparently different thematic properties of the (a) and (b) sentences? Let us consider the (a) sentence first. The Projection Principle, in particular the definition of "subcategorized position" discussed briefly in Chapter One (and cf. Chomsky 1981a, 33) prevents a structure like (216) at the level at which c-selection applies:

- (216)a. ja [_{VP} ^Vscitaju [_{NP} Mašu] [_{AP} p'janoj]]
 b. I [_{VP} consider [_{NP} Maša^V] [_{AP} drunk]]

In the structures in (216), the verbs scitat'/consider must θ -mark the NP they take as complement, in violation of c-selectional requirements, if these verbs c-select a clause.

If we define a clause roughly as a constituent containing a predicate phrase and a subject, then we are forced to the following structure for

(216)a-b, at least at LF, where α is some category:

- (217)a. ja [_{VP} ^Vscitaju [_{α} [_{NP} Mašu] [_{AP} p'janoj]]]
 b. I [_{VP} consider [_{α} [_{NP} Maša^V] [_{AP} drunk]]]

The properties of these structures have been discussed by Stowell (1981; forthcoming), who suggests that α is a "Small Clause", in something of

the sense of Williams (1975). The properties of these small clauses are discussed further by Chomsky (1981a). Chomsky notes that α must be a non-maximal projection in (217), in which respect Small Clauses resemble the S'-deletion infinitivals discussed in section 3 of this chapter.

We know that α may be a non-maximal projection because the subject of the small clause may be governed and properly governed by the higher verb. We know that the subject may be governed because it receives Case from the higher verb; if the higher verb is not a Case assigner, the subject of the small clause cannot be a lexical NP:

(218)a. * $[_S \underline{e} \text{ INFL } [_{VP} \text{scitaetsja } [_{\alpha} [_{NP} \text{Maša}^V] [_{AP} \text{p'janoj}]]]$
 is considered Maša drunk
 (fem acc) (fem instr)

b. * $[_S \text{it INFL } [_{VP} \text{is considered } [_{\alpha} [_{NP} \text{Maša}^V] [_{AP} \text{drunk}]]]$

Compare (219), where the subject of the embedded clause is assigned Case within its clause, by AGR:

(219)a. $[_S \underline{e} \text{ INFL } [_{VP} \text{scitaetsja } [_S, \text{cto Maša}^V \text{ INFL } [\text{byla p'jana}]]]]]$
 that Maša was drunk
 (fem nom)

b. $[_S \text{it INFL } [_{VP} \text{is considered } [_S, \text{that Maša}^V \text{ INFL } [\text{was drunk}]]]]]$

We know that the subject is properly governed because of the possibility of WH-movement:

(220)a. $\text{Maša}^V, [_S, \text{kotoruju}_i [_S \text{ja INFL } [_{VP} \text{scitaju } [_{\alpha-i} [_{AP} \text{p'janoj}]]]]]]]$
 whom I consider drunk
 (fem acc sing) (fem instr sing)

b. $\text{Maša}^V, [_S, \text{who}_i [_S \text{I INFL } [_{VP} \text{consider } [_{\alpha-i} [_{AP} \text{drunk}]]]]]]]$

We know that α must be a non-maximal projection because of the impossibility of PRO as the subject of a small clause complement of consider/scitat' and similar verbs. If α were maximal, PRO in (221) below would be ungoverned, as required by the Binding Theory, and we could rule (221) out:

(221)a. *ja [_{VP}^Vscitaju [PRO [_{AP}^P'janoj]]

b. *I [_{VP}consider [PRO [_{AP}drunk]]

If α is a non-maximal projection, what is it a projection of? Stowell suggests that α is a projection of the predicate, arguing that some verbs that take Small-Clause complements c-select the predicate of that Small Clause. Let us suppose that this suggestion is correct. The complement of consider/scitat' in the Small Clause constructions is thus an X^* , where X is some category, and $*$ represents some non-maximal number of bars:

(221)a. ja [_{VP}^Vscitaju [_{A*}^VMasu [_{AP}^P'janoj]]

b. I [_{VP}consider [_{P*}these books [_{PP}in good condition]]]

Chomsky (1981a, 105ff.) makes another point about Small Clauses, which is relevant here. Suppose a Small Clause were to be a maximal projection:

(222) I consider [_{AP}^{NP} [_Adrunk with power]]

We already know that the position of NP could not be Case marked: this rules out lexical NP and WH-trace as a value for NP. Can NP be PRO? No, since it is governed by drunk. The only possibility for NP is to be

NP-trace, but Chomsky suggests a revision of the definition of proper government which might make NP violate the ECP in any case (cf. also the penultimate section of Chapter Five). Supposing this revision is correct, there is no value for NP that can satisfy Case Theory, Government Theory (the ECP) and the Binding Theory at the same time. Hence structures like (222) cannot exist.

There is another possibility. Suppose a Small Clause were to be a maximal projection, but not a projection of the predicate:

(223) I consider [_S, NP [_{AP} drunk with power]]

The NP in (223) is unlicensed; hence it can and must be PRO. Nonetheless, it is striking that neither verbs meaning 'consider', nor any other predicate in English or in Russian, allows such a configuration:

(224) * . . . [_{VP} V [_S, PRO XP]]

Thus, the following seems so far to be true:

(225) Small Clauses are always non-maximal projections.

Let us now return to the (b) sentences of (211)-(213). Suppose, contrary to what we have argued, that c-selection is projected by the Projection Principle to each level of representation. It follows that a sentence like (211)b in English and Russian must have a representation in which the verb meet is assigning its θ -role to the object Masha^V at each level of representation. This implies a structure like (226):

(226)a. ja [_{VP} vstretil [_{NP} Masha^V] [_{β} . . . p'janoj]]

b. I [_{VP} met [_{NP} Masha^V] [_{β} . . . drunk]]

Notice now that the Projection Principle also tells us what the contents of ". . ." must be. Not only the verb meet/vstretit', but the secondary predicate drunk/p'janoj assigns a θ -role. By the θ -criterion and Projection Principle, this θ -role must be assigned to a category distinct from the object of meet/vstretit': a null category which takes as antecedent the object of meet/vstretit', which bears an independent role. Such a null category is, of course, PRO:

- (227)a. ja [_{VP}vstretil [_{NP}^VMasu]_i [_{β} PRO_i p'janoj]]
 b. I [_{VP}met [_{NP}^VMasa]_i [_{β} PRO_i drunk]]

But if β in (227) takes PRO as its subject, then β must be maximal, and a maximal projection must intervene between the predicate phrase and PRO:

- (228)a. ja [_{VP}vstretil [_{NP}^VMasu]_i [_S,PRO [_{AP}p'janoj]]]
 b. I [_{VP}met [_{NP}^VMasa]_i [_S,PRO [_{AP}drunk]]]

This structure, however, violates (225).

This contradiction is not terribly worrisome, of course, since (225) is not a plausible deep principle of grammar. It might be the case, as Chomsky suggests, that (225) should be revised as:

- (229) Subcategorized Small Clauses are always non-maximal projections.

(229) leaves non-subcategorized (c-selected) Small Clauses, like those in (228), free to be maximal. Other principles, like an adjacency condition on Case assignment, might require them to be maximal. (229), in turn, might derive from some deeper principle. Nonetheless, let us consider some facts from Russian that might suggest a different approach to the (b)

sentences of (211)-(213).

Recall that the analysis we have just presented was based on the assumption, which we have argued against, that c-selection falls under the Projection Principle, and applies at every level of representation. Since the verbs in question c-selected an NP, we had to assume that their object was not a bare small clause:

(230)a. ja [_{VP}vstretil [_{A*}Ma^Vsu [_{AP}p'janoj]]]

b. I [_{VP}met [_{A*}Ma^Vsa [_{AP}drunk]]]

Representations like (230)a-b violate c-selection at any level at which that theory applies. If c-selection applies at D-structure and at S-structure, then (230)a-b cannot be well-formed D- and S-structure representation. If c-selection does not apply at D- and S-structure, however, nothing prevents the representations in (230) at these levels. There is, in fact, some evidence that (230)a, at least, is the correct structure for the Russian secondary predicate construction in question at D-structure and S-structure.

4.4.2.1 Secondary Predicates and the D-structure [XP, VP] Restriction

Let us introduce some theory-neutral and construction-neutral terminology. We call the lexical NP Ma^Vsa in both consider-type sentences like (211)a and meet-type sentences like (211)b the understood subject of the secondary predicate. In (211)a, the understood subject is the subject of a small clause. If (211)b is analyzed as in (228)a, the understood subject is the controller of PRO.

In the preceding exposition, we have looked at secondary predicates in consider-type and meet-type constructions whose understood subjects

are accusative NPs. In certain circumstances, the understood subject of a secondary predicate may also be a nominative subject NP:

- (231)a. $\overset{V}{\text{Maša}}$ $\overset{V}{\text{kazetsja}}$ p'janoj
 $\overset{V}{\text{Maša}}$ seems drunk
 (fem nom sing) (fem instr sing)
- b. $\overset{V}{\text{Maša}}$ $\overset{V}{\text{prišla}}$ p'janoj
 $\overset{V}{\text{Maša}}$ came drunk
 (fem nom sing) (fem instr sing)

Similar examples can be constructed with other secondary predicates. The surface syntax of (231) is parallel to that of the object cases in (211)-(213). Once again, the secondary predicate agrees with its understood subject in number and gender, but bears instrumental Case.

Thematically, (231)a-b differ in the same way (211)a-b differed. In (231)a, the main verb meaning 'seems' c-selects a clause, as shown by the possibility of (232):

- (232) $\overset{V}{\text{kazetsja}}$, [_S, $\overset{V}{\text{cto Maša}}$ byla p'jana]
 (it) seems that $\overset{V}{\text{Maša}}$ was drunk

Evidence discussed by Burzio (1981) suggests that seem and its equivalents are simply ergative verbs that c-select a clause. Note, for example, that the -sja passive (but cf. footnote 14) of $\overset{V}{\text{scitat'}}$ 'consider' participates in constructions parallel to (231)a:

- (233) $\overset{V}{\text{Maša}}$ $\overset{V}{\text{scitaetsja}}$ p'janoj
 $\overset{V}{\text{Maša}}$ is considered drunk
 (fem nom sing) (fem instr sing)

The passive of $\overset{V}{\text{scitat'}}$ differs from the active, like all passives from actives, in not indirectly θ -marking its subject and in not assigning Case

to its object. Kazatsja 'seem' in (232) has the same properties. The D-structure of (232) is seen in (234)a below. The subject position in (234)a is non- θ , allowing movement into this position. The object position does not receive Case, forcing movement of the D-structure object, yielding the S-structure (234)b:

- (234)a. [_Se INFL [_{VP}kaz^Vetsja [_{A*}Maš^V [_{AP}p'janoj]]]] D-structure
 b. [_SMaš^V_i INFL [_{VP}kaz^Vetsja [_{A*}e_i [_{AP}p'janoj]]]] S-structure

In (231)b, however, it does not seem that the verb priš^Vla 'came' c-selects a clause:

- (235) *priš^Vlo, [_Scto Maš^V byla p'jana]
 (it) came that Maš^V was drunk

Rather, the understood subject of the secondary predicate p'jana appears to be θ -marked and c-selected by priš^Vla: If Maš^V arrived drunk, then Maš^V arrived. Following the logic we used with accusative understood subjects, we might motivate the following analysis of (231)b:

- (236) Maš^V_i INFL [_{VP}priš^Vla [_S,PRO p'janoj]]

Notice, however, that (236) is not necessarily an accurate representation, even if we assume the analysis with PRO. The verb priš^Vla is ergative; we might therefore expect (236) to look more like (237) at S-structure:

- (237) Maš^V_i INFL [_{VP}priš^Vla e_i [_S,PRO p'janoj]]

Speaking more generally, the understood subject of the secondary predicate in (231)b is a D-structure NP governed by VP.

Interestingly, it appears that only D-structure NPs governed by VP can be the understood subjects of secondary predicates in meet/come constructions. We have already seen the understood subject as the argument of an ergative verb. It may also be the subject of a passive:

(238) mjaso_i [_{VP}bylo kupleno a_i zamorožennym]
 meat was bought frozen
 (neut nom sing) (neut instr sing)

On the other hand, the understood subject may not be the subject of an agentive intransitive verb, a polyadic transitive verb, or of an adjective, in the sort of construction we are considering here:¹⁰

(239)a.??sobaka_i kusalas' golodnoj
 dog bit hungry
 (fem nom sing) (fem instr sing)

b.??Ivan ubil košku p'janym
 Ivan killed cat drunk
 (masc nom (fem acc) (masc instr sing)

c. *Maša^V čitaet ugrjumoj (from Neidle, forthcoming)
 Maša^V reads gloomy
 (fem nom) (fem instr)

d.??Maša^V ---- sčastliva p'janoj
 Maša^V happy drunk
 (fem nom) (fem) (fem instr)

On an analysis with PRO, exemplified in (237), we might argue that (239) shows a property of the theory of Control. Note however, that PRO as subject of an infinitive may be controlled by a D-structure subject, so long as there is no nearer object controller (in cases that obey

Rosenbaum's (1967) Minimal Distance Principle in the first place). Thus, the verb xotet' 'want' seems to fail the tests for ergative verbs (e.g. does not allow its subject replaced by the genitive under negation), yet controls PRO in example (240):

(240) ja xoc^Vu [PRO čitat'^V o betone]
 I want read about concrete
 (inf)

From another point of view, the restriction we find in (239) strongly resembles condition A. We may call it A':

(241) A': the understood subject of a secondary predicate in the meet/come construction must be an NP governed by V in D-structure.

We do not say an [NP, VP], for reasons that will be clear shortly.

4.4.2.2 Secondary Predicates and the Non-Obliqueness Restriction

Understood subjects show another familiar property:

(242) B': The understood subject of a secondary predicate in the meet/come construction may not occur in positions where oblique Case is required.

Recall, for example, that pomogat' requires a dative object:

(243) *Maš^Va pomog Ivanu p'janym
 Maš^Va helped Ivan drunk
 (nom fem) (masc dat) (masc instr)

Note that the gender of the secondary predicate would force it to take Ivan as its understood subject. Nonetheless, the sentence is bad.

Condition B' may be further exemplified with some minimal pairs. The verbs peredraznivat' 'mimic (in order to provoke mirth)' and podražat' 'imitate' are close in meaning. They differ in that peredraznivat' takes a non-oblique, accusative object, while podražat' takes a dative. Only peredraznivat', which takes accusative objects, allows its object to be the understood subject of a secondary predicate:

- (244)a. Maša^V peredraznivala Ivana p'janym
 Maša^V mimicked Ivan drunk
 (fem nom) (masc acc) (masc instr)
- b.??Maša^V podražala Ivanu p'janym
 (fem nom) (masc dat) (masc instr)

Similarly, the perfective verbs tronut' and kosnut'sja are also close in meaning; in the context below they both mean, roughly, 'touch'. (Tronut' suggests a more deliberate action; kosnut'sja, more accidental.) Tronut' takes an accusative object, while kosnut'sja requires a genitive:

- (245)a. Maša^V tronula portret mokrym
 Maša^V touched portrait damp
 (fem nom) (masc acc) (masc instr)
- b.??Maša^V kosnulas' portreta mokrym
 (fem nom) (masc gen) (masc instr)

These minimal, or near-minimal pairs suggest that the restriction in question is not semantic, but reflects the syntactic condition B' given above.⁷²

4.4.2.3 Secondary Predicates and Small Clauses

In the preceding two subsections, we saw that the understood subjects of secondary predicates in meet/come constructions obey two conditions

strikingly like those we observed QPs obeying earlier in this chapter.

It would, of course, be interesting to unify our accounts of these constructions. This is what we now proceed to do.

Notice that the predicates of consider/seem small clauses are free to occur in the complement of ergative and passive verbs that normally do not assign Case. In (231)a and (233), for example, the verb that θ -marks the small clause headed by p'janoj 'drunk (fem instr)' is not a Case assigner. Even when the predicate bears the features for N, the higher verb does not need to be a Case assigner:⁷³

- (246)a Igor'_i kazalsja [_{N*e_i} [_{NP}kruglym durakom] svoim družjam
 Igor seemed circular fool his friends
 (masc nom) (masc instr) (masc dat pl)
 'Igor seemed a perfect fool to his friends'
- b. Gidon_i s^vcitaetsja [_{N*e_i} [_{NP}velikim skripač^vem]]
 Gidon is considered great violinist
 (masc nom) (masc instr)

We may claim either that the predicates of Small Clauses satisfy the Case Filter in these constructions by virtue of their "inherent" instrumental Case marking (about whose source we remain vague), or else that predicates do not fall under the Case Filter. If we make the latter claim, we may go on to claim further that predicates of Small Clauses, and hence the clauses themselves, do not bear any Case Features at all, in the sense required by Case Theory.⁷⁴ If we make the former claim, we may simply say that predicates of Small Clauses bear no Case besides instrumental.

Recall now how we accounted for the Non-Obliqueness restriction on QPs (and infinitival free relatives). We claimed that oblique Case

requirements are linked to θ -marking, which, by the Projection Principle, applies at all levels. QPs and S's cannot bear Case features. Hence they cannot bear oblique Case features. Therefore, they cannot be θ -marked by a verb that requires oblique Case. Thus a QP or S' in a position where oblique Case is required violates the θ -criterion.

Suppose now that secondary predicate constructions with consider-type verbs differ from secondary predicate constructions with meet-type verbs in the manner suggested above -- the hypothesis we wish to argue against. Thus:

(247)a. ja [_{VP}^Vscitaju [_{A*}Mas^Vu [_{AP}P'janoj]]]

b. ja [_{VP}vstretil [_{NP}Mas^Vu]_i [_S,PRO_i [_{AP}P'janoj]]]]

There is no way to explain the Non-Obliqueness restriction with meet-type verbs on the analysis in (247)b, without bringing in some new assumptions. Of course, such assumptions are available. For example, we could assume, as suggested to me by J. Guéron and B. Schein (personal communications), that oblique Cases are actually PPs, so that an oblique NP will not c-command PRO in structures like (274)b. That c-command is necessary for control in these constructions has been argued by Williams (1980).

Now notice that if the predicate head of a Small Clause either cannot bear Case features or inherently takes instrumental Case, it cannot occur in a position where oblique Case is required. In other words, we never expect to find a Small Clause as the object of a verb taking oblique Case. With respect to consider-type verbs in Russian (of which there are not many)⁷⁵, this prediction is correct. There appear to be no verbs that c-select a Small Clause and require oblique Case of any sort.⁷⁶

At this point, let us ask whether the structural distinction between consider-type and meet-type secondary predicate constructions is actually correct. We noted earlier that the structural distinction seen in (247) is only motivated at D-structure and S-structure if c-selection applies at those levels. If c-selection does not apply at these levels, we are free to assume that the two constructions have identical structures there, as in (230), which we repeat:

(248)a. ja [_{VP}^Vscitaju [_{A*}Maš^V [_{AP}p'janoj]]]

b. ja [_{VP}vstretil [_{A*}Maš^V [_{AP}p'janoj]]]

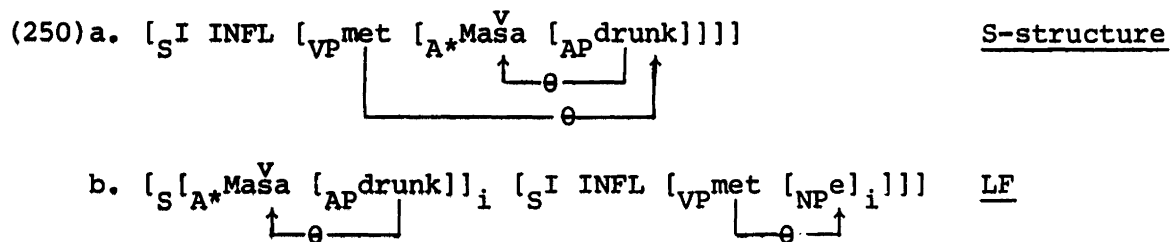
Similarly, in English:

(249)a. I [_{VP}consider [_{A*}Maš^V [_{AP}drunk]]]

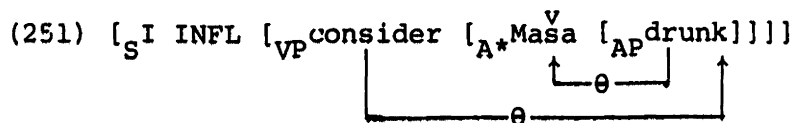
b. I [_{VP}met [_{A*}Maš^V [_{AP}drunk]]]

If the structures in (248) are correct, we can explain the Non-Obliqueness condition on the "understood subject" of the secondary predicate very simply. At D-structure and at S-structure, in constructions like (248)b and (249)b, a meet-type verb assigns its θ -role, not to a c-selected NP, but to a small clause headed by the secondary predicate, whose subject is the eventually c-selected NP. This is possible because c-selection does not apply at these levels, and the θ -criterion is consequently category-blind. Since Small Clauses cannot bear the oblique Case features required by certain verbs, these verbs cannot take Small Clause complements at D-structure and at S-structure, and do not participate in the meet-type secondary predicate construction. Hence condition B' follows.

Of course, the position occupied by the Small Clause in structures like (248)b and (249)b must be occupied by an NP non-predicate at LF, or else c-selection will be violated at LF. We wish to speculate that Small Clauses occupying positions in which an NP is c-selected undergo the movement rule QR, and leave an NP trace -- just like the QPs and S's we have considered above. In other words, a D-structure and S-structure like (250)a, where θ -roles are assigned as indicated, yields an LF like (250)b. (250)a does not satisfy c-selection, but does satisfy the category-blind Projection Principle. (250)b satisfies both c-selection and the Projection Principle:



In other words, meet-type verbs and consider-type verbs may have the same complementation at D-structure and at S-structure, due to the category-blindness of the Projection Principle. At LF, however they must differ, due to their different c-selectional properties. A Small Clause complement to a meet-type verb must undergo QR, to leave an NP trace that satisfies c-selection. On the other hand, a Small Clause already satisfies the c-selectional requirements of consider-type verbs, and no QR is called for. Thus, (251) is both S-structure and LF:



We return shortly to the question of what sense moving a Small Clause by QR might have. First, let us suggest how condition A' might be derived under this analysis. Extending our analysis to ergative and passive verbs, it follows that seem-type verbs do not contrast with come-type verbs in secondary predicate constructions as in (234)b/(236). Rather, at S-structure, they should be identical in structure:

(252)a. [_SMa^Va_i INFL [_{VP}seems [_{A*}e_{-i} [_{AP}drunk]]]]

b. [_SMa^Va_i INFL [_{VP}came [_{A*}e_{-i} [_{AP}drunk]]]]

Verbs like come, of course, c-select an NP, and not a clause, but this does not matter at S-structure, according to our hypothesis. Recall also that the subject positions of both seem and the ergative verb come are non- θ -positions. For this reason, both (252)a and (252)b satisfy the θ -criterion and Projection Principle.

Now let us see what might happen to (252)b at LF that will allow it to satisfy the Projection Principle. We might suppose, as before, that only QR is involved. This would yield an LF representation like (253):

(253) [_S[_{A*}e_{-i} drunk]_j] [_SMa^Va_i INFL [_{VP}came [_{NP}e]_j]]]

In order to make some sense out of this representation, we might suppose that the index j rewrites as i, yielding (254):

(254) [_S[_{A*}e_{-i} drunk]_i] [_SMa^Va_i INFL [_{VP}came [_{NP}e]_i]]]

While there might be some problem with the first occurrence of e_i in (254) under the Binding Theory, the Projection Principle and c-selection are satisfied by the LF representation in (254). Came assigns its θ -role to the chain ($\text{Masa}^V, [\text{NP } e_i]$), which is an NP chain, while drunk assigns its θ -role to an empty category, which we might assume to be PRO.

Alternatively, we might suppose that QR is preceded by reconstruction of Masa^V in (252)b into the position of its trace:

(255) [$\text{S } e$ INFL [$\text{VP } \text{came} [\text{A}^* \text{Masa}_i^V \text{ drunk}]$]]

-- in effect reproducing the D-structure representation of this sentence. Notice that problems should not arise with the Case Filter, if the Case Filter applies only at S-structure. Nor should problems arise with the ECP, if the subject empty category in (255) is expletive. QR will then yield (256):

(256) [$\text{S} [\text{A}^* \text{Masa}_i^V \text{ drunk}]_j$ [$\text{S } e$ INFL [$\text{VP } \text{came} [\text{NP } e]_j$]]]

What is important to us is the following: suppose a non-c-selected Small Clause were placed, not in the properly governed object position, as in our previous examples, but in subject position. Following QR, the ECP would be violated:

(257) [$\text{S} [\text{A}^* \text{subject predicate}]_j$ [$\text{S} [\text{NP } e]_j$ INFL VP]]

Condition A', like condition A for QPs, thus derives from the ECP at LF, supporting our analysis of secondary predicate constructions. Notice that the Projection Principle would rule out any sort of reconstruction to object position here, since the NP trace is indirectly θ -marked by VP.

This analysis predicts, of course, that a Small Clause should be grammatical as a D-structure, S-structure, and LF subject when it is in a position where a Small Clause is c-selected. On this question, there is much speaker variation, and it may be that other factors rule sentences like (258)a-b out for some speakers:

(258)a. % $[_{A^*} \text{men } [_{AP} \text{drunk}]]$ shocks me⁷⁷

b. % $[_{A^*} \overset{VV}{\text{musciny}} [_{AP} \text{p'janymi}]]$ potrjaslo menja
 men drunk shocked me
 (masc nom pl) (masc instr pl) (neut sing) (acc)

Finally, how are LF representations in which a clause has undergone QR to be interpreted? J. Higginbotham (personal communication) suggests that the subject of a clause that has undergone QR be interpreted roughly in the manner suggested by Chomsky (1976) for a focused NP (cf. discussion in Chapter Five, section 2), and that the predicate be interpreted as a restriction on the focus quantification. Thus, an LF representation like (259)a might be interpreted as in (259)b:

(259)a. $[_S [_{A^*} \overset{V}{\text{Maša}} [_{AP} \text{drunk}]]]_i$ $[_S \text{I met } [_{NP} \text{e}]_i]$

b. (for $x = \overset{V}{\text{Maša}}$: x drunk) (I met x)

Some support for this analysis might be provided by the following point. Secondary predicates in the meet/come construction obey a peculiar semantic constraint familiar from the existential construction: they must denote an "impermanent", "transient" property of the NP they indirectly θ -mark. By contrast, secondary predicates in the consider/seem construction do not have this restriction:

(260)a. I met Maša^V drunk

b. Oleg returned an accomplished virtuoso

c.#I met Maša^V intelligent

d.#Oleg returned the son of Maša^V and Paša^V

(261)a. I consider Maša^V drunk

b. Oleg seems an accomplished virtuoso

c. I consider Maša^V intelligent

d. Oleg is considered the son of Maša^V and Paša^V

(262)a. there were three men drunk

b.#there were three men intelligent

(260)c-d and (262)b are grammatical, but demand the odd presupposition that intelligence or parentage is a condition that varies with time. Thus, (260)c could be said if Maša^V's intelligence is known to fluctuate wildly, and (260)d could be said metaphorically if Oleg has just discovered, or just begun to pretend that he is the son of Maša^V and Paša^V. The point is that no such special interpretation is required in (261).

Following a suggestion of J. Higginbotham, we might relate the contrast between (260)c-d and (261)c-d to the assumption that the predicates in (260) are in a restricting clause at LF. Supposing a quasi-Gricean principle that restricting clauses may not restrict vacuously, it follows that only predicates denoting distinctive properties of the focused NP relative to the proposition in question may enter such a restricting clause. The same might be true in the existential construction of (262), if, as seems likely (cf. Safir 1982), QR is involved in the LF derivation of existential sentences.⁷⁸

Thus, there seems to be some evidence in favor of the hypothesis

that consider/seem secondary predicate constructions have the same D- and S-structure representations as meet/come secondary predicate constructions, an analysis which presupposes crucially that c-selection applies only at LF and that the Projection Principle is category-blind. As a final piece of evidence, based here on the principle of "guilt by association" we might mention an odd restriction that both secondary predicate constructions have in common. In this chapter, we will not be able to explain this restriction. In Chapter Four, however, we will provide an explanation for the consider/seem cases in terms of Path Theory, which will carry over to the meet/come cases if the two are analyzed as structurally identical at S-structure.

Both secondary predicate constructions do not allow the understood subject to be genitive under negation: the violation is weak, but noted in the literature on the subject (Ravič^V 1971):

(263)a.??ja ne sčitaju inostrannyx fil'mov interesnymi/*interesnym
 I NEG consider foreign films interesting
 (masc gen pl) (masc instr pl/sing)

b.??ja ne vstrečal ni odnoj devuški p'janoj
 I NEG met not one girl drunk
 (fem gen sing) (fem instr sing)

(264)a.?*ne sčitaetsja ni odnogo inostrannogo fil'ma interesnym
 NEG is considered not one foreign film interesting
 (masc gen sing) (masc instr sing)

b.?*ne prišlo ni odnoj devuški p'janoj
 NEG came not one girl drunk
 (neut sing) (fem gen sing) (fem instr sing)

In conclusion, we have examined two domains -- infinitival free relatives and secondary predicates -- where some of the restrictions we found on our QPs reappear, and we have shown how our analysis of

Russian QPs can be extended to these cases. There are doubtless other domains in which the assumption that c-selection is not part of the Projection Principle has consequences. For now, however, we will leave matters at this point. In the next section we take for granted our conclusions of the preceding sections and push the argument one step further, presenting evidence that not only is c-selection not part of the Projection Principle, but it also is not really a theory of syntactic category selection in the first place.

5.0 Categorical and Semantic Selection

5.1 In Chapter One, we observed that the Standard Theory subsumes most of both θ -theory and c-selection under the theory of subcategorization -- the exception being θ -marking and c-selection of subject position. We have argued that combining θ -marking and c-selection in this way is incorrect, that the two processes are separate and apply at different levels.

This conclusion is encouraging, because the primitives of θ -theory and the primitives of c-selection differ in their plausible epistemological status. The primitives of θ -theory -- notions like "agent", "patient", "goal", etc. -- probably meet the criterion of "epistemological priority" discussed briefly in Chapter One, and by Chomsky (1981a, 10). On the other hand, the primitives of c-selection -- syntactic categories like NP, S', Small Clause, etc. -- do not meet the condition of epistemological priority. They are not, in Chomsky's words, "concepts that can be plausibly assumed to provide a preliminary, prelinguistic analysis of a reasonable selection of presented data, that is, to provide the primary linguistic data that are mapped by the language faculty to a grammar". θ -theory, in

our conception, is that part of the theory of subcategorization whose primitives meet the epistemological priority condition; c-selection is that part of the theory of subcategorization whose primitives do not.

If this discussion is correct, it follows that we want to derive the theory of c-selection from some other theory, whose primitives are epistemologically prior. Such a theory would be a semantic theory -- specifically a theory of lexical semantics. The outlines of such a theory, which we may call semantic selection, or s-selection, have been sketched by Grimshaw (1979, 1981) in recent work. In the discussion below, we will sketch some elements of Grimshaw's theory.

What interests us is the following: Grimshaw argues that both c-selection (subsumed by subcategorization) and s-selection are autonomous subsystems of grammar. We will show that once her theory of s-selection is embedded in a general framework that includes Case Theory her argument for an autonomous theory of c-selection disappears. More importantly, a serious problem for her analysis, which she raises in Grimshaw (1979) and attempts to resolve in Grimshaw (1981) is immediately solved by Case Theory, if and only if c-selection does not exist at all as an independent theory. In so doing, we find linguistic evidence reflecting the condition of epistemological priority, which must in any case be correct, and we strengthen our argument that c-selection is distinct from θ -theory.

The cornerstone of Grimshaw's (1979) theory is an argument that predicates must bear features which select for the "semantic type" of their complements. Following in part observations of Baker (1968, 1970), she argues that the theory of s-selection allows predicates to select complements categorized as Q (question), P (proposition), or E (exclamation).

(We underline semantic categories, to distinguish them from syntactic categories like Q 'quantifier' and P 'preposition'.) Crucially, these semantic types are not in one-to-one correspondence with syntactic categories. In particular, while all these types may be associated with S', they may also be associated with NP (as a "concealed" question, proposition, or exclamation), or with nothing ("null complement anaphora"). For example, in each of the following sentences, the s-selectional requirement that ask takes a Q is satisfied:

(265)a. John asked me [_S what the time was]

b. John asked me [_{NP} the time]

c. Bill wanted to know what the time was, so I asked

In (265)a, Q is associated with a syntactic S'. In (265)b, it appears as NP. In (265)c, it does not appear at all in the syntactic structure, but is "filled in" at some later level, as discussed below.

From these and similar examples, particularly those involving null complement anaphora, Grimshaw correctly concludes that the theory of s-selection must be independent of theories that analyze syntactic categories. A predicate may bear selectional features like < Q >, < P >, or < E >, but these features are realized independent of any other lexical features involving strictly syntactic categorization.

As we have said, Grimshaw embeds this theory of s-selection in a standard theory of lexical entries. In particular, she argues that s-selectional features are needed in addition to subcategorizational features of the sort developed in Chomsky (1965), by which she means c-selection.⁷⁹

Grimshaw argues for the autonomy of s-selection and subcategorization, and thus for the existence of subcategorization, by noting that not all predicates that s-select a Q, P or E allow their complement to range over both NP and S'. For example, predicates like wonder, care, inquire, and give a damn s-select a Q, but prevent that Q from being realized as an NP. They do not allow "concealed questions":

(266)a. John wondered [_S,what the time was]

b.*John wondered [_{NP}the time]

(267)a. Mary cares [_S,where we are going]

b.*Mary cares [_S,our destination]

(268)a. Bill inquired [_S,how old I was]

b.*Bill inquired [_{NP}my age]

(269)a. I don't give a damn [_S,what your name is]

b.*I don't give a damn [_{NP}your name]

Similar observations can be made about predicates which s-select P:

(270)a. I'll assume [_S,that he is intelligent]

b. I'll assume [_{NP}his intelligence]

(271)a. I'll pretend [_S,that he is intelligent]

b.*I'll pretend [_{NP}his intelligence]

-- or about predicates which s-select E:

(272)a. Bill couldn't believe [_S,how incredibly hot it was!]

b. Bill couldn't believe [_{NP}the incredible heat]

(273)a. Bill complained [_S,how incredibly hot it was]

b.*Bill complained [_{NP}the incredible heat]

Grimshaw draws the relevant distinctions by appealing to the theory of subcategorization. Verbs like care and ask both s-select a Q; verbs like assume and pretend both s-select a P; verbs like believe and complain both s-select an E. The members of these pairs differ in their subcategorization. As far as S' and NP are concerned, four different subcategorization frames should be available (we return to null complement anaphora below):

(274)a. [_{__} ^{NP}{S'}] b. [_{__} S'] c. [_{__} NP] d. [_{__} ∅]

(274)a, if we restrict our attention to verbs s-selecting Q, is exemplified by ask; (274)b, by care or wonder. The assumption of two autonomous theories thus accounts elegantly for the differences and similarities among predicates.

Nonetheless, Grimshaw (1979, footnote 33) notes a problem with this theory, to which she returns in a more recent paper (Grimshaw 1981). This problem is the focus of our remarks here. There are no predicates (in English at least) which s-select a Q or an E and have the subcategorization frame in (274)c or (274)d. In other words, no predicates take only concealed or null questions or exclamations, and do not take clausal questions or exclamations. (Grimshaw does not discuss propositions; we return to them shortly.)

As Grimshaw notes, this gap might be accidental, but this should not be taken as the null hypothesis. A priori, we might explain the gap in one of two ways. One approach is to allow a theory of subcategorization (c-selection) to "overpredict" that the non-existent predicates should exist, and to find another theory which will rule them out. This is the approach taken by Grimshaw (1981) and by Woolford (1981) in comments on Grimshaw's paper. They both suggest that what rules out the non-existent predicates are certain properties of the Language Acquisition Device (LAD) of the child. We examine these proposals below.

Another approach is to abandon entirely the theory of subcategorization -- specifically, c-selection -- and to derive its effects from other subtheories of grammar, which will not overpredict the non-existent predicates. This is the approach we will argue for. Note that this proposal would preserve the essential elegance of Grimshaw's (1979) account of complement selection. Grimshaw's explanations rely on the interaction of a semantic theory with a syntactic theory. It is certainly possible that the syntactic theory in question is not subcategorization.

5.2 Let us now consider both Grimshaw's and Woolford's approaches.

Woolford's approach is based on Grimshaw's, but we present it first, for reasons of exposition. Woolford takes from Grimshaw the idea that epistemologically prior categories like "object"⁸⁰ and "action" have a Canonical Structural Realization (CSR) in syntactic categories. More precisely, CSR is a function mapping semantic categories onto syntactic categories. Thus, CSR(object)=N, and CSR(action)=V. As a general principle, "a word belongs to its CSP, unless there is evidence to the contrary". With Grimshaw, we believe that a notion like the function CSR is a sine

qua non for any theory that wishes to account for the projection of epistemologically prior categories onto syntactic categories.

As Grimshaw notes, it is plausible to assume that the CSR for P, Q and E is S'. (We question this assumption below.) Turning first to Woolford's theory, she proposes a principle that we might phrase as in (275);

(275) Learnability Principle

For all semantic categories C, LAD assumes that a predicate s-selects C if and only if LAD has evidence that the predicate c-selects (subcategorizes) CSR(C).

Given this principle, a child could never learn that a predicate s-selects P, Q, or E unless that predicate is found with S', the CSR for P, Q and E. Hence, no predicate could ever be learned that s-selects a P, Q or E and has a subcategorization frame that excludes S', as (274)c-d do. Hence the gap.

This explanation puts the cart before the horse, however. The Learnability Principle in (275) takes as its fundamental assumption that LAD can analyze syntactic categories, an assumption that is impossible if syntactic categories are not epistemologically prior. What is more, it makes CSR vacuous, except as a means of predicting the gap discovered by Grimshaw. If LAD can analyze syntactic (as opposed to semantic) categories, then there is no need for a theory mapping one set of categories onto the other.

There is other empirical evidence against the Learnability Principle proposed by Woolford. In fact, predicates that s-select a P, but never realize this P as an S', do exist, as Ken Hale (personal communication) has

noticed. This situation arises in English when a predicate requires a particular preposition in order to assign a θ -role to its complement.

Compare (276)-(277) with (278)-(279):

(276)a. we assume [_S,that unemployment will rise in the 80's]

b. we assume [_{NP}rising unemployment in the 80's]

(277)a. we noted [_S,that we were departing on Thursday instead of Friday]

b. we noted [_{NP}our departure on Thursday instead of Friday]

(278)a.*we approve (of) [_S,that unemployment will rise in the 80's]

b. we approve *(of) [_{NP}rising unemployment in the 80's]

(279)a.*we paid attention (to) [_S,that we were departing on Thursday instead of Friday]

b. we paid attention *(to) [_{NP}our departure on Thursday instead of Friday]

If approve of and pay attention to both s-select a P, and if CSR(P)=S', then (278)-(279) are counterexamples to Woolford's Learnability Principle.

These verbs select P, but never realize this P as a syntactic S'.

Of course, the ungrammaticality of (278)a and (279)a has an immediate (partial) explanation. Some syntactic principle, possibly Stowell's (1981) "Case Resistance Principle" must prevent propositional S' from functioning as the object of a preposition. Although we are confident that some deeper principle like Stowell's derives this consequence, for our purposes we may assume a filter like (280):

(280) * [_{PP} ^{S'} _{-WH}]

In the light of (280), we might try to revise the value of CSR(P). For

example, we might say that $CSR(\underline{P})=NP \ \& \ S'$. This idea fails immediately, within Woolford's theory, since LAD would have to hear both NP and S' in order to conclude that a predicate subcategorizing for these categories s-selects P. As we noted above, verbs like pretend (or claim, as Woolford notes) do select P and subcategorize for S' alone.

Alternatively, we may say that there is something special about the semantic feature P, such that $CSR(\underline{P})=NP \ \& \ S'$. This has a more subtle empirical problem. It misses the generalization that apparent instances of subcategorization (274)c arise only when there are independent explanations for the phenomenon, like filter (280).⁸¹ This suggests, of course, that all apparent cases of (274)c are actually cases of (274)a, with filter (280) rendering vacuous the subcategorization for S' in some cases. This, in turn, suggests that the Learnability Principle is on the wrong track. Apparently LAD can learn that a predicate s-selects P (and presumably Q or E) even if the only evidence comes from an NP complement. This conclusion leaves us where we started: we still do not know why only (274)a-b are possible subcategorization frames for verbs that s-select a P (or Q or E) complement. And in any case, as we have seen, Woolford's theory presupposes that LAD can analyze syntactic, as opposed to semantic categories, which is implausible. It is the meaning of a predicate that must determine its s-selection, and a theory of CSRs that determine the categories the predicate can take as complements.

Grimshaw (1981) suggests a principle that does not raise the objections we raised to Woolford's proposal. Grimshaw assumes that a child's LAD can deduce the s-selectional properties of a predicate without hearing the CSR for a given semantic type. In other words, Grimshaw's theory explicitly rejects the Learnability Principle in (275), and meets with none

of our a priori objections to that principle. Presumably, s-selectional properties of a predicate derive from the meaning of the predicate, something which must be available to LAD, since LAD does acquire a lexicon.

In place of the Learnability Principle (actually a subsequent proposal), Grimshaw proposes what she called the Context Principle, which we rephrase in our terminology:

(281) Context Principle

If a predicate s-selects a semantic category C, then it c-selects (subcategorizes) CSR(C).

If CSR(P, Q, E)=S', we have an immediate explanation for the gap noticed originally in Grimshaw (1979), which meets with none of the empirical objections that we have raised against Woolford's proposal. Grimshaw gives the following demonstration of how the Context Principle works (p. 179; we have renumbered her examples):

"The principle can be illustrated in the case where LAD receives evidence that a predicate takes NP questions but no evidence that it takes WH-questions: (282) but not (283) is in the data base.

(282) I asked John the time

(283) I asked John what the time was

Clearly, LAD can posit [NP] on the basis of (282). If the time is assigned the appropriate semantic representation, < Q > can also be associated with ask, giving the lexical entry in (284). But LAD also knows that the CSR(Q) is S'. Thus, by the Context Principle, LAD adds [S'] to the entry, giving (285).

(284) ask: [NP], < Q >

(285) ask: [{^{NP}S'}], < Q >"

Grimshaw goes on to note that no principle will add an NP to the subcategorization frame of a verb which is only found with S' complements (e.g. care, pretend). Thus, her theory correctly predicts that no verb which s-selects a P, Q or E will fail to subcategorize NP. Although Grimshaw does not discuss the matter, her Context Principle will also yield the correct results in (276)-(279). Predicates which need prepositions in order to assign a θ -role will always subcategorize an S', by the Context Principle, but the lexical realization of this subcategorization will always be blocked by filter (280).

Despite the explanatory adequacy of Grimshaw's Context Principle, it actually raises important doubts about the theory of subcategorization -- specifically, about c-selection. The Context Principle implies that in the unmarked case the mapping from s-selectional features to subcategorization is trivial. Given the s-selectional features of a predicate, we may always correctly deduce a subset of the syntactic categories in the subcategorization frame of that predicate. In the present case, by adding S' to the subcategorization frame for a predicate like ask, we are adding redundant information to the lexical entry. The interaction of the s-selectional properties of ask with the Context Principle automatically tells us that ask may take a sentential complement. Given the mechanism of strict subcategorization introduced in Chomsky (1965), it is, of course, technically necessary to enter this redundant information in the subcategorization frame, but one can imagine other theories that would not require this.

The theory of subcategorization is not at all necessary in order to determine whether a predicate may have a sentential complement. In

Grimshaw's theory, subcategorization is necessary in order to determine whether a predicate may take an NP complement. Suppose we were to eliminate the theory of subcategorization entirely, in favor of s-selection and the Context Principle. We would then need some other theory which could tell the child's LAD whether a predicate can take an NP complement or not. We could, of course, maintain a separate theory of subcategorization, or of c-selection, only for NPs, but this move clearly only obfuscates the essential problem.

Is there a subtheory of grammar which governs the possibility of an NP complement but not of a sentential complement? Put another way, is there a switch in the system of UG that can turn an NP on or off, but cannot affect an S'. There is just such a switch: Case Theory.

5.3 Suppose that strict subcategorization, specifically c-selection, does not exist. Instead, we have only Grimshaw's theory of s-selection, her Context Principle, and Case Theory. We immediately explain the non-existence of verbs that take only concealed, and not clausal, questions, propositions and exclamations. Recall that S's, but not NPs, may occur in Caseless positions at S-structure:

(286)a. it was proved [_S, that tomatoes are fruits]

b. *it was proved [_{NP} a theorem]

(287)a. John is curious (about) [_S, where I went]

b. John is curious *(about) [_{NP} life]

Now suppose that the CSR for P, Q and E is NP or S'. This is actually a reasonable assumption: since concealed propositions, questions

and exclamations exist, we know that NPs can be invested with "clausal content". If this is so, it is certainly a fact about CSRs, which match notions like "clause" with syntactic categories. It is not clear that NPs are less canonical bearers of clausal content than sentences are, particularly in view of the well-known parallels between the structure of S and the structure of NP.

It follows from our assumption about CSR(P, Q, E) that any predicate which s-selects P, Q, or E has, in effect, a "subcategorization frame" like (274)a, including both NP and S'. It remains useful to talk about subcategorization frames, even if they have no status in the theory: a subcategorization frame for a predicate is a list of the syntactic categories that the Context Principle allows the predicate to θ -mark, based on the predicates' s-selectional properties. We are, of course, departing from Grimshaw's usage.

Thus, as far as s-selection and the Context Principle are concerned, there will be no difference between (288)a and (288)b below (repeated from (266)):

(288)a. John wondered [_S, what the time was]

b. *John wondered [_{NP} the time]

Pairs like this differ, however, with respect to the Case Filter. We propose that the essential difference between a verb like ask and a verb like wonder is that ask allows objective Case to be assigned to its complement, but wonder does not. In other words, the lexical entries for wonder and ask look like (289):

(289)a. ask: <__ Q>, [+Case] b. wonder: <__ Q>, [-Case]

The Case features do all the work done by subcategorization features in Grimshaw's analysis. Case features, however, are independently motivated for verbs that s-select "objects", to distinguish transitive verbs like eat from intransitive verbs like stroll.

The theory of c-selection and the Context Principle, thus, allow any verb that s-selects P, Q, or E to take an NP or S' complement. Case Theory can turn off the possibility of NP. If we assume very crucially that no other theory referring to syntactic categories intervenes, we explain why nothing can ever turn off the S'.

Under this analysis, verbs like wonder show the same properties, for the same reasons, as the passive of verbs like ask (as far as their complementation is concerned). All forms of ask, like wonder, s-select a Q. Like wonder, but unlike the active form of ask, the passive form does not assign Case. As a result, it, like wonder, may take a sentential question as its surface object, but not a concealed question. (290) thus presents the same paradigm as (288):

- (290)a. it was asked [_S,what time it was]
 b.*it was asked [_{NP}the time]

The theory we have outlined eliminates the redundancy entailed by Grimshaw's discussion: the distribution of S' is redundantly predicted by s-selection and by c-selection; what remains of c-selection if this redundancy is eliminated falls naturally under Case Theory. Nonetheless, we believe that there is additional empirical evidence supporting our move.

As is well-known, English verbs that do not assign Case to their objects do not passivize. Thus:

(291)a.*John strolled NP

b.*it was strolled by John

(292)a.*Mary looked NP

b.*it was looked by Mary

If the behavior of verbs that do not take concealed questions and exclamations is consistent with that of other verbs that do not assign Case to their objects, they should not passivize. This is in fact the case. (293) contrasts with (294):

(293)a. it is not known whether John ever arrived (Q: NP and S')

b. it was asked when Mary would be there (Q: NP and S')

c. it has been guessed why you're here (Q: NP and S')

d. it has been determined who ate the tarts (Q: NP and S')

e. it was revealed what a fool he'd been (E: NP and S')

f. it was discovered how incredibly old he was (E: NP and S')

(294)a.*it is not cared what time it is (Q: S' only)

b.*it was inquired who killed Caesar (Q: S' only)

c.*it has been wondered where John went⁸² (Q: S' only)

d.*it was exclaimed what a fool he'd been (E: S' only)

e.*it was complained how incredibly old he was (E: S' only)

In languages like German, Dutch, and others, there is no restriction like that seen in (291)-(292), but I have been unable to find verbs with the properties of wonder and care. All things being equal, we predict that these languages should not show the paradigm in (293)-(294), since they do not show the paradigm of (291)-(292).

To account for (293)-(294), one might add to Grimshaw's account the stipulation that verbs that do not take concealed Q or E do not passivize, but this would yield a theory equivalent to our own, since subcategorization would then be entirely redundant. Alternatively, one might add a redundancy rule to the lexicon, which would prevent passive morphology from applying to verbs that do not have some expansion of their subcategorization frame that includes NP. This move, though possible, involves a special stipulation about the properties of wonder and care-type verbs. In the account we have outlined above, the properties of these verbs under passivization follow from a larger generalization about verbs that do not assign Case. A possible derivation of this generalization is proposed by Marantz (1981).

Our account also predicts that if a "dummy" preposition is available to a verb like wonder or care, these verbs will be able to take concealed questions and exclamations. Interestingly, all the verbs that s-select Q or E and do not assign Case appear able to combine with the preposition about, in which case they do allow NP complements:

- (295)a. John cares *(about) the time
 b. Mary inquired *(about) the murderer of Caesar
 c. Bill wondered *(about) John's whereabouts
 d. Lucy exclaimed *(about) the incredible fool she'd been
 e. Lear complained *(about) his incredible age

Perhaps about is in some way a "spell-out" of s-selectional features.

Notice that the preposition is semantically otiose: John cares (about) what time it is means the same with or without the preposition about. We

return to verb-preposition pairs below.

Finally, our theory predicts, correctly, that adjectives which take a Q or E will never take an NP complement, in English, at least:

(296)a. John is uncertain [_S,what time it is]

b.*John is uncertain [_{NP}the time]

Recall that English adjectives do not assign Case. But note once more the possibility of inserting a dummy preposition:

(297) John is uncertain about the time

Grimshaw also suggests that the ungrammaticality of (296)b need not be stipulated subcategorizationally. Rather, she proposes, the fact that adjectives do not take NP complements follows from the absence of an NP position after A in the base rules. In the next chapter, we sketch an argument from Chomsky (1981a, b; also Stowell 1981) that base rules do not exist: in the instance in question, the general absence of NP complements to A follows from Case Theory, so there is no need to stipulate this fact in the base.

5.4 If we are to defend the idea that subcategorization and c-selection do not exist as independent subsystems of grammar, we must face certain problems. The first problem concerns null complement anaphora. We believe that Grimshaw (1979) shows conclusively that null complement anaphora involves the "assumption" of an s-selected complement not present in syntax at a post-syntactic level. Consider a null complement anaphora response like I don't know, which answers a question like what time is it? We know that know s-selects a Q, among other things. When this Q is physically

present (as in I don't know what time it is), we know that the Q is associated with a θ -role assigned by the verb. What about the null complement anaphora case?

As Grimshaw observes, two sorts of analyses of null complement anaphora have appeared in the literature. Under one (stemming from Williams (1977b)), the complement in null complement anaphora sentences is structurally missing at S-structure:

(298) I don't [_{VP} [_V know]]

Alternatively, a null complement is syntactically present, as an empty category:

(299) I don't [_{VP} [_V know] [_{NP_e}]

In the framework we are assuming, we may reject (299). The empty category in (299) lacks an antecedent. Given the contextual definitions of empty categories sketched in Chapter One, the empty category must be [+pronominal, +anaphor] PRO. As PRO, however, (299) violates the Binding Theory, since it is governed. Alternatively, as discussed in Chomsky (1981b), the empty category in (299) might be [+pronominal, -anaphor], a category Chomsky calls pro; as such, however, it violates an "identification" requirement, which we discuss in greater detail in Chapter Four, section 4.4.

That leaves us with the analysis in (298). Under this analysis, there is no empty category, and hence no problems with the Binding Theory or any requirement of "identification". On the other hand, we now need to say something about how (298) satisfies the θ -criterion. If know

always assigns a θ -role to a complement, (298) violates the θ -criterion, since the θ -role remains completely unassigned.

The solution is obvious: in a theory with subcategorization, a verb like know subcategorizes optionally for S'. What this means is that parentheses are placed around S', indicating a legitimate expansion of the subcategorization frame for know omits S' entirely. In a theory without subcategorization, we can allow an analysis like (298) if the θ -marking properties of know are themselves in parentheses. In other words, verbs like know assign a θ -role optionally. As Grimshaw notes, there are other verbs which do not allow null complement anaphora, like discover:

(300) Statement: Guess what, John is telling lies again.

Response: (a) Oh, I didn't know

(b) *Yeah, I'd already discovered (Grimshaw 1979, (48))

Verbs like discover, unlike know, assign their θ -role obligatorily.

This analysis implies that s-selection is possible of a complement that is not θ -marked, and hence is absent from representations at D-structure, S-structure and LF. How are s-selectional restrictions satisfied in such a case? Clearly, s-selection cannot apply at D-structure, S-structure, or at LF. Therefore, it must apply at some further level past LF, perhaps a level called LF' by Chomsky (1981b). Chomsky argues, following suggestions of E. Williams, that certain processes like the coindexation of a relative word with the head of a relative take place at this level, and that certain requirements, such as a constraint on vacuous quantification, also apply at this level. N. Chomsky (personal communication) suggests that other sorts of "filling in" of missing arguments may take

place at LF'. For example, in (301), as noted by Manzini (1981), PRO is controlled by the "understood" indirect object of say:

(301) Bill said [PRO to leave]

If we handle null complement anaphora in the manner discussed above, we are in a position to shed some light on a technical problem in θ -theory and selection. Consider once more the verbs in (278)-(279), which require particular prepositions to precede their objects. It is generally assumed that some sort of c-selectional mechanism related to subcategorization accounts for the obligatoriness of these prepositions. It does not appear that a theory of s-selection has much to say about such cases. That leaves the theory of Case and θ -theory as possibilities. We wish to suggest that verbs which require a prepositional phrase headed by some particular preposition in English are analogous to verbs that require a particular oblique Case in a language like Russian.⁸³ Just as the lexical entry of a Russian verb might contain a statement like "assigns θ -role to a constituent headed by to". This seems the minimum necessary to capture the properties of such verbs.

This view has a nice consequence. Recall that we have claimed that null complement anaphora verbs are optional θ -assigners. This claim was important, because in the absence of a theory of subcategorization we cannot parenthesize S', since no theory refers to S' by name, except the universal Context Principle. Suppose now that θ -assignment were optional for a verb like approve or a collocation like pay attention, which require the prepositions of and to, respectively, in order to assign a θ -role. We predict that when no θ -role is being assigned, there will be no preposition.

This is, of course, what we find:

(302)a. Do you approve of rising unemployment? We approve.

b. Did you pay attention to our departure on Thursday instead of Friday? Yes, I paid attention.

On the other hand, we must say something, on this analysis, about the second gap in the typology of complements. Grimshaw notes that there are no predicates that s-select P, Q, or E and take only null complements. This is probably due to the fact, noted earlier in this Chapter, that c-selection, now subsumed by s-selection, is "parasitic" on θ -theory. That is, if a verb s-selects a category, then it always θ -marks a category. The fact that we allow θ -marking properties to be parenthesized does not affect this generalization.

We might also note the following, perhaps related, fact. Notice that null complement anaphora, as its name implies, is possible only in specific discourse contexts, where the content of the missing complement can be filled in from another sentence. In this respect, null complement anaphora constructions containing a bare verb are interpreted as if they contained a verb followed by a pronominal object. It is a fact that there are no predicates that require of one of their arguments that it be pronominal (John V'ed it/*John V'd the N).

5.5 Finally, we may address ourself to the translation of our analysis of Russian QPs, infinitival free relatives and secondary predicate constructions into a framework that lacks c-selection entirely, and in which c-selection is replaced by s-selection, a theory that applies at LF'.

For the most part, the translation is straightforward. Let us consider the QPs, since they raise the most interesting questions. We want to motivate QR of a QP whenever the QP occupies a position in which an NP is c-selected. We may assume trivially that verbs that c-select an NP are verbs that s-select some category $\underline{\alpha}$ such that $CSR(\underline{\alpha})$ is NP and not QP. The fact that s-selection applies at LF', and not at LF does not raise any serious problems, as long as no rule mapping LF representations onto LF' representations can replace a QP by an NP. In the absence of any such rule, QR, mapping S-structures onto LF representations, will still have to perform the job of replacing lexical QPs with NP traces, as before. A logical candidate for $\underline{\alpha}$ is some category like Grimshaw's "object", or perhaps "term". We may use the letter \underline{T} to denote this category. Thus, whenever QR occupies a position where \underline{T} is s-selected, QR will be motivated, since $CSR(\underline{T}) \neq QP$.

On the other hand, our discussion in this section raises anew the question of whether there is a semantic category that does correspond to QP. We might call this category \underline{K} (for Russian kvantor 'quantifier'). If $CSR(\underline{K})$ is Q, then no QR is indicated for a QP occupying a position where \underline{K} is also selected. Predicates that s-select \underline{K} may include the adjective ravno 'equal' which we considered at (170) above. Here we may suggest another class of predicates that might s-select \underline{K} . These would be the predicates that participate in the partitive genitive construction. Certain verbs allow partitive genitive objects. We may identify these objects as QPs of the type we assumed for the genitive of negation construction. An example, with the verb dobavit' 'add', is given in (303)a. Other verbs, like izobrazit' 'depict', do not allow the partitive genitive, as seen in (303)b.

- (303)a. $\text{povar dobavil } [_{QP} [_{Q-e}] [_{NP} \text{saxaru}]] \text{ v smes}'^{69,70}$
 cook added sugar to mixture
 (masc gen sing)
 'the cook added some sugar to the mixture'
- b. $\text{xudožnik izobrazil } [_{QP} [_{Q-e}] [_{NP} \text{saxaru}]] \text{ v natjurmorte}$
 artist depicted sugar in still life
 (masc gen sing)
 'the artist depicted some sugar in the still life'

If we accounted for this difference directly in terms of c-selection, saying that dobavit' takes a QP, while izobrazit' does not, we would not predict the following fact. Verbs that allow the partitive QP and also assign Case to their objects allow not only syntactic QPs, but also "concealed QPs" -- phrases like "a cup of sugar", where "cup" refers to a quantity, and not to a physical object:

- (305)a. $\text{povar dovabil } [_{NP} \text{stakan saxaru}] \text{ v smes}'$
 cup sugar to mixture
 (acc) (gen)
 'the cook added a cup of sugar to the mixture'
- b. $\text{xudožnik izobrazil } [_{NP} \text{stakan saxaru}] \text{ v natjurmorte}$
 'the artist depicted a cup of sugar in the still life'

(305)a is ambiguous. Since dobavit' also s-selects T, it may mean that the cook, for some reason, threw a container called a "cup" into the mixture. On its more natural interpretation, (305)a means that the cook added only a certain quantity (called a cup) of sugar to the mixture: no cup need be physically present in the scene. (305)b, on the other hand, is unambiguous. The artist's still life must show a physical object called a "cup", and not merely a certain quantity of sugar.

Clearly, these facts can be accounted for only by assuming a theory of s-selection like Grimshaw's, in which dobavit', but not izobrazit', s-selects a K. And, in an obvious way, this selection arises from semantics: one may add X to Y in varying quantities, but one may not depict X in varying quantities. The crucial assumption is the CSR(K) includes both NP and QP, while CSR(T) includes both NP.

On the other hand, this analysis predicts that the partitive genitive, as a QP that satisfies s-selection in situ, should be found equally in properly governed and non-properly governed positions. According to Babby (1980, 83), this is not true. The effect may, however, be due to the null Q, in a fashion discussed with reference to the genitive of negation in Chapter Four, section 2.4.

5.6 To summarize: we showed in the first four sections that subcategorization should be regarded as two independent theories -- θ -theory and c-selection. Only the former falls under the Projection Principle. In this section, we noted that of these two theories, only θ -theory has as primitives categories that meet the criterion of epistemological priority. Thus, in separating out θ -theory and c-selection from the theory of subcategorization, we have in fact separated a theory with plausible primitives from a theory with implausible primitives. This suggests that we have made the correct move.

We investigated the consequences of eliminating entirely the theory of c-selection as an autonomous subtheory of grammar. Much of the explanatory burden of c-selection can be transferred trivially to Grimshaw's theory of s-selection -- distinguishing the complement of kick from the complement of conjecture, for example. Most significantly, the effects

that Grimshaw claimed could not be captured by s-selection alone turned out to follow from Case Theory. Furthermore, using Case Theory instead of c-selection accounts for the non-existence of certain types of predicates, which c-selection predicts should exist.

At this point, an absolutely legitimate question may be posed: are the primitives of Case Theory any more epistemologically prior than the primitives of c-selection? The answer, it seems, is no. Notice, however, that Case Theory is much wider in its explanatory scope than c-selection. Not only does Case Theory account for gaps in the list of possible predicates, but it has innumerable other consequences, some of which have been discussed at various points in our discussion. Clearly, we hope to reduce the properties of Case Theory to other subsystems of grammar, whose primitives may be more plausible than the Case features we have discussed here. Burzio's generalization takes a step in this direction. By collapsing what is left of c-selection with Case Theory, we are reducing the number of subsystems of grammar that rely on unlearnable primitives. In so doing, we may be moving closer to a true theory of UG, but we do not arrive at one.⁸⁶

Finally, the elimination of c-selection in favor of s-selection and Grimshaw's Context Principle should lead naturally to a deeper investigation of the properties of the function CSR. Can we make some deeper sense out of the many-many relations between syntactic and semantic categories that we have assumed to make our analyses work out?

(305) $CSR(\underline{P}, \underline{Q}, \underline{E}) = NP \text{ or } S'$

$CSR(\underline{K}) = NP \text{ or } QP$

$CSR(\underline{T}) : NP$

The study of CSR may help us to understand the semantic content of syntactic categories, and lead naturally into a study of Lexical Semantics -- the cooccurrences of s-selectional features, Case features and θ -features in lexical entries.

We will end our discussion of θ -theory, c-selection and s-selection here. In the next part of this study, relying somewhat on our results in Part I, we take up a new topic: Path Theory.

CHAPTER TWO: FOOTNOTES

1. The particle ni seems to indicate that the word to which it is attached does not refer, independent of whatever quantification takes place at LF. In this we differ from Klima (1964), who viewed ni in Russian as syntactically inserted in the domain of the negation ne, and agree with the traditional view that ni can "intensify" negation. This is because of sentences like (i) below, where ni occurs without sentential negation:

- (i) my zabludilis'. my — nigde!
 we are lost we (are) nowhere

This issue is largely irrelevant to us, but is considered here simply because ni occurs so frequently with the genitive of negation, and positively disposes sentences with negation towards using the genitive of negation.

2. Impersonal (weather) verbs take this form, e.g.:

- (i) xolodalo
 got cold (neut sg)

Russian verbs are overtly marked for person and number in the present and future tenses, and for number and gender in the past tense. Gender distinctions are always neutralized in the plural.

3. The pedigree of this analysis extends at least to Ružička (1963), as discussed below in the text.

4. We use the term monadic to refer to a predicate that θ -marks (directly or indirectly) a single argument. See footnote 7 below.

5 . Babby (1980), in an important study of the genitive of negation construction, proposes that the nominative subjects of intransitive verbs that can be replaced by the genitive are those with a "close semantic link" to their predicates. In particular, he suggests that which verbs allow genitive phrases corresponding to nominative subjects depends as much on the particular subject as on the particular verb: the verb must "denote the subject's most typical 'anthropocentric' action, and that this tendency may be real-world (e.g. rain:fall) or 'idiomatic' (attention:pay)."^V He proposes a notation based on work by Mel'čuk to capture these dependencies.

While it is true that constructions like rain falls in Russian (lit. 'rain goes') are particularly comfortable homes for the genitive of negation construction, this fact does not appear to be central. As we see later in the text, the classification of verbs relevant for the genitive of negation construction is somewhat fuzzy, and the closer a verb gets in a given context to being a pure predicate of existence, the less agentive it is and the more easily it will admit the genitive of negation. Many examples of the genitive of negation can be produced in which there is no particular dependency between the genitive phrase and the verb, even when the genitive phrase corresponds to a nominative subject. This is the case with all the passive examples, where the θ -structure is determined by their active counterparts, and semantic factors are almost non-existent. For example, in (8)b, which is a perfectly natural sentence, one can hardly claim that "showing up" is the most typical 'anthropocentric' action" of a lecturer, nor is "being taken by the enemy" a predicate closely linked to "city" in (4)b from Chvany (1975).

Thus, while there is some effect of the sort discovered by Babby, it is probably derivative of the considerations discussed in the text, and not a primary determinant of the possibility of the genitive of negation.

6. I am grateful to O. Vinogradova for suggesting these examples.

7. We use the term polyadic here to mean verbs that assign one θ -role directly, and another indirectly. This is a slight departure from the more normal usage, which does not distinguish direct from indirect θ -roles. Our usage and normal usage conflict with forms like the passive of a double object verb (e.g. was given NP NP), which is not polyadic in our sense.

8. We assume that AGR, and not TNS, assigns Case and governs in INFL, since [-tense, +agreement] clauses like subjunctives (cf. Picallo, 1982; also Chapter Five, section 3 below) and Portuguese inflected infinitives (Rouveret 1980; Zubizarreta 1980) take nominative subjects. Cf. also Chomsky (1981a, 52).

9. We assume the categorial features of Chomsky (1970):

[+N, +V] = A	[-N, +V] = V
[+N, -V] = N	[-N, -V] = P

The Case assigners are thus V and P. Following Kayne (1981d), we assume that English does not distinguish objective Case (assigned by V) from oblique Case (assigned by P), though nothing hinges on this, for us.

Notice that (25)i-iii are filters on Case assignment, rather than Case assignment rules, strictly speaking. We may add the following rule:

(i) Assign Case

The output of (i) will be governed by principles like (25)i-iii.

10. Vergnaud's work on Case dates from 1977. (25)i-iii is our formulation.

11. Surprisingly, given Case theory, S' cannot be the subject of an infinitival S'. This issue is discussed further in Chapter Three, section 2, but remains mysterious:

(i) *it is difficult [_S, that John came to suggest an answer to our question]

12. This is true in other languages with rich Case systems. Van Riemsdijk (1980) presents a theory of Case features that explains this phenomenon. We discuss oblique Case in greater detail below in the text.

13. Recall from footnote 7 that we use the term polyadic in a nonstandard sense, to mean a verb that assigns at least one indirect θ -role and at least one distinct direct θ -role. A verb like vydat 'issue', 'hand out', is polyadic in the active, where it takes a subject and two objects:

(i)	zamdekan	vydal	dokumenty	studentam
	assistant dean	issued	documents	students
	(masc nom sg)		(masc acc pl)	(masc dat pl)
	'the assistant dean issued documents to the students'			

In the passive, it is not polyadic, since it assigns no external θ -role:

(ii)	nikakie dokumenty _i	ne	byli vydany	e _i	studentam
	no	documents	NEG were issued		students
	(masc nom pl)				(masc dat pl)

The subject in (i), since it is a D-structure subject, cannot be replaced by the genitive under negation, but the subject in (ii), which is a D-structure object, may be.

14. Comparison of (8) with (12) shows that verbs with the "reflexive" suffix -sja (-s' after vowels) may be either ergative or intransitive. This contradicts the analysis of Babby and Brecht (1975), who propose, in effect, that all -sja-verbs are ergative, and that -sja is a "spell-out" of the trace of NP-movement. This analysis is intuitively wrong for cases like kusat'sja 'bite', in (12), an intransitive verb transparently related to the transitive verb kusat' 'bite':

- (i) sobaka kusalas' 'the dog bit'
 (ii) sobaka kusala košku^V 'the dog bit the cat'

We probably wish to say that the subject sobaka is an agent in both sentences, and is indirectly θ -marked by its verb, to capture the obvious parallels between (i) and (ii).

Many of the properties of -sja remain mysterious, if indeed there is any regularity to -sja's behavior beyond the fact that -sja-verbs (in standard Russian) can never assign Case. Sometimes addition of -sja, like addition of passive morphology, dethematizes the subject position as well as removing the ability of the verb to assign Case; other times, addition of -sja dethematizes the object, as in (i). Still other verbs, like pojavit'sja 'show up' in (8), are -sja-verbs tantum, at least in their meaning with -sja. It may be the case that all -sja-verbs tantum are ergative, although this is not certain. (I am grateful to L. Babby (personal communication) for raising the issue of -sja in this context.)

15. Exceptions to Burzio's generalization in English include strike and impress, if sentences like (i) have the S-structure indicated:

(i) John_i strikes me e_i as obnoxious

Russian contains a host of what appear to be counterexamples to Burzio's generalization, discussed by Babby (1975) -- verbs that take null, neuter subjects and accusative objects. These include verbs of uncontrollable physical sensation, like tošnit' 'nauseate' or znobit' 'chill':

(i) tošnilo Mašu 'Maša felt nauseated'
nauseated Maša
(neut sg) (fem acc sg)

(ii) ego znobilo 'He was chilled'
him chilled
(acc) (neut sg)

-- and also certain cases of "spontaneous demotion" with verbs that "denote a physical action that can occur spontaneously, without the initiation or intervention of an animate agent" (Babby 1975, note 4):

(iii) veter unes lodku
wind carried away boat
(masc nom sg) (masc sg) (fem acc sg)

(iv) uneslo lodku (vetrom)
(neut sg) (fem acc sg) (masc instr sg)
'the boat was carried away by the wind'

The subject of all of these verbs may be a θ -position, however,

if we accept Chomsky's (1981a) suggestion that impersonal subjects like "weather it" in English are actually "quasi-arguments" that receive a special θ -role of some sort, perhaps a θ -role "natural cause". Notice that if it is not "nature" or "fate" carrying off the boat in (iii)-(iv), "spontaneous demotion" is impossible:

(v) Ivan unes lodku
 Ivan carried away boat
 (masc nom sg) (masc sg) (fem sg)

(vi) *uneslo lodku Ivanom
 (neut sg) (fem sg) (masc intr sg)

16. We differ from Anderson, whose rules only apply to transitive polyadic verbs (in our terminology). Our rule is essentially identical to the universal proposed by Williams (1981b): "if there is an Actor [= our agent], it must be external [= indirectly θ -marked] for V".

17. This point is also made by Williams (1981b). Cross-linguistically, attributive and identificational predicates fail tests for ergativity. For example, in Italian, post-verbal subjects of such predicates never allow ne-cliticization, which is restricted to post-verbal subjects of ergative verbs (Belletti and Rizzi 1981; Burzio 1981):

(i) ne sono arrivati molti 'many of them have come'
 of-them are come many

(ii) *ne sono buoni molti 'many of them are good'
 of-them are good many

In French, such predicates never participate in the il-impersonal construction:

- (iii) il est arrivé un homme 'there arrived a man'
 there is come a man
- (iv) *il est enivré un homme 'there was drunk a man'

In Russian, a handful of adjectives allow their sole argument to be genitive under negation, suggesting that the argument is not an "attribute", if our rule in (44) is right. All these adjectives optionally allow their argument to be marked accusative instead of nominative. The adjectives in question are all either adjectives of perception (vidno 'visible'; slyšno 'audible'; oščutimo 'perceptible') or modal adjectives (nužno 'needed'; neobxodimo 'essential'):

- (i) ^Vženšćina ^{VV} byla vidna
 woman was visible
 (fem nom sg) (fem sg)
- (ii) vidno bylo ^Vženšćinu ^{VV}
 visible was woman
 (neut sg) (fem acc sg)
- (iii) ne bylo vidno ni odnoj ^Vženšćiny ^{VV}
 NEG was visible not one woman
 (neut sg) (fem gen sg)

The accusative marking in (ii) violates our Case marking conventions for Russian. As we shall see in the next section in text, we cannot claim that accusative Case is oblique here, in view of the possibility of the genitive in (iii). We may make some ad hoc statement about these predicates -- possibly they are "verbal" in some sense -- but no interesting account of their properties will be attempted here. Perhaps the problem is related to the impersonal verbs discussed in footnote 15.

18. Not only studenty, but duraki may also not be replaced by a genitive phrase under negation, suggesting perhaps that both arguments are indirectly Θ -marked, and are perhaps daughters of S.

19. We disagree here with Babby (1979, 1980). He claims (citing Russian grammarians in support of his position) that "indefiniteness" is biuniquely correlated with inverted order in all sentences without the genitive of negation, while sentences with the genitive of negation are "indefinite" with any order. We find, in working with informants, that inverted order tends to favor an indefinite reading in all cases. Nonetheless, sentences without the genitive of negation, like (55)a, are ambiguous with any order, while sentences with the genitive, like (55)b are indefinite with any order. As noted earlier, and contra Babby, it seems that the genitive of negation is always more natural with inverted order, a fact which we do not derive directly from anything having to do with logical interpretation, as shall be apparent later.

20. The pas-de construction in French (which we discuss briefly in section 2 of this chapter and in Chapter Four), analyzed by Kayne (1975, 1981a), is very similar to the genitive of negation construction in Russian, and may admit the same analysis. Isabelle Haïk (personal communication) observes that a problem similar to that discussed in text exists with indefinite NPs in object position in French, when they do not participate in the pas-de construction. Thus, (i) is odd with the natural sense "Jean hasn't got a sister", while (ii) is acceptable:

- (i) ??Jean n'a pas une soeur
 Jean NEG hasn't a sister
- (ii) Jean n'a pas de soeur
 of

Note that (i) is odd despite the overt article, which presumably marks indefiniteness.

The Russian problem is also connected to the well-known near obligatoriness of the genitive of negation with phrases corresponding to accusative objects when the phrase includes the "intensifying" negative particle ni discussed in footnote 1 (cf. Ravič^V 1971):

- (iii) ??ja ne polučal^V ni odno pis'mo
 I NEG received not one letter
 (neut gen sg)

Not surprisingly, given our observations in text, but still mysteriously, no such contrast appears when nominative subjects and genitive phrases corresponding to them are considered (data due to M. G. Sinicyn and O. Vinogradova):

- (v) ne pojavilsja ni odin student
 NEG showed up not one student
 (masc sg) (masc nom sg)
- (vi) ne pojavilos' ni odnogo studenta
 NEG showed up not one student
 (neut sg) (masc gen sg)

Finally, it should be noted that there is no problem with ni in oblique positions, where condition B excludes the genitive of negation, as seen in (53)a, nor with accusatives in PPs, as in (51)a. We have

no explanation to offer for these facts. See Boguslavskij (1979) for some discussion of the problem.

21. Logical Forms like (56)b and (58)b for the genitive of negation construction are also given by Padučeva (1974, 143-159). As I. Boguslavskij (personal communication) has pointed out, similar contrasts are found with accusative expressions of duration, as in (i)-(ii) below:

(i) ja odin ^Včas ne spal
 I one hour NEG slept
 (masc gen sg)

(ii) ?ja odnogo ^Včasa ne spal
 I one hour NEG slept
 (masc gen sg)

(i), with accusative Case, tends to mean "there was a one-hour period in which I did not sleep", while (ii) means "I didn't even sleep for an hour" -- i.e. slept for less. (i) can also have the latter interpretation. In sentences like (ii), as also when a genitive phrase corresponds to a nominative subject, acceptability is increased if a particle like i 'even' is added:

(iii) ja i odnogo ^Včasa ne spal

Particles like i may provide the context necessary to interpret the quantification imposed by the genitive of negation in these cases.

22. Crockett (1976b, 352) gives some examples of po-phrases with plural verbal agreement:

- (j) [po dvadcat' ^Včelovek] priezžali s nim
po twenty people came with him
(pl)
'twenty people came with him [each time]'

Given our theory presented in section 2, we expect such constructions to exist, and also predict that they should violate condition A. My informants accept phrases like (i), where po precedes a numeral (which itself involves various Case-marking complications), but do not accept sentences like (ii), where no numeral is involved:

- (ii) *[po devuške] priezžala/priezžali s každy^Vm mal'čikom
po girl arrived with each boy
(fem dat sg) (fem sg) / (pl)

The only good version of (ii) has a neuter singular verb. See discussion of the "QP hypothesis" in section 2.

23. Crockett (1976a) cites an article by Sidorov and Il'inskaja (1949) (which I have not consulted), who observed that po-phrases can only occur as subjects of intransitive verbs or as direct objects, with intransitive given its traditional meaning -- including both ergatives and intransitives. Only the distinction between ergatives and intransitives separates Sidorov and Il'inskaja's formulation from our condition A.

24. The rule of absorption is discussed in greater detail in another context in Chapter Five, section 2.

25. We distinguish the term numeral (for words like six, two, etc.) from grammatical number (singular, plural).

26. Observations close to our own have appeared elsewhere in the literature. Suprun (1959, 91) notes that the use of the plural with numeral phrases is "especially preferable in constructions with nouns denoting active personae, when the active character of the action is emphasized, and not the number of the personae...The plural form is used in constructions with an adjective in the nominative case..." (translation mine). The restrictions on usage with adjectives are also noted by Corbett (1980). See also Crockett (1976b, 349) for possibly related observations; she notes that in clauses with no-agreement numeral phrases the verb carries "little communicative weight. The verbs in such sentences are therefore essentially existential, or verbs used in an existential sense, unless they are verbs which cannot possibly provide any new information in the given context." The somewhat vague term "existential" is also used by Babby (1980) to characterize the monadic verbs that participate in the genitive of negation construction.

It should be noted further that certain intransitive verbs, like rabotat' 'work', which do not allow the genitive of negation or po-phrases (cf. (11)a-b), are occasionally found in constructions with no-agreement numeral phrases, as are some polyadic transitive verbs, as in (i), from Crockett (1976b, 359, note 16):

- (i) \check{V} suby kupilo [pjat' turistov]
 furs bought five tourists
 (fem acc pl) (neut sg) (masc gen pl)
 'five tourists bought furs'

Such examples are also occasionally found in the French il-impersonal

construction, which also obeys condition A (Pollock 1981), and may indicate a reanalysis of the verb and its direct object into a single, ergative verb. In any case, such examples are exceptional. Crockett notes that if the numeral phrase is preverbal in (i), the result is less acceptable, to a greater degree than usual.

27. For example, we have seen that the "paucal" numerals 2-4 and numerals ending in 2-4 are followed by a noun in the genitive singular in non-oblique positions. If this noun is modified by an adjective and is of masculine or neuter gender, the adjective will be genitive plural. If the noun is feminine, the adjective may be genitive plural or nominative plural, unless, according to Borrás and Christian (1971), the nominative plural of the noun, normally homonymous with the genitive singular, differs from the genitive singular in stress, in which case the adjective is normally in the genitive plural. If the modifying adjective is in the nominative plural, the verb must agree, while if the modifying adjective is in the genitive plural, verbal agreement is apparently optional. Remarks of Iomdin (1980, 66) suggest that the choice of Case for the modifier has some effect on logical interpretation, possibly related to our condition C. We will make no attempt to account for all these conditions on numeral phrases.

28. We ignore some complexities of Polish numerals. For example, no-agreement numeral phrases, which include all uses of the higher numerals, do not obey A, while they behave like their Russian equivalents in oblique positions. Numerals 2-4, when they do not

trigger agreement, appear in the genitive Case, again not obeying A. Thus, pending further investigation, our discussion of Polish should be treated with caution. I am grateful to Wayles Browne and to Hanna de Hackbeil for discussion of Polish numerals.

29. More precisely, the LF representation resulting from QR is (i):

(i) [_S six mathematicians_i [_S e_i proved the theorem]]

From (i), further rules of conversion construct a representation like (79).

30. We can tell that razlu^Vcit'sja is ergative by the grammaticality of a sentence like (i), with the genitive of negation:

(i) ne razlu^Včilos' ni odnoj gruppy
 NEG parted-company not one group
 (neut sg) (fem gen sg)

Notice that after QR, (i) contains a variable which ranges over groups. Cf. footnote 31 below.

31. Sentences like (82)b and d have, of course, a grammatical "science fiction" sense, in which mathematicians and other people are treated as collectivities, like clouds or crowds. Also, we predict that no-agreement numeral phrases with verbs like razlu^Vcit'sja should be acceptable if the numeral quantifier binds a variable which itself ranges over groups, as in:

(i) šest' grupp razlu^Včilos' na mostu

This prediction seems correct: (i) means that each of six groups individually dispersed. The corresponding agreement version of (i) is, apparently, ambiguous, as predicted. Simple plural NPs also participate in the ambiguity discussed in the text, but do not cooccur with non-agreeing verbs. Sometimes conjoined NPs do, as discussed by Crockett (1976b). Whether conditions A-C obtain in this domain is unknown, although Crockett does remark that agentivity appears to favor agreement, suggesting condition A.

32. If, as we argue in 2.3, Q does not bear Case features, we might relate the failure of these "adnominals" to modify Q to the fact that they will not be able to agree with Q and receive Case. If adnominals are subject to some sort of Case Filter, they will violate it in this instance. This suggestion, of course, rests on the assumption that a pre-Q adnominal cannot be a scrambled modifier of N.

Revzin (1978, 269) discusses no-agreement numeral phrases (QPs for us) with a genitive adnominal preceding Q. This adnominal must belong to a very small class, e.g. kakie-to 'some', but not vse 'all':

(i) kakix-to / *vsex ^V sest' studentov prišlo / *prišli
 some all sex students arrived
 (masc gen pl) (neut sg)/ (pl)

We have no account of this phenomenon.

33. This phenomenon is limited to direct Θ -marking. Why? Recall that lexical items are responsible for direct Θ -marking, while indirect Θ -marking is a compositional property of a predicate together with its complements. It is a frequent observation that idiosyncratic lexical

properties are lost under composition. P. Kiparsky has noted that the irregular alternation between singular leaf and plural leaves is lost in the name of a sports team Toronto Mapleleaves. Similarly, Carlson and Roeper (1980) note that addition of the prefix re- to a verb eliminates various idiosyncratic subcategorizations, e.g. for a particular preposition. The same observation probably carries over to Θ -role assignment: any idiosyncratic requirement a verb might decide to make of its subject will be lost under the rule of composition that yields indirect Θ -marking. On the other hand, we avoid a number of interesting questions concerning apparent dative subjects in Russian: are these arguments directly or indirectly Θ -marked?

34. On the other hand, Θ -marking may depend on Case assignment in some instances, if we accept Chomsky's reduction of the Case Filter to the Θ -criterion, discussed below in text.

35. Actually, we may say that oblique Case marking is also formally optional, but is made obligatory whenever it can occur by the Θ -criterion.

36. Such expressions of duration occur even with verbs that take quirky objects, and they remain, of course, accusative:

- (i) ja pomogal `etoj devuške celuju nedelju
 I helped this girl all week
 (fem dat sg) (fer acc sg)

37. We phrase the restriction as we do because of the possibility of an accusative expression of duration even when governed by an ergative or passive verb:

- (i) on byl priglasen každyju nedelju
 he was invited each week
 (fem acc sg)

38. We do not exclude the possibility that accusative might function as an oblique Case for some verbs, which would then be stipulated to Θ -mark a [+accusative] constituent. We might detect such verbs by examining verbs that do not allow passivization. For most speakers, Russian verbs that take oblique objects do not passivize (some verbs requiring dative Case being an exception, for some speakers). This is probably due to a convergence of three factors:

(1) Oblique Case must be present on the directly Θ -marked argument at all levels, because of the Projection Principle.

(2) The passive of a verb requiring oblique Case does not allow that Case to be assigned to its object, just like any other passive verb.

(3) AGR may not govern an oblique NP in Russian (pace possible dative subjects; cf. note 33)

(3) appears to be a parameter among languages. Icelandic, for example, allows any sort of oblique NP to fill subject position (cf. Levin 1981, and references therein; also Ingria 1981).

As noted by Švedova (1970, 352) there are verbs taking accusative objects that do not passivize, where there is no plausible semantic explanation. These may be oblique accusatives, which cannot passivize because of (1)-(3) above. On the other hand, it is hard to imagine what evidence a Russian child detects that can distinguish a structural accusative object from an oblique accusative object.

39. Recall from our Case assignment conventions that [+N] categories only assign oblique Case to their complements (in the unmarked instance,

genitive). Extending this suggestion to English, we might account for the lack of "raising to subject" in NPs. Consider the contrast between (i) and (ii):

- (i) John's_i consideration e_i for the job
(ii) *John's_i consideration [e_i to be a fool]

In both (i) and (ii) John can bear no other Case but oblique, since it is not governed by [-N] or by AGR. In (i), oblique Case is required by the noun consideration (possibly by a lexical redundancy rule stating that [+N] requires genitive to assign its θ -role, unless otherwise specified).

Supposing, as we suggest below in text, that Case features are shared by all members of a chain, both John and e_i in (ii) will bear genitive Case. This Case may be required by the noun consideration, but consideration is not the θ -marker of the chain (John, e_i). Rather, the phrase to be a fool θ -marks this chain, and the phrase to be a fool does not require genitive Case. (ii) thus violates the Oblique Biconditional.

40. One instance in which a WH-construction and a construction with a lexical NP appear to differ with respect to the Case Filter is noted by Kayne (1980):

- (i) John, who_i I assure you [e_i to be the best]
(ii) *I assure you [John to be the best]

We have no account of this.

41. On Chomsky's reduction of the Case Filter to the θ -criterion, we should say "if our QPs did have to belong to chains to receive a θ -role".

We will continue to say "QP is not subject to the Case Filter", at least as an informal way of speaking.

42. Sinicyn (1982b) also notes the following examples, with "scrambling" of an in situ WH into a higher clause, from Colloquial Russian:

- (i) *on kto_i dumaet [e_i pridet pervyj]?
 he who thinks will come first
 (nom) (nom) (3 sg)
- (ii) ?(*)ty kto_i dumaes'^v [e_i pridet pervyj]
 you who think will come first
 (nom) (nom) (2 sg)
- (iii) ty kogo_i dumaes'^v [on ljubit e_i]
 you whom think he loves
 (nom) (acc) (2 sg) (nom) (3 sg)

The greater ungrammaticality of (i) than (ii) probably results, as Sinicyn suggests, from a "Feature Coincidence Filter", perceptually based, that blocks scrambling into a higher clause when the scrambled material can be confused with an argument of the higher clause, as in (iv)-(v):

- (iv) ty komu_i dumaes'^v [ja peredal knigu e_i]
 you who think I delivered book
 (nom) (dat) (2 sg) (nom) (acc)
- (v) **ty komu_i skazal emu [ja peredal knigu e_i]
 you who said him I delivered book
 (nom) (dat) (dat) (acc)

Nonetheless, we suggest, contra Sinicyn, that the contrast between (i) and (ii) results from the fact that (i) violates both the Feature Coincidence Filter and the ECP (cf. below in text), while (ii) violates only the ECP. In other words, we suggest that the second person ty and the third person kto (ii) do not count as the "same" for Sinicyn's

filter, while the third person on and kto in (i), and komu and emu in (v) do.

Certain apparent acceptable long movements from subject position of a tensed sentence are probably left dislocations with a suppressed subject pronoun, as in (vi)-(vii) (from Sinicyn op. cit.):

(vi) [muž^v ee]_i, ne proxodilo dnja, čtoby^v (on_i) ne vspominal jego
 husband her NEG passed day that (he) NEG thought of him
 (nom) (gen) (gen) (subjunc)

'her husband_i there didn't pass a single day but that (he_i) thought of him' (Gogol', Dead Souls)

(vii) on_i vidno [čto^v (on_i) tronut]
 he visible that (he) moved
 (nom)

'him, it's visible that (he) is moved' (Tolstoj, Anna Karenina)

(109)d is not amenable to such an analysis.

As discussed in Pesetsky (1982), long movement of WH-words from a tensed S' is generally rather unacceptable in Standard Russian, though judgments vary, and subjunctive clauses generally allow extraction more freely than indicative clauses (cf. Comrie 1973 and Haiman 1974). The same contrast between long movement of a subject and object does show up in indicative sentences -- a point relevant in Chapter Five, section 3.

43. As is well known, these effects disappear in English when the complementizer is not phonologically present. This correlation is discussed in Pesetsky (1982) (cf. also Taraldsen 1980, Kayne 1980, 1981a), and will be discussed in Chapter Three. Colloquial, but not literary Russian, allows empty COMPs, as seen in footnote 42, but I do not know whether an empty COMP causes CTP effects in WH-movement to disappear in Russian.

44. We will discuss the revisions of the definition of government in greater detail in the penultimate section of Chapter Five.

45. Picallo (1982) discusses some evidence suggesting that AGR is a proper governor in Italian. I defer discussion of her work until Chapter Five, section 3.

46. With respect to structurally based conditions like the ECP, Binding Theory, and with respect to Θ -marking, Russian acts as if it were a language with fixed structure, like English. For this reason, we take scrambling, at least clause-internally, to be a relatively "superficial" phenomenon. Put another way, Russian does not appear to be a W*" language of the sort envisioned by Hale (1979), in which phrase structure is flat at all levels.

In this connection, remember that Russian speakers, although they allow a significant freedom of word order, have relatively clear intuitions about unmarked vs. marked word orders. Thus, our treatment of Russian does not derive from fitting it into a configurational Procrustean bed. Rather, empirical investigation of the properties of Russian with respect to principles of grammar leads to the conclusion that at some level of representation Russian is configurationally well-articulated, just like English.

Finally, note that the data from Sinicyn (1982b), discussed in footnote 42, suggest that inter-clausal "scrambling" is a movement rule of syntax, since its output appears subject to the ECP (with the caveat that this is not Sinicyn's explanation of the facts).

I am grateful to R. Kayne and H. Borer, who both predicted that

things would turn out this way in Russian.

47. But cf. Pesetsky (1982b, 3.3), where we present an argument against this view, based on the influence of surface linear order on quantifier scope. Since the connection between linear order and scope (challenged in Huang 1982) is not made in the present theory in the first place, the relevance of our earlier argument is unclear.

48. English and French do allow postverbal subjects with indefinite NPS:

(i) they arrived a cow

(ii) il était arrivé une vache

Safir (1982) suggests that these languages do allow a special sort of transmission of nominative Case in these constructions, which entails the indefiniteness restriction in his theory. Modern Russian lacks even this possibility, although Old and North Russian appear to have similar constructions:

(iii) prišlo^v korova
 arrived cow
 (neut sg) (fem nom sg)

49. Actually, Pesetsky (1982) asserts this fact based on data involving the negated existential predicate net. As we show later, net has some special properties which make it not the best predicate to use in testing the QP hypothesis.

50. These results, obviously, should be duplicable with the other QPs

considered in section 1. The problem is that QPs like po-phrases do not naturally include WH-elements (cf. English: *[what book each] did you give to the men), necessitating examples involving topicalization, which also appear to be rather odd in all cases. With numerals, however, the relevant contrast appears to hold with skol'ko 'how many', although I have found great variation among informants:

- (i) *skol'ko žurnalov_i ty xočeš', čtoby byli vybrošeny
 how many magazines you want that be thrown away
 (pl)
- (ii) ?skol'ko žurnalov_i ty xočeš', čtoby bylo vybrošeno
 (neut sg)

In Chapter Four, section 2.4, we present some extremely powerful evidence that QP numeral phrases (that do not trigger agreement) are S-structure objects of ergative and passive verbs. The evidence comes from constructions in which skol'ko (like French combien; Kayne 1981a) is extracted from its numeral phrase.

51. Note that QPs are excluded from replacing NPs as the complements of adjectives and nouns by the Non-Obliqueness Restriction, since [+N] categories that take NP complements must assign oblique Case to these complements. There may be instances of QP complements to N in phrases like:

- (i) [_{NP} šest' knig po 25 rublej]
 six books po 25 rubles
 'six books @ 25 rubles apiece'

52. Kayne explained his observation in terms of the Nominative Island Condition (NIC) of Chomsky (1980), but his explanation is easily rephrased in terms of the ECP. Picallo (1982) notes that Kayne's examples all involve subjunctive verbs, a fact she considers significant. See Chapter Five, section 3, for discussion.

53. The relationship between a negative ne and a word with the particle ni in Russian appears to be similar. Since ne cannot control a ni in a subordinate clause unless that clause is infinitive, Kayne's paradigm cannot be reproduced in Russian.

54. Rizzi (1982) notes as an argument that the ECP applies in Italian that Kayne's paradigm reproduces exactly with Italian nessuno 'no one'. Nessuno may have wide scope as an object or postverbal subject, but not as a preverbal subject. See Chapter Five, section 3.

55. Insofar as proper binding is relevant for the definition of government, clause (ii) requiring c-command is actually redundant with clause (iii) of the definition of government in (132).

56. Suppose we added Case as a type of feature relevant to possible antecedency, and claimed that possible antecedency was relevant only for the proper binding relation necessary for proper government of a subject. It might follow that a WH-word would have to match the Case of the category it binds when that category is a subject, but not otherwise. This might explain certain asymmetries in "Case attraction" phenomena: for example, in Classical Greek. According to Goodwin (1894, §1033), a WH in a relative construction may assume the Case of the head of the relative when that WH binds an accusative position, but seldom when a nominative

position is bound.

57. We must assume, given the obligatoriness of QR for QP expressions of duration (cf. footnote 21), that such expressions of duration are "inherently" c-selected to be NPs. We leave open how this "inherent c-selection" works.

58. An exception is provided by certain refinements of the definition of governing category introduced by Chomsky (1981a, 58). We discuss these cases briefly in Chapter Four, 4.5, but they do not affect the thrust of our discussion here.

59. One might provide a notational variant of subjacency which analyzes traces in COMP (cf. Koster 1978b; Bresnan and Grimshaw 1978) which would motivate the existence of e_i^1 in (173).

60. Recall that we have not shown that they must move. We discuss this below in text.

61. It is often proposed (cf. Babby 1980 and references therein) that the availability of constructions we have analyzed as QP's depends on discourse factors -- in particular on the QP's being in the comment or rheme of the sentence. This intuition, which appears regularly in the literature, has never been made precise, but has usually been assumed to condition the possibility of the construction. For example, Babby (1980) (1980) states a rule (his (160) which stipulates, in effect, that an NP may participate in the genitive of negation construction in the environment:

(i) [_R V ____]

"R" here is identified with rheme: subjects of transitive verbs, he

suggests, are never part of the rheme, while "subjects" of intransitive verbs and passives may be part of the rheme, thus accounting for condition A.

If our results are correct, however, we need not assume that condition A is derived from discourse notions like "theme" or "rheme". We may identify R with the syntactic node VP, and motivate (i) by formal principles of some generality, leading to the requirement that QPs (including genitives of negation) must remain in VP at all levels. Babby's remark that sentences with QPs (our analysis) are in some sense "unpartitioned" in the usual way into subject and predicate (constituting "thetic judgments", in the sense of Kuroda 1972) follows in the same way: a QP argument of a passive or ergative verb is always in the syntactic predicate -- in the VP.

As noted by Halle (1980), in choosing between "formal" and "functional" explanations of linguistic facts, it is a mistake to reject out of hand the functional explanation. When faced with "solid facts" like condition A, however, or phonological processes of the sort Halle considers, it generally turns out that the formal explanation is superior to the functional alternative. Judgments about theme/rheme structure tend to be soft: intonation, context, and a host of other factors influence these judgments. Condition A, however, -- at least and particularly as it applies to genitive phrases under negation -- is not a soft fact. No changes in intonation or context can legitimate the genitive of negation as the subject of a predicate like reads books. Thus, condition A probably has nothing to do with notions like theme and rheme, at the level of explanation. This is not to say that there is no residue of rhematic influence on judgments about our facts, which should be accounted for. For example, if a QP (e.g. the genitive of negation) is used by a speaker instead of an NP

in a given construction, the sentence will always be interpreted such that the QP is in the rheme, while if an NP is used, the rhematic place of the NP is uncertain. Thus, as Babby notes, a speaker will use word order to signal the rhematic status of an NP, while word order in the case of a QP generally expresses a different distinction -- that between old and new information. These facts, however, have a clearly different status from condition A: they are "soft", express tendencies, and can be overridden by all sorts of other signals. These facts, and their explanation, probably belong to discourse theory.

62. We might mention here a construction that may provide striking support for our general approach to the Russian QPs we have been discussing. This is the construction discussed extensively in Chapter Five of Crockett (1976b), which she calls "type (7) sentences", and we might call Crockett-sentences. These are constructions like (i):

(i) mal'^Včikov ostalos' sem'
 boys remained seven 'there remains seven boys'
 (masc gen pl) (neut sg) (Crockett, 1976b, 318)

(ii) [punktov etix] my nametim tri
 points these we will note three 'we will note three of these
 (masc gen pl) points'
 (Peškovskij, 1956, 224)

In these constructions, a noun phrase in the genitive is associated with a quantifier, but the two are separate. That this is not simple scrambling is shown by Crockett (1976b). First, the quantifier tri 'three' takes the genitive singular when in construction with a noun, not the plural as in (ii). Second, perhaps the one true limitation on scrambling in Russian is the impossibility of separating a genitive from the phrase it is

dependent on:

- (iii) *Ivana ja kupil knigu
Ivan's I bought book 'I bought Ivan's book'
(masc gen) (fem acc sg)

Rather, we might suggest the following structure for these examples:

- (iv) [_S [_{QP} [_Q^e] mal'^Vcikov]_i [_S e INFL [_{VP} ostalos' [_{NP} sem'^V]_i]]]
(v) [_S [_{QP} [_Q^e] punktov etix]_i [_S my INFL [_{VP} nametim [_{NP} tri]_i]]]

Examples (vi)-(vii) show clearly that the "adjoined" QP in Crockett-sentences may bind an NP:

- (vi) [_S [_{QP} [_Q^e] družej]_i [_S e INFL [_{VP} priš^Vlo [_{NP} odna Maš^Va]_i]]]
 friends arrived only Maš^Va
 (masc gen pl) (neut sg) (fem nom sg)
'of my friends, only Maš^Va arrived'
(vii) [_S [_{QP} [_Q^e] družej]_i [ja INFL [_{VP} priglasil [_{NP} odnu Maš^Vu]_i]]]
 friends I invited only Maš^Va
 (masc gen pl) (nom) (masc sg) (fem acc sg)
'of my friends, I invited only Maš^Va'

Crockett-sentences are remarkable in a number of ways, and deserve further study. First, (vi) represents the only place in modern standard Russian (cf. footnote 48) where nominative Case is found with a non-agreeing verb. We speculate that the NPs in (vi)-(vii) have the properties of the QPs that seem to bind them, hence they are not subject to the Case Filter and remain in object position of the ergative verb in (vi).

What is particularly striking is the fact that these Crockett-sentences, if we have given them the right analysis, represent exactly the configurations we attribute at LF to normal Russian sentences with QPs: the QP is in an \bar{A} -position binding an NP. In our text examples of

this configuration, however, this NP was phonologically null and a creature of our analysis, while in Crockett-sentences the NP is phonologically realized and uncontroversial. It is especially interesting, in this connection, that something very like condition A holds in Crockett-sentences. We have placed odna Maša^V in object position in (vi) because when it is an uncontroversial subject -- if the verb agrees and/or is not ergative -- the result is bad, as noted in part by Crockett (p. 330):

(vii) *druzej prišla^V odna Maša^V
 friends arrived only Maša^V
 (masc gen pl) (fem sg) (fem nom sg)

(viii) *druzej odna Maša^V otvetila na vopros
 friends only Maša^V answered to question
 (masc gen pl) (fem nom sg) (fem sg)
 'of my friends, only Maša^V answered the question'

If in some sense (cf., perhaps, Chapter Five, section 2), the bound NP in Crockett-sentences has the properties of an empty category, the contrast between (vii)-(viii) and the other Crockett-sentences we have discussed can be related to the ECP, as instances of the general contrast between (ix) and (x) which derived condition A for us in the text:

(ix) *_S QP_i [_S NP_i INFL VP]]

(x) [_S QP_i [_S subject [_{VP} V NP_i]]]]

It is also worth mentioning that Crockett-sentences also obey condition B, suggesting again that the NP appears to have the Case properties of the QP that binds it, although how this comes about is very unclear:

(xi) *druzej ja pomog odnoj Maše^V
 friends I helped only Maša^V
 (masc gen pl) (fem dat sg)

Finally, uninteresting semantic equivalents to Crockett-sentences

with the preposition iz 'from', 'of' show none of the properties discussed above:

- (xii) iz moix družej, ja priglasil odnu Mašu^V
of my friends I invited only Maša^V
(masc gen pl) (fem acc sg)
- (xiii) iz moix družej, prišla/*prišlo odna Maša^V
of my friends arrived only Maša^V
(masc gen pl) (fem nom sg)
- (xiv) iz moix družej, ja pomog odnoj Maše^V
of my friends I helped only Maša^V
(masc gen pl) (fem dat sg)

I wish to thank V. Schiller for these data.

63. Obviously, further interpretation of the LF representation in (198) takes place. Perhaps a rule of "conversion" (cf. May 1977) of the following sort applies:

$$(i) \quad [{}_S, WH_i [...e_i...]] \longrightarrow [{}_S, \exists x_i: ...x_i...] / \text{ in an } \bar{A}\text{-position}$$

The relation between the WH-operator and the existential quantifier is obscure. Rule (i) stipulates this relation, but does not explain it. We might compare the interpretation of S as a restriction on the quantification in (i) to the interpretation of the predicate of a small clause as a restriction on quantification below in text; section 4.4.2.3.

64. Thus, we have no "matching effects" (discussed by Bresnan and Grimshaw 19878) with respect to Case in these constructions, but there are matching effects with respect to semantic type (see section 5 below). The WH-word must be an operator appropriate for the c-selected variable left after QR:

- (i) *ja kupil [_S, kogda_i [_S PRO pisat' moj roman]]
 I bought when write my novel
 (infin)

-- but cf:

- (ii) ?ja kupil [_S, gde [_S PRO pisat' moj roman]]
 I bought where write my novel
 'I bought a place in which to write my novel' (i.e. a dacha)

This suggest, perhaps, decomposition of gde 'where' into in + what place followed by reconstruction of the preposition in.

65. Actually, (199)a demands a clarification of the Oblique Case Biconditional in (92). (92) allows a constituent to bear Oblique Case if and only if that Case is needed by a θ -marker of that constituent. If θ -roles are shared only by members of A-chains, then a WH in COMP will not bear a θ -role and should not be able to bear Oblique Case. We thus need to stipulate that for the purposes of the Oblique Case Biconditional, WH does inherit a θ -role from its trace. Alternatively, we might stipulate as a convention that WH inherits features from its trace, and allow this inheritance to override the Oblique Case Biconditional.

66. Of course, if an indirect question is c-selected in subject position, the result should not be ungrammatical, since no QR is motivated. Russian does not like very much preverbal sentential subjects (that are not clearly nominal, introduced by to, cto 'the fact that'), but (i) appears to be grammatical:

- (i) [_S, ^Vcem_i [PRO pisat' e_i]] volnovalo nas
 what write disturbed us
 (instr) (infin) (neut sg) (acc)

'what to write with disturbed us'

67. A minority of English speakers find a subject/object asymmetry in tensed free relatives with who, insofar as these are minimally acceptable in the first place:

- (i) ??I punched who you were talking to
- (ii) *who you were talking to punched me

68. That the construction in (210) involves the same predicate as that in (208) is particularly clear in the past or future tense:

- (i) ne bylo/budet knjig
NEG was/will be books
(3 neut sg) (fem gen pl)
- (ii) ne bylo/budet ^Včem pisat'
NEG was/will be what write
(instr) (infin)
'there was nothing to write with'

Interestingly, the genitive is not as obligatory in (i) as in (208), due presumably to the AGR morpheme in the past and future tenses (cf. below in text). For some discussion, see Peškovskij (1956, 366-367).

69. We beg the question of what net c-selects. Our suspicion is that it doesn't -- that is, that it makes no categorial requirements of its complement. Recall that when a QP is found in a position where an NP is c-selected, QR is obligatory, in order to supply an NP trace. For this reason, proper names may not participate in the genitive of negation construction, except in a very special usage:

- (iii) *ne priš^Vlo Maš^Vi
NEG came Maš^Va
(neut sg) (fem gen)

- (iv) #ne prišlo (nikakix) Maš^V
 NEG came no Maš^Vas
 (neut sg) (fem gen pl)
 'nobody by the name of Maša came'

Proper names are never quantificational. On the other hand, a proper name is normal in the genitive of negation with net:

- (v) Maš^Vi net
 (fem gen)
 'Maša isn't here'

We can explain the possibility of (v) if the proper name, contained in a QP, does not have to undergo QR. It will not undergo QR if QPs do not violate the c-selectional requirements of net. Since it is hard to believe that net does not c-select an NP for anything other than Case reasons (see footnote 68), and since NP, QP and S' do not form a natural class, we might conclude that net simply lacks c-selectional requirements entirely. Cf. Ružička 1961, for some discussion of the semantic issues raised in this footnote.

70. Two extremely important points are in order about this data. First, for some speakers, the stars and question marks in (239) are rather strong, although these examples are the fuzziest we have considered so far. For most speakers, sentences like those in (239) are acceptable, but only with what Peškovskij (1956, 244) calls a "special temporal shade of meaning" (osobyj vremmenoj ottenok). This special meaning is almost always available to instrumental secondary predicates, but Peškovskij cites certain examples which have only this meaning:

- (i) kupil ja ego [ešce mičmanom]
 bought I it still midshipman
 'I bought it [some cloth] while still a midshipman' (Gogol)

(ii) ja pomnju, ty ditej s nim ^Včasto tancevala
 I remember you child with him often danced
 (fem nom) (instr) (fem sg)
 'I remember, you often danced with him when you were a child'
 (Griboedov)

(iii) rebenkom on uprjam byl i rezov
 child he stubborn was and playful
 (masc instr) (nom) (masc sg)
 'As a child, he was stubborn and playful' (Ogarev)

Note that the "special temporal meaning" Peškovskij observes constitutes, in effect, an interpretation as a temporal clause denoting an extended period of time. As we note later in the text, the secondary predicate construction that we are interested in always denotes a "transient" state, which is limited to the time-frame of the sentence's tense. Note that those cases in which the "transient" reading is excluded, and in which only the "temporal" reading is present -- and we have quoted all of Peškovskij's examples -- are all instances of the structures in (239). In other words, even when choosing an understood subject that is not governed by V in D-structure does not lead to ungrammaticality, it disambiguates.

There is another interpretation available in some cases for sentences like those in (239). Here the instrumental secondary predicate has what we might call an "indirect" predicative meaning (or, with Borrás and Christian (1971, 4), we might call this the "instrumental of comparison"). Our examples come once more from Peškovskij (loc. cit.):

(iv) vy mne otvetili takim rycarem
 you me answered such knight
 (nom) (dat) (instr sg)
 'you answered me like such a cavalier' (Dostoevskij)

Unlike (v), below, (iv) cannot entail that "you" are a knight. Merely, you are like a knight. In this respect, (iv) differs from (v):

(v) Artur vernulsja rycarem
 Arthur returned knight
 (nom) (instr)

In (v), on its natural interpretation, Arthur is a knight. In sentences like (iv), the complementizer (?) kak 'as' may easily be inserted before the instrumental. Note once more, that the understood subject of the instrumental in (iv), like those in (239), is a D-structure subject. By contrast, vernut'sja 'return' is an ergative verb; Artur is thus a D-structure object.

What is the structure of the constructions in (i)-(iv)? Suppose we maintain the claim that all Small Clauses are non-maximal projections. It is still possible for PRO to be the subject of a Small Clause, if the Small Clause is itself ungoverned. If temporal expressions like (i)-(iii) and "indirect" comparisons like (iv) are taken to be adjoined to S', then they are ungoverned, and PRO should be possible. Alternatively, we could allow Small Clauses to have COMP when COMP is interpreted as filled by a temporal or "comparative" operator, optionally null, but sometimes real, as in (i) or (iv), if kak 'as' introduces the secondary predicate. If Small Clauses are maximal only when temporal or "indirect", it is not so surprising that maximal Small Clauses are never c-selected (the problem which led us to try to maintain the view that Small Clauses of any type appear never to be c-selected. No verb requires a when-clause or an as-clause as a lexical property (except, perhaps English regard in I regard X as Y).

The second important point that (239) raises concerns English. Why does English not show the restriction discussed in the text? For most (but not all) speakers, (vi)-(vii) are acceptable:

(vi) John ate nude

(vii) Mary is happiest drunk

I suspect that the possibility of (vi)-(vii) in English is related to the existence in Russian of nominative secondary predicates. We again draw on Peškovskij (p. 249-252):

(viii)	Aleksandra Ivanova	sidela	bagrovaja	ot	jarosti
	Aleksandra Ivanova	sat	crimson	from	rage
	(fem nom sg)	(fem sg)	(fem nom sg)		

(F. Sologub)

(ix)	(ja)	govorju	eto	ne	sedoj	i	izmucennyj,	a	polnyj	sily	
	I	say	this	not	gray-haired	and	worn-out,	but	full	of	force
	(nom)		(acc)		(masc nom sg	-----)					

(Vrubel')

As (ix) shows, and as Peškovskij notes, such nominative secondary predicates may take as their understood subject the subject of polyadic transitive verbs, and, presumably adjectives and the other predicates in (239). We do not have an analysis of this construction (which, as Peškovskij observes, is often regarded as substandard, though common in the literary language); perhaps the nominative is "in apposition" to the subject in some way. We want to suggest that the possibility of English constructions like (vi)-(vii) may be due to the English equivalent of the Russian nominative construction, and not the Russian instrumental construction. Only in a language with Oblique Case marking, of course, can these constructions be visibly distinguished.

71. Similar observations are made by Nichols (1981), although, interestingly, her examples involve the "temporal" secondary predicates discussed in footnote 70, and she notices some variation among informants with respect to B' in these constructions. My own experience with the non-temporal secondary predicates indicates that B' is a rather strong effect. Condition B' is discussed for similar constructions in Icelandic by Levin and Simpson (1981).

Notice that there is no comparable restriction on control of PRO in infinitivals. It is true that there are no control verbs with a genitive or instrumental controller, but dative controllers abound:

- (i) ja pozvolil Borisu_i [s, PRO pojti v kino]
 I permitted Boris go to movies
 (masc dat) (infin)

For an excellent discussion of control in Russian, cf. Bergel'son (1979).

72. (243) might also be odd in English, but the English equivalent of (244) seems fine.

73. This observation was made independently by Safir (1982).

74. Actually, there may not be any issue here. We might claim that predicates fall under the Case Filter, but are assigned instrumental Case by some element, like a preposition, which does not fall under the Case Filter. A proposal along these lines has been suggested by Schein (forthcoming). Something like this is probably necessary, if instrumental predicates are not to violate the Oblique Case Biconditional.

75. Others include naxodit' 'find', polagat' 'suppose', priznavat' 'deem', 'admit'.

76. It would be possible to detect a verb that takes a Small Clause and requires oblique Case in the following way. All verbs that take a Small Clause also take a tensed clausal complement:

- (i) ja polagaju, [_S, ^Vcto Maša p'jana]
 I suppose that Maša (is) drunk

Tensed clauses may also be headed by the semantically empty demonstrative to:

- (ii) ja polagaju [_{NP} to [_S, ^Vcto Maša p'jana]]
 'I suppose it that Maša is drunk'

When an oblique Case is required, the demonstrative shows that Case:

- (iii) on byl dovolen [_{NP} tem [_S, ^Vcto Maša p'jana]]
 he was satisfied "it" that Maša drunk
 (instr)

As far as I can tell, there are no verbs that take a Small Clause, and show, by the Case of to when they take a tensed complement, that they require oblique Case.

77. But cf. (i), which is extremely ungrammatical, suggesting that (258)a might not be an actual subject Small Clause:

- (i) *[it obvious that John was drunk] shocked me

78. Higginbotham (1981) develops an approach to perception verbs complements in English based in part on our analysis of secondary predicate constructions in this section. He notes a similar "confinement to the transient" in the complements to perception verbs, and suggests a somewhat different explanation.

79. Grimshaw does argue that s-selection eliminates the need for the non-local subcategorization features introduced by Bresnan (1970; 1972).

80. In the sense of "object in the world" not "direct object".

81. Almost. Steriade (1981) notes that the verb welcome, which appears to take a concealed proposition, but not a sentential proposition:

- (i) John welcomed [_{NP} Mary's departure on time]
- (ii) *John welcomed [_S, that Mary departed on time]

She notes the same property in understand, when it has the meaning 'sympathize with'.

Other cases might include the "obligatory extraposition" verbs with factive complements, like resent:

- (v) I resent *(it) that John is here
- I offer no explanation for these cases.

82. The passive of wonder is the least unacceptable of the examples given. B. Schein (personal communication) notes that the passive is rather good with a temporal adverb like often:

- (i) it has often been wondered where you got your funny accent

R. Kayne (personal communication) notes (ii), which suggests that wonder is a Case assigner in some circumstances, if WH-trace needs Case:

- (ii) what_i I'm wondering e_i is why you came

But compare:

- (iii) *what_i I (care, inquired, exclaimed, complained...) e_i is ...

83. A similar approach is taken by Borer (1981). We take no position on whether the preposition is really the head in these constructions, a question with implications for θ -theory and the definition of government. We speak below in text of "a constituent headed by to" for convenience only.

84. In some cases, as with this verb, dictionaries (e.g. Ožegov^V) list the verb as lexically assigning genitive Case (optionally). This just obscures the matter: the genitive Case has partitive meaning. It is also, however, indicative of the fact that the partitive genitive is taken to be selected by the predicates that allow it.

85. The partitive generally takes the endings of the "second" or "-u" genitive in the masculine, rather than the more normal ending in -a. This is also characteristic of the genitive of negation, although the second genitive may occur less frequently with the genitive of negation. The choice of -a or -u genitive is largely stylistic, as remarks of Borras and Christian (1971, 20) suggest. One might suppose that the second genitive is a specific Case assigned by the category Q.

86. In any case, the theory of c-selection (embedded in subcategorization) vastly overpredicted a variety of lexical entries that do not exist. D. Steriade (personal communication) notes, for example, that there are many verbs that require a particular preposition (like those we considered above in text), or a semantic class of prepositions and adverbials (like put, which requires a directional expression), but there are no verbs that subcategorize for PP in general. That is, there are no verbs whose complement can be drawn from an unrestricted set of PPs. Similarly,

particularly given the analysis of seem and similar verbs as subcategorizing small clauses in constructions like John seems sad, there are no predicates that subcategorize AP. Thus, it is fairly clear that c-selection is simply the wrong theory to explain the complementation of predicates.

PART II: PATHS

CHAPTER THREE: CROSSING EFFECTS, THE ECP, AND THE SUBJECT CONDITION

0. Introduction

In Part II of this study, we shift our attention from using the ECP as a probe into the nature of selection and Θ -marking to the study of the etiology of the ECP itself. The ECP, after all, is a rather curious condition, for reasons that we discuss in section 2 of this chapter. In the remainder of this study, we hope to show that a number of the central effects of the ECP actually derive from a much more general principle, which governs the interaction of "paths" in a tree. We will demonstrate that this general principle has also a much wider explanatory scope than the ECP.

In this chapter, we unite the central effects of the ECP for \bar{A} -bound empty categories with two apparently independent conditions: an often proposed constraint on the crossing and nesting of dependencies (Kuno and Robinson 1972; Bordelois 1974; Fodor 1978; Hendrick 1979; Reinhart 1981; among others), and the Subject Condition of Chomsky (1973).

In Chapter Four, we apply our theory to constructions with multiple gaps. Adapting recent work of Kayne (1982) to our framework, we explain a pattern of violations of the Subject Condition in parasitic gap constructions that Kayne has noted. Additionally, we show how our general framework can explain the interaction of the properties of parasitic gaps with crossing and nesting effects. We then proceed to derive most of Ross's (1967) Coordinate Structure Constraint. Applying the theory of multiple gap constructions to coordinate structures, we derive the pattern of "across the board" exceptions to Ross's constraint discussed by Ross and Williams (1978), including constraints on these

constructions that have been attributed to "parallelism" requirements. Our theory extends to constraints on the conjunction of tensed and infinitival clauses, and to a number of other, seemingly unrelated cases. In addition, our analysis may suggest an analysis of parasitic gap phenomena that differs from Chomsky (1981b) and that bears on the nature of the relationship Move α .

Finally, in Chapter Five, we return to the ECP. We discuss an analysis of various constraints on WH-in-situ that have been variously attributed to the ECP or to a Superiority Condition (Chomsky 1973), and suggest an analysis in our terms. Additionally, we consider a number of different topics that have attracted attention in connection with the ECP, in particular the difference between indicative and subjunctive clauses discussed by Picallo (1982).

Our constant goal shall be to propose that a new subsystem of UG, which we might call "Path Theory", incorporating a definition of paths, a constraint on their interaction, and possibly some additional "visibility" conditions, replaces the "Government Theory" of Chomsky (1981a), whose central principle is the ECP. Our work owes much to recent investigations by Kayne, particularly Kayne (1981a; 1982), who first suggested that "paths" in a phrase marker had important properties, and proposed to derive ECP and Subject Condition effects from conditions on these paths. We differ from Kayne in treating the interaction between different paths as crucial to the ECP and the other phenomena that we derive. Nonetheless, our debt to Kayne will be clear, particularly in Chapter Four, where we build on his analysis of parasitic gap constructions.

Before turning to our main topics, however, we present a number of arguments for a structure for S that differs from that assumed in recent

work in the Government-Binding Theory. This structure will be crucial to us when we discuss our derivation of the ECP and Subject Condition; we therefore feel it important to motivate it independently. Part of our discussion derives from Stowell (1981) and from recent suggestions of Chomsky (class lectures, 1982). We will show that the structure for S that we must assume for our theory to work has a number of other nice consequences, predicting the position of the subject of S and providing an extremely simple solution for wanna-contraction and other notorious contraction facts that have been discussed in the recent literature. In section 2 we discuss the ECP in general terms, and in section 3 we begin presenting our derivation of the ECP.

1. The Structure of S

In recent work, Farmer (1980), Stowell (1981), and Chomsky (1981a) consider the possibility of reducing the phrase structure rewrite rules of the base component (as presented in Aspects) to other, independent principles of grammar. For example, Chomsky (1981b) considers a traditional phrase structure rule expanding VP:

(1) $VP \longrightarrow V NP S'$ (Chomsky's (4)ii)

He notes that the fact that the V dominates a head called VP need not be specified in such a rule, since this follows from X-bar theory. We may therefore simplify (1) as (2):

(2) $VP \longrightarrow \text{Head } NP S'$

We now note that the fact that the head of VP, a verb, may govern an NP and/or an S' is redundantly specified both by (2) and by the Projection

Principle: the Projection Principle will also require a verb that θ -marks an NP or an S' to take such a constituent as a complement.¹ On the other hand, the Projection Principle says nothing about the relative order of NP and S' complements, which is specified in (2). As Stowell discusses at length, however, the relative order of complements of V seems to derive from Case theory. As noted by Chomsky (1981a), Case assignment appears to require string-adjacency in English, a requirement which may be slightly relaxed in languages like French, or relaxed completely in some "non-configurational languages" like Russian:

- (3)a. John ate some meat with relish
 b. ??John ate with relish some meat

Since NP, but not S', requires Case, in order to receive a θ -role from V, it follows that if a verb has both a nominal and a sentential complement, the nominal object will be adjacent to V. (See Stowell for discussion of more complicated cases.)

Thus, given the Projection Principle, which determines the nature of the complements of V, and Case theory, which determines their order, we may simplify (2) to (4):

- (4) VP \longrightarrow Head Complements

Finally, we may ask about the relative order of the head of NP and its complements. Note that in all categories of English, heads precede complements (although specifiers may precede heads). Greenberg (1963; cf. also Huang 1982) notes a general uniformity across categories in the relative ordering of heads and complements in many languages. Thus, a language in which V is last in VP will tend to show P last in PP and

have relative clauses precede N in NP. In English, V is first in VP, P is first in PP, and N is first in NP (modulo specifiers). Thus it looks as if the ordering shown in (4) should be derived from the more general English-specific value for the parameter Head First/Head Last. What is left of (4) is (5):

(5) NP \longrightarrow {Head, Complements}

-- which itself is but one instantiation of the more general X-bar schema :

(6) Phrase \longrightarrow {Head, Complements}

In short, nothing is left of the rule (4), which has decomposed into independent principles of Θ -theory and the Projection Principle, Case Theory, X-bar theory, and the parameter Head First/Head Last.

This derivation raises the question of eliminating entirely the base component of Chomsky (1965) -- a task undertaken in Stowell (1981). If we assume that this is a desirable move, which it surely seems to be, certain problems immediately arise with the rule expanding S given by Chomsky (1981):

(7) S \longrightarrow NP INFL NP

Given X-bar theory, we may eliminate INFL from rule (7), assuming INFL to be the head of S, as we eliminated V from rule (1):

(8) S \longrightarrow NP Head VP

At this point, however, it is not immediately obvious how to proceed. Chomsky (1981a) proposes as a general principle that clauses

have subjects -- let us suppose for the moment that subjects are always NP's. This plus the Projection Principle constitutes the Extended Projection Principle. Nonetheless, this principle by itself does not allow us to eliminate the subject node from (8), because we must still stipulate the relative order of the subject with respect to the other constituents of S.

The ordering problem for subjects arises only in the case of tensed S, however.² In so-called S'-deletion cases -- untensed complements to verbs like consider -- the requirement of adjacency for Case assignment will force the subject to be first in S:

- (9)a. consider [_S John to have left]
- b. *consider [_S to John have left]
- c. *consider [_S to have left John]

Assuming that consider is the only possible Case assigner for John, only in (9)a is John adjacent to the Case assigner. That the Case assigner is to the left of S in (9) follows, of course, from the Head First parameter discussed above, applying in the higher VP.

In tensed sentences, however, where NP receives Case from AGR, a constituent of INFL, we might expect either of two orders (cf. Harlow (1981), who suggests that (10)b obtains in some superficially VSO languages, like Welsh):

- (10)a. NP INFL VP
- b. INFL NP VP

Related to this question is the position of INFL relative to the other constituents. If INFL is the head of S, why do we not find (10)b, rather than (10)a? We suggest that we may answer these questions by assuming a more articulated structure for S:

- (11)a. $S (= INFL'') \longrightarrow NP INFL'$
 b. $INFL' \longrightarrow INFL VP$

$INFL'$ resembles the Predicate Phrase node of Chomsky (1965). Suppose we stipulate that the Projection Principle tells us that the complement of $INFL$ is VP (perhaps also AP , NP and PP , for predicative constructions), then by the Projection Principle and X-bar theory (as in (8)), (11)b reduces to (12) (though this step is slightly ad hoc) :

- (12) $Head' \longrightarrow Head Complement$

We may now place curly brackets around the right side of (12), since the relative order of Head and Complement follows here, as in (4), from the Head First Parameter. Indeed, in (12), we have reduced (11)b to (4).

As for (11)a, where NP requires Case, the relative order of subject and $INFL'$ will derive straightforwardly from the adjacency requirement. Since $INFL$ is first in $INFL'$, only by preceding $INFL'$ will NP be Case-marked. Hence we may reduce (11)a once more to X-bar theory, Case theory, and the Extended Projection Principle (which requires that clauses have subjects):

- (13) $INFL'' \longrightarrow \dots Head' \dots$

But we have now voided our phrase structure rules of all independent content, as desired. Notice that the expansion of S represented in (11) was crucial, since without it we could predict neither the position of the subject nor the position of the head of S .

We now discuss two other consequences of this more articulated structure for S , a structure that will be important to us in future discussion.

1.1 The Definition of Subject

Chomsky (class lectures, 1982) has noted that the assumption of an INFL', or Predicate Phrase constituent of S allows a uniform configurational definition of subject that extends across categories and that applies equally to tensed S, S'-deletion cases, small clauses, and subjects of NPs. A subject is always the sole sister of a Predicate Phrase (cf. Rouveret and Vergnaud 1980 for a similar proposal):

- (14) [_{αN} subject [_α α complements]]
- (15) a. I consider [_{A*} John [_A angry [at Mary]]]
 b. I consider [_{INFL*} John [_{INFL'} to [have arrived]]]
 c. [_{INFL"} John [_{INFL'} INFL [has arrived]]]
 d. [_{NP} John's [_{N'} destruction [of everything holy]]]

The INFL' hypothesis renders transparent the cross-categorial similarities shown in (15) (see Stowell 1981 and forthcoming).

1.2 Contraction

The INFL' hypothesis and the results we can derive from it make an interesting prediction. We derived the S-initial position of the subject from Case theory: subjects which need Case must be adjacent to INFL (in tensed sentences) or to a higher Case assigner (in S'-deletion cases) to receive Case. Notice that this use of Case theory does not in fact derive all the contents of (11)a, since it makes no prediction about the position of a subject that does not need Case. All things being equal, we expect subjects that do not need Case to be able to occur on either side of INFL'.

Such cases fall into two groups. The first group requires a stipulation on our part, which does, however, raise a number of

interesting questions. The second group requires no stipulation and leads to the interesting results we will concentrate on.

We have assumed in preceding discussion that non-NPs -- S's, QPs, etc. -- do not need Case. We predict therefore, that they should be able to occur as post-INFL' subjects in English. This prediction is false:

- (16)a. that John is here shocks me
 b. *shocks me that John is here

But notice that, for mysterious reasons, subject S's act as if they are NPs: they only occur in case-marking positions. This is true even as subjects of small clauses, where a topicalization analysis along the lines of Koster (1978a) seems untenable. While (17)a is very awkward, perhaps ungrammatical (as claimed by Stowell 1981), it contrasts clearly with (17)b, where an S' subject is not in a Case-marking environment:

- (17)a. ??I consider that John is incompetent (to be) a calamity for
 us all
 b. *it was considered that John is incompetent (to be) a calamity
 for us all

Note the contrast with (18):

- (18) it was considered that John is heartless

We may stipulate a special Case filter for subject, requiring all lexical subjects, regardless of category, to be Case marked, but the reasons for this effect remain obscure.³ What is important for us is this: given that all lexical subjects appear to need Case, it is not a surprise that all lexical subjects, including clauses, must occur preverbally, where

they can receive Case.

The second group of cases where our theory falls short of deriving (11)a consists of empty categories that do not need to be in a Case-marking position. Chomsky (1981a, 334) suggests that the Case filter be derived from the θ -criterion in the following way:

- (19) Suppose that the position P is marked with the θ -role R and $C = (\alpha_1, \dots, \alpha_n)$ is a chain. Then C is assigned by R by P if and only if for some i, α_i is in position P and C has Case or is headed by PRO.

The Projection Principle, of course, makes the assignment of θ -role R obligatory at all levels. (19) is assumed by Chomsky to apply at LF; we have given some evidence that it applies at S-structure.

From the Binding Theory it follows that a lexical NP or variable will always head its chain (or else it would be A-bound, and violate principle C). By Burzio's generalization and the Binding Theory, it follows that it will generally be the only element in its chain to be in a Case-marking position.⁴ It follows that a variable, in particular, will always (or almost always) be in a Case-marked position.

On the other hand, an A-bound NP-trace will not head a chain, and hence will not be in a Case-marked position, by the same reasoning. The empty category PRO, by virtue of (19), need not be in a Case-marked position. If Case is assigned only under government, then it will never be in a Case-marking position, since the Binding Theory requires PRO to be ungoverned.

Thus, among the empty categories, only variables must be in a position to which Case is assigned, while PRO and NP-trace do not have

to be, and perhaps may not be in such a position.

If the position of the subject in its clause is determined exclusively by Case theory, a surprising conclusion follows: no principle will determine the position of a subject NP-trace or PRO in S, but variables, like lexical NP's, will be forced by the adjacency requirement on Case assignment to occur to the left of INFL'. In other words, the following paradigm will obtain:

- (20)a. John, who I want [_S t [_{INFL}' to come]]⁵
 b. *John, who I want [_S [_{INFL}' to come] t]
- (21)a. I want [_S' [_S PRO [_{INFL}' to come]]]
 b. I want [_S' [_S [_{INFL}' to come] PRO]]
- (22)a. John is certain [_S t [_{INFL}' to come]]
 b. John is certain [_S [_{INFL}' to come] t]

In (20), t is a variable, \bar{A} -bound by who; hence it must occur to the left of INFL' to receive Case. In (21), PRO needs no Case, and may occur either before INFL' or after. Similarly in (22): t is an NP-trace, A-bound by John, and, needing no Case, may occur on either side of INFL'. Since in each case the subject in question is phonologically null, one might think that no evidence could tell us whether the paradigm of (20)-(22) is correct. Fortunately, this is not the case.

Notice that the examples in (21) and (22) allow the well-known contraction of the matrix predicate with to, while (20) does not:

(20)c. *John, who I wanna come

(21)c. I wanna come

(22)c. John is cert'na come (cf. Emonds 1977, 242, example (7))

For minimal pairs distinguishing NP-traces from variables, cf.:

- (23)a. John was saidda be a fool (said + to)
 b. *John, who I saidda be a fool
- (24)a. there has gotta be an answer (got + to)
 b. *John, who I gotta sing for us⁶

Suppose the position of the subject in S were fixed by a phrase structure rule like the one given in (7) or in (11)a. Sentences like those of (20)-(22) would have only the (a) structures. How then could we distinguish structures like (20) where contraction is impossible, from (21)-(22), where it is possible? Jaeggli (1980a) draws a distinction, adopted also by Chomsky (1980b, 1981a) between Case-marked empty categories (variables) and non-Case-marked empty categories (PRO, NP-trace). According to this solution, an empty category intervenes between the verb and the to which is to be cliticized to it. The intervening empty categories block contraction if they have Case features. Under such a solution, what does the child know when he knows the facts of (20)-(22)c? Clearly the child knows two things:

- (25)a. want + to \longrightarrow wanna, said + to \longrightarrow saidda,
 certain + to \longrightarrow cert'na ...
- b. Case-marked empty categories intervening between the two left-hand terms of (25)a block contraction; non-Case-marked empty categories do not.

As suggested by Aoun (1979), (25)b may be reduced to a more general visibility condition on the Phonetic side of the grammar (PF): an element is visible to a PF rule if it has features that make it visible. Thus,

a Case-marked trace, visible in PF, will block contraction as surely as a lexical NP, since the rules in (25)a will not analyze a string want + to, said + to, etc. A non-Case-marked trace, invisible in PF, will simply not be seen by the rules of (25)a, and contraction will not be blocked.

The Jaeggli-Chomsky-Aoun solution is interesting, because it allows us to maintain the optimal theory about what the child learns about contraction -- namely, the minimum, expressed in (25)a. The restrictions in (25)b clearly must derive from a principle of UG, since there is no plausible way the child can learn these facts from what he hears. Nonetheless, this solution is maintained at the cost of a rather odd universal principle. It is natural that the rules of the different components should differ in the features that they can analyze, but it is unclear why PF should analyze the distinction between Case-marked categories and non-Case-marked categories. This is particularly true if the Case filter reduces to the θ -criterion, as in (19). The Case filter, the one principle which clearly refers to Case features, then applies at S-structure or LF, but presumably not in PF.

Now suppose that the position of the subject in S is not fixed by any phrase structure rule, but rather depends, as we have suggested, on the adjacency condition on Case assignment and on the requirements of the Case filter (19). Under this hypothesis, we may take into consideration both the (a) and (b) structures of (20)-(22). It is immediately apparent that we may dispense with the visibility hypothesis of Aoun (1979), and with the principle (25)b derivable from it. We may suppose that PF analyzes all empty categories. Structures (20)a, (21)a and

(22)a all block contraction, because a category (by chance, empty) intervenes between the contracting elements. Structures (20)b, (21)b and (22)b, however, do not block contraction, because the contracting elements are, as required, adjacent. The subject of the lower clause is to the right of INFL'. In (20)b, it happens that the empty category, which must receive Case, cannot, by virtue of the adjacency requirement on Case assignment. Therefore, the only structure in (20) to which contraction can apply is ruled out by the Case filter. In other words, the only empty categories that do not block contraction are the ones that don't have to be present at the contraction site: clearly the simplest hypothesis.⁷

Lightfoot (1977) raises one remaining question with respect to contraction. We have been assuming that the complement of want is a reduced S, like the complement of consider (contra Chomsky and Lasnik 1977, Chomsky 1981a, see note 5). This was not crucial to our argument: so long as the subject of the complement to want can only get Case from a left-hand Case assigner (an abstract complementizer, Case assignment by want across S', etc.), our argument stands. Note that if the complement of want is an S, and not an S', it will lack a COMP. Suppose, however, that it is an S', and has a COMP. Then a problem for our analysis may be posed by cases where movement is assumed through COMP, as in (26):

(26) who do you want [_S, [_{COMP}] [_S PRO to visit t]]

If COMP contains a trace of who, given the optimal hypothesis that all empty categories block contraction, contraction should be blocked. But

it is not:

(27) who do you wanna visit

But recall from our discussion in Chapter 2 that trace of an object in COMP is the one trace that is required to exist neither by the Projection Principle, nor by the Binding Theory, nor by the ECP. If we suppose, as we did, that traces are obliged to occur only when independently mandated by some principle of grammar, then no trace need be assumed in the COMP of (26), and the possibility of contraction raises no problem. Alternatively, we might appeal, with Chomsky and Lasnik (1977) to some rule freely deleting elements in COMP to allow this case. As indicated, we prefer to analyze (26) as not containing a lower COMP node at all. In any case, it is clear that such examples do not pose a problem.

As noted by Postal and Pullum (1978), and discussed by Jaeggli (all following Selkirk 1972), liaison in French, like contraction in English, appears to be blocked by variables, but not by NP-trace or by PRO (+ indicates liaison):

(28)a. *la somme que nous donnerons + à l'OXFAM

the sum that we shall give to OXFAM

b. nous voulons + y aller

we want there to go

c. Jeanne paraît + aimer les huîtres

Jeanne seems to like the oysters

Whether liaison involves optional non-deletion of a final consonant before a vowel, or the optional insertion of such a consonant (cf.

Rotenberg 1978), we may account for (28) by assuming that an empty category intervening in the potential liaison environment is seen by the rule of liaison, and does not count as a vowel. The paradigm of (28) is structurally identical to that of (20)-(22) and admits the same explanation: only categories that are there block liaison.⁸

It is interesting to compare our proposal with the Jaeggli-Chomsky-Aoun proposal. Both proposals rely crucially on Case theory: the appropriate distinction between contraction-blocking empty categories and contraction-permitting categories is the distinction between categories that need Case and those that don't (but cf. note 7). In the earlier proposal, however, this distinction is made a postulate: Case-marked traces block contraction because of a specific statement that has this result. Incorporating the approach to phrase structure of Chomsky (1981b) and Stowell (1981), however, under which the order of constituents itself derives from Case theory, we are able to preserve the distinction made by the earlier theory of contraction without making it a primitive.

Note, however, that this approach to phrase structure extends to the node S only if we assume the articulated structure shown in (11). Without such a structure, we cannot predict the position of the subject, and we lose our solution to the contraction problem. The contraction phenomena thus provide strong evidence for this structure, and the structure in turn will be important to us in future sections (cf. 4.4 of this chapter).

2. Remarks on the ECP

A major role in our discussion of Russian was played by the ECP, as introduced by Chomsky (1981a). We repeat below the relevant principles and definitions:

- (1) ECP: [α e] must be properly governed
- (2) α properly governs β if and only if α governs β [and $\alpha \neq \text{AGR}$]
- (3) Consider the structure (i):
 - (i) [$\beta \dots \gamma \dots \alpha \dots \gamma \dots$], where
 - (a) $\alpha = X^0$ or is coindexed with γ
 - (b) where ϕ is a maximal projection, if ϕ dominates γ then ϕ dominates α
 - (c) α c-commands γ

In this case, α governs γ . (Chomsky 1981a, 280)

As has been often noted (e.g. Kayne 1981a), the ECP and the definitions of government and proper government necessary to make it work have a number of odd features. At least three questions may be raised:

1. Why is AGR not a proper governor? Chomsky suggests that AGR is not "lexical", and that only lexical governors are proper governors. But why should the ECP "tighten" the definition of government in this way. Picallo (1982) suggests that AGR might be a proper governor in some languages (a suggestion to which we return in Chapter Five), but it still remains a mystery why there should be any distinction between government and "proper" government.

2. Why is the definition of government satisfied by either one or two apparently unlike configurations? An empty category may be lexically governed, for example by V, and lack an antecedent entirely;

or else it may lack a lexical governor, if it is locally coindexed with an antecedent. This question is first raised by Kayne (1981a) and discussed again in Stowell (1980, 1981). Kayne suggests that the notion of antecedent should be the primary notion of the ECP, and that the distinctions between lexical and non-lexical government made by Chomsky's ECP limit the ability of an empty category to be linked to its antecedent. Stowell similarly suggests that antecedency is the central notion behind the ECP, but suggests generalizing the notion "antecedent" with the notion of government (more exactly, θ -role assignment), so that V itself acts as an antecedent for an object trace. We will return to Kayne's proposal in Chapter Four, and to Stowell's in the last section of this chapter.

3. Finally, why should there be an ECP in the first place? More generally, why should empty categories have to be in a particular relation with other elements in a phrase marker? Chomsky (1981a) suggests that the ECP is closely linked to the Recoverability condition on deletions. Recently, Chomsky (1981b) extends this observation, suggesting that conditions like the ECP and the Binding Theory are means of "identifying" empty categories (cf. Jaeggli 1980b on "identification"). We saw in Chapter Two that the ECP, the Binding Theory, and the Projection Principle conspire to determine where empty categories are located in a phrase marker. Chomsky (1981a, b) note that the ECP and Binding Theory establish a near partition of the possible location of empty categories, forming a "Generalized Empty Category Principle" that determines what kind of Empty Category is to be found in each position. Roughly, traces must be properly governed, and PRO must be ungoverned. We have suggested in Chapter Two that this near partition be made a full

partition: expletive empty categories exist which are governed but not properly governed (a stipulation we will derive in Chapter Four). Notice, however, that this addition to the Generalized ECP does not follow from any other principle. Rather, we built it into our statement of the ECP in Chapter Two. Assuming our arguments were correct, we are left with two questions: (a) why should expletive and non-expletive empty categories occupy different places in the partition of empty categories? (b) why is the partition what it is? Question (b), of course, is a more general version of our original question: why is there an ECP?

Ultimately, questions like "why is there an ECP?" are questions about the evolution of the language faculty in man. More immediately, however, we may answer questions like (3) by reducing them to other questions. Recent work by Brody (1982) and by Aoun (1982), for example, seeks to derive the ECP from modified versions of the Binding Theory. They would claim, therefore, that the question "why is there an ECP?" reduces to the question "why is there a Binding Theory?"

In the following sections, I suggest a different reduction of question (3), which will answer question (1) and shed some light on question (2). In essence, I will unify the "ECP effect" with the phenomena that have been attributed to constraints on the nesting of dependencies, and extend this account to the Subject Condition of Chomsky (1973).

A word of caution: The ECP of Chomsky (1981a) was motivated to account for certain ungrammatical cases of \bar{A} -binding: "Complementizer-Trace Phenomena" (henceforth CTP) and allied phenomena in LF, that we discussed in Chapter Two. It is this aspect of the ECP that we shall

derive. Chomsky (1981a) shows that certain cases of ungrammatical A-binding (or chain formation) also fall under the ECP. We shall not account for these phenomena here, but will have some suggestions to make in Chapter Five. In this respect, then, our theory will fall short of Chomsky's ECP.

We begin our discussion with "Crossing" or "Nested Dependency" effects, and then show, at an informal level, how they can be generalized with CTP effects. We then develop the theory that accomplishes this.

3. Crossing and Nesting

3.1 Crossing Effects

Sentences like (29) are often taken to be ungrammatical, as a result of the so-called "WH-island" effect, subsumed by the Subjacency Condition in Chomsky (1977) and elsewhere:

(29) what subject_i do you know [_S who_j [_S PRO to talk to t_j about t_i]]

For many speakers of English (as well as Italian (Rizzi 1978), French (Sportiche 1981), Russian (Sinicyn 1982a)), WH-island violations like (29) are almost completely acceptable, particularly when the WH-island is infinitival. Nonetheless, it has been frequently noted that some WH-island violations are sharply worse than (29):

(30) *who_j do you know [_S what subject_i [_S PRO to talk to t_j about t_i]]

Strictly speaking, (30) is ungrammatical as a violation of selectional restrictions. It has a nonsensical interpretation in which the indices on the traces are reversed, and one is talking to a subject about a person. What is unavailable is the reading indicated by the indices in (30).

It has been suggested repeatedly (first, perhaps, by Kuno and Robinson 1972, who credit the observation to A. Berman) that the linear sequence of WHs and extraction sites is what distinguishes acceptable from ungrammatical WH-island violations. In particular, it has been claimed that when a phrase contains more than one "dependency" between two elements, like the WH-trace dependencies in (29)-(30), these dependencies must be nested, and not intersect. Thus, (31)a represents the dependencies in (29); (31)b, the dependencies in (30). The dependencies are shown by lines in (29)'-(30)':

(31)a. $WH_i \dots WH_j \dots t_j \dots t_i$

b. $*WH_i \dots WH_j \dots t_i \dots t_j$

(29)' what subject do you know [_S, who [_S PRO to talk [to t] [about t]]]

(30)' *who do you know [_S, what subject [_S PRO to talk [to t] [about t]]]

In support of this position, we note that judgments on (29) and (30) reverse when the order of the PPs is reversed (cf. Culicover and Wexler 1977, 27):

(32) *what subject do you know [_S, who [_S PRO to talk [about t] [to t]]]

(33) who do you know [_S, what subject [_S PRO to talk [about t] [to t]]]

In each of the bad sentences, the dependencies cross, when we examine the linear string at S-structure.

"Crossing effects" of this sort recur with every construction that Chomsky (1977a) argues to involve WH-movement:

Questions (additional examples)

(34)a. what books do you know [_S, who [_S PRO to persuade t [_S, PRO to read t]]]

b. *who do you know [_S, what books [_S PRO to persuade t [_S, PRO to read t]]]

Relative Clauses

(35)a. chess, which I wonder [_S,who [_Syou believe [_St to play t well]]]

b.*John, who I wonder [_S,what game [_Syou believe [_St to play t well]]]

Topicalization

(36)a. this problem, Mary knows [_S,who [PRO to consult t [about t]]]

b.*this specialist, Mary knows [_S,what problems[PRO to consult t [about t]]]

Infinitival Relatives

(37)a. I finally found a subject [_S,[PRO to ask [_S,who to talk [to t] [about t]]]

b.*I finally found a subject [_S,[PRO to ask [_S,who to talk [about t] [to t]]]

Tough Movement

(38)a. what balalaika are these partitas easy [_S,[_SPRO to play t [on t]]]

b.*what partitas is this balalaika easy [_S,[_SPRO to play t [on t]]]

(39)a. this instrument is difficult to guess [_S,what piece [PRO to play t [on t]]]

b.*this piece is difficult to guess [_S,what instrument [PRO to play t [on t]]]

(40)a. this book is impossible to find out [_S,who [to persuade t [to read t]]]

b.*Sue is impossible to find out [_S,what book [to persuade t [to read t]]]

Too Movement (Enough Movement)

(41)a. what viola are these concerti too dissonant [_S,[PRO to play t [on t]]]

b.*which concerti is this viola too fragile [_S,[PRO to play t [on t]]]

(42)a. these bonds, which he is naive enough [for [us to persuade t [to buy t]]]

b. *John, who these bonds are attractive enough [for [us to persuade t [to buy t]]]

Comparatives

(43)a. he wrote more articles [than I could imagine [_S, who [to ask t [to read t]]]

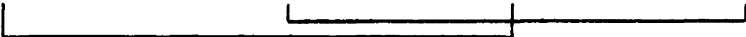
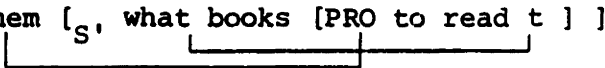
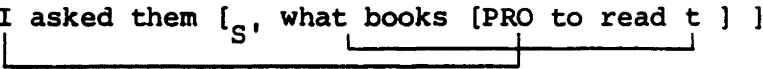
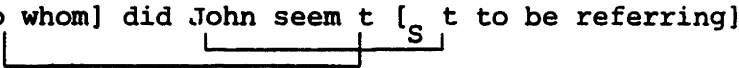
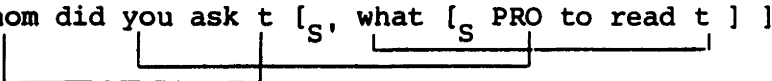
b. *more people came [than I could imagine [_S, what articles [to ask t [to read t]]]

While many of the (a) sentences are very awkward, perhaps even ungrammatical, the (b) sentences, to my ears, are completely incomprehensible. The paradigm demonstrated in (29)-(43) may be duplicated in French and in Italian, and presumably in a number of languages (but see our discussion of Dutch below).

We have ended our lines in a number of the diagrams above in COMP (for example, the upper line in (38)a), rather than at the lexical NP apparently binding the trace, in keeping with the analysis of Chomsky (1977a). Chomsky notes that the constructions listed above form a "family", having a number of properties in common: they leave "gaps"; they allow long distance movements, in apparent violation of the subjacency constraint; they observe the Complex NP Constraint of Ross (1967); they observe (weakly, we claim) WH-island constraints. Chomsky (1981b) notes that all these constructions also allow "parasitic gaps", a phenomenon to which we shall return. In the framework of Chomsky (1981a) we may characterize the common feature of these constructions rather simply: the gap in a θ -position is locally \bar{A} -bound. Apparent violations of subjacency are due to movement through COMP, an \bar{A} -position; the possibility of parasitic gaps also follows from the assumption that

these constructions involve \bar{A} -binding. In other words, each of the COMPs at which we ended our lines in (29)-(43) contains an operator locally binding a trace.

Chomsky (1976) presents a number of counterexamples to a constraint prohibiting crossing dependencies (we omit an example involving Czech clitics):

- (44)a. what books have those men written t [about each other]

- b. I told them [_S, what books [PRO to read t]]

- c. I asked them [_S, what books [PRO to read t]]

- d. [to whom] did John seem t [_S t to be referring]

- e. whom did you ask t [_S, what [_S PRO to read t]]


Notice that in each case in which two dependencies cross in (44), one of the dependencies involves A-binding. On the other hand, all the crossing dependencies in (29)-(43) which do create violations involve \bar{A} -binding.¹¹

Let us call the lines we have drawn to indicate dependencies "dependency lines", borrowing here from Fodor (1978; cf. our discussion below). A descriptively adequate Crossing Constraint must limit itself to \bar{A} -dependency lines, if it is to capture the facts of (29)-(43), without mistakenly assigning stars in (44):

(45) Crossing Constraint

\bar{A} -dependency lines may not cross.

Note that (45), as stated, is a linear constraint; that is, it is not structure-dependent. We will be presenting many arguments to show that a

constraint like (45) is wrong, and we will replace it by a structure-dependent principle shortly. For the moment, however, (45) suits our purposes. Even with (45), we may show how such a Crossing Constraint can explain Complementizer-Trace Phenomena, a central empirical motivation for the ECP.

3.2 Complementizer-Trace Phenomena

The surface order of the major constituents of S' and S in English are given in (46). The structural organization of these constituents is not important yet:


(46) COMP NP INFL VP

It is well-known that there is a special relationship between COMP and INFL. For example, Chomsky (1981a), assuming that COMP expands to include +WH, -WH (that) or for (p. 53), notes (p. 305) the necessity of conventions having the following effect:

- (47)a. INFL is [+Tense] only if S' contains COMP
- b. INFL is [+Tense] if COMP is [-WH]
- c. INFL is [-Tense] if COMP is for

Den Besten (1978) and Stowell (1981) both argue that the relationship between COMP and INFL is the result of movement of the tense feature from INFL to COMP (Stowell) or from COMP to INFL (Den Besten). We will discuss the details of this relationship in detail later. For now, suppose there exists a binding relationship created by movement between some element in INFL and COMP. Since COMP is an \bar{A} -position, this

binding relationship is one of \bar{A} -binding. Therefore, an \bar{A} -dependency exists between COMP -- at least when tensed -- and INFL:

(48) COMP NP INFL VP


If (48) is right, we expect the relationship between COMP and INFL to obey the Crossing Constraint (45). We will now demonstrate that Complementizer-Trace Phenomena (CTP) derive from this interaction. We have briefly discussed the CTP in Chapter Two. The most basic contrast is that of (49):

- (49)a. who_i did John say [_S [_{COMP} t_i that] [_S Mary liked t_i]]
 b. who_i did John say [_S [_{COMP} t_i] [_S Mary liked t_i]]
 c. * who_i did John say [_S" [_{COMP} t_i that] [_S t_i liked Mary]]
 d. who_i did John say [_S [_{COMP} t_i] [_S t_i liked Mary]]

Recall how this paradigm is explained by means of the ECP of (1). The object trace in (49)a satisfies the ECP, since it is properly governed by the verb liked. Similarly in (49)b. The subject trace in (49)d is properly governed by the coindexed trace in COMP. In (49)c, however, the presence of that somehow prevents the trace in COMP from properly governing the subject trace. In Pesetsky (1982), where the paradigm of (49)' was explained using the Nominative Island Condition (NIC), instead of the ECP, we suggested that a version of Chomsky and Lasnik's filter barring doubly filled COMPs requires the trace in COMP in (49)a and (49)c to delete at S-structure:

(50) * [_{COMP} α β]

Thus, at LF, (49)c has the structure of (51) :

(51) who_i did John say [_S [_{COMP} that] [_S t_i liked Mary]]

In (51), the subject trace is not properly governed, and thus violates the ECP. Alternatively, we might assume, as in much recent work (cf. Kayne 1981a), that the trace in (49)c is not required to delete, but that it nonetheless does not properly govern the subject trace because it fails to c-command it. For simplicity of diagramming, we will assume the hypothesis of Pesetsky (1982) -- that is, we will not write traces in COMP that do not properly govern the subject position, as if they were deleted.

Let us now see how the Crossing Constraint, combined with the same assumptions about binding from COMP as we have just made, can account for paradigm (49). In (52), we give a fuller structure for (49)a-d, and we mark the \bar{A} -dependencies. We assume as in (48) that there is an \bar{A} -dependency between COMP and INFL. For ease of diagramming we end our dependency lines from INFL at the lexical complementizer, but we assume that the actual relation is with the COMP node :

(52)a. who did John say [_S [_{COMP} that] [_S Mary [INFL liked t]]]

b. who did John say [_S [_{COMP} t] [_S Mary [INFL liked t]]]

c. * who did John say [_S [_{COMP} that] [_S t [INFL liked Mary]]]

d. who did John say [_S [_{COMP} t] [_S t [INFL liked Mary]]]

In (52)a-b we see that the \bar{A} -dependency between a WH and an object will never cross the dependency between COMP and INFL: the dependencies will always nest. In (52)c-d, however, we see that an \bar{A} -dependency

between a WH and the subject of a tensed sentence will cross the dependency between COMP and INFL of that sentence unless the subject is bound from that sentence's COMP.

Recall that Rizzi (1982) (also Kayne 1980a) showed that Italian lacks an apparent CTP effect with WH-movement because of the possibility of free inversion of the subject. Notice that if the CTP is derived from the Crossing Constraint, as suggested here, Italian presents no problems. An \bar{A} -dependency between COMP and a subject inverted to the right of VP will never cross the dependency between COMP and INFL: the situation will be parallel to (52)a-b.

Also note that various language-specific "escapes" from the CTP will be analyzable identically under a Crossing account and under an ECP account, just as NIC accounts of these escapes carried over to ECP accounts. For example, the change of complementizer que to qui in French allows long movement of a subject, as discussed in Pesetsky (1982) (cf. also Kayne (1981c)) :

- (53)a. *qui est-ce que tu crois [_S, [_{COMP} qu'] [_S t est venu]]
 'who do you believe that has come'
 b. qui est-ce que tu crois [_S, [_{COMP} qui] [_S t est venu]]

We assume, following our previous analysis, that a special rule copies the index of a WH or trace in COMP onto the complementizer que, yielding an indexed allomorph qui, when WH or trace locally binds a nominative position:

- (54) [_{COMP}^{α_i} que] → [_{COMP}^{α_i} qui_i], where α_i locally binds [_e]¹²
 +NOM

Given (54), (53)b will have the same status as (52)d, while (53)a will have the same status as (52)c:

(53)a'. * qui_i est-ce que tu crois [_S, [_{COMP} qu'] [_S t_i [INFL est venu]]]

b. qui_i est-ce que tu crois [_S, [_{COMP} qui_i] [_S t_i [INFL est venu]]]

Note that under the ECP analysis, the indexed qui properly governs the subject trace, while unindexed que does not, accounting for the contrast.

Similarly, we may assume an indexed complementizer that in simple "short" relative clauses like (55) in English. As noted in our earlier paper, the English rule, unlike the French rule, must be restricted to copy the index from WH, but not trace of WH, onto that. If the English rule were like the French rule, we would have no CTP effects at all, since that would always be a proper governor:¹³

(55) [WH_i that] → [WH_i $that_i$]

(56) the man [_S, $that_i$ [_S t_i [INFL came]]]

Finally, note that if we use the Crossing Constraint instead of the ECP to explain the CTP, examples like (57)a are ruled out in two ways by the Crossing Constraint (cf. remarks in Chomsky 1980a, fn. 43 on this point). (57)b is, correctly, not excluded:

(57)a. *who did you wonder [_S, what [_S t [INFL saw t]]]

b. what did you wonder [_S, who [_S t [INFL saw t]]]

In (57)a, the dependency between who and its trace crosses both the dependency between COMP and INFL and the dependency between what and its

trace, although the COMP-INFL dependency and the what-t dependency do not cross each other (cf. what did John see).

Thus the Crossing Constraint can be extended naturally to cover the central cases for which the ECP has been motivated. This raises the question of eliminating entirely the ECP for \bar{A} -binding, and with it the notion of proper government, which is only relevant to the ECP for \bar{A} -binding.

Nonetheless, we will show in the next sections that the Crossing Constraint, as a linear constraint, is inadequate, and -- in some languages -- empirically incorrect. We will propose a structure-dependent version of the Crossing Constraint, replacing the one-dimensional notion "dependency" with a two-dimensional notion "path", similar to the concept developed in Kayne's recent work. We will use the path analog of the Crossing Constraint to explain a number of phenomena which fall under neither the Crossing Constraints of (45) nor the ECP of (1), while preserving the basic features of our analysis so far.

4.0 Crossing Dependencies vs. Paths

4.1 Crossing in Two Dimensions

Suppose we rephrase the Crossing Constraint as (58):

(58) Crossing Constraint

The further to the right an \bar{A} -bound trace occurs, the further to the left its binder must occur, in relation to other trace- \bar{A} -binder dependencies.

(45) or (58) have been interpreted as plausible constraints on language parsing. For example, if the parser seeks to match "fillers", such as

WH, with appropriate "gaps", such as traces, and if this mechanism has the properties of a "push-down store device", then it might follow, as Bach (1977, 150) suggests, that the last filler passed by the parser is always linked to the first gap passed, the second to last filler to the second gap, and so on, given the "last in - first out" operation of a push-down store device.

Suppose (58) does conform to the nature of the human parser, viewed as having push-down store properties. We must still ask whether (58) is a principle of linguistic competence, which has evolved to ease the burden on the parser, or a principle of performance, directly expressing limitations on the parser. In other words, are the starred sentences of the previous sentence ungrammatical, or merely unacceptable, like cases of center embeddings discussed by Miller and Chomsky (1963) and by Chomsky (1965)?

While firm answers cannot be given, it seems that the starred sentences of the previous section are ungrammatical, and not merely unacceptable. This is clear in the case of the CTP violations which we showed to fall under the Crossing Constraint. Cases of unacceptability generally improve in felicitous circumstances: if one is provided with pencil and paper to figure out the links, if the meaning is made clear by context, etc. The meaning of CTP violations is perfectly clear without pencil and paper; nonetheless the sentences resist acceptance. The crossing sentences of (29)-(43) present a similar picture. Here, the complexity of the examples may require drawing lines of dependency, as we have. Nonetheless, even after the intended meaning of the sentences is clear, the sentences remain impossible to accept, in

mark 1 contrast to garden path and center-embedding sentences (the horse raced past the barn fell, and the like). Thus, it is extremely probable that Crossing effects and CTP effects derive from principles of competence, not of performance.

It follows that if the Crossing Constraint can be linked to ease of parsing -- if the parser has some push-down store property -- the linkage is at the level of human evolution. It is not surprising if man's linguistic competence has molded itself to the limitations and needs of the parser. (Suggestions to this effect regarding the Binding Theory and subadjacency are made by Marcus 1980; Berwick and Weinberg 1982.)

Nonetheless, Fodor (1978) notes that there is considerable evidence suggesting that the human parser does not have the characteristics of a push-down store device. Discussing her Crossing Constraint (a "Nested Dependencies Constraint"), she suggests that such a Constraint arises as an attempt to eliminate potential ambiguities in the proper linkage of gaps to fillers. Apart from the fact that it is extremely difficult to characterize what counts as an ambiguity (if the parser knows that one cannot talk to a subject about a person, sentences like (30) should be unambiguous), it is unclear why the parser should pick a constraint against crossing as a means of avoiding ambiguity, as opposed to any other constraint. Fodor offers a number of arguments that the constraint is related to parsing, some of which¹⁴ may reflect other parsing-related constraints, but not necessarily the Crossing Constraint, others of which merely reflect the contrast between A- and \bar{A} -binding discussed above. The main point is, however, that once one rejects the "push-down store" explanation for the Crossing Constraint, the reason for this

constraint is up for grabs. Any pattern of ungrammatical sentences can be described as a parsing constraint, but this does not obviate the need for an explanation of why the parser, or grammar, should have the properties it has. Why not favor crossing over nesting? In fact, we shall see later that some languages do just that.

Finally, a parsing explanation is inevitably both too strong and too weak. Why should the Crossing filter ignore some dependencies, those involving A-binding, which are clearly of relevance to the parser, while paying attention to dependencies that are of dubious relevance to the parser, such as the dependency between INFL and COMP?

Suppose, then, that we abandon (perhaps temporarily) the search for an explanation of Crossing effects in terms of parsing strategies, and concentrate on the status of (58) as a principle of competence: a principle of grammar. Recall that crossing effects hold, not only in English, but in French and Italian as well. The sentences below, for example, are the equivalents of (34)a-b:

(34)a' quels livres est-ce que tu sais [_S qui [_S PRO persuader t [_S de PRO lire t]]]

b' *qui est-ce que tu sais [_S quels livres [_S PRO persuader t [_S de PRO lire t]]]

(34)a". che libri sai [_S chi [_S PRO persuadere t [_S da PRO leggere t]]]

b". *chi sai [_S che libri [_S PRO persuadere t [_S da PRO leggere t]]]

English, French and Italian are all "Head First" languages: V precedes its complements, P precedes its object, INFL precedes VP, etc. Notice now that a Head First language will be predominantly right-branching,

with respect to major constituents, at least. This is because branching complements (e.g. PP or S' governed by V) will occur to the right of the head of their phrase. In a Head First language, the farther to the right a node is, the "lower" it is in the tree, as a rule. (The generalization may be violated as the result of rightward movement rules, or in double complement structures where the first complement branches farther than the second.)

This observation raises the question of whether the Crossing Constraint (58) should be phrased in linear terms of left and right, or in two-dimensional terms of "higher" and "lower" in a tree. This issue is raised by Reinhart (1981), who cites unpublished work by J. Maling and A. Zaenen. We might state such a two-dimensional constraint informally as in (59):

(59) Crossing Constraint - 2 Dimensions (informal)

The lower an \bar{A} -bound trace occurs in a tree, the higher its binder must occur, in relation to other trace- \bar{A} -binder dependencies.

For the time being, we assume a simple notion of "higher" and "lower", counting nodes from the root: the more nodes, the lower.

An immediate problem now arises for (59) in the contrast (29)-(30) and (32)-(33). In these sentences we find crossing effects generated by two traces, each of which is the object of a preposition in PP governed by V. Since each trace is separated by one node (PP) from the VP in which the PP is governed, they are equally distant from the root node, and there should be no crossing effects.

This problem may be solved by adopting a proposal from Hornstein

and Weinberg (1981), though we abandon part of their motivation for this proposal. They claim that a reanalysis rule exists in English, which can incorporate post-verbal material into the verb, creating a "complex verb", which functions as a single lexical item. They argue that a surface filter barring oblique traces forces this reanalysis in cases both of \bar{A} -binding and of A-binding. The absence of the reanalysis rule in languages like French or Russian, they suggest, explains the absence of preposition stranding in these languages. Crucial to their theory is the Case system of Chomsky (1980), under which V assigns objective Case, while P assigns oblique. As Hornstein and Weinberg note (their fn. 9), this proposal may create problems in an inflected language like Russian, where, as we have seen, both verbs and prepositions may require oblique Case of their objects, yet "verb stranding" (under \bar{A} -binding, at least) is always possible, while preposition stranding is impossible.

Notice that reanalysis of V with P will be motivated by Burzio's generalization when the object of a preposition is A-bound.

(60) John_i [was [_V looked at] t_i]

In (60), the subject of look must be a non- θ -position, in order to allow the chain (John, t), under the θ -criterion. By Burzio's generalization, then, the verb must absorb its ability to assign Case to an argument object. As we saw in the last chapter, this is impossible in English when the verb does not assign Case to begin with. Suppose that a verb, reanalyzed or not, of the form [_V V...X], can assign Case to an object argument if and only if X is a Case assigner (cf. Williams' (1981) claim that the "head of a word" is the rightmost constituent -- in a syntactic Head First language, presumably). Thus, while look will not

be a Case assigner, and thus cannot passivize, the complex verb [_V look at] will be a Case assigner, and can passivize. Hence the reanalysis.

The obligatoriness of reanalysis of V with following material under A-binding is supported by other evidence. R. Higgins first noted that the string between V and a trace A-bound by the subject of V is an "anaphoric island" for \bar{A} -binding (in the sense of Postal 1969; cf. Stowell 1981):

- (61)a. we paid close attention to John
 b. how much attention_i did you pay t_i to John
 c. who_i did you pay close attention to t_i
 d. close attention_i was paid t_i to John
 e. John_i was paid close attention to t_i
 f. who_j was close attention_i paid t_i to t_j
 g. *how much attention_j was John_i paid t_j to t_i¹⁵

If (61)e is only possible when the sequence V-NP-P is reanalyzed, we can explain the ungrammaticality of (61)g as the result of binding into a word:

- (61)e'. John_i [was [_V paid close attention to] t_i]
 g'. *how much attention_i [John_j [was [_V paid t_i to] t_j]]

Of course, we can rule out (61)g if only the NP t_i is analyzed. Suppose that NP-movement has taken place, as in (61)d-g, into the subject position of pay. It follows that the subject is non- θ , from which it follows (by Burzio's generalization) that the verb cannot assign Case to an argument: in this case, the NP close attention, how much attention. Suppose this argument itself moved into subject position; we would have (61)d. Suppose, however, that the object of to moves into subject

position, as in (61)e,g. Then, in order for the Case Filter (19) not to rule out the argument close attention, how much attention, this argument may be incorporated inside the verb itself. Once this occurs, it follows that it may not be extracted, hence the star on (61)g.

We know that the P must be reanalyzed as well, by the logic of two paragraphs ago. A reanalyzed sequence [_V V NP] does not end in a Case assigner; hence it will not assign Case; hence it cannot bear passive morphology. If we add the P to the reanalyzed verb, however, the verb will end in a Case assigner, and will be able to bear passive morphology. Note that adverbial material may not precede the P in sentences like (61)e, another piece of evidence for reanalysis:

- (62)a. ?we pay close attention usually to John
 b. *John is paid close attention usually to

Thus, given apparently correct assumptions about passive and reanalysis, the obligatoriness of reanalysis in the passive construction follows. We now ask if any independent principles we have discussed can make reanalysis obligatory in cases of \bar{A} -binding. It seems that the answer at the moment is no, if we reject Hornstein and Weinberg's oblique Case filter. Consider again (29):

(29) what subject_i do you know [_S who_j [_S PRO to talk to t_j about t_i]]

Suppose reanalysis were obligatory with preposition stranding under \bar{A} -binding as a general rule (Hornstein and Weinberg's, as well as Stowell's and Kayne's (1981a) claim). Then (29) would perforce have the structure of (63):

(63) what subject_i do you know [_S who_j [_S PRO to [_V talk to t_j about] t_i]]]¹⁶

Structure (63) is motivated by no considerations we have discussed, and furthermore violates the very anaphoric island condition that explained (61)g for us. Hence, *prima facie*, not only is reanalysis not forced in cases of \bar{A} preposition stranding, but in some cases, like (63), it must not apply at all.

Additional cases, that would involve violations of the anaphoric island principle by A-binding, if reanalysis were obligatory for \bar{A} preposition stranding, are credited by Hornstein and Weinberg to H. van Riemsdijk and E. Williams (fn. 21):

- (64)a. which problems_i has Harry_j [been [talked to t_j about t_i]]
 b. who_i do you_j like [PRO_j to [be [sung to t_j by t_i]]]

We find these completely acceptable, though they may be somewhat awkward, as suggested by Hornstein and Weinberg. Again, were (64)a-b to involve complex verbs [_V talked to t_j about] and [_V sung to t_j by], plausible anaphoric island constraints on binding into a word would be violated.

Finally, to our ears, (65) has the status of (62)a, and not (62)b :

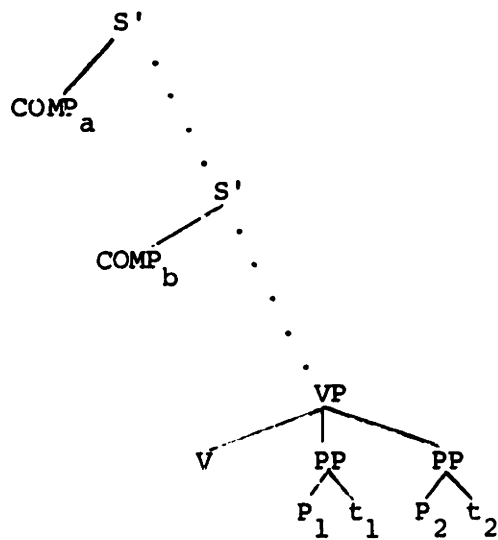
(65) ?John, who_i we pay close attention usually to t_i

If this is correct (and there is much variation on examples like this), then we have more evidence that reanalysis is not required for \bar{A} preposition stranding. For the time being, we leave open the question of what rules out preposition stranding -- A or \bar{A} -- in languages like French, Italian and Russian.

Despite this demonstration that reanalysis is not always required in \bar{A} preposition stranding constructions in English, it does not follow that it is impossible. The facts about passive discussed above show that English does have a rule of reanalysis. This rule is formally autonomous from the other processes involved in the passive construction (dethematization, absorption of Case, NP-movement, etc.). There is no reason why it should be prohibited from applying in cases of \bar{A} preposition stranding, where other constraints are not violated. We may even find cases where other principles require this reanalysis, much as reanalysis was required by Burzio's generalization in the case of passive. So far, we have no such principles.

Recall our problem. We suggested that the linear Crossing Constraint of (45) and (58) might be replaced by a two dimensional, structure-dependent Crossing Constraint, informally presented in (59). This constraint required that \bar{A} -bound traces lower in the tree be bound higher in the tree than \bar{A} -bound traces higher up the tree are bound. We defined "lower" and "higher" as functions of the number of nodes from the root. From this it followed, incorrectly, that there should be no Crossing effects when two prepositions in VP are \bar{A} -stranded (cf. (29)-(30), (32)-(33)). That is, the two traces in the following tree are equidistant from the root:

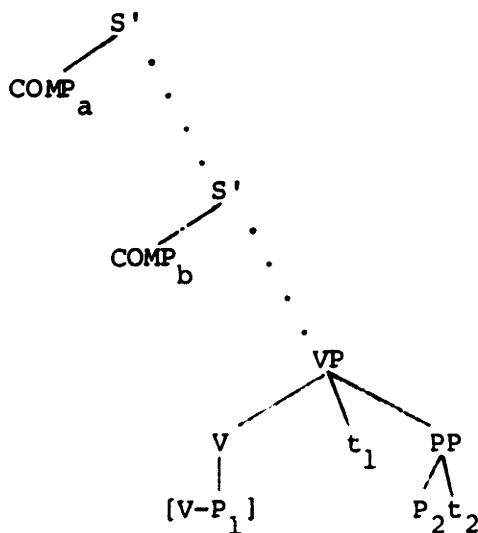
(66)



Given the two-dimensional Crossing Constraint, we should be able to bind t_1 or t_2 from either $COMP_a$ or $COMP_b$, while in fact t_1 must be bound from $COMP_b$ and t_2 from $COMP_a$.

Now suppose some principle required P_1 to reanalyze with V . (Notice that P_2 could not reanalyze with V , or else anaphoric island principles would be violated.) Then (66) would look like (67):

(67)



Now t_1 is 3 nodes from the root in our diagram, while t_2 is 4 nodes from the root. t_2 is thus lower than t_1 and must be bound higher than t_1 is. As desired, t_2 must be bound from $COMP_a$ (which is 1 node from the root),

and t_1 must be bound from $COMP_b$ (2 nodes from the root). Thus, if reanalysis of P_1 with V is obligatory, we capture the Crossing effects, given the two-dimensional Crossing Constraint of (59).

What principle forces reanalysis in (67)? And why should Universal Grammar contain a principle which requires node counting, a rather odd principle to assume? We shall suggest that the answers to both these questions lie in a deeper, simpler version of the Crossing Constraint in two dimensions. As things stand, the linear version of the constraint seems simpler and more plausible. We shall be able to demonstrate that the simpler version of the two-dimensional constraint is actually more plausible and a more nearly correct principle of Universal Grammar.

Notice first that, as stated, both (58) and (59) (as opposed to our original statement in (45)) are empirically inadequate in a trivial way. Consider a structure like (68), which is not ungrammatical:

(68) [who_i [you tell t_i [_S, what_j [_S PRO to do t_j]]]]

In (68), t_j is lower and farther to the right than t_i . It is, however, bound farther to the right and lower than t_i is bound. To avoid ruling (68) out with a two-dimensional Crossing Constraint, we must be more specific about the conditions under which the constraint applies. In particular, we need a tree analog to the notions "nesting" and "crossing" that we expressed earlier with dependency lines. Clearly, the proper notion is that of "path between trace and binder".

4.2 Paths

Suppose we assume the following definition of a "path" between an

empty category and an \bar{A} -binder:

(69) Definition of Paths¹⁷

Suppose \underline{t} is an empty category locally \bar{A} -bound by \underline{b} . Then

- (i) for α the first maximal projection dominating \underline{t}
- (ii) for β the first maximal projection dominating \underline{b}
- (iii) the path between t and b is the set of nodes P such that

$$P = \{x \mid (x=\alpha) \vee (x=\beta) \vee (x \text{ dom. } \alpha \ \& \ \neg x \text{ dom. } \beta)\}$$

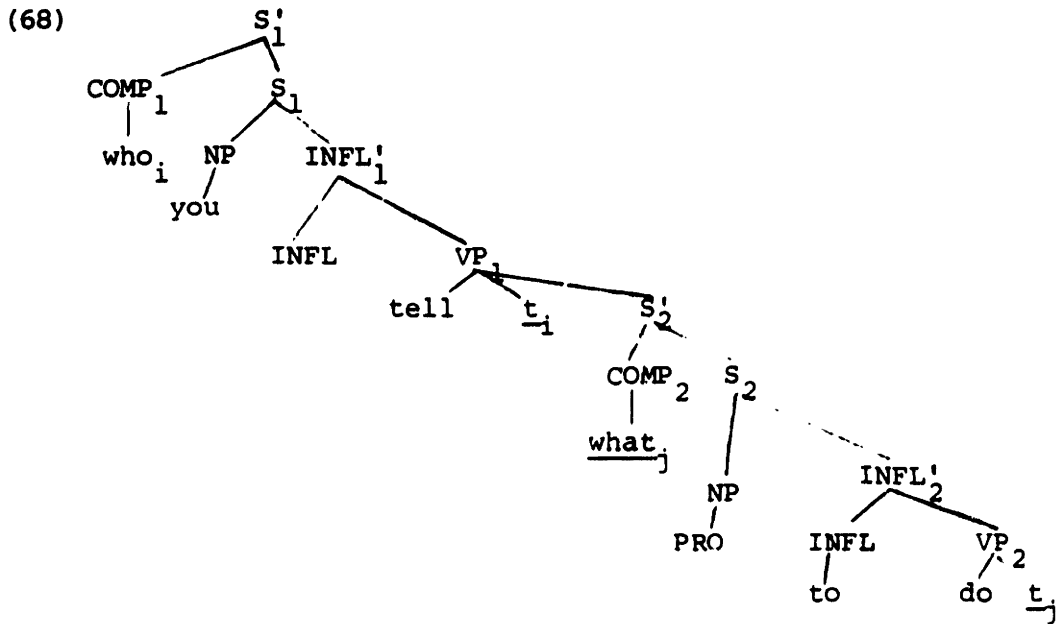
A path is essentially a line segment in a tree, which runs from the first maximal projection dominating a trace and the first maximal projection dominating its local \bar{A} -binder. This notion of path, except for the reliance on maximal projections, is very much the two-dimensional analog of the "dependency lines" we considered above. Replace the predicate dominate with be to the left of, and we have a working definition of dependency lines, similar to the one we tacitly assumed earlier.

Now let us restate the Crossing Constraint, to eliminate the problem of (68). (70) is a first approximation:

(70) Crossing Constraint (first approximation)

If the intersection of two paths is non-null, one path must contain the other.

Consider now (68) in tree form:



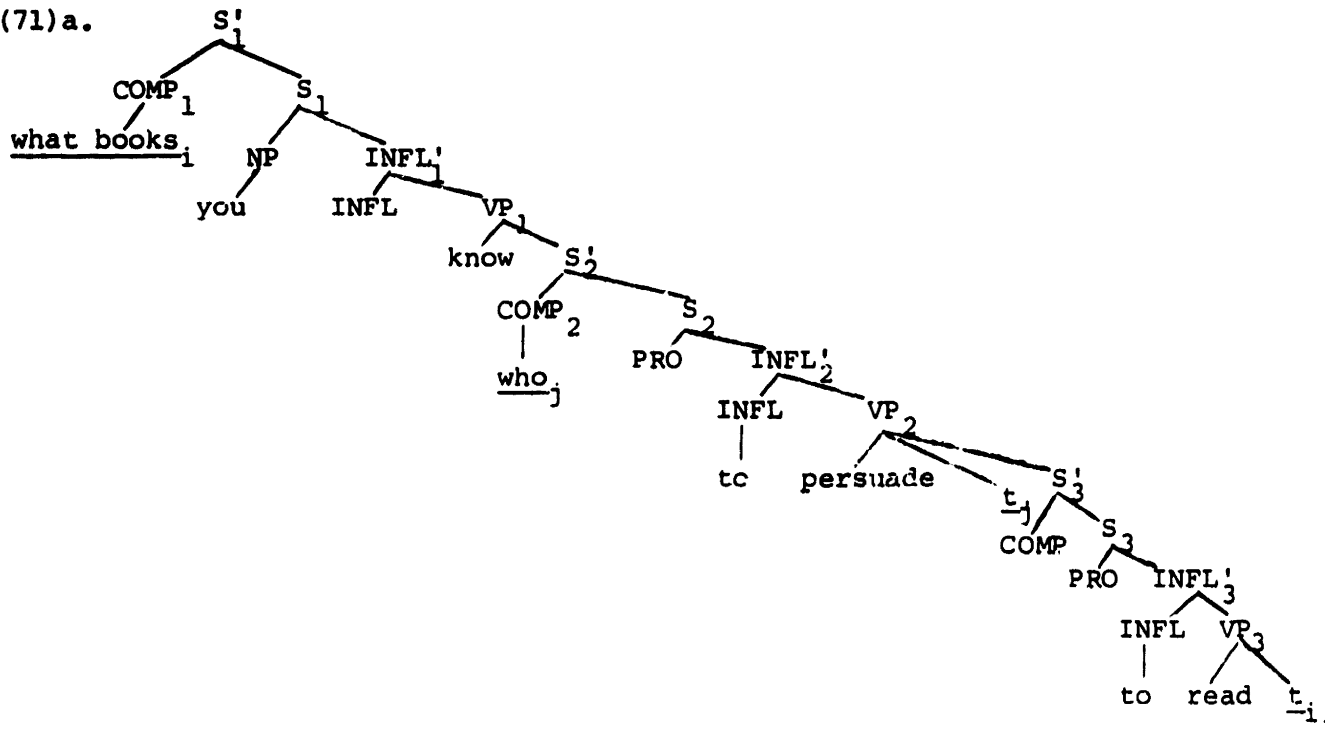
The path between \underline{t}_i and \underline{who}_i is: $\{VP_1, INFL'_1, S_1, S'_1\}$: a path from the first maximal projection dominating the trace (VP_1) to the first maximal projection dominating its \bar{A} -binder. The path between \underline{t}_j and \underline{what}_j is: $\{VP_2, INFL'_2, S_2, S'_2\}$. Notice that the intersection of these paths is null. Hence no Crossing Constraint is invoked.

Now consider a simple Crossing effect like the contrast (34)a-b, repeated below as (71)a-b:

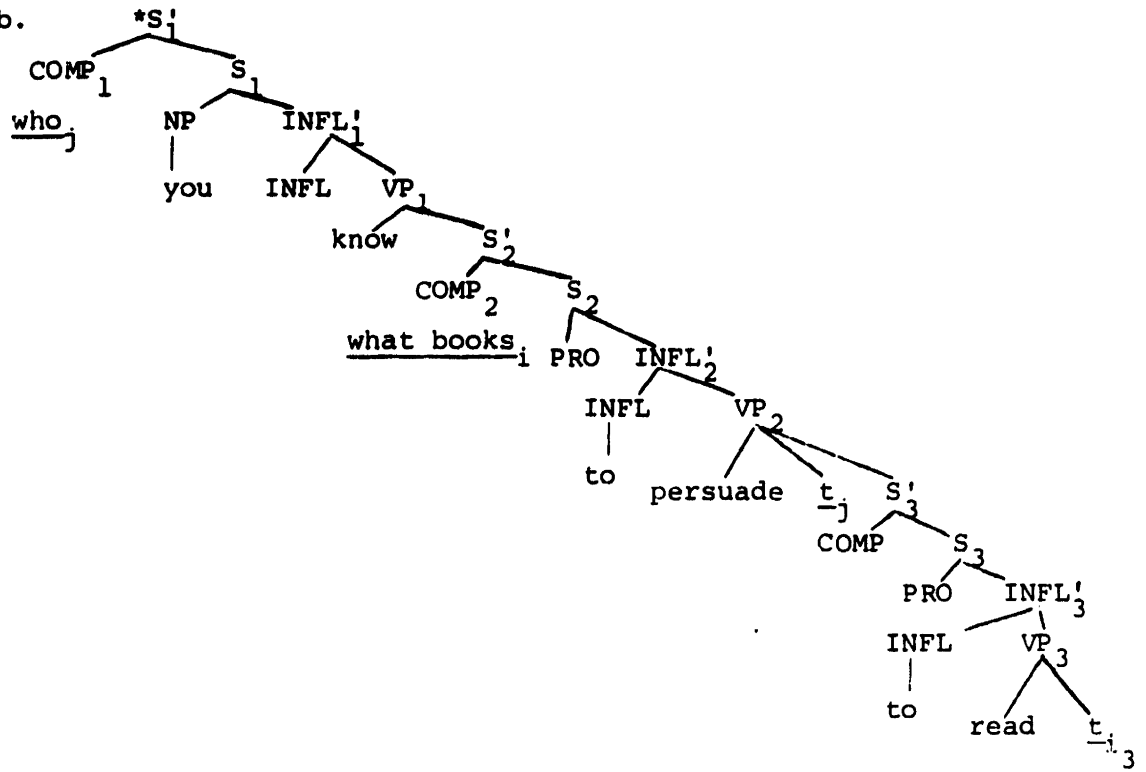
- (71)a. what books_i do you know [_S, who_j [_S PRO to persuade t_j [_S, PRO to read t_i]]
 b. *who_j do you know [_S, what books_i [_S PRO to persuade t_j [_S, PRO to read t_i]]

Consider the phrase markers associated with (71)a-b:

(71)a.



b.

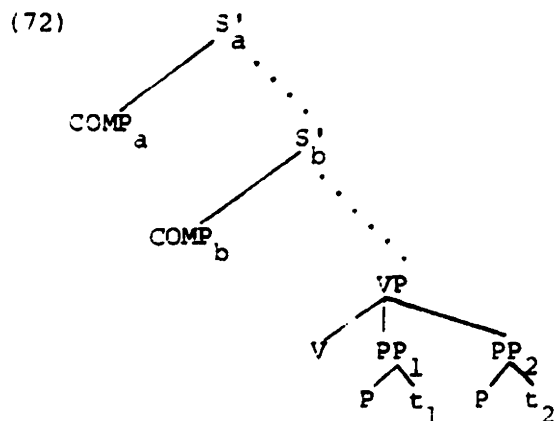


Now let us examine the paths in each of these phrase markers:

- (71)a. (i) Between t_i and what books $_i$:
 $\{VP_3, INFL'_3, S_3, S'_3, VP_2, INFL'_2, S_2, S'_2, VP_1, INFL'_1, S_1, S'_1\}$
- (ii) Between t_j and who $_j$:
 $\{VP_2, INFL'_2, S_2, S'_2\}$
- (71)b. (i) Between t_i and what books $_i$:
 $\{VP_3, INFL'_3, S_3, S'_3, VP_2, INFL'_2, S_2, S'_2\}$
- (ii) Between t_j and who $_j$:
 $\{VP_2, INFL'_2, S_2, S'_2, VP_1, INFL'_1, S_1, S'_1\}$

Each phrase marker contains two paths with which we are concerned at the moment (we discuss INFL and COMP shortly). In (71)a, path (ii) is contained by path (i); that is, (ii) is a subset of (i). In (71)b, however, (i) does not contain (ii) and (ii) does not contain (i). (i) contains various nodes of S'_3 that (ii) does not, while (ii) contains various nodes of S'_1 that (i) does not. If the Crossing Constraint requires that one path contain the other, (71)b violates the constraint, while (71)a obeys it. (71)a obeys it non-trivially, while (70) obeys it trivially.

Now let us return to our problem with doubly stranded prepositions, which we left after examples (66)-(67). Consider again a structure like (66), repeated in (72):



Recall that our earlier 2-dimensional Crossing Constraint (59), which counted nodes from the root, predicted incorrectly that structures like this should be grammatical no matter where the \bar{A} -binders of t_1 and t_2 are. In other words, there should be no crossing effects. Let us see what predictions the paths theory makes about this structure. Suppose t_2 is bound by a binder in $COMP_b$, and t_1 is bound from $COMP_a$ (example (30)). The paths will be as follows (irrelevant nodes omitted, as in (72)):

- (73) (i) Between t_2 and binder of $COMP_b$:
 {PP₂, VP, S'_b}
- (ii) Between t_1 and binder in $COMP_a$:
 {PP₁, VP, S'_b, S'_a}

Neither path contains the other: path (i) contains PP₂, which path (ii) lacks; path (ii) contains PP₁ and S'_a, which path (i) lacks. The structure is thus ruled out, as desired (*who do you know what subject to talk [to t] [about t]).

Now let us see what happens if t_2 is bound by a binder in $COMP_a$, and t_1 by a binder in $COMP_b$. Linearly, the dependencies are now nested (example (29)):

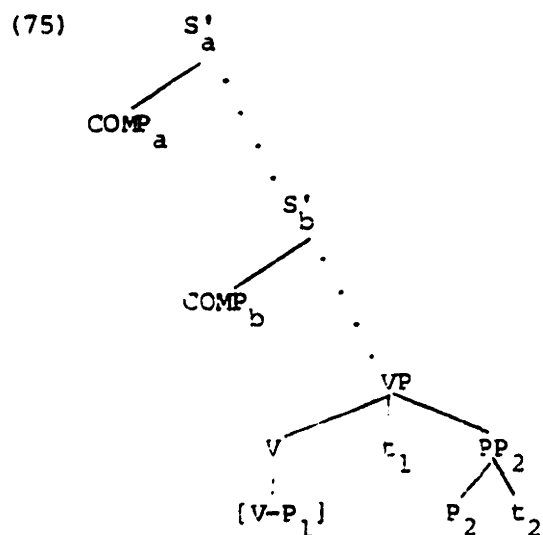
- (74) (i) Between t_2 and binder in $COMP_a$:
 {PP₂, VP, S'_b, S'_a}
- (ii) Between t_1 and binder in $COMP_b$:
 {PP₁, NP, S'_b}

The situation now is no better. Neither path contains the other: path (i) contains PP₂ and S'_a, which path (ii) lacks; path (ii) contains PP₁,

which path (i) lacks. This time, the ill-formed structure corresponds to a well-formed string (what subject do you know who to talk [to t] [about t]). On the surface, it would appear that the paths approach fares just as poorly as the approach of (59). We will show that this is not the case.

Recall that under (59), we were led to expect no Crossing effects, given a structure like (72). We noted that if something were to force the first preposition to reanalyze the V we could capture the actual Crossing effects that do obtain. The logical problem was that, since all binding possibilities in (72) satisfied (59), there was nothing to force the reanalysis.

Under a paths approach, however, (72) will never yield a grammatical structure (where the traces are \bar{A} -bound), as we have just seen. If reanalysis of the first P with V does yield a grammatical structure, reanalysis will then have been motivated, by the Crossing Constraint on paths. Let us examine how the paths approach fares if reanalysis does take place, as in (67) (repeated in (75)):



Suppose t_2 is bound from $COMP_b$ and t_1 is bound from $COMP_a$ (example (30)). The paths will be as follows:

- (76) (i) Between t_2 and binder in $COMP_b$:
 { PP_2, VP, S'_b }
- (ii) Between t_1 and binder in $COMP_a$:
 { VP, S'_b, S'_a }

Neither path contains the other: path (i) contains PP_2 , which path (ii) lacks; path (ii) contains S'_a , which path (i) lacks. The structure is correctly ruled out (*who do you know what subject to [talk to] t [about t_1]).

Suppose now that t_2 is bound from $COMP_a$, and t_1 from $COMP_b$ (example (29)). The paths will now be as follows:

- (77) (i) Between t_2 and binder in $COMP_a$:
 { PP_2, VP, S'_b, S'_a }
- (ii) Between t_1 and binder in $COMP_b$:
 { VP, S'_b }

In (77), the Crossing Constraints on paths is at last observed. Path (i) contains path (ii): path (i) contains all the nodes in path (ii), plus PP_2 and S'_a (what subject do you know who to [talk to] t [about t_1]).

Thus, on a paths account of Crossing effects, we will motivate reanalysis (which, we showed, would otherwise be optional) in order to yield a grammatical sentence from a structure like (74). After reanalysis, the only binding configuration that does not violate Crossing as a constraint on paths is the one just described, where "nesting" is observed as well.

What have we shown? We have not so far shown that a paths account is in any way superior to a linear account of the basic Crossing phenomena. We have shown that a paths account is at least as descriptively adequate in the basic cases as the linear account. We have also shown that the paths account, but not the linear account, motivates reanalysis in cases like the one just discussed. As it turns out, there is some slight evidence indicating that reanalysis is indeed indicated.

In (62) and (65) we showed that adverbial material may not intervene in a reanalyzed verb:

- (62)a. ?we pay close attention usually to John
 b. *John_i is paid close attention usually to t_i
 (65) ?John, who_i we pay close attention usually to t_i

Now consider :

- (78)a. ?I don't know who_i you talked usually to t_i about my problems
 b. these problems, which_i I don't know who_j you talked (*usually) to t_j about t_i

Only where reanalysis is forced by the Crossing Constraint (70) on paths is "niching" of an adverbial between V and the following P impossible. The facts are not outstandingly clear, but if the judgments are as indicated, they support the paths account of Crossing over the other accounts, particularly the linear account, which do not force reanalysis in this case. (Remember that (65) and (78)a show that reanalysis is not required in other cases of \bar{A} preposition stranding.)

Before turning in 4.4 to a paths account of Complementizer-Trace

Phenomena -- essentially a translation of our linear analysis -- we consider in the next section some further empirical evidence against the linear theory of Crossing effects.

4.3 Paths vs. Linear Crossing

We began the discussion in Part 4 by noting that in a Head First language, nodes that were linearly further to the right were also generally further down the tree. The linear crossing constraint told us that if a sentence contains two intersecting \bar{A} -dependencies, they must be nested: in other words, the one that begins further to the left must end further to the right. The paths account of these phenomena given in (70) says that when a phrase marker contains two paths created by \bar{A} -binding, one path must contain the other, if the intersection of the paths is non-null: in other words, the one that begins higher up the tree must end further down the tree.

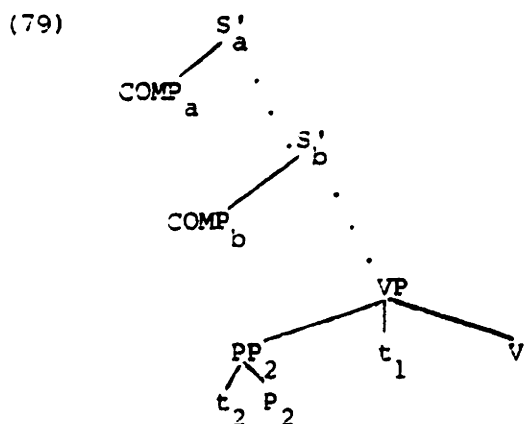
Suppose we were to examine a pure Head Last language: a mirror image of English, French and Italian. In such a mirror-image language, COMP would be last in S', INFL final in INFL', P final in PP, and the basic order would be OVS. If \bar{A} -binding worked much as in English, such that traces in A-positions were bound from COMP, we would expect the linear theory and the paths theory of Crossing effects once more to make nearly identical predictions. We do not have any data on such languages; since they do not distinguish between our competing theories, however, we may leave them aside for now.

Now consider a language that would be mixed Head First and Head Last. In particular, suppose that V were final in VP, but COMP initial in S'. In such a language, a linear Crossing Constraint and a constraint

on paths make opposite predictions. In VP, "left" and "right" would correspond to structural relations opposite from English. For \bar{A} -binders in COMP, however, "left" and "right" would correspond to the same relations as in English. On a linear theory, therefore, dependencies between COMP and a position in VP that cross in English should be nested in this mixed language; dependencies that are nested in English should cross in this language. Thus, the translations of English sentences like (29)-(43) should be matched with opposite judgments, according to a linear theory.

On the other hand, under a paths account, this language's judgments on sentences (29)-(43) should be identical to English. The parameter Head-First/Head-Last has no effect on paths, as can be readily seen. Paths are defined entirely in terms of dominance relations. (The only element of our paths account that is linear is the rule of reanalysis, but this too does not distinguish right from left: it only cares about contiguity.) Such a mixed language as we are considering, therefore, would provide an ideal test of the linear and paths theories of Crossing effects.

To make matters clearer, we may use a diagram similar to (75). Note that V is last in VP, but COMP is first in S':



The reader may reread the discussion of (75), where we showed that the paths account forces t_1 to be bound from $COMP_b$, and t_2 from $COMP_a$. That conclusion follows equally in (79). Linearly, however, the paths account predicts the following judgments:

- 80) a. $[S_a [COMP_a \text{ binder} [S \dots [S_b [COMP_b \text{ binder} [S \dots [VP [PP_2 t_2 P_2] t_1 V]]]]]]$
 b. * $[S_a [COMP_a \text{ binder} [S \dots [S_b [COMP_b \text{ binder} [S \dots [VP [PP_2 t_2 P_2] t_1 V]]]]]]$

The Germanic languages belong to this class. The predictions made by each of the Crossing theories are very difficult to test, because of quite a number of extraneous factors. Nonetheless, a crucial case can be devised. One of the problems is that word order is relatively free in VP in Germanic languages. Lacking a good theory of Germanic phrase structure, it is often hard to tell where a trace is, in the linear representation of a structure, making testing the linear theory difficult. Another problem arises with the analogs of examples like (34)a-b. If we extract one WH from a higher clause than another, we can be sure that the lower trace is within the confines of the lower clause. Unfortunately, embedded clauses are obligatorily extraposed to the right of V, leading to the prediction that both linear and non-linear theories of crossing will work in these cases (they do).

The crucial case seems to arise in Dutch, which allows preposition stranding to a limited degree, and under special circumstances, as discussed at length by van Riemsdijk (1978a,b): in particular, when a special inanimate WH-pronoun waar is moved, or in constructions involving a null operator, such as Tough Movement.

Even here, however, there are problems. As shown by van Riemsdijk (1978a, 96-112), Dutch has a rule of incorporation, under which a pre- (or post-)-position becomes morphologically part of V, and the former object of the preposition acts like a verbal object. In particular, it can appear in more than one place in \P, like a verbal object:¹⁹

- (81)a. ...om [met de auto] [in Amsterdam] te gaan. (no incorporation)
 in-order-to with the car to Amsterdam to go
- b. ...om [met de auto] Amsterdam [in te gaan] (incorporation)
- c. ...om Amsterdam [met de auto] [in te gaan] (incorporation)

(81)c is judged somewhat better than (81)b.

Clearly we are not going to be able to test our predictions in such a case, since, were we to extract both NP's in structures like (81), we would have no way of knowing the relative linear order of the traces. Now consider structures like (82):

- (82) ... om [met de auto] [naar Amsterdam toe] te gaan
 in-order-to with the car to Amsterdam to go

The "circumposition" naar-toe¹⁹ does not undergo the incorporation rule seen in (81):

- (83) *...om Amsterdam [met de auto] [naartoe te gaan]

Nonetheless, met or naar-toe may be stranded under \bar{A} -binding, as in the following examples with Tough Movement (cf. van Riemsdijk 1978b):

- (84)a. de auto is gemakkelijk om [t mee] [naar Amsterdam toe] te gaan
 the car is easy for with to Amsterdam to go

- b. ?Amsterdam is gemakkelijk om [met de auto] [t naartoe] te gaan
 Amsterdam is easy for with the car to t to go

(84)b is odd, probably for semantic reasons of some sort, much like its English equivalent: ?Amsterdam is easy to go in the car to.

Suppose now that the circumposition naar-toe may undergo a reanalysis rule, applying late enough (after oblique Case marking?) to rule out (83). It would be, perhaps, an LF correlate of incorporation. This rule would allow (84)b to have the structure:²⁰

- (85) Amsterdam is gemakkelijk om [met de auto] t [naartoe te gaan]

Assuming this rule is supportable, (85) has a post-reanalysis structure like that of (79), except that the object of the left-hand PP is lexical. Recall that the paths theory of Crossing effects predicts that, in a structure like (79), with double \bar{A} -binding, the trace to the right should be bound further to the right than the trace to the left. In other words, the paths account predicts that crossing of dependencies (in the linear string) should be better than nesting.

For reasons probably having to do with subadjacency (cf. van Riemsdijk 1978b), double preposition stranding is always ungrammatical, or at least highly unacceptable, in Dutch. Nonetheless, within the confines of ungrammaticality, the prediction of the paths account is borne out, for all speakers I have asked:

- (86)a. *** Amsterdam, waar de auto gemakkelijk is [s, om [t mee] t naartoe te gaan]
 Amsterdam, which the car easy is for with to to go
- b. *de auto, waar Amsterdam gemakkeiljk is [s, om [t mee] t naartoe te gaan]
 the car, which Amsterdam easy is for t with t to to go

While (86)b is ungrammatical, (86)a is incomprehensible.²¹ Recall that the better of the two contains a Tough Movement that was, in (84), actually the worse of the two. Thus the Crossing Constraint has overridden the judgments in (84).

It is worth comparing the English equivalents of (86), linearly speaking:

- (87)a. ?Amsterdam, which this car is easy [S, [S PRO to [ride in] t [to t]]
 b. *This car, which Amsterdam is easy [S, [S PRO to [ride in] t [to t]]

If we think in terms of a linear theory, the judgments are just the opposite of the Dutch judgments in (86). In terms of paths, however, the judgments are identical. Clearly, only the paths theory can account for the judgments in English and in Dutch in a unified way. The one stipulation we have had to make is that Dutch has a more abstract correlate of the English reanalysis rule. Everything else follows from independent and obvious facts about the order of constituents in Dutch: Dutch is Head Last in VP, but COMP-first in S'.

With the linear theory, of course, go the attempts to explain it in terms of parsing strategies, unless the term "parsing strategies" means the same thing as "principle of grammar". In Dutch, it appears, the first filler passed is matched with the first gap passed, while in English the last filler passed is matched with the first filler passed: we would hardly expect parsing strategies to differ so radically (or at all) from language to language. Notice that even should it turn out, as is likely, that the analysis of (86)a-b is more complicated than represented here, and even if paths do not play a crucial role²², this last point

still stands. Even restricting the Crossing Constraint to \bar{A} -binding, a linear Crossing Constraint clearly makes exactly the wrong predictions about Dutch. As it happens, the paths account appears to make exactly the right predictions.

Another empirical argument against a linear theory comes from phrases in English (and French) that can appear on either side of V. Consider (88), assuming it has the structure given:

(88)a. ?John, [_{S₁} who [_{S₁} I know [_{S₂} what book [_{S₂} I must publish t₁ [_{S₃} in order to impress t₂]]]]]]

b. *John, [_{S₁} who [_{S₁} I know [_{S₂} what book [_{S₂} I [_{VP} must impress t₁ [_{S₃} in order to publish t₂]]]]]]]]

(88) 'a. ??Jean, que je sais quel livre je dois publier t, pour impressionner t
b. **Jean, que je sais quel livre je dois impressionner t, pour publier t

(88) and its French equivalent are straightforwardly captured by both a linear Crossing Constraint and a paths account. The linear dependencies are shown above. Consider now the paths (irrelevant nodes omitted, as in (88)):

(89) In (88)a:

(i) Between t₂ and COMP of S₁':
{S₃', VP, S₂', S₂', S₁', S₁'}

(ii) Between t₁ and COMP₁:
{VP, S₂', S₂'}

In (88)b:

- (i) Between t_2 and COMP of S'_2 :
 $\{S'_3, VP, S_2, S'_2\}$
- (ii) Between t_1 and COMP of S'_1 :
 $\{VP, S_2, S'_2, S_1, S'_1\}$

In (88)a, as we see above, path (i) contains path (ii). In (88)b, neither path contains the other; hence, Crossing is violated.

Now suppose the in order to (pour) clause is placed before the verb. A trace in that clause will now precede the object of the verb that governs it. Since dominance relations are, we assume, preserved, the paths will remain the same. Since the linear order of the traces is reversed, (88)a will now show linear crossing, while (88)b will show linear nesting. Thus, the linear theory predicts that judgments will reverse, all things being equal, while the paths theory predicts no change in judgments. To my ears, the paths theory makes the right prediction:

(90)a. ?John [_{S₁} who [_{S₁} I know [_{S₂} what book [_{S₂} I [_{VP} must [in order to impress t |
 publish t]]]]]

b. *John [_{S₁} who [_{S₁} I know [_{S₂} what book [_{S₂} I [_{VP} must [in order to publish t |
 impress t]]]]]

(90)'a. ?*Jean, que je sais quel livre je dois, pour impressionner t,
 publier t

b. **Jean, que je sais quel livre je dois, pour publier t,
 impressionner t

None of the sentences of (90) are sterling, but the judgments of

(89) do seem to be preserved. Notice that the proponent of a grammatical, but not a performance-based, linear Crossing theory might account for (90) by claiming that the fronting observed here is a stylistic rule, applying after the level at which Crossing has applied. (Note that we have not yet discussed the level at which Crossing applies.) For this reason, the argument is not definitive. We take the definitive arguments to be forthcoming: a path-based account of Crossing effects derives automatically a number of other constraints, which a linear account is hopeless to subsume.²³ Nonetheless, we have presented this argument, and the argument from Dutch, to show that a linear Crossing theory must be patched in a number of ways to survive in domains in which the paths theory works perfectly. In the next section, we return to our derivation of the ECP from the Crossing Constraint, now assuming paths.

4.4 CTP Effects

To capture the Complementizer-Trace phenomena in a paths framework, we will have to make one stipulation, to handle a contradiction that will arise later on. Recall that we make the Complementizer-Trace phenomena (and, potentially, other ECP effects) fall under the Crossing Constraint by exploiting the special relationship between INFL and COMP. Recall also that we argued in the first part of this chapter for an articulated structure of S: $[_S \text{ NP } [_{\text{INFL}'} \text{ INFL VP}]]$. This structure will finally become important here.

If we are to claim that there is a path between INFL and COMP, there must be a path from some node to S', the first maximal projection dominating COMP. Let us stipulate that this node is INFL', although INFL'

does not at present seem like a maximal projection. Let us also assume that S, when INFL is [+tense], is a maximal projection. We might suggest that INFL' is the maximal projection of AGR in INFL, while S is the maximal projection of TNS, as a rationalization of these decisions. Nonetheless, they remain stipulations for now. It follows that the path between INFL and COMP will be {INFL', S, S'}. It also follows that any path from the subject of a tensed S will begin: {S ...}.

Notice that we would not have to stipulate anything if paths began, not with the maximal projection dominating a trace, but with the first non-lexical node dominating the trace (as suggested to me by N. Chomsky). That is, if we replaced (69)ii with (91):

(91) for β the first non-lexical node dominating \underline{t} ²⁴

then a path from INFL would begin with INFL', a path from the subject would begin with S, a path from the object of P with PP, from the object of V with VP, etc., exactly as desired. Problems with small clauses and cases of S'-deletion that will arise later cause us to reject that solution, in favor of the stipulation that INFL' and tensed S are, in effect, maximal projections. Clearly something is being missed, but we will leave the problem as is, pointing out where it arises at the appropriate time.

To introduce some nomenclature: we will identify paths in the text either by their endpoints (like a line segment), as "the INFL'-S' path", "the VP-S' path:", or by the nodes that cause the path to exist, as "the path between INFL and COMP", "the path between the object and COMP". We will be consistent in this convention, using "-" to name a path by its endpoints, and "between" to name a path by the nodes that generate it.

a non-null intersection, and path (ii) contains path (i), as required.

(92)b, on the other hand, raises an interesting technical problem:

(92)b. (i) Between INFL and COMP of S_2 :

$\{INFL_2', S_2, S_2'\}$

(ii) Between t_1^2 and COMP of S_2' :

$\{VP_2, INFL_2', S_2, S_2'\}$

(iii) Between t_1^1 and COMP of S_1' :

$\{S_2', VP_1, INFL_1', S_1, S_1'\}$

If we look just at paths (i) and (ii), everything works: their intersection is non-null, and (ii) contains (i). The problem arises with path (iii). As things stand, path (iii) is not in a proper relation to either path (i) or path (ii), since it intersects each of these paths at one node, S_2' . This problem will arise wherever COMP-to-COMP movement leaves a trace, and is a technical problem that will arise under a precise version of either a linear or paths theory of Crossing effects. In the present case, we might argue, as we did when discussing contraction in section 1, that the trace in the COMP of S_2' is actually optional. If it did not exist, (92)b would have the paths of (92)a. On the other hand, we will shortly see that such traces may exist, and must exist, in order to save (92)d. For this reason, it is worthwhile handling the problem now.²⁵

Clearly, we must replace the notion "non-null" in (70) by "non-singleton". If two paths intersect at just a single point, then there is no containment requirement, and (92)iii is acceptable. For the moment, let us make this a matter of definition. We may define overlapping as a special case of non-null intersection:

(93) Overlapping (def)

Two paths overlap iff their intersection is non-null and non-singleton.

We can now rephrase our Crossing Constraint in terms of overlapping.

Let us rename it the Path Containment Condition (PCC):

(94) Path Containment Condition (PCC)

If two paths overlap, one must contain the other.

As formulated, the PCC of (94) is the basic principle of "Path Theory".

We shall reformulate the definition of paths a number of times, but the PCC will remain the principle that governs their interaction.

The prime virtue of the PCC is its simplicity, therefore it is worth asking about the odd status of our definition of "overlapping". Why should "Path Theory" single out non-null, non-singleton intersections? That is, what is special about the numbers 2 and above, such that only intersections with two or more members count as "overlaps"?

An answer to this question has been suggested to me by R. Kayne and by I. Haik (personal communications). They point out that the number 2 holds a special place in geometry, where two points form a line segment. We have been conceiving of paths as sets of nodes, which together form a line segment in a tree. We might equally well conceive of paths as sets of line segments, i.e. as sets of pairs of nodes. On such a definition, paths (ii) and (iii) in (92)b above would be as follows:

(92)a.(ii) Between the object of VP₂ and COMP of S'₂:

{(VP₂, INFL'₂), (INFL'₂, S₂), (S₂, S'₂)}

- (iii) Between COMP of S'_2 and COMP of S'_1 :
 $\{(S'_2, VP_1), (VP_1, INFL'_1), (INFL'_1, S_1), (S_1, S'_1)\}$

The interaction of paths (ii) and (iii) is now null. In general, if paths are sets of pairs of nodes (line segments), we can eliminate the need for the notion "overlapping" and revert to the most intuitive formulation of the PCC: "if two paths have a non-null intersection, one must contain the other". I believe that this solution is more than likely to be correct. On the other hand, for ease of presentation and to save space, I will continue to represent paths as sets of nodes, and to assume the definition of overlapping and the PCC as in (94). At present, there does not seem to be any empirical difference between the two approaches; our continued use of (94) thus appears to be innocuous.

We now turn to (92)c (*who did you say [S'_2 that [S_2 t liked Mary]]):

- (92)c.(i) Between INFL and COMP of S'_2 :
 $\{INFL'_2, S_2, S'_2\}$
- (ii) Between subject of S_2 and COMP of S'_1 :
 $\{S_2, S'_2, VP_1, INFL'_1, S_1, S'_1\}$

(92)c violates the PCC, just as it violated the linear Crossing Constraint. Path (ii) lacks the $INFL'_2$ nodes of path (i); path (i) lacks the nodes above S'_2 that are part of path (ii). Complementizer-Trace Phenomena thus derive from the PCC.

Finally, consider (92)d (who did John say [S'_2 t¹ [S_2 t² liked Mary]]):

- (92)d.(i) Between INFL and COMP of S'_2 :
 $\{INFL'_2, S_2, S'_2\}$
- (ii) Between t² and t¹ in COMP of S'_2 :
 $\{S_2, S'_2\}$
- (iii) Between t¹ and COMP of S'_1 :
 $\{S'_2, VP_1, INFL'_1, S_1, S'_1\}$

(92)d satisfies the PCC, given our definition of overlapping. (i) and (ii) overlap, and (i) contains (ii). (iii) does not overlap with any path. In other words, we correctly predict that extraction of the subject of a tensed sentence is possible if it is bound from the nearest COMP. This analysis extends, of course, to the French que/qui alternation discussed at (53)-(53)', and to English that-relatives discussed at (56), with the stipulation of an indexed complementizer indicated earlier.

Note the crucial role played here by the INFL' node argued for in section 1. Suppose S had the ternary structure suggested in Chomsky (1981)a: [_S NP INFL VP]. A path from the first maximal projection that dominates INFL would include only S and S'. Long movement of the subject would generate a path like path (ii) in (92)c, which would not violate any version of the PCC. Given the ternary structure for S, the only way we could derive the CTP effects from the PCC would be to assume a path from VP to COMP, which would include VP, S and S'. All the required results would follow. Such a path would, however, be entirely unmotivated, since there is no particular connection between the choice of verb and COMP. In Chapter Four, we will present detailed arguments that the path really does run between INFL and COMP, and is motivated by the movement of a [+tense] TNS node. Thus, the binary structure for S seems the simplest in this context, and we consider the evidence presented in section 1 to be additional compelling evidence for it.

Finally, it is worth pausing for a moment to compare our theory at this point to the ECP theory of the CTP. Recall that the ECP requires that empty categories be properly governed. Proper government is a

concept that supplies a list of environments in which empty categories may occur: they may occur where governed by $X^0 \neq \text{AGR}$, or where governed by coindexation from COMP. Insofar as the ECP applies to \bar{A} -bound empty categories, we may now dispense with the notion "proper government". The PCC implies a set of environments in which an \bar{A} -bound empty category may not occur, namely when the closest dominating maximal projection is in the middle of a path, and the \bar{A} -binder is outside of that path.

The PCC account of the CTP thus does not need to say that there is something special about V, as opposed to AGR, nor to make an odd linking of coindexation with lexical government. Thus, the first two questions we asked about the ECP in section 2 -- (1) Why is AGR not a proper governor? and (2) Why does the definition of government link coindexation with lexical government? -- both disappear. Without the notion "proper government", neither question need be posed. Notice that we may eliminate coindexation entirely from the definition of government, since that clause is necessary only for cases of long movement of subjects. We return to this question in Chapter Five. Since neither (1) or (2) has been successfully answered, the disappearance of these questions is desirable. In their place, however, we have two new questions.

1. Why is there a path between INFL and COMP?
2. Why does this path begin at INFL', and why does the path from the subject of a tensed sentence begin at S?

We shall provide an answer to the first question in Chapter Four, which will have some surprising consequences with respect to coordination. The second question, as we have indicated, is more difficult. We

suspect that the answer will lie either in a reformulation of the definition of path, or in a reanalysis of the X-bar structure of S. The question thus remains open.

In section 2 of this chapter, we asked a third question: (3) Why is there an ECP in the first place? Clearly, if our proposal is correct, this question (for \bar{A} -binding, at least) reduces to the question of why there is a PCC. We consider this a good reduction, since the PCC is more general than the ECP of Chomsky (1981a), and seems simpler. We shall end this study with some speculations about the origins of the PCC. This new third question, however, will also remain open.

We shall return, in 4.6 of this chapter, to a comparison of the PCC account of the CTP with Stowell's (1981) version of the ECP. In the next chapter, we shall be comparing the PCC with Kayne's account. Finally, remember that the PCC, whose domain we have restricted to \bar{A} -binding, does not handle the A-binding facts grouped by Chomsky under the ECP. We return to this question in Chapter Five.

4.5 The Subject Condition

In this section, we shall see that the PCC automatically derives the Subject Condition of Chomsky (1973), in tensed sentences. In the next section (4.6), we extend the definition of path to capture the Subject Condition in S'-deletion infinitivals and small clauses (cf. Kayne 1982). In Chapter Four, we extend our discussion of the Subject Condition to incorporate apparent exceptions in parasitic gap constructions analyzed by Kayne, adapting his analysis to our framework.

In the previous section, we showed that the absence of long

movement of the subject of a tensed S is predicted, if we assume a path between INFL and COMP of the type discussed, and also that tensed S is a maximal projection. From these two assumptions, it followed that the subject of a tensed S must be bound from the nearest COMP, to avoid "crossing" the path between INFL and COMP. The possibility of binding the subject position from COMP thus helps save an \bar{A} -bound subject trace from the PCC.

Now suppose we consider a trace that is not itself a subject, but is contained within a subject. As is well known (since at least Chomsky 1955-1975), such traces are always ungrammatical except as parasitic gaps; cf. Chapter Four: we ignore these cases for now):

(95)a. [_S who_i [_S (did) John [_{INFL}, INFL [_{VP} hear [_{NP} stories
[_{PP} about t_i]]]]]]]

b.*[_S who_i [_S (did) [_{NP} stories [_{PP} about t_i]] [_{INFL}, INFL
[_{VP} terrify John]]]]]

(96)a. John, [_S who_i [_S it [_{INFL}, INFL [_{VP} surprised me
[_S' Mary saw t_i]]]]]]]

b.*John, [_S who_i [_S [_S' that Mary saw t_i]] [_{INFL}, INFL
[_{VP} surprised me]]]]]

(97)a. Bill, [_S who_i [_S it [_{INFL}, INFL [_{VP} surprised me
[_S' to see pictures of t_i]]]]]]]

b.*Bill, [_S who_i [_S [_S' to see pictures of t_i]] [_{INFL}, INFL
[_{VP} surprised me]]]]]

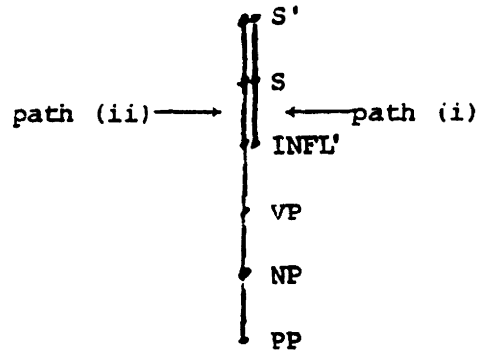
Chomsky (1977) notes that (95)b can be ruled out by subjacency, if S is a bounding node in English. Movement of WH from the position of t_i crosses at least two bounding nodes: NP and S. We will see later that this observation is empirically correct: nonetheless (95)b is also straightforwardly excluded by the PCC. We will see in Chapter Four, section 2 that this redundancy is correct: there are environments in which the effects of PCC or of subjacency can be neutralized; the interaction of the two principles yields exactly the desired results.

(96)b and (97)b, however, do not fall under subjacency, since the sentential subject contains a COMP through which WH can move.²⁶ If S is a bounding node, movement to COMP of S'_2 crosses one bounding node, S_2 . Movement to COMP of S'_1 crosses one bounding node, S_1 . Similarly, if S' is bounding; each movement crosses one node. Let us now see how the PCC rules (95)b, (96)b, and (97)b out.

The PCC rules out extraction of a piece of a subject very simply. The path between COMP and INFL includes INFL', which no path from subject to COMP will ever include. The path from within a subject to COMP always includes the subject node itself, as well as nodes dominated by it, which the path between INFL and COMP will never include. Hence, the PCC can never be satisfied. Below, we give the paths for (95)a, an extraction from inside an object, and (95)b, an extraction from inside a subject. To help visualize what goes wrong in (95)b, we have added diagrams: the paths are represented as line segments, with the nodes indicated. Think of these segments as pieces of a tree. Thus:

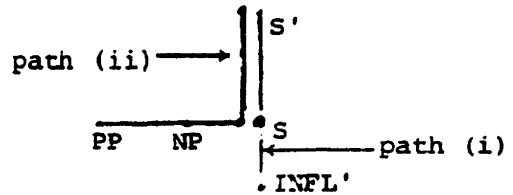
(95)a.(i) Between INFL and COMP:
 {INFL', S, S'}

(ii) Between t_i and COMP:
 {PP, NP, VP, INFL', S, S'}



(95)b.(i) Between INFL and COMP:
 {INFL', S, S'}

(ii) Between t_i and COMP:
 {PP, NP, S, S'}



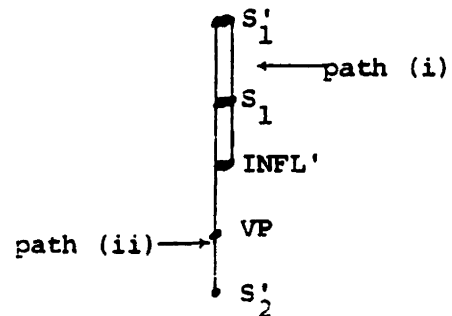
(95)a satisfies the PCC: path (ii) contains path (i). (95)b violates the PCC: paths (ii) and (i) overlap, but neither one contains the other.

Similarly for (96)-(97). We have eliminated irrelevant nodes in the embedded S':

(96)a./ (97)a.

(i) Between INFL' and COMP:
 {INFL', S₁, S'₁}

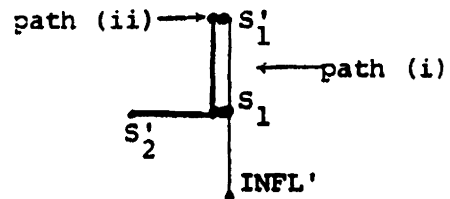
(ii) Between t_i and COMP of S'₁:
 {S'₂, VP, INFL', S₁, S'₁}



(96)b./ (97)b.

(i) Between INFL' and COMP:
 {INFL', S₁, S'₁}

(ii) Between t_i and COMP of S'₁:
 {S'₂, S₁, S'₁}



(96)a and (97)a satisfy the PCC: path (ii) contains path (i). (96)b and (97)b do not, however. Path (ii) contains S'_2 (and various nodes dominated by S'_2 , which we have omitted), which path (i) does not. Path (i), on the other hand, contains INFL', which path (ii) does not. Examples like (96)a and (96)b can be constructed with long movement as well. As expected, the presence or absence of that is irrelevant: the structure violates the PCC whether or not there is a local binder in the nearest COMP:

(98) *Bill, who I think (that) to see pictures of t would surprise me

The PCC thus derives Chomsky's (1973) Subject Condition for \bar{A} -binding, in tensed sentences. The relevant part of Chomsky's condition is:

(99) No rule can involve X, Y in the structure

... X ... [_{α} ... Y ...] ...

where ... α is a subject phrase properly containing the minimal major category dominating Y ...

The PCC goes beyond the Subject Condition, however. Note that effects of the type noted should be found with any category immediately dominated by a tensed S. In general we should be unable to \bar{A} -bind into any category dominated by a tensed S. In this way we may derive the well-known fact that prepositions may not be stranded when dominated by S. The facts are particularly clear when the PP is preverbal, but have been claimed to hold (Chomsky 1965) when post-verbal:

(100)a. ?I wonder [_S, what war_i [_S John [_{INFL}, INFL [_{VP} died]] [_{PP} during t_i]]]

b. *I wonder [_S, what war_i [_S [_{PP} during t_i] John [_{INFL}, INFL [_{VP} died]]]]]

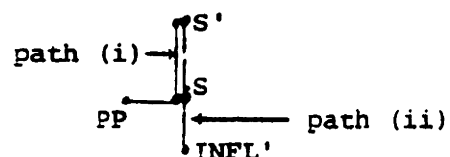
The paths for (100)b are:

(100)b. (i) Between INFL' and COMP:

{INFL', S, S'}

(ii) Between t_i and COMP:

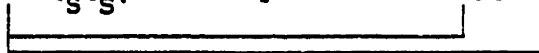
{PP, S, S'}



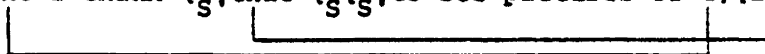
The paths for (100)a-b, with the structures shown, violate the PCC in the same way as the Subject Condition violations discussed above. The relative acceptability of sentences like (100)a for speakers such as myself (cf. Rothstein 1981) may be attributed to an optional inclusion of sentential PPs in the VP.^{27 28}

Notice that the Subject Condition can in no way be derived from a linear Crossing Constraint. Since both the subject and pieces of the subject are to the left of INFL, a Crossing Constraint stated linearly predicts no contrast between the two cases. Rather, a movement of a piece of a subject, like movement of the whole subject, should show an asymmetry between long and short movement:

(101)a. *Bill, [_S who [_S [_S to see pictures of t] [INFL would surprise me]]]



b. *Bill, who I think [_S that [_S [_S to see pictures of t] [INFL would...]]]



In (101)a, the linear dependencies are nested; in (101)b, they cross. Nonetheless, both are ungrammatical. To rule out (101)a, a linear theory would need a separate constraint, such as the Subject Condition. The paths theory, however, captures (101) the same way it captures the basic CTP facts.

4.6 The Subject Condition in Small Clauses and S'-deletion Infinitives

4.6.1 Chomsky's (1973) condition, given in (99), extends naturally to the subjects of infinitivals that have undergone S'-deletion. Chomsky himself notes this fact, and uses it as an argument against a rule of Raising to Object:

- (102)a. [_{S₁} who_i [_{S₁} (do) you [_{INFL₁} INFL [_{VP₁} believe [_{S₂} John [_{INFL₂} to [_{VP₂} like [_{NP₃} stories [_{PP₃} about t_i]]]]]]]]]]]
- b. * [_{S₁} who_i [_{S₁} (do) you [_{INFL₁} INFL [_{VP₁} believe [_{S₂} [_{NP₂} stories [_{PP₂} about t_i]] [_{INFL₂} to [_{VP₂} have terrified me]]]]]]]]]

Given the extended definition of subject proposed by Chomsky, discussed at the beginning of this chapter, Chomsky's Subject Condition also covers analogous facts in small clauses (cf. Kayne 1982, note 2):

- (103)a. ? [_{S₁} who_i [_{S₁} (do) you [_{INFL₁} INFL [_{VP₁} consider [_{A₂} Bill [_{AP₂} angry [_{PP₂} at [_{NP₃} some friends [_{PP₃} of t_i]]]]]]]]]]]
- b. * [_{S₁} who_i [_{S₁} (do) you [_{INFL₁} INFL [_{VP₁} consider [_{A₂} [_{NP₂} friends [_{PP₂} of t_i]] [_{AP} dull]]]]]]]]]

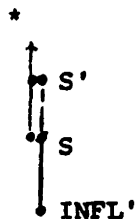
Nothing in our paths account so far distinguishes (102)a/(103)a from (102)b/(103)b. Notice that it would not do to imagine that S'-deletion infinitivals and small clauses have a COMP and INFL at some level of representation, like tensed clauses, since the subjects of these infinitivals and small clauses show no CTP effects. Long movement of

the subject is as perfect as movement of an object:

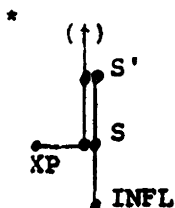
(104)a. $\{S_1, who_i \{S_1 (do) you \{_{INFL_1} INFL \{_{VP_1} consider \{S_2 t_i \{_{INFL_2} to$
 $\{_{VP_2} be dull\}\}\}\}\}\}$

b. $\{S_1, who_i \{S_1 (do) you \{_{INFL_1} INFL \{_{VP_1} consider \{A_2^* t_i \{_{AP_2} dull\}\}\}\}\}$

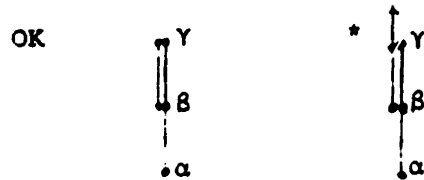
Let us review for a moment how we derived the basic CTP and subject condition facts in tensed sentences. We proposed that there was a path $\{INFL', S, S'\}$ in tensed sentences. We also supposed that tensed S was a maximal projection -- perhaps the maximal projection of AGR. Thus, extraction of the subject of a tensed S yields a path beginning with the maximal S; if the subject trace is bound outside S', the PCC is violated:



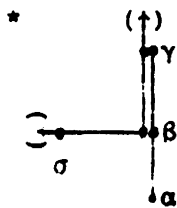
Extraction from within the subject of a tensed S will also violate the PCC. Since the path from the trace of this extraction begins with a maximal projection dominated by the S, the maximality of tensed S does not matter here. The only feature relevant to Subject Condition effects is the existence of the path between INFL and COMP (XP is the subject of S):



Let us be more abstract about what causes "CTP-like" effects and "subject condition effects". CTP-like effects are a general result of paths of the form $\{\alpha, \beta, \gamma\}$ when β is maximal. Extraction of material immediately dominated by β (e.g. the subject, when β is a tensed S) creates a path that begins with β . This path is acceptable if it ends at γ . If it does not end at γ , however, the PCC is violated:



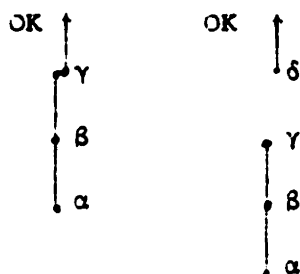
Extraction of material within the first maximal projection dominated by β -- call this node σ -- (e.g. the subject of a tensed S) -- creates a path that includes or begins with σ and goes on to include β and γ , regardless of where it ends: $\{(\dots) \sigma, \beta, \gamma(\dots)\}$. This path, regardless of where it begins or ends, will always violate the PCC if there is another path $\{\alpha, \beta, \gamma\}$, where β dominates α , as before:



Notice that whether β is maximal or not does not matter here, since σ is maximal. This is the point we made above, where β was a tensed S.

Suppose now that a given structure contains a path $\{\alpha, \beta, \gamma\}$, where β is not maximal. They either γ is maximal, or some δ dominating γ is maximal. Suppose we extract, as before, the constituent dominated by β , to an \bar{A} -position. There will be a path between the trace of this constituent

and its \bar{A} -binder, but the path will not begin with β , since it is not maximal. Rather it will begin with γ , if it is maximal, or else with δ . Either way, the path will not overlap the path $\{\alpha, \beta, \gamma\}$, and the PCC will not be invoked:



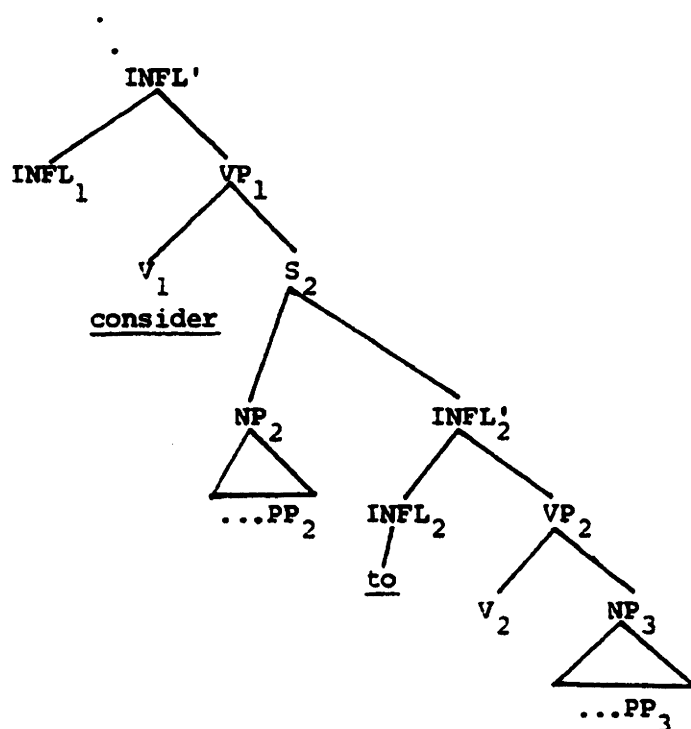
Thus, if β is not maximal, there will be no "CTP-like" effects arising from the extraction of material immediately dominated by β . On the other hand, whether or not β is maximal, there will be "Subject-Condition-like" effects if we extract material contained by a maximal projection dominated by β , as we have just seen.

This pattern recalls, of course, our problem with small clauses and S'-deletion infinitivals. The node dominating the small clause or infinitival is non-maximal. Extraction of the subject of the small clause or infinitival is grammatical: there are no CTP-like effects. Extraction from within the subject is ungrammatical: there are Subject Condition effects. If we identify the non-maximal projection dominating the small clause of S'-deletion infinitival with β in the abstract diagrams above, and the subject of the small clause of infinitival with σ , and if there is a path from some α dominated by this β to some γ that dominates this β , we predict that there will be Subject Condition effects, but no CTP effects, in small clauses and S'-deletion infinitivals.

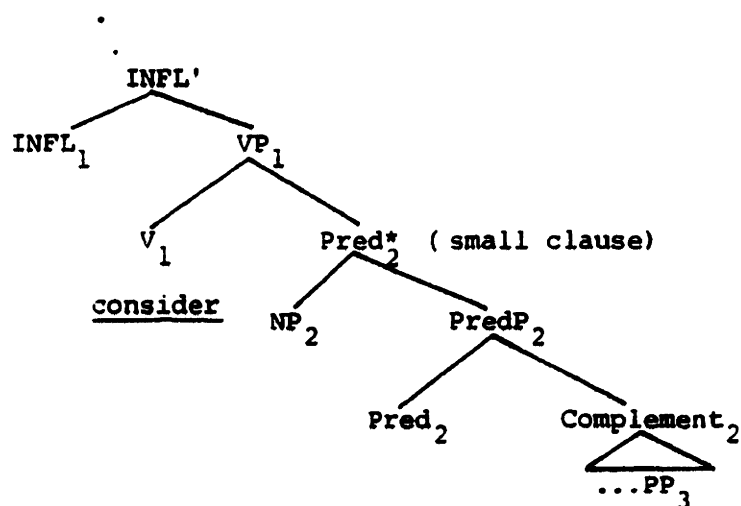
The obvious question that now arises is this: if β is the X^* dominating a small clause with an XP predicate, and the S dominating an S'-deletion infinitival, what are α and γ , and why does a path between α , β , and γ exist?

Let us examine the structures in question, referring back to (102)-(103):

(102) .



(103) .



In (102), the node we have called β is S_2 . In (103), it is Pred^* . These nodes immediately dominate the subject, which is σ . It follows that γ is VP_1 in both examples. It is also reasonable to identify α with the predicate phrase INFL' in (102) and with PredP , the predicate phrase of a small clause, in (103).

Thus, suppose that S' -deletion infinitival sentences have a path between INFL and the V that governs the infinitival: $\{\text{INFL}'_2, S_2, \text{VP}_1\}$. Small clause structures contain a path from the predicate phrase of the small clause to the VP that contains the V that governs the small clause: $\{\text{PredP}_2, \text{Pred}^*_2, \text{VP}_1\}$. Let us see, concretely now, how the existence of these paths derives Subject Condition effects in S' -deletion infinitivals and small clause structures, while not mistakenly ruling out simple extraction of the subject or complement of these clauses.

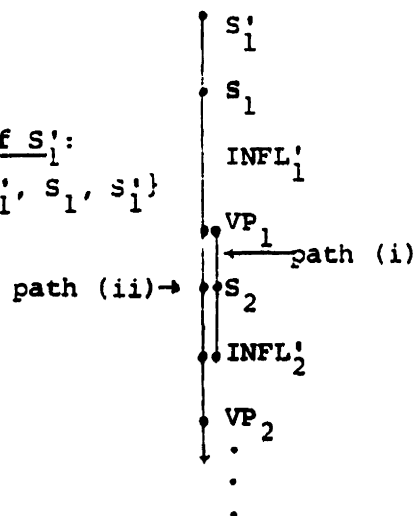
Consider again (102)a-b, which demonstrates Subject Condition effects in S' -deletion structures. Suppose there is, as we postulate, a path between the INFL of the infinitival and the V that governs the infinitival clause. The relevant paths for (102)a (who do you believe John to like stories about), with extraction from the complement of the infinitival, and for (102)b (*who do you believe stories about to have terrified me) are given below, along with schematic representations of the paths:

(102)a. (i) The INFL' - VP path:

$\{\text{INFL}'_2, S_2, \text{VP}_1\}$

(ii) Between the object of PP_3 and COMP of S'_1 :

$\{\text{PP}_3, \text{NP}_3, \text{VP}_2, \text{INFL}'_2, S_2, \text{VP}_1, \text{INFL}'_1, S_1, S'_1\}$

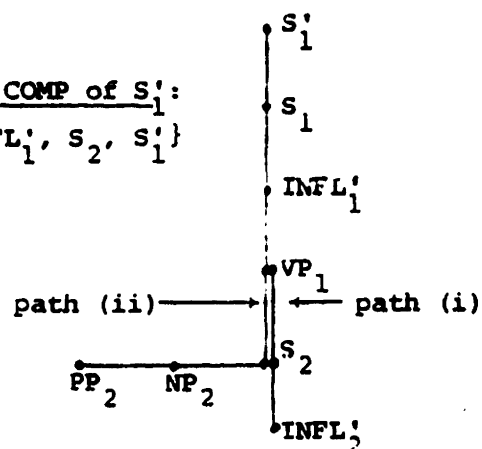


(102)b. (i) The INFL'-VP path:

{INFL'₂, S₂, VP₁}

(ii) Between the object of PP₂ and COMP of S'₁:

{PP₂, NP₂, S₂, VP₁, INFL'₁, S₂, S'₁}



In (102)a, paths (i) and (ii) overlap, and (ii) contains (i), satisfying the PCC. In (102)b, paths (i) and (ii) overlap, but neither path contains the other: the PCC is violated.

Consider now the paths in (104)a (who do you consider to have come).

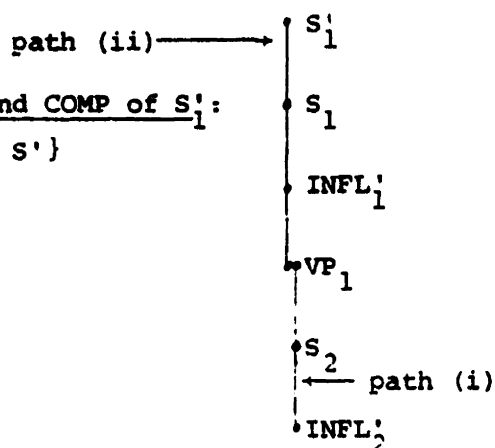
Remember that infinitival S is non-maximal; hence the subject's path begins at VP, the first maximal projection that dominates it:

(104)a. (i) The INFL'-VP Path:

{INFL'₂, S₂, VP₁}

(ii) Between the subject of S₂ and COMP of S'₁:

{VP₁, INFL'₁, S₁, S'}



In (104)a, paths (i) and (ii) do not overlap, since they do not share two nodes. Therefore the PCC is not violated. Notice that the PCC interacts here with the independently motivated assumption that the node dominating infinitival complements to verbs like English consider is non-maximal: only if it is non-maximal can the higher verb govern its subject and assign Case to it.

Thus, the PCC, in tandem with various independently necessary assumptions predicts that S'-deletion infinitivals will show Subject Condition effects, but not any general constraints on extracting subjects like the CTP. We may quickly demonstrate the same for small clauses. Recall that we are stipulating for the moment that there is a PredP-VP path in small clause constructions.

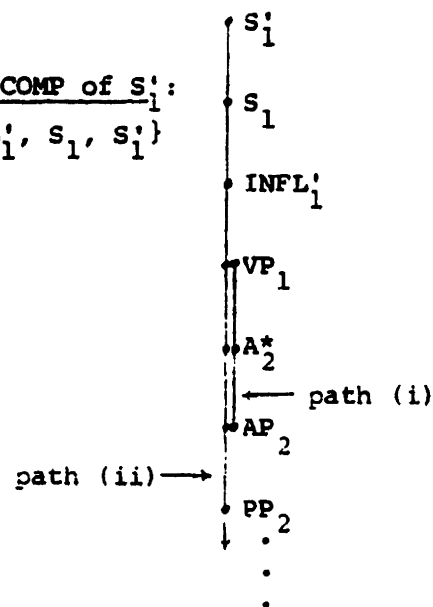
The paths relevant for (103)a (?who do you consider Bill angry at some friends of) and for (103)b (*who do you consider friends of dull) are as follows:

(103)a. (i) The PredP-VP Path:

{AP₂, A*₂, VP₁}

(ii) Between the object of PP₃ and the COMP of S'₁:

{PP₃, NP₃, PP₂, AP₂, A*₂, VP₁, INFL'₁, S₁, S'₁}

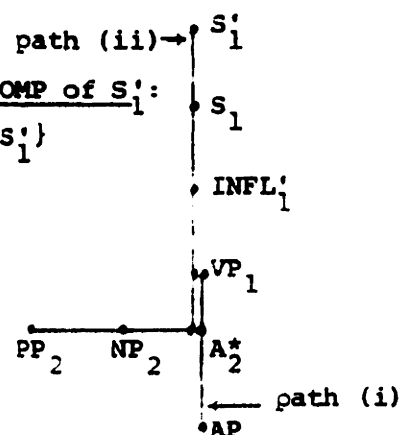


(103)b. (i) The PredP-VP Path:

{AP, A₂^{*}, VP₁}

(ii) Between the object of PP₃ and the COMP of S₁':

{PP₂, NP₂, A₂^{*}, VP₁, INFL₁', S₁, S₁'}



The containment relations are the same as in (102). (103)a satisfies the PCC, since path (ii) contains path (i). (103)b violates the PCC, since paths (i) and (ii) overlap, but neither contains the other.

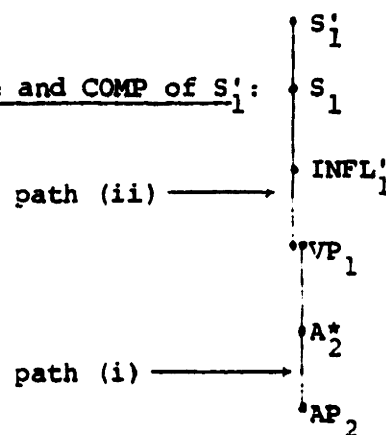
Finally, the PCC also predicts that the subject of the small clause should be extractable as a whole, just like the subject of an S'-deletion infinitival. Remember that the node we are calling Pred* (A* here) is a non-maximal projection of PredP. The following are the relevant paths in (104)b (who do you consider dull):

(104)b. (i) The PredP-VP Path:

{AP₂, A₂^{*}, VP₁}

(ii) Between the subject of small clause and COMP of S₁':

{VP₁, INFL₁', S₁, S₁'}



Paths (i) and (ii) do not overlap, and the PCC is not violated.

4.6.2 Now we must ask why the paths {INFL', S, VP} in S'-deletion constructions and {PredP, Pred*, VP} in small clause constructions exist. Using our nomenclature from 4.4, we ask what these paths run between, and why.

That the path ends with VP suggests that there is something about V that causes the path to exist. The fact that the path's other end is in the complement to V suggests that there is something about the relationship between V and the complement that creates the path. The obvious answer is subcategorization: that is, selection and Θ -role assignment. Notice that the path runs from VP -- the first maximal projection dominating V, and Θ -assigner -- to the first maximal projection of the complement that V Θ -marks. This suggests more generally that Θ -marking, like \bar{A} -binding, creates a path.

If this is true, we might revise our definition of path given in (69) to (105):

(105) Definition of Paths (revision of (69))

Suppose \underline{t} is an empty category locally \bar{A} -bound by \underline{b} , or a category directly assigned a Θ -role R by \underline{b} . Then

- (i) For α the first maximal projection dominating \underline{t}
- (ii) For β the first maximal projection dominating \underline{b}
- (iii) the path between \underline{t} and \underline{b} is a set of nodes P such that

$$P = \{x \mid (x=\alpha) \vee (x=\beta) \vee (x \text{ dom. } \alpha \ \& \ \neg x \text{ dom. } \beta)\}$$

We assume that Θ -roles are assigned to lexical heads -- that is, to the minimal projection of a complement. In particular, when a predicate

assigns a Θ -role to a small clause, it assigns the Θ -role specifically to the lexical head of the small clause, its predicate. Similarly, when a Θ -role is assigned to an S'-deletion infinitival, it is assigned specifically to the INFL that heads it.

Notice that a path between a Θ -marker and an argument it Θ -marks will run from the first maximal projection dominating the Θ -marker to the first maximal projection dominating the recipient of the Θ -role, the head of the Θ -marker's complement. In all cases except small clauses and S'-deletion infinitivals, the maximal projection dominating the head of the complement is also the highest projection of the head. The non-maximal projections that are dominated by no maximal projection, found in small clauses and S'-deletion infinitivals, have a unique status in this respect. Consider the more usual situation:

- (106)a. John [_{VP} [_V saw] [_{NP} the man]]
 b. John [_{VP} [_V thinks] [_S, that [_S I am here]]]

In (106)a, V assigns a Θ -role to N. By (105), this creates a path {NP, VP}. In (106)b, V assigns a Θ -role to the head of S'. By (105), since S' is maximal, there will be a path {S', VP}.

Given the usual situation, in which maximal projections of X dominate all other projections of X, and all non-maximal projections of X are dominated by maximal projections of their heads, and given (105), most paths created by Θ -marking will contain exactly two members. The paths in (106), for example, had two members each. Now consider the PCC and the definition of overlapping given in (93)-(94). We can prove as a simple theorem that a path with two members can never cause a PCC

violation with any other path:

Suppose we have two paths: $P = \{a, b\}$, and Q , such that P and Q violate the PCC. Then P and Q overlap. They overlap if their intersections are non-null and non-singleton. Suppose $a \notin Q$, and $b \notin Q$; then P and Q do not overlap and do not violate the PCC. Similarly if $a \in Q$, and $b \notin Q$, and if $a \notin Q$ and $b \in Q$. If $a \in Q$ and $b \in Q$, then Q contains P , and P and Q still do not violate the PCC. Since this exhausts the possible intersections of P and Q , Q does not exist.

Thus, extending the notion of path to include Θ -marking will be innocuous and undetectable in the vast majority of cases: only in the odd case where a Θ -marking path will have three members will effects be detectable. But this in turn suggests that we should find some way of simplifying (105), so that the odd disjunction of \bar{A} -binding and Θ -role assignment makes sense. Subject condition effects in small clauses and S' -deletion infinitives is certainly a topic we do not expect Universal Grammar to have a special paragraph about.

We can eliminate the disjunction from (105) by adapting an idea proposed by Stowell (1981), as a way of eliminating a similar disjunction in the ECP. Interestingly, his proposal works without problems in the framework we are developing, while it raises serious difficulties in a framework assuming the ECP. Stowell suggests that when a Θ -marker T assigns a Θ -role to an X it governs, T actually \bar{A} -binds X . This is because each Θ -marker has associated with it a thematic grid: essentially a series of slots, one for each Θ -role assigned by T . Θ -marking X consists in entering the referential index of X in the appropriate slot in the thematic grid of T . Thus, (106), for example has roughly

the representation in (107):

(107) John [_{VP} [_V saw
[_(theme)]_i] [_{NP} the man]_i]

In (107), the object NP is \bar{A} -bound by V, by virtue of being bound from the thematic grid of V. Stowell notes that if there is a one-to-one correspondence between Θ -roles and slots in the thematic grid, the Θ -criterion can be rephrased in terms of binding from slots in the grid:

- (108)a. Each slot in a Θ -grid is associated with exactly one argument.
b. Each argument is associated with exactly one slot in a Θ -grid.

Stowell (1980, 1981) shows a number of applications of this idea. In particular, Stowell claims that the assumption of a Θ -grid allows a simplification of the definition of government in (3), necessary for the notion of proper government used in the ECP. Recall that the ECP (cf. (1)-(3) of this chapter) requires that an empty category be either coindexed within a maximal projection or governed by a lexical node (\neq AGR). Recall that our second question about the ECP asked why the ECP should contain this disjunction. Stowell suggests that the Θ -grid allows this disjunction to be eliminated from the definition of government and proper government. Since most categories lexically governed by a governor G also receive a Θ -role from G, it follows that these categories will be locally \bar{A} -bound by a slot in the Θ -grid of G. These cases thus already fall under the coindexation clause of the definition of government (or proper government). This in turn suggests that lexical government might be removed from the ECP, in favor of local \bar{A} -binding within a

maximal projection -- thus eliminating the disjunction. In other words, Stowell suggests that the notion of lexical government be excluded from the notion of proper government: proper government would require only local binding. Extraction of the man in (107) (what man did John see?) would be allowed for the same reason as extraction of the subject is allowed (who saw the man?): local \bar{A} -binding within a maximal projection.

The problem with Stowell's suggestion, in an ECP framework, is that G may lexically govern a category C without G assigning a θ -role to C. This happens, for example, when VP contains a non-argument, like the accusative expressions of time in Russian, which we considered in Chapter 2:

- (109)a. Ivan spal celuju nedelju
Ivan slept a whole week (acc.)
- b. [skol'ko nedel']_i Ivan spal t_i
how-many weeks Ivan slept
'how many weeks did Ivan sleep?'

It is unlikely that spal 'slept' assigns a θ -role to the expression of time, which in (109)b seems to be properly governed. (Recall also that these expressions appear in the genitive of negation, which we discuss again in 2.4 of Chapter Four.) One might propose that a special θ -role is assigned here, but this would be ad hoc. No verb lexically requires such an expression, suggesting that such expressions do not fall under the θ -criterion.

A more serious problem, that Stowell discusses, concerns subjects of small clauses and S'-deletion infinitives. As we saw in (104)a-b, these may be "long-moved" by WH-movement. In an ECP framework, this shows that they are properly governed. The only candidate for proper

governor, however, is the verb that governs the small clause or S'-deletion clause. This verb assigns a Θ -role to the whole clause that it governs, but surely not to the subject of the clause. If the index of the subject of the clause is not entered in the Θ -grid of the higher verb, it will not be properly governed, contrary to the facts of (104)a-b. Thus Stowell is forced to assume that the index of the subject is entered in the Θ -grid of the higher verb; he suggests that the small clause or S'-deletion infinitive lacks an index itself, forcing the subject to supply one for it.

It thus appears that in order to use the idea of a Θ -grid to eliminate the disjunction between government and binding in the ECP the notion of Θ -grid must be considerably expanded, to the point where all categories governed by G are entered in G's Θ -grid.²⁹ The Θ -grid thus becomes a government grid instead. If the Θ -grid is replaced by a government grid, however, then the disjunction we have eliminated from the ECP and the definition of government simply pops up in another place: we must again ask what the purely structural relation of government and the notion of binding have in common. Linking Θ -marking and \bar{A} -binding has some plausibility: each relationship involves the assignment of a referential property from one element to another. Giving an argument a Θ -role or binding a variable are operations that allow the argument or variable to refer in a proposition. Government does not have this property.

We argue that the problems raised by Stowell's proposal reveal a deficiency in the ECP, rather than a deficiency in the notion of Θ -grid. In this connection, it is useful to compare the ECP with the PCC on an

abstract level, as they relate to \bar{A} -binding. Looking at the raw facts, we want to account for the observation that complements may be \bar{A} -bound either "long" or "short" -- from within their clause or from outside of it. Subjects of tensed sentences, however, can only be bound "short" in a non-pro-drop language. In stating the constraint that encompasses these observations, there are two possible approaches. One may stipulate the environments in which "nothing special" happens: say where an empty category must be in order to be "long" bound. Alternatively, one may stipulate where "something special" does happen: say where an empty category cannot be if it is "long" bound. The ECP takes the first approach. It states that an empty category must be in a certain position (lexically governed) if "long" vs. "short" binding does not matter, and says that if it is in another position (not lexically governed), only "short" binding is possible. We are abstracting from issues related to PRO, but the characterization is essentially correct. The PCC, on the other hand, states that there are certain positions where "something special" does happen -- namely, when the path from a certain position overlaps another path.³⁰

On a PCC account, therefore, we do not need to say anything special about "normal" cases, where a category generates a path that does not overlap any paths it does not contain. It happens that extraction of a complement will generally fall in this category (given our theorem about two-member paths). Using Stowell's notion of a Θ -grid, however, we can characterize certain environments in which something special does happen -- namely, positions inside the subject of a small clause or S'-deletion infinitival. Since we do not need to extend the notion of Θ -grid in the way Stowell has to, under the ECP, the Θ -grid retains its

initial plausibility as a unification of Θ -marking with \bar{A} -binding.

We can thus return to the definition of paths in (69), where nothing special is stipulated about Θ -marking. We assume crucially that the binding relation involved in Θ -marking is between a non-empty slot in the Θ -grid of a Θ -marker and an empty position in the head of the recipient of the Θ -role. Thus, all directly Θ -marked categories are "empty" and " \bar{A} -bound" in the sense relevant to the definition of paths.

Summing up this discussion, we have suggested that the paths {INFL', S, VP} and {PredP, Pred*, VP} are paths between VP and the category it Θ -marks. These paths exist by virtue of the Θ -grid proposed by Stowell: Θ -marking is \bar{A} -binding. From these paths, as we have seen, follow the Subject Condition effects in small clauses and S'-deletion infinitivals.

Finally, let us consider briefly some problems raised by our analysis, and by the notion of Θ -grid. First, we ask whether only direct Θ -marking of a complement by a lexical node is to be seen as \bar{A} -binding (as stipulated), or whether indirect Θ -marking of a subject is also \bar{A} -binding. Stowell assumes that the Θ -grid replaces the sub-categorization frames of Chomsky (1965): from this it follows that only direct Θ -marking is represented in a grid. Suppose we were to treat indirect Θ -marking as \bar{A} -binding. That is, suppose phrasal nodes like VP also contain a Θ -grid, when they assign a Θ -role. Then a Θ -marked subject would be \bar{A} -bound by the Θ -grid associated with the phrasal node. Suppose the subject of a tensed sentence is assigned a Θ -role by VP, and hence is \bar{A} -bound by VP. It would follow that a path would begin with the first maximal projection that dominates VP, and end with the first maximal projection that dominates the subject that receives the Θ -role.

lines: perhaps the relevant chain is actually (John, [[]_i, V]), where it is the slot in the verb's Θ -grid that bears the Θ -role for the chain. The problem here, of course, is that the verb-internal slot that binds the object position is not an A-position, under the definition given. Clearly we want the verb-internal slot not to count as an intervening \bar{A} -binder for cases of A-binding. How to effect this is, however, unclear.

Thus, while certain problems arise when we try to integrate Stowell's proposal of a Θ -grid into GB theory, it seems clear that the Θ -grid, when combined with the paths theory and the PCC, predicts the behavior of subjects of small clauses and S'-deletion infinitivals. Note finally the crucial role played by the fact that paths begin with the first maximal projection dominating the node responsible for the paths, rather than with the first branching node. If extraction of the subject of a small clause or S'-deletion infinitival led to a path beginning with Pred* or S, we would predict that such an extraction would be ungrammatical, just as long movement of the subject of a tensed S is ungrammatical. As we noted earlier, the use of maximal projection in the definition of a path raises problems with respect to the status of INFL' and S in X-bar theory. As we see it, these problems can be solved for now by stipulation, while the problems that would arise if we redefined paths in terms of branching nodes would be insoluble: we could not capture the Subject Condition facts discussed in this section.

4.7 Conclusion

In this chapter we have proposed a condition on paths created by \bar{A} -binding, the Paths Containment Condition, that accounts for "Crossing Dependencies" effects, Complementizer-Trace phenomena, and Subject Condition

effects. In accounting for the CTP, we have taken a step toward subsuming ECP effects under the PCC. It should be clear that other ECP effects that we considered in the previous chapter -- governing movement in LF -- can be made to fall under a PCC account. In the next chapter we will briefly discuss how our Chapter Two analysis of Russian QPs might be translated into a PCC approach. The bulk of the next chapter will be concerned with applying the PCC to a wider class of cases: constructions with multiple gaps. We will also explore more deeply certain issues touched upon in this chapter -- in particular, the reasons for the path between INFL and COMP that derives CTP and Subject Condition effects in tensed clauses.

FOOTNOTES: CHAPTER THREE

1. More accurately, given our previous conclusions, the Projection Principle will require that a verb have some number of complements at D- and S-structure, while c-selection (or s-selection) will require at LF that the complements be of some particular categorial type.

A number of interesting questions arise with respect to constraints on the distribution of complements. For example, it has often been taken as an argument for a base component consisting of context-free rewrite rules that such rules provide substantive constraints on possible lexical entries. For example, given (2), or some appropriate complication to accommodate PP's, etc. (cf. Jackendoff 1977), we can capture the fact that no verb can subcategorize more than two complements -- i.e. that there exist no verbs with subcategorization features like +[NP PP S'], where all three constituents are obligatory. Nonetheless, as Chomsky (1981a, 31) notes, this argument is false. Although rule (2) states this generalization, it does not simplify the grammar at all, and is, in fact, otiose. Verbs that take one or two complements must still be listed as such, and no explanation is provided for the absence of verbs with three or more complements.

A more serious argument for rules like (2) might call attention to the fact that (2) limits the types of complements a verb may take, and their order. On the question of order, see text below. On the question of type, one might argue that (2) captures the generalization that if a verb takes two complements and one of them is sentential (or clausal), the other will be an NP. That is, there are no verbs taking two sentential complements, meaning, for example, 'cause to imply':

(i) *John caused-to-imply that John came home that he was drunk

With respect to this generalization, (2) does allow us to simplify lexical entries, since any verb with a subcategorization frame like +[NP S'] could be assigned a reduced frame like +[XP S'], where rule (2) will tell us that X = N. This argument too seems to be without force, however, since the restrictions on the first complement of a double-complement verb go beyond categorial identity. In fact, the first complement of a verb taking a second sentential complement must be an animate (or perhaps human) NP, and not just an NP. We can see this by comparing English and French verbs meaning to force. English allows the following sentences:

- (ii)a. Mary forced Bill to work
- b. Mary forced the car to work
- c. ?Mary forced there to be a confrontation with the police

French, however, allows only a sentence corresponding to (ii)a:

- (iii)a. Marie a forcé Pierre à travailler
- b. *Marie a forcé la voiture à marcher
- c. *Marie l'a forcé (à) y avoir une confrontation avec la police

This is because the English verb is ambiguous between a control subcategorization for an NP and an S' and a uniclausal subcategorization with S'-deletion:

- (iv) a. Mary forced NP [_S, PRO to VP]
- b. Mary forced [_S NF to VP]

(ii)a-c may have the structure of (iv)b; only (ii)a may have the structure of (iv)a, since NP in (iv)a must be animate. Since French entirely

lacks structures like (iv)b (cf. *considérer Jean avoir gagné 'consider John to have won'), only the sentence acceptable under structure (iv)a is possible in French.

As Chomsky showed in Aspects, it is highly unlikely that a base component should be allowed to introduce features for animacy. Given that all animate constituents are NPs, whatever component of the grammar explains the generalization discussed here, it will need to refer to animacy, not to noun-ness: this component will not be the base component.

2. The derivation of the position of the subject presented here derives from a suggestion in Stowell (1981, 281).

3. An answer might be found in the mechanism by which clauses escape the Case filter in Chomsky (1981a). As indicated below in the text, Chomsky suggests that a chain must contain a member in a Case-marked position or be headed by PRO to bear a θ -role. By stipulation, NPs can only receive a θ -role by membership in a (possibly singleton) chain. Hence an NP or its trace must bear Case or be PRO. A clause, on the other hand, may bear a θ -role without being a member of a chain; hence we expect the Case or PRO requirement not to hold.

Suppose that a subject is coindexed with its predicate by a rule of Predication, as argued in Williams (1980). If we take this coindexation to form a chain containing the subject and predicate, then this chain will always have to bear Case or be headed by PRO, regardless of whether the subject is NP or S'.

The facts discussed in the text might also be related to the impossibility of Case transmission into a non-Case-marked subject position,

both in English and in pro-drop languages like Italian:

- (i) a. there were considered to have arrived some men
b. *there were considered some men to have arrived.
- (ii) a. è ritenuto essere arrivato Giovanni
is believed to have arrived Giovanni
b. *è ritenuto Giovanni essere arrivato

Perhaps coindexation of a subject with its predicate makes the predicate the local binder of the subject, impeding Case transmission from a higher INFL.

4. The argument is as follows. Assume the Binding Theory, and the claim that a position in VP is filled by an argument if and only if it is a Θ -position (preventing Raising to Object). Consider an element L which is a member of a chain C. If L is in governing category G, it can only be bound by a subject in G -- i.e. by the subject of L's clause. This is because the subject of L's clause is the only position that satisfies both Principle A of the Binding Theory (is in the domain of the nearest TNS and subject to L) and the Θ -criterion (can be a non- Θ -position).

Now suppose L is indeed locally A-bound in chain C by a subject in G. By the Θ -criterion, this subject must be a non- Θ -position. If L is governed by the V, and if the VP containing V does not assign a Θ -role to the subject that binds L, then by Burzio's generalization it will not assign Case to L. Since by the Θ -criterion and the Binding Theory, VP will not assign a Θ -role to the subject, it will indeed not assign Case to L.

Now consider the subject that locally A-binds L. Suppose that it is also locally A-bound. Then the story begins all over again, and this subject itself is an L and may not be assigned Case. Only if the subject is not locally A-bound -- i.e. if it is the head of C -- may it receive Case in situ, and, unless it is PRO, it must receive Case in situ to satisfy (19).

Thus, by induction, we see that only the head of C may be in a Case-marking position, for all L governed by V. The only problematic cases are elements L that might be governed and assigned Case by a preposition P. If PP is not a governing category (certainly the case in English, though perhaps not in Romance languages), then Burzio's generalization predicts nothing about the status of L as an object of P. As it happens, there is evidence (Hornstein and Weinberg 1981) that some principle forces P to be absorbed by the verb in passive constructions -- possibly the ECP, but see our discussion in section 3. This would make apparent prepositional objects into verbal objects, bringing them under Burzio's generalization and into the argument sketched above.

5. We assume that the subject of the complement of want is assigned Case by want, and not by an abstract or deleted complementizer for. We argue for this in Chapter 5. For now, note that, in any case, the variable bound by who will be to the left of INFL'. The presence of a complementizer for in COMP will also force a Case-marked subject to be preverbal. Similar examples can be constructed with uncontroversial S'-deletion verbs. See below.

6. Cf. Postal and Pullum (1978). Note I got there to be a gloss on the furniture: get seems to be an S'-deletion verb with an optional θ -role assigned to the subject. Thus, it has the properties characteristic both of consider and of seem. One might ask whether aspectual get (= become, come) allows contraction: ?John gotta be a bigshot on his own merits.

Also, does say contract when it is a control verb? Cf: ??John saidda stop at the corner. If the structure is John said [_S to stop ... PRO], we predict contraction to be possible. Note that PRO is obligatorily not controlled by the subject of say here (cf. Perlmutter 1971; Manzini 1981). If contraction is impossible here, it would support Pullum and Postal's theory of contraction phenomena over ours. On their analysis, which translates naturally into Government-Binding terminology (but is a stipulation on any theory), contraction is possible iff the subject of the lower clause is locally A-bound in the next clause up, which PRO is not, in our example with say. We might note, with Manzini, that PRO here is not arbitrary, but rather is controlled by the understood indirect object of say (cf. *John said PRO to kill oneself, which should be acceptable if PRO is arbitrary). If so, we might suppose that the indirect object is present as some sort of empty category in PF, blocking contraction.

7. We might also ask if there are empirical arguments distinguishing our theory from the visibility theory. For example, if it could be shown that NP-trace, which by (19) must belong to a chain headed by a Case-marked NP or PRO, receives Case features by transmission, then it should block contraction, under the visibility hypothesis. A distinction between the visibility hypothesis and our own lies in the distinction

between "bearing Case features" and "receiving Case features in situ from a local Case assigner". The former is relevant to the visibility theory, the latter to ours.

If NP-trace can receive Case features by transmission, then our theory is empirically superior. There is some evidence that PRO may in some languages inherit Case features from its controller, which is reflected in various agreement phenomena in the clause containing PRO (cf. Neidle 1982; Simpson, forthcoming; Franks 1982). If such PROs do not block contraction, they too would provide evidence in support of our proposal.

8. Note that the adjacency requirement on Case assignment in French, though relaxed to overlook certain adverbs, still holds in the important cases (cf. also discussion in Stowell):

- (i) j'ai persuadé Jean de venir 'I persuaded Jean to come'
- (ii) *j'ai persuadé de venir Jean (except as Heavy NP Shift)

9. Postal and Pullum (1982) raise the question of various cases where the structural description of the simplest contraction rule appears to be met, but where contraction seems to be blocked:

- (i) I don't want [_S PRO to flagellate oneself] to become standard practice in this monastery (*wanna)
- (ii) I don't want [_{NP} anyone [_S who continues to want]] to stop wanting (*wanna)

About such examples, we agree with Chomsky and Lasnik (1978) that "any

reasonable theory of contraction will be restricted to within an intonational phrase", a condition surely not met in (i) and (ii).

Nonetheless, one might note that certain syntactic locality conditions which may be relevant for phonology are not met by (i)-(ii) and other examples that Postal and Pullum have cited. Note, for example, that in each case, two L-containing nodes (Chomsky 1973), S and S' in (i), NP and S' in (ii), separate want from to. We might suppose that the two items to be contracted cannot be separated by more than one L-containing node.

N. Chomsky has also pointed out that similar restrictions obtain with Romance restructuring (Rizzi 1978b; Burzio 1981), which only applies between a restructuring verb and its immediate complement.

10. Another argument for the INFL' constituent, in S'-deletion infinitivals, results from some points raised by G. Carden. If INFL' is a constituent, we might expect it to move like a constituent. In particular, consider the possibility of INFL' undergoing a process analogous to Heavy NP Shift. Suppose, for clarity's sake, that this involves adjunction to VP:

(i) I [_{VP} [_{VP} consider [_S John [_{INFL'} e]_i]] [_{INFL'} to be a fool]_i]

If such structures are allowed, certain purported arguments for Raising to Object might disappear, e.g. the possibility of interpolating adverbial material belonging to the higher clause between the subject and predicate phrase of an S-adverbial (Postal 1974, 146 ff., developing an argument of S. Kuno's; (ii) b-c are my examples):

- (ii) a. Jane believes Bob, if I am not mistaken, to be Hungarian
 b. I consider John with all my heart to be mad
 c. I might consider Mary tomorrow to have killed John yesterday

Postal notes that adverbs cannot normally lower into a subordinate clause:

- (iii) a. *Jane believes that Bob, if I am not mistaken, is Hungarian
 b. *I consider that John with all my heart is a genius
 c. *I might consider that Mary tomorrow has killed John yesterday

If we rule out INFL' movement from tensed clauses (by the Binding Theory?) but allow it from infinitivals, we explain the contrast:

- (iv) I [_{VP} [_{VP} consider [_S John [_{INFL'} e]_i] with all my heart]
 [_{INFL'} to be mad]_i]

Presumably, the sequence to-VP could only move if it were a constituent.

Further evidence that "Heavy NP Shift" of INFL' is possible comes from the fact that parasitic gaps (see Chomsky 1981b, our Chapter Four, and references therein) bound by INFL' are almost possible:

- (v) ??John considered Mary __, without really believing Bill __, to be crazy.

Parasitic gaps must be licensed by a real gap bound from an \bar{A} -position, as we shall see, motivating movement of INFL' in (v).

11. Hendrick (1979) discusses the sentences in (44) as counterexamples to his Association Principles, which include a version of the Crossing Constraint as a condition on rule application. He suggests that some

examples of crossing A-binding are unacceptable (stars his):

- (i) *these writers_i, the teachers_j recommend t_i to each other_j
- (ii) *those two books_i, the teachers_j recommended t_i to each other_j

He suggests that the "referential antecedent" for trace in the topicalizations of (i)-(ii) creates the crossing violation. I do not detect any unacceptability in these examples, which seem to have the same status as Chomsky's sentences.

Chomsky's Czech example, which I have omitted, contains crossed dependencies between clitics and their traces. Bordelois (1974) and Hendrick (1979) both extend their versions of the Crossing Constraint to clitic configurations. I will have nothing to say about clitics in this thesis. It is not clear that the theory I will develop can handle the cases discussed by Bordelois and Hendrick.

12. As noted in Pesetsky (1982), this rule extends to simple relatives:

- (i) l'homme qui (*que) aime Marie
the man that loves Marie
- (ii) l'homme que (*qui) Marie aime
the man that Marie loves

As before, however, it remains mysterious why the rule is limited to binding of a nominative trace -- that is, why (i) is ungrammatical with que. Under an ECP explanation, we could say that qui is the allomorph of que used when COMP properly governs in S, since a VP node blocks proper government. Even this remains a stipulation, however.

13. Kayne (1981a) suggests that that inherits its index from the head of the relative, and not from WH in COMP. This theory captures the appropriate distinctions without stipulation, since COMP will contain relative WH at S-structure only when locally linked to the head. However, Kayne's theory raises some questions if this "local linking" or "predication" (Chomsky 1977) does not take place until a level past LF (LF'), as argued by Chomsky (1981b). Chomsky uses this assumption to explain the absence of weak crossover effects in relative clauses. If the ECP applies before LF', then an unindexed that will, of course, violate the ECP. If the ECP is replaced by a Crossing Condition, the same argument obtains. Thus the question remains open.

14. E.g. the tendency to interpret a gap in a too/enough structure as bound by the subject of the construction:

- (i) *those nails_i are too blunt for any wood_j to be soft enough for you to hammer t_i into it_j (Fodor's (137))

15. We assume that (60)f is not to be ruled out by any version of the Crossing Constraint, which seems not to apply to A-binding. We discuss this question in greater detail in Chapter Four, section 3.3, and in Chapter Five.

16. Sentences like (29) are no worse than other violations of the WH-island condition, such as (i):

- (i) what subject do you know who to persuade to discuss

Thus, their perhaps reduced acceptability cannot be blamed on problems with reanalysis or preposition stranding. Many of the sentences in (34)-(43) make the same point.

17. (69)iii contains a redundancy: if nodes dominate themselves, we can eliminate the expression " $(x = \alpha)$ ", since this case will fall under the third disjunct, " $(x \text{ dom. } \alpha \ \& \ \neg x \text{ dom. } \beta)$ ". Similarly, if nodes do not dominate themselves, we can eliminate the expression " $(x = \beta)$ ", since this case will now fall under the third conjunct.

Note the similarity between (69) and the definition of c-command argued for by Aoun and Sportiche (1982) and discussed in Chapter One. If a node γ c-commands all nodes dominated by the first maximal projection that dominates γ , then if nodes do not dominate themselves:

for $T = \{x \mid \underline{t} \text{ c-commands } x\}$
 for $B = \{x \mid \underline{b} \text{ c-commands } x\}$
 for $C = \{x \mid \underline{x} \text{ c-commands } \underline{t}\}$
 for β the first maximal projection dominating \underline{b}
 the path between t and b is a set of nodes P such that
 $P = \{x \mid x \in \{B \cup C \cup \bar{T}\} \vee (x = \beta)\}$

If nodes dominate themselves, replace β with α , the first maximal projection dominating \underline{t} , as in the text.

18. I am extremely grateful to Lec Wetzels, and also to Annie Zaenen and Hilda Koopman for help with Dutch data. They are not responsible for the use to which I put their sentences.

19. Van Riemsdijk (1978a, 108-112) shows that naar is the head of the phrase. toe may be incorporated, and naar + NP reordered, much as in

(81)c. The crucial point is that naar-toe as a whole does not undergo the incorporation seen in (81).

20. This reanalysis must apply after verb raising, which treats incorporated prepositions as part of the verb to be moved, but never moves naar-toe. Note that we have not discussed where the Crossing Constraint applies, as yet.

21. Some further examples:

(i) het boek is gemakkelijk om [t voor] [naar de bibliotheek] te gaan
 the book is easy for for to the library to go
 'the book is easy to go for to the library'

(ii) de bibliotheek is gemakkelijk om [voor het boek] [t naartoe] te gaan
 'the library is easy to go to for the book'

(iii)***de bibliotheek, waar_i het boek_j gemakkelijk is om [t_j voor]
 [t_i naartoe] te gaan

(iv) *het boek, waar_j de bibliotheek_i gemakkelijk is om [t_j voor]
 [t_i naartoe] te gaan

Examples with WH-violations in questions appear to be uniformly bad in Dutch, making examples like (29)-(30) impossible to test.

22. For example, it is unclear what happens to the paths account of these phenomena if we accept Van Riemsdijk's argument that P-stranding is only possible when the object of P is locally bound from a so-called R-position, located, in our terms, at the left margin of INFL' (cf. Stowell 1981 for the suggestion that this is a head of INFL'). We leave these questions for a more detailed analysis of Dutch preposition

stranding in a paths framework.

23. Engdahl (1981) notes that parasitic gaps provide evidence for a different structure for purpose clauses, in which they are outside the VP dominating the object. This is because parasitic gaps must be in an anti-c-command relation with the movement-derived gap that licenses them, and sentences like (i) are possible:

(i) what book did you buy t in order to review e

(where e is the parasitic gap). The parasitic gap will not be c-commanded by t only if the following structure, or something like it, is assumed:

(ii) What book did you [_{CP} [buy t] [in order to review e]]

Note that under this structure double WH-movement of the type discussed in the text should be completely impossible, given a paths account of Crossing phenomena. The structure of (ii) has the same properties, vis-a-vis paths, as does double preposition stranding if reanalysis does not occur, as discussed in connection with (72). The path from a trace in the in order to clause to its \bar{A} -binder, wherever it may be, will include the nodes of the in order to clause, which the path from an object trace will not. On the other hand, the path from the object trace will include VP, which the path from the trace in the in order to clause will not. Hence containment will not obtain. Thus sentences like (88)a could not exist at all, with a structure as in (ii).

This problem can be resolved, however, if we note that the argument from parasitic gaps shows only that in order to clauses may be outside of VP, as in (i), not that they must be (when parasitic gaps are not involved). Similarly, the argument from double extractions

like (88)a shows only that in order to clauses may be inside of VP, not that they must be (unless double extractions are involved). If we suppose that in order to clauses are structurally ambiguous, occurring both inside and outside of VP, there is no problem with either double extractions or parasitic gaps.

24. We could similarly alter (69)i, since S' immediately dominates COMP, though we would either have to call COMP a lexical node, or allow it to inherit an index from its contents, to avoid calling COMP β , which would prevent COMP-to-COMP movement, among other things.

25. Alternatively, we might allow only empty categories in A-positions to generate paths, eliminating (iii) entirely. Our analysis of subjunctive complements in Chapter Five will suggest that this move is not desirable. Additionally, the "overlapping" convention will prove essential in Chapter Four, where we discuss parasitic gaps inside COMP.

26. This seems true, contra Chomsky (1977), unless sentential subjects are dominated by NP, which we do not assume.

27. If, as we believe (cf. Chapter Five) infinitives do not generally contain a path between INFL and COMP, we predict, perhaps falsely, that contrasts like (100)a-b should not be found in infinitives. We cannot vouch for cases like (100)a, but the absence of sentences like (i) below might be attributable to the absence of sentences like (ii):

(i) *I wonder what war [during t] to die

(ii) ??I wonder how [during this war] to die

28. We shall show later that LF movements are also subject to the PCC (recall the ECP, Chapter 2). If so, we can explain Rizzi's (1980) observation that quantifiers in preverbal PPs may not take wide scope: i.e., where non is a scope marker, (i) is ungrammatical:

- (i) ?non credo [che a nessuno Gianni l'abbia detto]
'I do not believe that to nobody Gianni said it'

Assuming there is pied-piping in LF, (i) has the paths status of (98).

We make an even stronger prediction, however: a quantifier in a preverbal PP should be confined to its PP, or else the resulting structure would have the status of (100)b. This should be true even without a wide scope marker:

- (ii) credo [che a nessuno Gianni l'abbia detto]

As Rizzi indicates, (ii) is grammatical. We do not know how to test whether nessuno has sentential scope or not; so the question remains open.

29. Jaeggli's (1980b) proposal that the ECP involves government by a subcategorization feature ("s-government") rather than by a category ("c-government") faces the same difficulties as Stowell's proposal, and for similar reasons. Jaeggli suggests that the subject of a small clause or S'-deletion infinitive is the head of the construction, and therefore s-governed. This seems unlikely, particularly if one adopts our analysis of secondary predication in Chapter 2.

30. Note that the NIC of Chomsky (1980) is like the PCC in this respect: it states where "something special" does happen -- namely in positions

marked with nominative Case. Arguments that the CTP should not be subsumed by the NIC are given in Chomsky (1981a) (and references cited there, particularly Freidin and Lasnik (1981)).

CHAPTER FOUR: MULTIPLE GAPS AND PATHS

1.0 Introduction

In this chapter, we will show how the PCC helps to explain a variety of phenomena involving multiple gaps.

In 2.1, we will take up the "parasitic gap" construction studied by Taraldsen (1981), Engdahl (1981) and Chomsky (1981b). In particular, we will consider certain apparent violations of the Subject Condition, treated by Kayne (1982). To explain these phenomena, we will adapt Kayne's analysis to our framework: we suggest, with Kayne, that the paths between an \bar{A} -binder and multiple gaps locally bound by it are united in a single "forked" path. In this section we will discuss only briefly the further properties of parasitic gaps, reserving more detailed consideration for section 4.

In 2.2, we will compare our adaptation of Kayne's theory with the actual analysis proposed by Kayne. We present a case in which the two analyses can be distinguished, involving parasitic gaps in COMP. In 2.3, we consider parasitic gaps in left branches in French, English and Russian, in the light of a comparison of our approach with Kayne's. In this connection we will reconsider our analysis of Russian quantificational phrases from Chapter Two in 2.4, and suggest a translation into the PCC framework. The translation introduces infinite paths, rays with only a lower bound, as the equivalent of the unbound NP traces in Chapter Two. The reanalysis involves some difficulties, which will be briefly noted. In 2.5, we show that the PCC theory of parasitic gaps makes correct predictions about a class of examples in which Crossing effects interact with parasitic gaps.

In section 3, we turn our attention to coordinate structures. In 3.1 we show that the PCC can derive that part of Ross's Coordinate Structure Constraint that does not fall under the A/A condition. In 3.2, the "Across-the-Board" exceptions to the Coordinate Structure Constraint, studied by Ross (1967) and Williams (1978), are seen to derive naturally from the interaction of our PCC theory of coordinate structures with our theory of multiple gaps, motivated in section 2. In 3.3 we reconsider the path between INFL to COMP and its interaction with our derivation of the Coordinate Structure Constraint. We motivate this path as the result of TNS movement, and show its interaction with other aspects of our PCC theory. We suggest an explanation for our Chapter Two exclusion of expletives and PRO from the ECP (now subsumed by the PCC). This explanation, in turn, suggests a way to eliminate the stipulations about A and \bar{A} -binding in the definition of paths.

Finally, in 4.0, we bring our theory of multiple gaps together with the analysis of parasitic gaps in Chomsky (1981b). We discuss the interaction of multiple gap phenomena with subjacency. We suggest that this interaction argues for an approach that allows simultaneous application of Move α from several positions, or equivalently, which states subjacency as a condition on representations at S-structure and postulates an additional level, SS', intervening between S-structure and LF. On this latter approach, the PCC applies at S-structure, SS' and at LF, but subjacency only applies at S-structure.

2.0 Paths and Parasitic Gaps

2.1 Parasitic Gaps and Subject Condition Violations

In the previous chapter, we have considered constructions where each

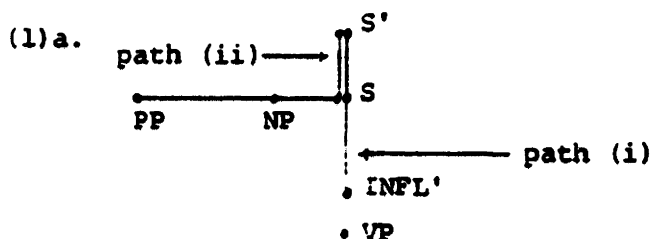
\bar{A} -binder locally binds a single empty category. (Recall that X locally binds Y if there is no distinct Z which also binds Y, such that X c-commands Z. The local binder is thus the "first binder up the tree".) Let us now consider how the PCC could apply to constructions in which several empty categories are locally \bar{A} -bound by the same element: in this section, so-called "parasitic gap" constructions. For present purposes, we may remain vague about the source of these gaps, and of their indices. We will return to these questions in great detail in section 4 of this chapter. For now, we simply stipulate that the structures presented exist, with the bindings as indicated, and we examine their properties.

In this section, we will motivate a treatment of parasitic gaps by analyzing certain facts about parasitic gaps and the Subject Condition effect which were discovered by Kayne (1982). Our analysis adapts Kayne's proposal to our framework. This adaptation will be the key to all our results in this chapter.

Let us consider tensed clauses first. Taraldsen (1981) and Kayne (1982) notice that Subject Condition violations like (1)a often improve when a second gap is added, as in (1)b:

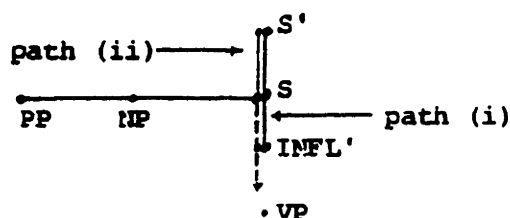
- (1)a. *a person [_S, who_i [_S [_{NP} close friends [_{PP} of e_i] [_{INFL}, [_{VP} admire me]]]]]
 b. ?a person [_S, who_i [_S [_{NP} close friends [_{PP} of e_i¹] [_{INFL}, [_{VP} admire e_i²]]]]]

Recall what goes wrong in (1)a: a path (i) runs between INFL and COMP and a path (ii), between e_i and who in COMP. They overlap and violate the PCC:



Crucially, path (ii) is missing the INFL' node that path (i) has, and path (i) is missing the PP and NP nodes that path (ii) has.

Now suppose that some operation on path (ii) between e_i and who in COMP of S' could add the INFL' node to it. The addition of the INFL' node could add the INFL' node, and perhaps others, to it. This addition would cause it to properly contain path (i) between INFL and COMP, and violate the PCC no longer:



The dotted line shows the addition of INFL' to path (ii). We will propose that the second gap in (1)b accomplishes exactly this operation: it adds nodes to path (ii).

Consider now (1)b. Suppose -- contrary to what we will suggest -- that each gap in (1)b creates its own path to COMP. Consider the paths that would result:

(1)b. Paths

(i) Between INFL and COMP:

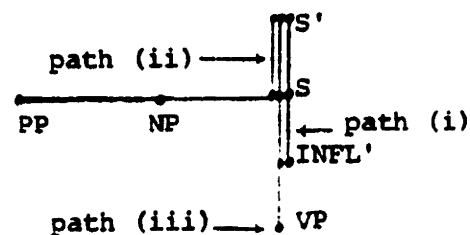
{INFL', S, S'}

(ii) Between e_i^1 and COMP:

{PP, NP, S, S'}

(iii) Between e_i^2 and COMP:

{VP, INFL', S, S'}



Paths (i) and (ii) are the paths already found in (1)a: they overlap, but neither one contains the other. Contrary to fact, we expect (1)b to be as bad as (1)a. (1)b, if the paths are as above, contains a further violation of the PCC. Paths (ii) and (iii), between the gaps and COMP, also overlap. Here too, neither one contains the other, and the PCC should be violated.

On the other hand, notice that path (iii) and path (i) overlap, and the PCC is satisfied, since path (iii) contains (i). Notice also that path (iii) contains the crucial INFL' node, which, if added to path (ii), would allow it to satisfy the PCC with respect to path (i).

(1)a-b thus present two related problems. First, why is the violation of the PCC caused by the path from inside the subject and the path between INFL and COMP apparently nullified in (1)b, where there is a second gap? Second, why does this second gap itself not create an additional violation of the PCC, as the interaction between paths (ii) and (iii) leads us to expect? As these two questions show, we expect (1)b to be, if anything, worse than (1)a, since it contains two violations of the PCC, while (1)a contains only one.

Each of these questions can be resolved in the same way. Recall that we remarked that the addition of INFL' to the path from inside a subject (e_i^1) could eliminate the PCC violation. Recall also that the path from the lower gap in (1)b includes INFL'. Both gaps are bound by the same \bar{A} -binder. Suppose that in such a case the two gaps and the \bar{A} -binder do not generate two distinct paths, but rather one common path, which is the union of what would be the separate paths from each. Notice that clauses (i) and (ii) of our earlier definition of paths in (69) of Chapter Three do not exclude the possibility that an \bar{A} -binder might locally bind two categories. What

is necessary is to change clause (iii), so that paths include the nodes dominating any of a set of \underline{t} s locally \bar{A} -bound by \underline{b} .

The change required in the definition of paths is minimal:

2) Definition of Paths (revised)

Consider T , such that:

$T = \{ \underline{t} \mid \underline{t} \text{ is an empty category locally } \bar{A}\text{-bound by } \underline{b} \}$

(i) for a set $A = \{ \alpha \mid \exists t \in T: \alpha \text{ is the first maximal projection dominating } \underline{t} \}$

(ii) for β the first maximal projection dominating \underline{b}

(iii) the path between the members of T and \underline{b} is the set of nodes

P such that:

$P = \{ x \mid (x \in A) \vee (x = \beta) \vee (\exists \alpha \in A: x \text{ dom. } \alpha \ \& \ \neg x \text{ dom. } \beta) \}$

By defining paths over sets of empty categories, and by introducing the existential quantifier into (iii), we in effect allow the "path union" of the sort described below.

What (2) says is that when a number of empty categories share an \bar{A} -binder, there is one big path that runs between all these empty categories and the \bar{A} -binder. There are no individual paths from each empty category to the \bar{A} -binder. As we shall see, the "one big path" that a set of identically bound empty categories creates is "forked": it has one upper terminus, but a number of lower termini in the tree.

More specifically, (2) asks us to consider a set T of empty categories that share an \bar{A} -binder \underline{b} . We now consider the set A consisting of the first maximal projection dominating each of the members of T . The path between the members of T and \underline{b} consists of all the members of A , the first maximal

projection dominating b, and all the nodes in between.

Let us now show explicitly how (1)b does not violate the PCC.

There are two e's in (1)b: e_i^1 and e_i^2 . Consequently there are two α 's:

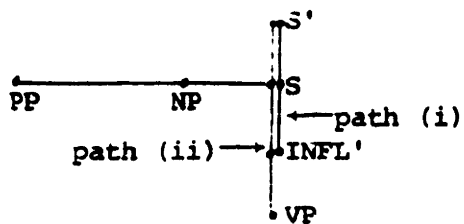
PP and VP. Thus, the paths are as follows:

(1)b. (i) Between INFL and COMP:

{INFL', S, S'}

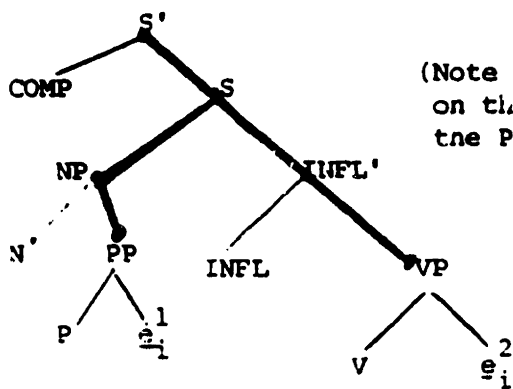
(ii) Between e_i^1 , e_i^2 and COMP:

{VP, INFL', PP, NP, S, S'}



The singly underlined nodes in (ii) are the exclusive contribution of e_i^2 to the path; the doubly underlined nodes, of e_i^1 ; and the other nodes, of both. Notice that path (ii) properly contains path (i); hence the PCC is not violated. Unlike the paths we have considered before, paths like path (ii) are not line segments. Rather, they take the form in a tree of an upside-down Y, which forks at the first node dominating both gaps. (We might construct examples with more than two gaps, in which case the path will fork several times.) Thus, path (ii) above looks like (3) below:

(3)



(Note that INFL', S and S' lie on the path indicated, satisfying the PCC.)

Kayne notes, as we expect, that not every "second gap" serves to nullify a Subject Condition violation. Rather, the second gap must be

lower in the tree than the subject that contains the erstwhile Subject Condition violation. Thus, (4)a contrasts with (1)b, and (4)b in turn contrasts with (4)a:

(4)a. *a person [_{S₁} who_i [_{S₁} you [_{INFL₁} [_{V₁}* [_{VP₁} admire e_i¹] [_{S₂} because
 [_{S₂} [_{NP₂} close friends [_{PP₂} of e_i²]] [_{INFL₂} [_{VP₂} became famous]]]]]]]]]

b. ?a person [_{S₁} who_i [_{S₁} you [_{INFL₁} [_{V₁}* [_{VP₁} admire e_i¹] [_{S₂} because
 [_{S₂} [_{INFL₂} [_{VP₂} know [_{NP₂} close friends [_{PP₂} of e_i²]]]]]]]]]]]]]

(4)b shows that a second gap is possible in constructions like this. (4)b does not violate the PCC:

(4)b. (i) Between INFL and COMP of S₂ⁱ:

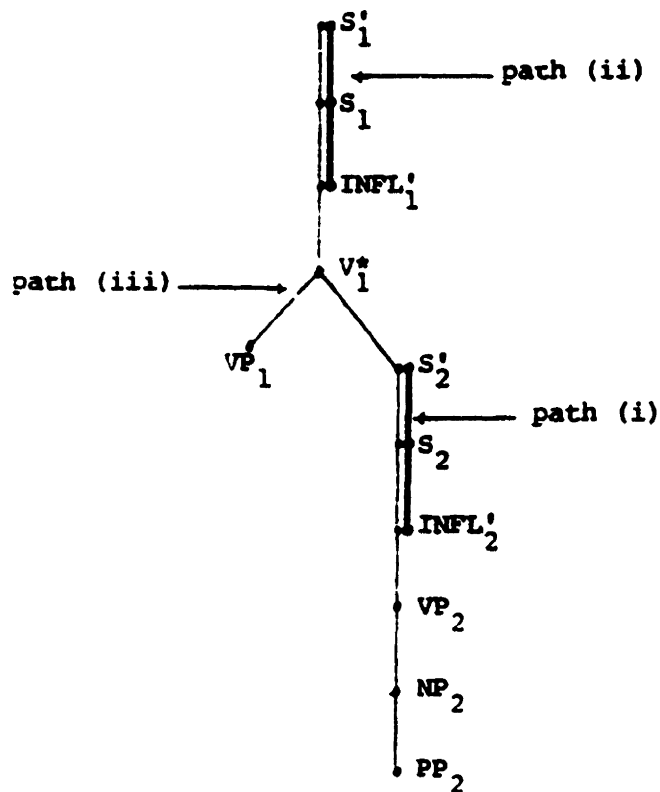
{INFL₂ⁱ, S₂, S₂ⁱ}

(ii) Between INFL and COMP of S₁ⁱ:

{INFL₁ⁱ, S₁, S₁ⁱ}

(iii) Between e_i¹, e_i² and COMP of S₁ⁱ:

{PP₂, NP₂, VP₂, INFL₂ⁱ, S₂, S₂ⁱ, VP₁, V₁ⁱ, INFL₁ⁱ, S₁, S₁ⁱ}



(4)a, on the other hand, does violate the PCC. As in (1)a, there is an $INFL'$ missing from the path from one of the empty categories, which is, of course, a member of one of the paths between $INFL$ and $COMP$. In turn, the path between the empty categories and their common binder contains a number of nodes not contained by one of the paths between $INFL$ and $COMP$:

(4)a. (i) Between $INFL$ and $COMP$ of S'_2 :

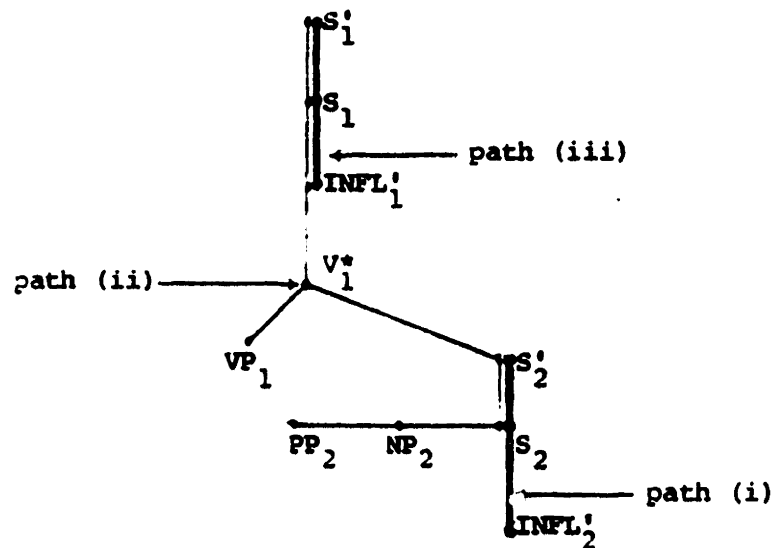
$\{INFL'_2, S_2, S'_2\}$

(ii) Between e_1^2, e_1^1 , and $COMP$ of S'_1 :

$\{PP_2, NP_2, S_2, S'_2, VP_1, V^*_1, INFL'_1, S_1, S'_1\}$

(iii) Between $INFL$ and $COMP$ of S'_1 :

$\{INFL'_1, S_1, S'_1\}$



Paths (i) and (ii) overlap and violate the PCC. Looking at the diagram above, it is obvious that S'_2 contains the basic configuration that characterizes Subject Condition effects. What happens above S'_2 is irrelevant, in fact. The violation is essentially the same as that in (1)a.¹

It follows also, of course, that the "lower" gap that nullifies Subject Condition effects cannot itself be a Subject Condition violation, unless there is a third gap lower still, which is not itself a Subject Condition violation:

- (5)a. *a person who [close friends of t] think that [enemies of t] hate Bill
 b. ?a person who [close friends of t] think that [enemies of t] hate t
 c. *a person who [close friends of t] think that [enemies of t] claim that [the parents of t] like Mary
 d. ?a person who [close friends of t] think that [enemies of t] claim that [the parents of t] like t

Also, we expect phenomena similar to those we have been considering to show up in small clauses and S' -deletion infinitivals. They do:

(6) ? $\{_{S'} \text{who}_i \{_{S'} \text{you} \{_{INFL'} \{_{VP} \text{consider} \{_{A*} \{_{NP} \text{friends} \{_{PP_2} \text{of } e_i^1 \} \{_{AP} \text{angry} \{_{PP_1} \text{at } e_i^2 \} \} \} \} \} \} \}$

(i) The VP-AP Path (θ -grid):

(AP, A*, VP)

(ii) Between e_i^1 , e_i^2 and COMP of S' :

{ PP₁, AP, PP₂, NP, A*, VP, INFL', S, S' }

The singly underlined members of path (ii) are exclusively due to e_i^2 ; the doubly underlined members to e_i^1 . Path (ii) contains path (i), by virtue of the singly underlined members of the path. Thus we see how Subject Condition effects can be nullified in small clause constructions. The same analysis can be given for sentences with S' -deletion infinitivals (?who do you consider friends of t to admire t).

As before, the second gap must be lower than the Subject Condition violation, in order to add the crucial AP (or INFL') to the path:

(7)a. *a person who I admired t because I believed friends of t angry
at you

b. ?a person who I admired t because I believed you angry at t

And, as before, the lower gap may not itself be a Subject Condition violation, unless there is a third gap still lower, which is not a Subject Condition violation (cf. (5)):

(8)a. *a person who I believe [close friends of t] to consider [enemies
of t] angry at Mary

b. ?a person who I believe [close friends of t] to consider [enemies
of t] angry at t

We have not discussed the source of two gaps in the sentences we have been considering. In the terminology of Taraldsen (1981) and Engdahl (1981), one of the gaps is a parasitic gap, and one is a gap created by movement. More precisely, in the analysis of Chomsky (1981b), one gap is an empty category at D-structure, while the other contains the future \bar{A} -binder at D-structure. The first gap is thus a null pronominal at D-structure, an element which one expects to exist, given the null hypothesis that phonetic content is optional. At a later level, this empty category becomes locally \bar{A} -bound in English, since English has no other way to "identify" it at LF, as discussed in Chapter Two. How it becomes \bar{A} -bound, and at what level, are questions we return to in section 4, where we reexamine the analysis of parasitic gaps given in Chomsky (1981b) in the light of other multiple gap structures. For the moment let us assume that these null pronominals must be locally \bar{A} -bound at some later level, and review a consequence that follows.

What follows is an explanation of the observation of Taraldsen that the gaps in multiple gap configurations must not c-command one another. For example:

- (9) a. an article that_i I [_{v*} [_{vp} like e_i¹][without PRO reading e_i²]]
 b. *an article that_i [e_i¹ [_{v*} [_{vp} pleased me] [without PRO reading e_i²]]]
- (10) a. a person that_i you [_{v*} [_{vp} hired e_i¹][before I could object to e_i²]]
 b. *a person that_i [e_i¹ [_{v*} [_{vp} hired me][before I could object to e_i²]]]

As Taraldsen notes, e_i² in (9)b and (10)b is not locally \bar{A} -bound by that, as required. Rather, it is locally A-bound by e_i¹. Being locally A-bound, it forms a chain with e_i¹, which bears two θ -roles and violates the θ -criterion. Much the same explanation can be given for other cases in

which Chomsky (1981) invokes Principle C of the Binding Theory for variables, requiring them to be A-free (as noted by Chomsky (1981b), who cites D. Sportiche; cf. Lasnik (1982)):

(11) *a person who_i [_She_i thinks [_Sthat Mary likes e_i]]

In this example of Strong Crossover, e_i is locally A-bound, and thus is not a variable and violates the θ -criterion. (Problems arise when e_i is locally \bar{A} -bound from the nearest COMP, discussed by Chomsky (1981b).)

From the anti-c-command requirement on parasitic gaps as well as the very definition of paths we have assumed, it follows that certain other cases in which a second gap might nullify a PCC violation do not arise. In particular, since a subject c-commands everything below it on the tree (unlike a piece of a subject), second gaps cannot salvage a CTP sentence:

(12) *John, who_i I think that e_i¹ said that Mary liked e_i²

The first gap violates the PCC because its path lacks INFL'. If e_i² added nodes to the path from e_i¹, it would add this INFL'. It cannot form a path, however, because it is locally A-bound by e_i¹, and not \bar{A} -bound. It also violates the anti-c-command requirement independently imposed on parasitic gaps. We will not deal with this redundancy here, but see 1.6 of this chapter for some discussion.

2.2 Parasitic Gaps and "Connectedness"

2.2.1 We have noted several times that our analysis of parasitic gaps in the previous section is an adaptation of the analysis of Kayne (1982) in a PCC framework. In this section, we will present what we believe to be the essential features of Kayne's analysis, and will show some cases

that support our approach over Kayne's. This comparison will lead us to discover some new examples of the phenomena discussed in the previous section.

In the approach we have been outlining, a path can always be formed from a trace to its \bar{A} -binder. Crossing effects, CTP and Subject Condition effects were derived from the PCC, which functions as a sort of "traffic rule" modulating the interaction of overlapping paths.

Kayne (1981a, 1982) proposes that the construction of a path (of a somewhat different sort) from a trace to its binder be constrained structurally, CTP and Subject Condition effects (as well as certain other constraints -- e.g. on preposition stranding in Romance) arise when a coherent path cannot be constructed from the bound element to its binder. Kayne thus views these effects as arising from a "connectedness" requirement. Elements must be connected to their antecedents by a well-formed path; in certain cases, such a path cannot be constructed.

So far, the differences between Kayne's approach and our approach are matters of point of view. We might adopt Kayne's approach, and require a path to link elements with their antecedents, by viewing the PCC as a constraint on path formation. In other words, we might say that the path constructable under the definition in (2) is the maximal path that does not violate the PCC with respect to some other path. The PCC could thus intervene to block the linking of antecedents with the elements they bind.

Empirical differences between Kayne's approach and ours seem to arise in two domains. First, there are differences between the PCC and the constraints Kayne places on path construction; these differences have empirical content. Second, while we have so far restricted our definition

of paths to cases where a category is locally \bar{A} -bound, we might change this to "not locally A-bound", allowing for theoretically infinite paths from unbound empty categories to be governed by the PCC; we will take precisely this step in later discussion. In such a case, it might be possible for Kayne's theory to speak of the need for theoretical connectedness to a "point at infinity", but the intended content of the connectedness requirement would seem to be weakened.

In this section, we deal mainly with a difference of the first sort. Kayne (1981a) puts a proper government constraint on path construction; Kayne (1982) proposes a constraint based on the left/right asymmetry created by the Head First/Head Last parameter: a sort of Left Branch Condition in English. Subject Condition effects in small clauses and S'-deletion infinitivals argue for Kayne's later approach over the former; we will therefore concentrate on the proposals in Kayne (1982).

Kayne (1982) argues, following Kayne (1981a), that the ECP should require both a governor for an empty category and a connection to an antecedent. The first requirement derives from the second: Kayne proposes that the connection to an antecedent is possible when there is an appropriate path from a governor. Thus, if there is no governor, there is no connection to the antecedent. Elements entering into the path that make this connection are g-projections of the governor. The ECP is thus stated as in (12) -- we will call this the K(ayne)ECP, to distinguish it from Chomsky's formulation:

(12) KECP (Kayne 1981a, 1982, 4)

An empty category β must have an antecedent α such that

- (i) α governs β or
- (ii) α c-commands β and there exists a lexical category X such that X governs β and α is contained in some g-projection of X

(13) G-Projection (definition)

Y is a g-projection of X iff

- (a) Y is a projection of X (in the usual sense of X-bar theory) or of a g-projection of X
- or (b) X is a structural governor and Y immediately dominates W and Z, where Z is a maximal projection of a g-projection of X, and W and Z are in a canonical government configuration.

(14) Canonical Government Configuration (definition)

W and Z (Z a maximal projection, and W and Z immediately dominated by some Y) are in a canonical government configuration iff

- (a) V governs NP to its right in the grammar of the language in question and W precedes Z
- or (b) V governs NP to its left in the grammar of the language in question and Z precedes W.

To see how the KECP works, let us consider long movement of an object:

(15) [_{S₁} who_i [_{S₁} you [_{VP₁} think [_{S₂} that [_{S₂} Bill [_{VP₂} [_{V₂} likes] t_i]]]]]]]]

Kayne, like Chomsky (cf. 3.(3)), allows government by X^0 or by coindexation. Who in (15) does not govern t in either way, since many maximal projections intervene. Who is, however, the antecedent for t:

they are α and β in (12). V_2 is a lexical category X which governs \underline{t} . To satisfy the KECP, who must be contained in some g -projection of V_2 .

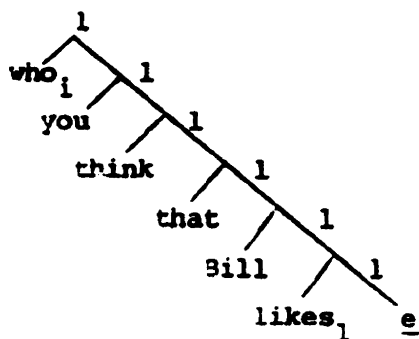
VP_2 is a g -projection of V_2 , since it is an X -bar projection of V_2 . Kayne seems to assume, following Kayne (1981a), that S is an X -bar projection of V : hence S_2 is a g -projection of VP_2 , hence of V_2 . In this case, even if S were a projection of INFL, it would still be a g -projection of VP_2 ,² but the assumption that it is a projection of V is necessary later.

As for S'_2 , if it too is a projection of V , then it is automatically a g -projection of S_2 , VP_2 and V_2 . If it is a projection of COMP, as is perhaps necessary for Kayne elsewhere (see below), it is still a g -projection of S , by (13)b: S'_2 (y) immediately dominates COMP (W) and S (Z), a maximal projection of the structural governor V (X). In English, verbs govern to the right (English is Head First), hence W must precede Z . Since it does, S'_2 (Y) is a g -projection of S_2 , hence of VP_2 and V_2 , by transitivity.

As for VP_1 , it should be a g -projection of S'_2 by a similar argument, if S' is the maximal projection of a structural governor (COMP or V). S_1 will then be a g -projection of VP_1 , since it is an X -bar projection of VP_1 . S'_1 in turn will be a g -projection of S_1 , by one of the arguments given in the preceding paragraph for S'_2 . Since the antecedent of \underline{t} -- who -- is contained by S'_1 , a g -projection of V_2 by transitivity, the KECP is satisfied, as desired.

The diagram in (15) shows the set of g -projections clearly. Following Kayne, we omit the node labels, to emphasize the independence of his notion of "g-projection" from the labelling of categories:

(15)



The numeral 1 labels each node that is a g-projection of likes, the governor of e.

Notice that because of the notion "canonical government configuration" the KECP sets up a left/right asymmetry in the construction of g-projections. In English, the result is a "Left Branch Condition" of sorts. The "g-projection path" from an element to its antecedent proceeds from projection to projection and from a right branch to the category dominating it, but never from a left branch to the category dominating it. We can see this by examining Subject Condition cases, the main motivation for Kayne's theory:

(16) * $[_S \text{who}_i [_S \text{you} [_{VP} \text{consider} [_{A^*} [_{NP} \text{friends} [_{PP} [\text{of } t]] [_{AP} \text{dull}]]]]]]]$

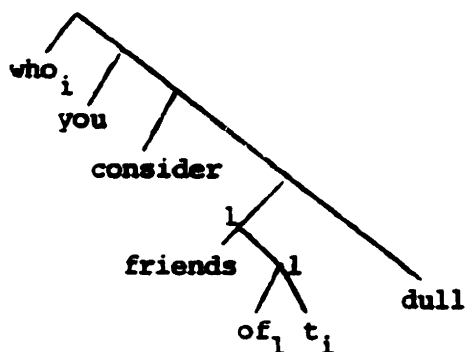
Who is once more the antecedent of t, but does not govern it. To satisfy the KECP, who must be contained in a g-projection of the governor of t, the preposition of.

PP is a projection of P, hence a g-projection. NP (Y) immediately dominates N (W) and PP (Z). Since A is a maximal projection of P (X), which is a g-projection of itself, and since PP is in the canonical government configuration with respect to N, NP is a g-projection of PP, hence of P.

A*, however, is not a g-projection of NP. A* (Y) immediately dominates NP (Z) and AP (W), where NP is a maximal projection of a g-projection of P. NP is not in the canonical government configuration with AP, since it (Z)

follows AP (W). Since A^* is not a g-projection of P, it follows that no node dominating A^* will be a g-projection of P. Thus no g-projection of P will contain the antecedent of the trace t , which is contained by S' . (16) thus violates the KECP:

(16)

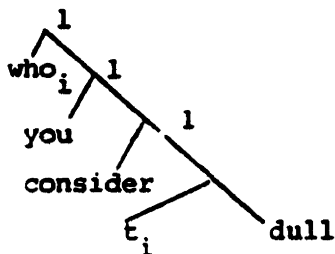


As Kayne notes, there is no prohibition against t itself being a left branch:

(17) [S , who_i [S you [VP consider [A^*t_i [AP dull]]]]]]

The governor of t is consider. From consider to the S' containing who, there is an unbroken chain of right branches and projections. Hence S' is a g-projection of the V consider. Since the antecedent for t is contained in a g-projection of a governor of t , it satisfies the KECP:

(17)



Now consider CTP effects:

(18) * [S_1 , who_i [S_1 John [VP_1 said [S_2 that [S_2 t_i INFL [VP_2 came]]]]]]]]

If INFL is a structural governor, then (18) should not be ruled out. If INFL is the head of S, then S_2 will be a g-projection of INFL. Since a chain of right branches and projections links S_2 with S'_1 , which contains who, who will be contained in a g-projection of the governor of t, and not violate the KECP.

If INFL is not the head of S, but structurally governs the subject, S_2 will once again be a g-projection of INFL. S_2 (Y) will immediately dominate the subject (W) and INFL (Z), where INFL is a maximal (and minimal) projection of itself. Since the subject precedes INFL, the canonical configuration obtains, and S_2 is a g-projection of INFL. This too will cause (18) not to violate the KECP, incorrectly.

It follows from Kayne's theory, therefore, that INFL is not a structural governor. If it is not a structural governor, then S_2 will not be a g-projection of INFL, and the KECP will be violated, as desired. Kayne takes the additional step of ruling INFL out entirely as a governor. From this it would follow that there are no g-projections of a governor of the subject, and the KECP will be trivially violated.

Finally, consider a case in which the antecedent and governor are the same:

(19) [_Swho_i [_St_i INFL left]]

As Kayne (footnote 14) notes, not only must who govern t across S, but S' must be a g-projection of who -- not unreasonable, if COMP is the head of S', as argued by Stowell (1981) (and references cited by Kayne, especially Chomsky (1981a, 274)).

Before turning to Kayne's analysis of grammatical Subject Condition violations, let us briefly compare the KECP and PCC accounts of the phenomena

we have so far considered. We believe that the KECP treatment of extraction of and from a subject shows a conceptual disadvantage of the theory, in comparison with the PCC account. On a PCC theory, the range of CTP effects and Subject Condition effects are explained in exactly the same way. Both types of effects follow from the interaction of a "WH-movement path" with a path running from below the subject to above the subject. In fact, once Subject Condition effects are explained, nothing more need be added to handle other extraction phenomena. The contrast between long and short movement of the subject of a tensed S, in particular, arises without any recourse to the requirement of a governor or to the disjunction between lexical and coindexing government stemming from Chomsky (1981a).

On a KECP theory, however, CTP and Subject Condition effects, though both technically derived from the KECP, arise in different ways. Subject Condition effects derive from the inability to construct a chain of *g*-projections from the governor of a trace to a category containing its antecedent. If one reconsiders the discussion just completed, it appears that the KECP could explain all the Subject Condition effects if the chain of *g*-projections began with the trace itself, and not with the governor. The need for a governor is invoked to capture CTP effects: the contrast between a structural and non-structural governor yields subject/object asymmetries with long WH movement, while the definition of government allowing both lexical and coindexing government permits short movement of the subject. Having invoked the need for a governor, to be sure, one may suggest, as Kayne has, that various island effects (e.g. the absence of preposition stranding in various languages) are due to the incapacity of various categories (like prepositions) to act as structural

governors. Although the PCC account does not extend in this direction, we argue that its natural extension to the domain of Crossing effects (and Coordinate Structure effects, discussed later), as well as its non-stipulative coverage of the basic WH-movement facts suggests that it has a conceptual advantage over the KECP account.

Returning to (16), we notice that it is a general consequence of Kayne's theory that a piece of a left branch may never contain an empty category bound outside that left branch (given the canonical government configuration of English), unless the binder is also a governor of that category. Kayne thus rules out (20) the same way he rules out (16):

(20) *John, [_S who_i [_S [_{NP} friends [_{PP} of t_i]] INFL came]]

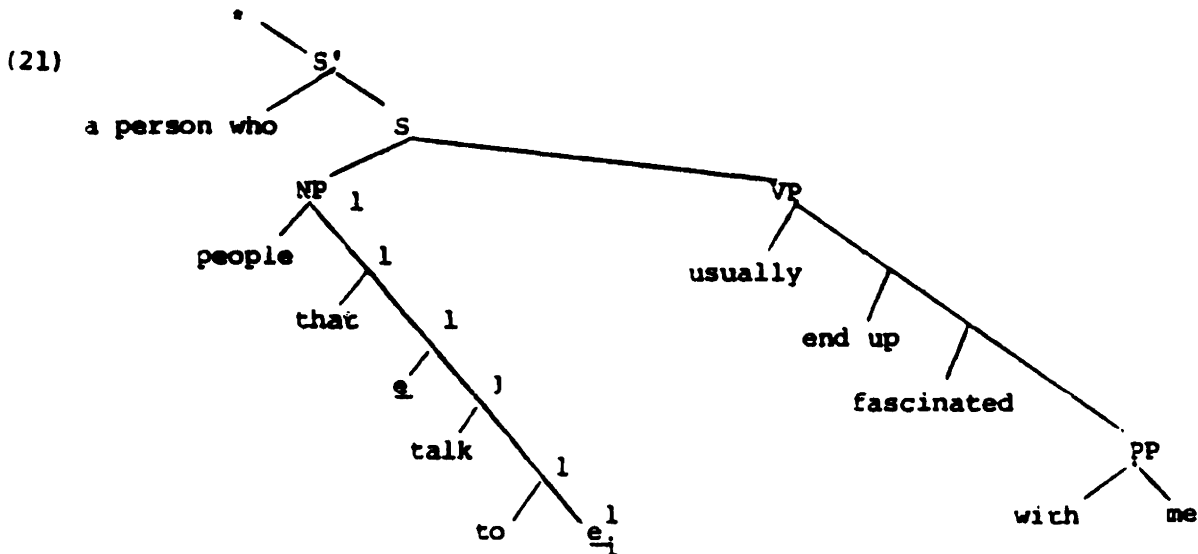
S is not a g-projection of the preposition of, thus who is contained in no g-projection of the preposition, and the KECP is violated.

Now let us examine how Kayne accounts for the violations of the Subject Condition discussed in the last section. Recall that structures like (2) (or (16)) improve if a second gap bound by the same \bar{A} -binder is added lower in the tree. A second gap higher in the tree does not improve structures like (20).

Our own analysis was an adaptation of Kayne's analysis. Kayne defines a g-projection set consisting of a gap, the g-projections of its governor and any nodes that may dominate the gap but not its governor (for cases of cross-boundary government, as in small clauses). This set roughly corresponds to our paths, except that our paths do not include nodes below the maximal projection of the governor. He then suggests that where multiple gaps share a local \bar{A} -binder, we take a union of their

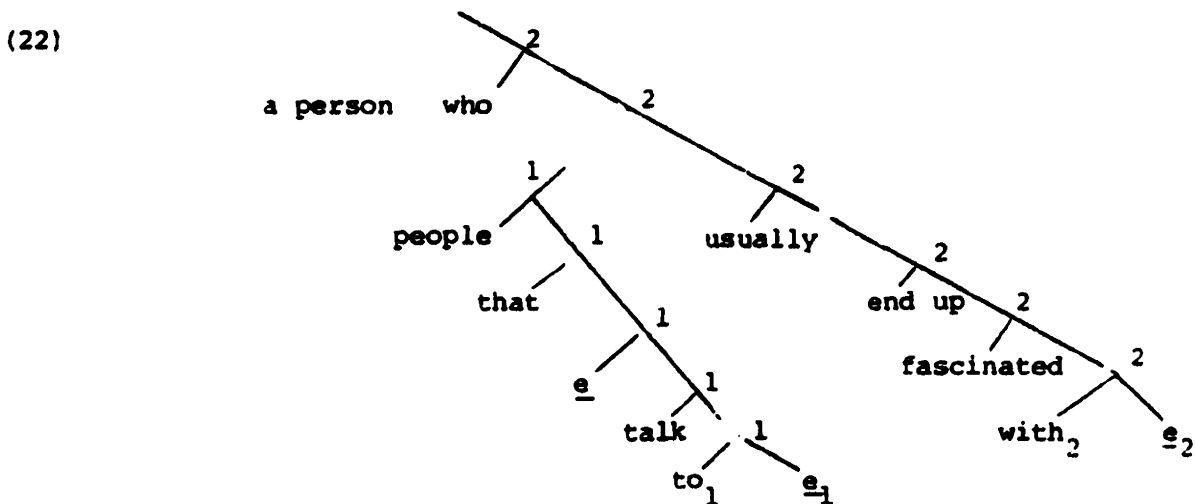
g-projection sets. The result of this union must constitute a coherent subtree of its phrase marker.

We illustrate with some examples from Kayne:



In (21), as in (20), who is contained in no g-projection of the governor of e_1 . The numeral 1 is attached to all nodes which are members of the g-projection set of e_1 . Note that who is not dominated by such a node. The KECP is restated by Kayne to require that the antecedent be dominated by a member of the g-projection set of the element it locally binds. This is equivalent here to (12)^{3,4}.

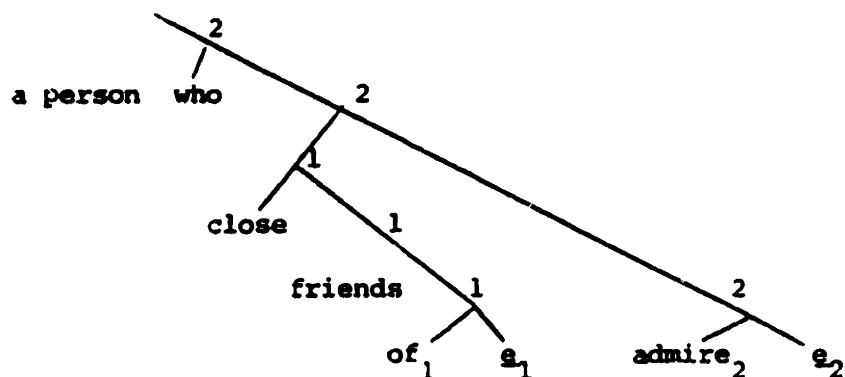
Now consider (22):



We have marked the g -projection sets of e_1 and e_2 here. Now consider the union of these sets. Note two things about this union: (1) it forms a coherent subtree of the phrase marker, and (2) who, the antecedent of both gaps in question, is contained within a g -projection of the union of the two g -projections. The nodes labelled 2 thus help "complete the circuit" from e_1 to its antecedent, while completing its own circuit from e_2 to the same antecedent. Recall that in the PCC theory outlined in the previous section, the union of two paths allowed the first path to avoid a PCC violation. The logic of the solutions is the same: the constraint invoked is not.

Consider now (23) from Kayne:

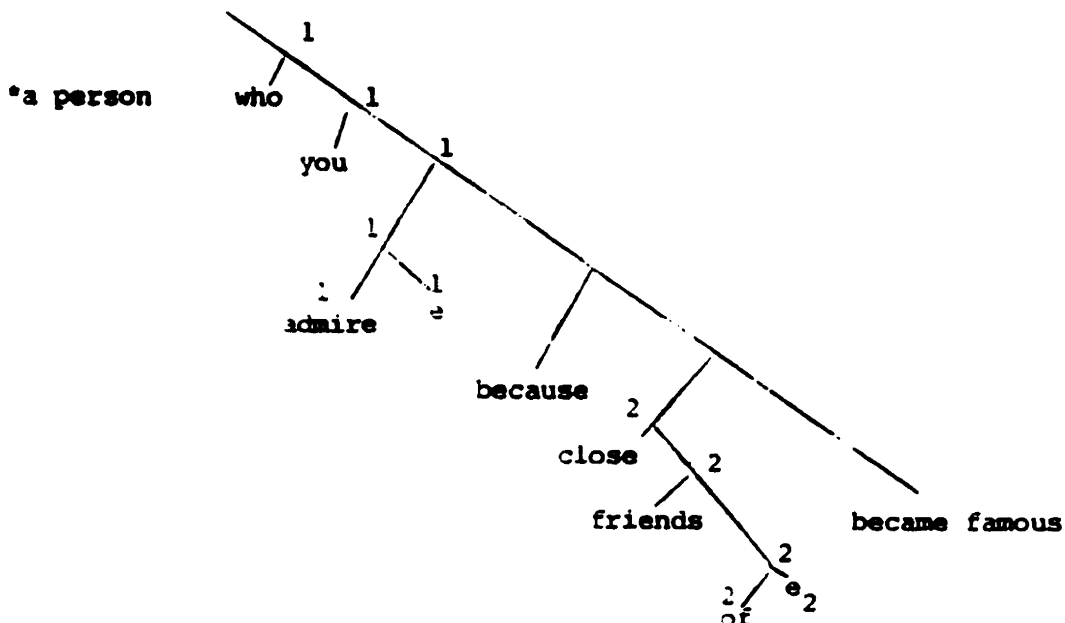
(23)



(23) is parallel to (22). Were e_2 not a gap, so that there was no second g -projection set, the antecedent who for e_1 would not be a member of the g -projection set of e_1 and would violate the KECP. Given the second gap, however, and the union of the two g -projection sets, which forms a subtree, who is contained in the union of the two sets, and does not violate the ECP.

Now consider (24) (= (4)a):

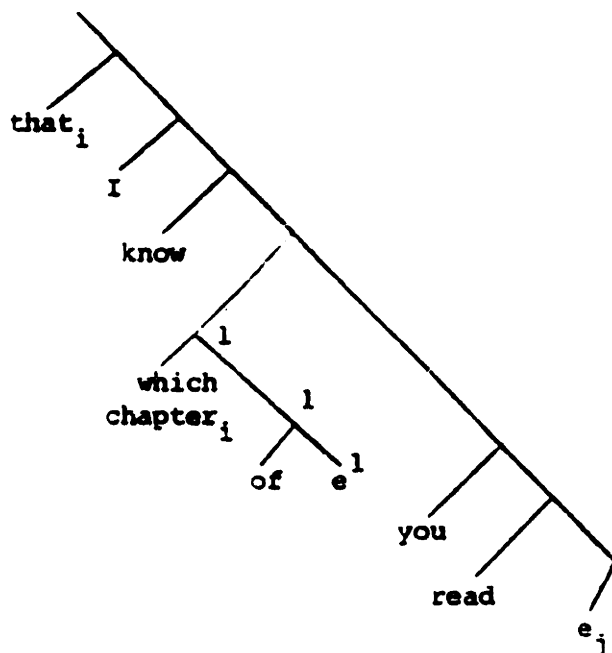
(24)



If we take the union of the two g -projection sets in (24), we note that the antecedent of both e_1 and e_2 is contained by a node in the united set. The structure is ruled out, however, because the union of the two g -projection sets does not form a coherent subtree of the representation in (24). This will always occur when the gap that cannot be connected by itself with its antecedent is lower in the tree than the gap that can. The first set in (24) does nothing to "complete the circuit" for the second. Kayne thus derives the fact that the gap that saves a Subject Condition violation must be lower in the tree than the violation.

To sum up, both Kayne's theory and the PCC theory postulate that the paths from empty categories that share a local \bar{A} -binder are united. To capture Subject Condition effects, Kayne postulates that extraction from inside a left branch (in English) does not yield a connection between the trace in that left branch and its antecedent. The PCC theory derives the Subject Condition effects from the interaction between a path beginning inside the subject and a path running between one node lower and one node

(25)



The PCC, however, does not rule (25) out. Consider all the paths that might be relevant (we omit the path between INFL' and COMP of S'_1):

(25) (i) Between e_j and COMP of S'_2 :

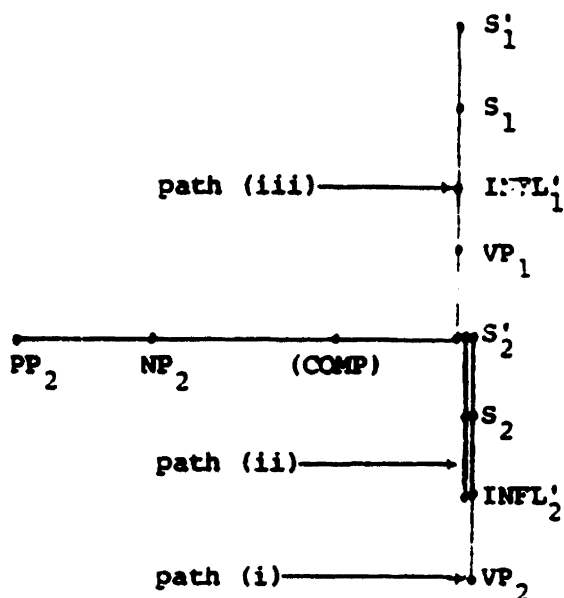
{ VP_2 , INFL' $_2$, S_2 , S'_2 }

(ii) Between INFL $_2$ and COMP of S'_2 :

{INFL' $_2$, S_2 , S'_2 }

(iii) Between e_i and COMP of S'_1 :

{PP $_2$, NP $_2$, S'_2 , VP $_1$, INFL' $_1$, S_1 , S'_1 }



The only paths that overlap are paths (i) and (ii), which do not violate the PCC. Note that the definition of overlapping (non-null and non-singleton intersection) is crucial here, since path (iii) shares a single node with paths (i) and (ii) (cf. our earlier remarks in 4.2 of Chapter Three on eliminating the overlapping condition).

We shall now show that, contrary to appearances, the PCC theory makes the correct prediction. Notice that (25) is also a subjacency violation, if NP is a bounding node, as well as either S or S'. The link between e_i crosses NP_2 and either S'_2 or S_1 , without any intervening COMP landing sites. In other words, it violates subjacency in the same way a violation of the Complex NP Constraint does (*who do you know [NP someone [S who [S t wrote about t]]]).

We will discuss the relationship between the subjacency constraint and multiple gap constructions in detail in 4.5. For now, note a simple consequence of the analysis of parasitic gaps of Chomsky (1981b), sketched above. A parasitic gap is a null pronoun at D-structure, coindexed with a movement-derived variable at S-structure. The movement-derived variable

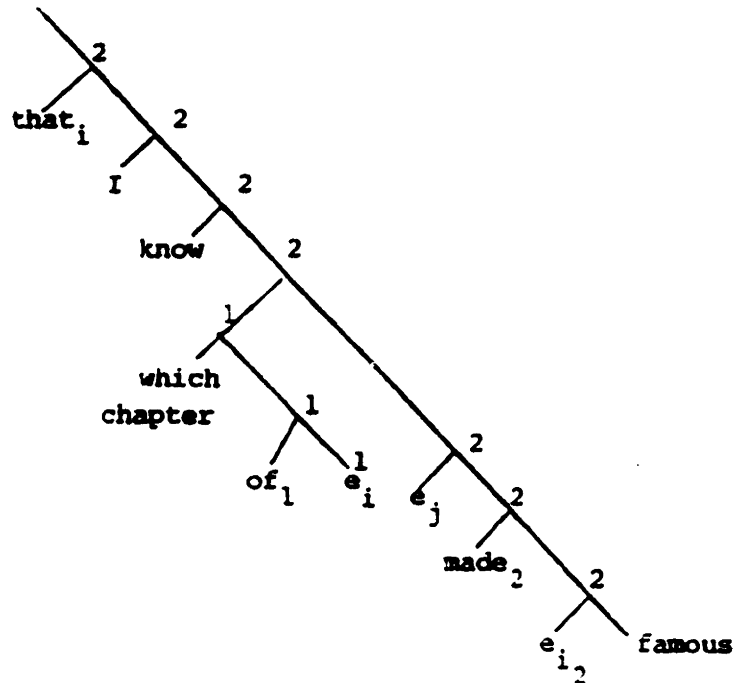
will show subjacency effects as a property of the rule Move α . The parasitic gap will not show subjacency effects, since it is a D-structure gap, linked to an \bar{A} -position by coindexing. Stipulating for now that a parasitic gap is only found when a movement-derived gap is also present, it follows that at least one empty category of a set of empty categories sharing a common local \bar{A} -binder will obey the subjacency condition. Thus, in (26)a, where only one gap violates subjacency, the structure is acceptable. In (26)b, where both of two gaps violate subjacency, the structure is unacceptable. This follows if one of the two gaps must derive from Move α :⁵

- (26)a. a man who_i [_{NP} people [_S, that_j [_{S_j} e_j know e_i]]] admire e_i
 b. *a man who_i [_{NP} people [_S, that_j [_{S_j} e_j know e_i]]] admire [_{NP} the things [_S, that [_S I write about e_i]]]

Returning to (25), if we add a second gap to (25), we should be able to eliminate the effects of subjacency, since we will be able to treat e_i as a parasitic gap. Thus, if (25) is bad because of subjacency, the addition of a second gap should improve it.

On the other hand, the addition of a second gap lower in the tree than the subject should improve (25) on Kayne's KECP account as well as on a subjacency account; so such a gap does not distinguish between the theories. Indeed, as expected on both theories, a second gap lower in the tree does improve the structure. Compare (25) with (27), diagrammed according to Kayne's account:

- (27) ?this book, [_{S_i} that_i [_S I know [_S, [which chapter of e_i]_j [_S e_j made e_i famous]]]]_i



(27) is, of course, no worse on a PCC account than (25). (Add VP, INFL' and S of the lowest S' to path (iii) of (25), to yield the paths for (27): path (iii) now overlaps and contains both path (i) and path (ii).)

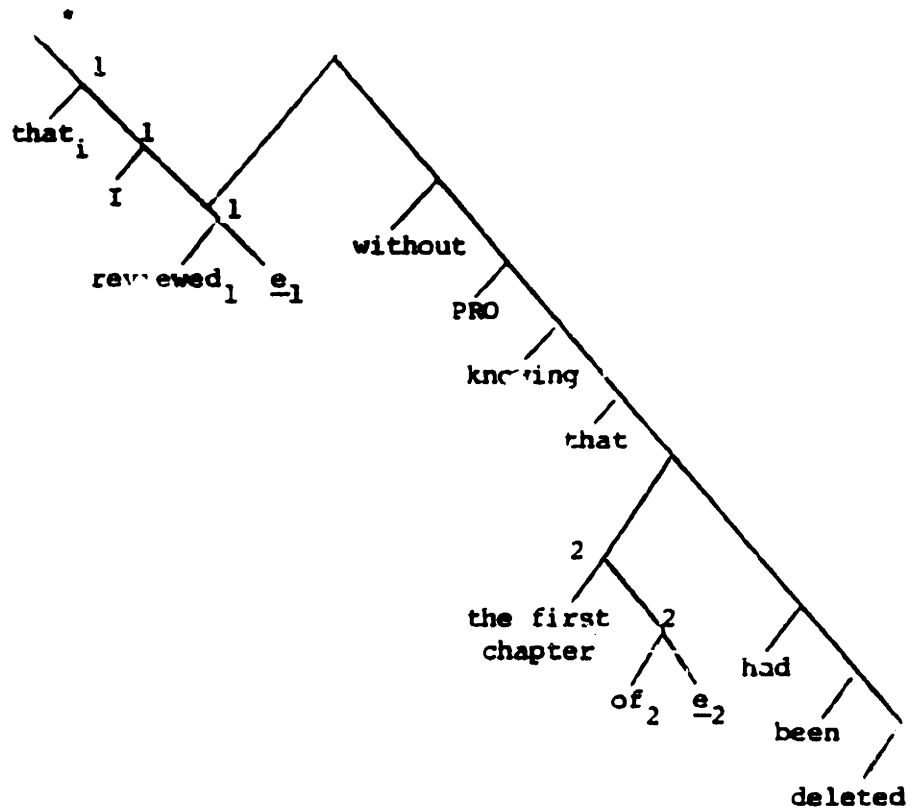
The crucial test, then, involves the addition of a second gap to (25) which is higher in the tree than COMP. Under Kayne's account, the result should be as bad as (24). Under a PCC account, the PCC should again play no role, if the second gap does not independently violate it. Subjacency should not rule the sentence out, since the gap in COMP is a parasitic gap, and the structure should be much better than (24). The results, I believe, favor the PCC account. Compare (28)a, a violation of the same type as (24), with (28)b:

(28)a. *a book that_i I reviewed e_i [without knowing [_S that [_S [_{NP} the⁶ first chapter of e_i] had been deleted]]]

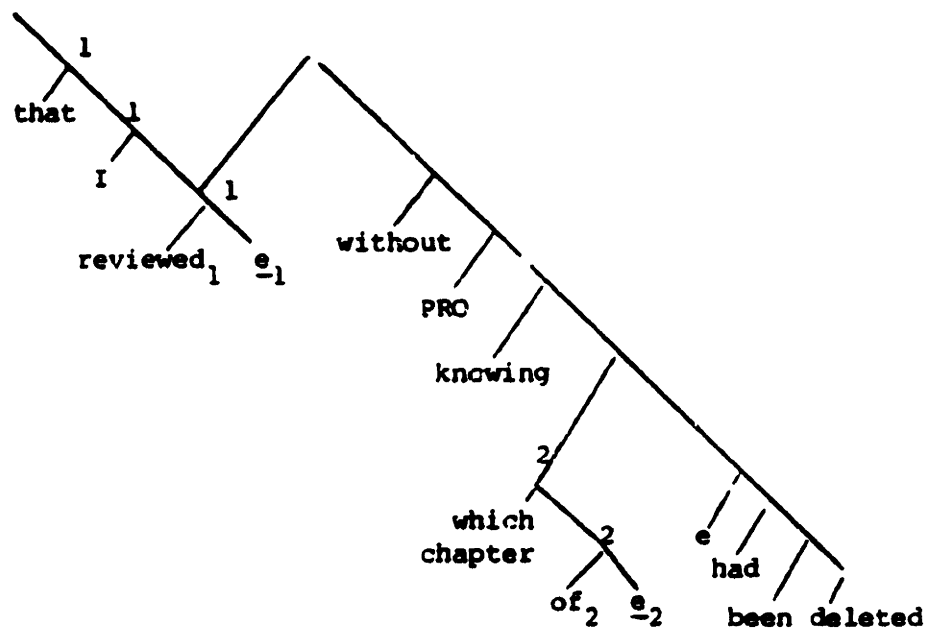
b. ?a book that_i I reviewed e_i [without knowing [_S [_{COMP} [_{NP} which chapter of e_i]]]_j [_S e_j had been deleted]]]

While (28)b has the usual oddness of parasitic gap constructions, it contrasts clearly with (28)a. Notice that the status of the g-projection sets in each case is the same, under Kayne's theory:

(28)a. (cf. (24))



(28)b.



In each case, the antecedent that is not contained in any g-projection of the lower of the two traces. Yet (28)b is acceptable and (28)a is not.

Summarizing the argument, Kayne predicts that the status of gaps inside COMP and gaps inside a subject should be the same: they should be ungrammatical unless a second gap is added, and they should remain ungrammatical unless the second gap is lower than the COMP or subject in question. The PCC theory predicts that extraction from the subject position should be ungrammatical unless a second gap is added lower than the subject. Extraction from COMP, however, does not fall under the PCC at all. It is ruled out by subjacency unless a second gap is added, but the position of this second gap is irrelevant. The second gap merely suffices to render the first gap parasitic, allowing it to violate subjacency. The acceptability of (28)b thus argues for the PCC theory against Kayne's.⁷

2.3 Left Branch Extractions

As far as we can tell, empty categories in COMP are the only configurations in English where the PCC theory and Kayne's make contradictory predictions. There are, of course differences in the scope of the theories. We have already noted some of these: The PCC account, unlike the KECP account, has nothing to say about the absence of preposition stranding in some languages. The KECP, on the other hand, does not deal with Crossing effects. Although one could build a Crossing constraint into the KECP, this would allow us to eliminate (13)b and (14), making the KECP equivalent to the PCC. As we extend the PCC theory to cover other cases of multiple gaps, WH-in-situ, and other problems, we will continue to compare the PCC account with what might be said on a KECP account, as well as on an ECP account.

In this section, we wish to focus on a number of cases for which both the PCC and KECP theories are descriptively adequate, but where the adequacy of the KECP account rests on an asymmetry not found in the PCC account. While the argument against the KECP is slight, discussion of this point will allow us to consider a number of interesting examples not previously mentioned in the literature, and will return us briefly to a discussion of Russian quantificational constructions, which will suggest a stronger (though problematic) argument against the KECP and "Connectedness", and will allow an empirical improvement over the ECP account given in Chapter Two.

Kayne's account of Subject Condition effects and the parasitic gap constructions that nullify them rests on a left/right asymmetry introduced by the notion of "canonical government configuration". Since the asymmetry is induced into the definition of g -projection by reference to the Head First parameter, it is not in itself an undesirable stipulation. On the other hand, it is odd that this asymmetry appears only in the definition of non-initial g -projections given in (13). That is, why does the governor of β in the KECP not have to be in a canonical government configuration with α (since this first governor is the first g -projection, being a projection of itself), while all other links in the g -projection set require the canonical government configuration?

As it happens, this asymmetry makes the right predictions. In many languages, subject to ill-understood restrictions, an empty category may be non-canonically governed, so long as the maximal projection of its governor is canonically governed. In Head First languages, this situation arises when a specifier to NP is extracted from that NP: where this is possible, the NP must be on a right branch. The most familiar case is the

extraction of combien 'how many' from NP in French (Obenauer 1976).

Analogous data obtains for skol'ko 'how many' extraction in Russian, was extraction from the was fur NP 'what kind of NP' construction in German (Obenauer 1976, den Besten 1981), and Sub-C ("Subdeletion") in comparative constructions in English and other languages.

To begin with French, the following paradigm is familiar (Kayne 1981a, 97):

(29) [_S, [_{QP} combien]_i [_S est-ce que Paul [_{INFL}' [_{VP} aime [_{NP} e_i de femmes]]]]]
 how many does Paul love of women
 'how many women does Paul love'

(30) * [_S, [_{QP} combien]_i [_S est-ce que [_{NP} e_i de femmes] [_{INFL}' [_{VP} aime Paul]]]]]
 how many do of women love Paul
 'how many women love Paul'

(30) is a straightforward Subject Condition effect of the type we have been considering, and falls under either Kayne's analysis or a PCC account. On a PCC account, (29) has the paths indicated below:⁸

(29) (i) Between INFL and COMP:
 {INFL', S, S'}

(ii) Between e_i and COMP:
 {NP, VP, INFL', S, S'}

Paths (i) and (ii) overlap, and (ii) contains (i), as required by the PCC.

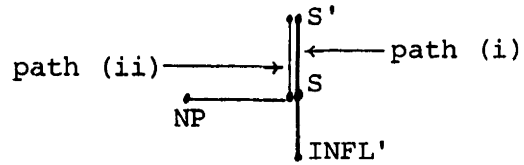
(30), on the other hand, has the following paths:

(30) (i) Between INFL and COMP:

{INFL', S, S'}

(ii) Between e_i and COMP:

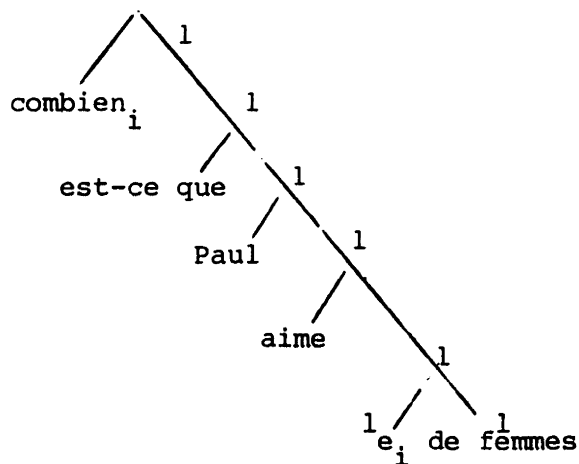
{NP, S, S'}

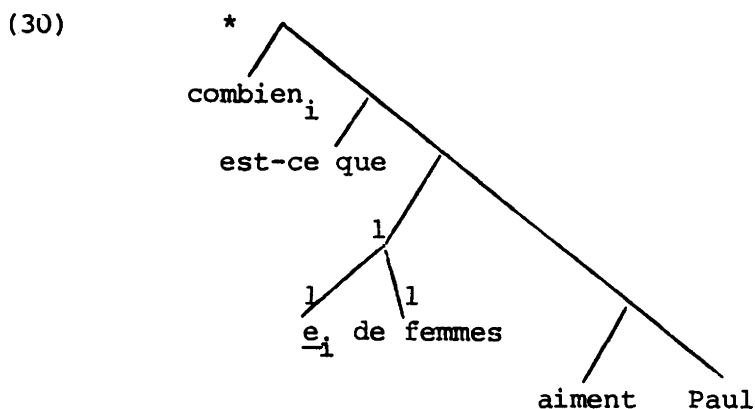


Paths (i) and (iii) here overlap, but neither contains the other. The violation is a straightforward Subject Condition effect, subsumed by the PCC.

In a KECP framework, NP in both (29) and (30) is a g-projection of N, the governor of the trace of combien, even though N is not in a canonical government configuration with the trace of combien. Linear order is only important for non-initial g-projections. In (29), a sequence of right branches from NP to S' makes S' a member of the g-projection set of the trace of combien. Since S' contains the antecedent for this trace, the KECP is satisfied. In (30), however, NP is on a left branch of S: it is not in a canonical government configuration with the other constituent(s) of S. S is thus not a g-projection of NP, and not a member of the g-projection set of the trace of combien. Since S is not, S' is not. No member of the g-projection set of the trace of combien contains its antecedent: hence the KECP is violated:⁹

(29)





Similar, but weaker, effects are found in the subject of a small clause or non-maximal infinitival, as expected:

(31)a. [combien d'argent]_i est-ce que tu considères [_{A*}e_i suffisant]
 how much of money do you consider enough

b.?[combien]_i est-ce que tu considères [_{A*} e_i d'argent]_i suffisant]

(32)a. [combien de femmes]_i as-tu laissé [_S Paul embrasser e_i]
 how many of women have you let Paul kiss

b. [combien]_i as-tu laissé [_S Paul embrasser [e_i de femmes]]

c. [combien de femmes]_i as-tu laissé [_Se_i [m'embrasser]]
 how many of women have you let kiss me

d.??[combien]_i as-tu laissé [_S [e_i de femmes][m'embrasser]]

On a PCC account, these data follow from the interaction of the path between combien and its trace with the path created by the θ -grid of considérer, laisser. On a KECP account, the small clause or S would not be a g-projection of the N governor of the QP trace, because the subject of the small clause or S is not in a canonical government configuration with the predicate.

Finally, as predicted, sentences like (30)b, (32)d, and presumably (31)b (though clear examples are hard to construct) improve if a second empty category bound by combien is added lower on the tree:

(33)a.?_S [_S [combien]_i [_S est-ce que [_{NP}₁ e_i d'hommes] [_{INFL}' [_{VP} aiment
 [_{NP}₂ e_i² de femmes]]]]]]
 'how many men love how many women'

b.?_S [_S [combien]_i [_S as-tu [_{INFL}' [_{VP} laissé [_S e_i¹ d'hommes] embrasser
 [e_i² de femmes]]]]]]]]
 'how many men have you let kiss how many women'

On the PCC account, the addition of a lower gap in (33)a adds the INFL' node to the path between the traces and COMP that is missing in (30):

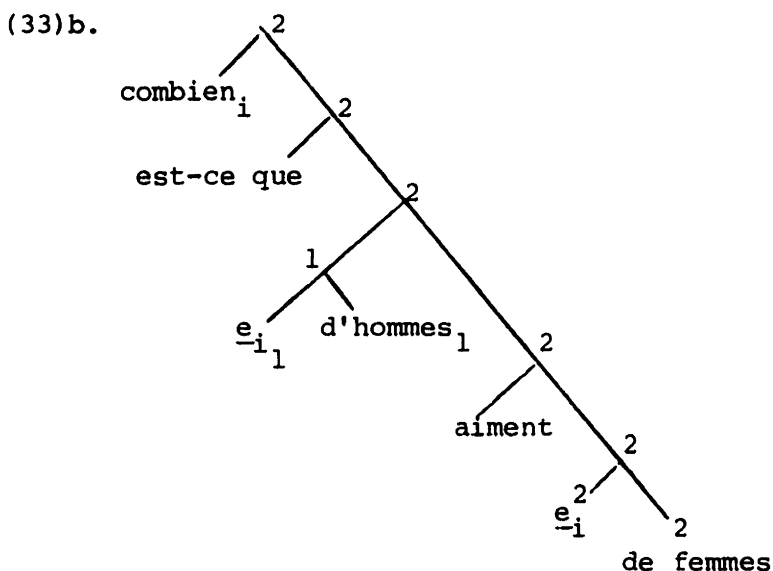
(33)b. (i) Between INFL' and COMP:

{INFL', S, S'}

(ii) Between e_i¹, e_i² and COMP:

{NP₂, VP, INFL', NP₁, S, S'}

On a KECP account, the addition of the lower gap makes the connection between the upper gap and its antecedent possible. The resulting representation is the union of the representations of (29) and (30) above:



The analysis of (33)b on both approaches follows similar lines. Instead of the path between INFL and COMP, the θ -grid-induced path between VP and the predicate phrase of the small clause or S'-deletion structure will be relevant on a PCC account. On a KECP account, the effects will follow from the definition of canonical government configuration.

The second gap that saves sentences with combien extraction from the Subject Condition must, as always, be below the subject in the tree, and must not itself be a Subject Condition violation:

(34)a. [_{S_i} [combien]_i est-ce que tu as persuadé Jean [_{S₂} que [_S Paul
 [_{INFL}' [_{VP} aime [_{NP_{-i}} de femmes]]]]]

'how many women did you persuade Jean that Paul loves'

b.*[_{S_i} [combien]_i est-ce que tu as persuadé Jean [_{S₂} que [_S NP_{-i}
 d'hommes] [_{INFL}' [_{VP} t'aiment]]]]]

c.?[_{S_i} [combien]_i est-ce que tu as persuadé Jean [_{S₂} que [_S [_{NP_{-i}}^{e₂}
 d'hommes] [_{INFL}' [_{VP} aiment [_{NP_{-i}}^{e₁} de femmes]]]]]]]

d.?[_{S_i} [combien]_i est-ce que tu as persuadé [_{NP_{1-i}}^{e₂} d'hommes] [_{S₂} que
 [_S Jean [_{INFL}' [_{VP} aime [_{NP_{2-i}}^{e₁} de femmes]]]]]]]

e.*[_{S_i} [combien]_i est-ce que tu a persuadé [_{NP_{1-i}}^{e₂} d'hommes] [_{S₂} que
 [_S [_{NP_{2-i}}^{e₂} de femmes] [_{INFL}' [_{VP} m'aiment]]]]]

f.*[_{S_i} [combien]_i est-ce que tu as persuadé Jean [_{S₂} que [_S [_{NP_{1-i}}^{e_i}
 d'hommes] [_{INFL}' [_{VP} pensent [_S que [_S [_{NP_{2-i}}^{e_i} de femmes] [_{INFL}' [_{VP} l'aiment]

The relevant examples are (34)e and (34)f. In (34)e, the relevant paths are as follows (simplifying as above):

(34)e. (i) Between INFL and COMP of S'_2 :

{INFL', S, S'_2 }

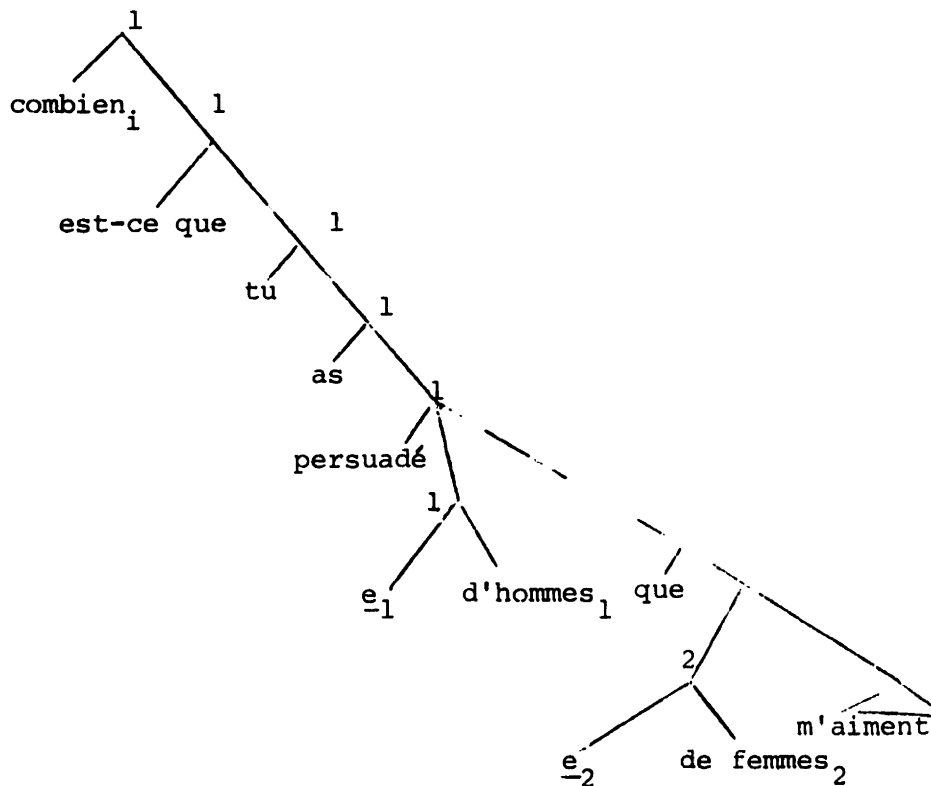
(ii) Between e_i^2 , e_i^1 and COMP of S'_1 :

{NP₂, S, S'_2 , NP₁, S'_1 }

Paths (i) and (ii) overlap, but neither one contains the other. Hence the PCC is violated. Crucially, the addition of e_i^2 does not add the INFL' node that e_i^1 adds in (34)c.

Similarly, in a KECP approach, the higher gap does not make the appropriate connection:

(34)e.



In (34)f, as can be easily seen, the lower gap does "save" the higher gap, but is itself in violation of the PCC or KECP. The "saving" gap cannot itself be a Subject Condition violation, unless it itself is saved by some still lower gap:

- (35) ?[combien]_i est-ce que tu as persuadé Jean [_S,que [e_i d'hommes]
 pensent [_S,que [e_i de femmes] aimaient [e_i d'enfants]
 'how many men did you persuade Jean that how many women liked how
 many children'

(35) is hard to process, but seems better than (34)f.

Thus it appears that extractions of left branches have exactly the properties of extractions of right branches. In particular, they show Subject Condition effects in exactly the same distribution as any other extraction. On a PCC theory this follows automatically: the notion of "canonical government configuration" plays no role in the theory of paths. On the KECP theory this does not follow automatically. Instead, it follows from an asymmetry in the definition of g-projection and in the KECP itself. The first member of the g-projection, the local governor of an empty category, may be in any relation whatsoever with the empty category. Further g-projections, however, must be projections of nodes that are in the canonical government configuration with the other g-projections.

Before turning to Russian quantificational constructions, let us note that the "Subdeletion" construction in English appears to show the same contrasts as combien extraction. In this construction, which we may theory-neutrally call Sub-C(omparative), a null QP appears as the specifier of N. Bresnan (1975) argued that this null QP arises as the result of a deletion operation. She noted that it appeared to obey the Complex NP

Constraint, subsumed by Chomsky (1973) under the Subjacency Condition:

- (36)a. I'll have to give as many F's as you've proposed to give^e[e A's]
 b.*I'll have to give as many F's as you've discussed a proposal to
 give [e A's]

Since subjacency effects appear to be a property of the rule Move α , Chomsky (1977a) argues that Sub-C constructions involve movement -- in particular, of a null QP to COMP (or of a lexical QP followed by obligatory deletion). That movement is to an \bar{A} -position is suggested by the fact that the movement is apparently unbounded: the empty QP is not subject to principles A or B of the Binding Theory:

- (37) I bought more books than you said that Mary wanted Bill to
 buy [e records]

Chomsky thus argues for a structure like (38) for (36)a, where Q is an operator:

- (38) I'll have to give as many F's [_S, as Q_i [_S you've proposed to give
 [e_i A's]]]

The similarity between (38) and combien extraction is clear (cf. Kayne 1981a, 115-20). The following paradigm thus provides further evidence that \bar{A} -binding is involved in the Sub-C construction. The Sub-C construction exhibits two other hallmarks of \bar{A} -binding: PCC effects and parasitic gaps. Example (39)b was pointed out to me by L. Rizzi:

- (39)a. I saw as many women [as I think (that) you saw [e men]]
 b.*As many women saw me [as I think (that) [e men] heard you]

- c.?As many women bought this book [as I think (that) [e men] bought [e records]]
- d.?As many women bought this book [as I persuaded [e men] (that) I would buy [e records]]
- e.*As many women bought this book [as I persuaded [e men] (that) [e records] would be on sale]
- f.?I consider as many books interesting [as I consider [e pamphlets boring]]

Judgments on C-sub constructions are always delicate; as is well-known, judgments tend to decrease in reliability as C-sub sentences become longer. Nonetheless, I believe the judgments in (39) to be reasonably clear. (39)b shows a Subject Condition effect in a tensed sentence; (39)f, in a small clause (cf. (30)b, (31)b). (39)c shows a parasitic gap construction; a second gap lower than the Subject Condition violation nullifies the Subject Condition effect; similarly in (39)d (cf. (33)b, (34)d). (39)e shows that the second gap cannot be higher up the tree (cf. (34)e). Note that because of the nature of the Subject Condition effect the deletion of that plays no role in the grammaticality judgments. (Except, as Bresnan 1975 notes, that Sub-C constructions are always more natural when that is deleted.) Whether the trace of QP inside a subject is bound from the nearest COMP or not, it will violate the PCC (or KECP) (but cf. footnotes 7 and 9).

The paradigm of (39) repeats itself in French:

- (40)a. j'ai vu autant de femmes [que je crois que tu as vu [e d'hommes]]
- b.*autant de femmes m'ont vu [que je crois que [e d'hommes] t'ont écouté]
- c.?autant de femmes ont acheté ce livre [que je crois que [e d'hommes] ont acheté [e de disques]]

- d. ?autant de femmes ont acheté ce livre [que j'ai persuadé [e d'hommes] que j'achetèrais [e de disques]]
- e. *autant de femmes ont acheté ce livre [que j'ai persuadé [e d'hommes] que [e de disques] seraient à vendre
- f. ??je considère autant de livres intéressants que je crois [e de brochures] ennuyantes

We have considered the French equivalents of (39) because of a problem that arises in English. Notice that both the PCC and KECP accounts of Subject Condition effects predict no asymmetry between long and short WH-movement (cf. the irrelevance of that deletion discussed above). Nonetheless, short Sub-C movement in English seems much better than long movement. Compare (41) with (39)b:

(41) as many women saw me as men heard you

This is unexplained on any PCC account that explains (39).¹⁰ In French, however, short Sub-C does show a slight subject/object asymmetry:

(42)a. j'ai vu autant de femmes que tu as vu [e d'hommes]

b. ?autant de femmes m'ont vu que [e d'hommes] t'ont entendu

French conforms to the PCC and KECP predictions; English does not. We might suggest, unhelpfully, that (41) is ungrammatical, but acceptable by analogy, given the apparent general difficulties with Sub-C judgments. We leave (41) as an unresolved problem (but cf. Kayne 1981a for a suggestion).

On the PCC theory, we expect Sub-C constructions to show other types of Crossing effects -- in particular, to interact with WH-movement in the fashion discussed in the previous chapter. To my ears, all WH-island

violations in Sub-C constructions result in incomprehensibility. The PCC predicts a contrast between (43)a and (43)b, as indicated in parentheses, which may exist. Given the problems, however, we do not attach any value to these examples:

(43)a.(OK) John read more memoirs [than \bar{O}_j I know [who_i [to persuade \underline{e}_i [to read [\underline{e}_j novels]]]]]

b.(*) John knows more authors [than \bar{O}_j I know [what book_i [to persuade [\underline{e}_j editors] to read \underline{e}_i]]]

(\bar{O} is a null operator that \bar{A} -binds instances of \underline{e} ; cf. Chomsky 1977.)

Not wishing to flog a dead horse, we nonetheless turn to examples of Sub-C constructions in Russian. Not surprisingly, the paradigms found French and in English repeat themselves in Russian. The paradigm in Russian, however, provides additional evidence for our analysis of Russian quantificational constructions given in Chapter Two, and will lead us into a reconsideration of our analysis, in light of the PCC. The results will be mixed.

Russian Sub-C constructions with comparatives of equality transparently involve \bar{A} -binding created by movement. They are essentially relative clauses of quantity, whose head is the demonstrative stol'ko 'so many', 'as many', formed with the WH-word skol'ko 'how many', 'as many'. Skol'ko is also an interrogative.

Phrases with stol'ko and skol'ko have the properties of numeral phrases discussed in Chapter Two. Like numerals ending in five and above, stol'ko and skol'ko trigger genitive plural marking on their dependent N in non-oblique contexts. Like numeral phrases, the phrases [stol'ko/skol'ko + N] trigger verbal agreement optionally with ergative and passive verbs:¹¹

- (44)a. priš^Vlo [stol'ko^V že devušek^V], [skol'ko mal'čikov^V] uš^Vlo¹²
 came as many girls as many boys left
 (neut sing) (gen plur) (neut sing)
- b. [stol'ko^V že devušek^V] priš^Vli, [skol'ko mal'čikov^V] uš^Vli
 as many girls came as many boys left
 (gen plur) (plur) (gen plur) (plur)
- (45)a. [skol'ko mal'čikov^V] uš^Vlo?
 how many boys left
 (gen plur) (neut sing)
- b. [skol'ko mal'čikov^V] uš^Vli?
 how many boys left
 (gen plur) (neut sing)

Recall how we accounted for the difference between structures without agreement (the (a) sentences) and those with agreement (the (b) sentences). Stol'ko/skol'ko phrases, like numeral phrases, have the form [Q + N]. We argued in Chapter Two that when the phrase [Q + N] triggers agreement, the N is the head, and the phrase is an NP. When [Q + N] does not trigger agreement, the Q is the head, and the phrase is a QP.

The D-structure object position of an ergative or passive verb, we claimed (following Burzio 1981; Chomsky 1981a), does not allow a Case-marked element. Hence, any constituent that is subject to the Case filter will have to move to a Case-marked position. We argued that NPs are subject to the Case filter, but QPs are not. It followed that D-structure NP objects of ergative and passive verbs must be S-structure subjects, to satisfy the Case filter. D-structure QP objects of such verbs, however, may remain S-structure objects, and possibly must remain objects, as we discussed. Recall that the unmarked position for APs in such constructions is postverbal, while NPs tend to be preverbal, supporting this hypothesis. On this analysis, therefore, the skol'ko-clause of (44)a - (44)b and

sentences (45)a and (45)b have the following S-structures:

(46)a. [_S, QP_i [_S e [_{INFL}, INFL [_{VP} V [e]_i]]]] (= (44)a/(45)a)

b. [_S, NP_i [_S [e]_i [_{INFL}, INFL VP]]]

The subject e in (46)a is expletive. (See Section 4 of this chapter for discussion.)

Like French combien, skol'ko may be extracted from its phrase. (47) shows such extraction from an uncontroversial direct object. The long movement is responsible for the slight unacceptability, but guarantees that the movement of skol'ko is not mere scrambling:

(47)a. ?devuski^V citajut^V [stol'ko^V že knjig], [skol'ko]_i ja xoću^V, [_S, čtoby^V
girls read as many books as many I want that

mal'čiki^V smotreli [e]_i fil'mov]]
boys watch films
(nom plur) (subjunct) (gen plur)

'the girls are reading as many books as I want the boys to watch films'

It comes as no surprise that long movement of skol'ko is impossible from the NP subject in (44)b/(45)b. Interestingly, long movement of skol'ko is possible from the QPs in (44)a/(45)a. This is straightforwardly accounted for if the QPs are S-structure objects, as we have argued. The contrast in (48)-(49) thus provides more evidence for our analysis in Chapter Two:

(48)a. ?priš^vlo [stol'ko ž^ve devušek], [skol'ko]_i ja xoč^vu, č^vtoby uš^vlo
 came as many girls as many I want that leave
 (neut sing) (gen plur) (neut sing)

[_{QP^e-i} mal'čikov]
 boys
 (gen plur)

b. *[stol'ko ž^ve devušek] priš^vli, [skol'ko]_i ja xoč^vu, č^vtoby [_{NP^e-i}
 as many girls came as many I want that
 (gen plur) (plur)

mal'čikov] uš^vli
 boys leave
 (gen plur) (plur)

(49)a. ?[skol'ko]_i ty xočeš^v, č^vtoby uš^vlo [_{QP^e-i} mal'čikov]?
 how many you want that leave boys
 (neut sing) (gen plur)

b. *[skol'ko]_i ty xočeš^v, č^vtoby [_{NP^e-i} mal'čikov] uš^vli?
 (plur)

The contrasts in (48)-(49) (unlike their French counterparts, for some speakers) are extremely sharp, regardless of word order or other factors. The contrasts follow straightforwardly as a subject/object asymmetry in a PCC or KECP framework, as discussed for French above. A similar contrast is found, as expected, in the subject position of small clauses:

(50)a. ?ja kupil [stol'ko ž^ve grammatik], [skol'ko jazykov]_i ja xoč^vu,
 I bought as many grammars as many languages I want

č^vtoby ty scital [_{A^e-i} interesnymi]
 that you consider interesting

b.?*ja kupil [stol'ko^v že grammatik], [skol'ko]_i ja xoću, ^včtoby ty
 I bought as many grammars as many I want that you
^vsčital [_{A*}[e]_{-i} jazykov] interesnymi]
 consider languages interesting
 (gen plur)

'I bought as many grammars as I want you to consider languages interesting'

(51)a.?[skol'ko jazykov]_i ty xočeš', ^včtoby ja ^vsčital [_{A*}[e]_{-i} interesnymi]?
 how many languages you want that I consider interesting

b.?*[skol'ko]_i ty xočeš', ^včtoby ja ^vsčital [_{A*}[e]_{-i} jazykov] interesnymi]?
 how many you want that I consider languages interesting
 (gen plur)

'how many languages do you want me to consider interesting?'

One might claim that the asymmetries seen above, particularly (48)-(49) are not specifically subject/object asymmetries of the type covered by the PCC or KECP, but rather derive from some more basic difference between NP and QP. We can dispose of this objection by noting that extraction of a modifier of an NP is quite generally grammatical in Russian. For example, the quantifiers stol'ko and skol'ko cannot be the head of their constituents in oblique positions, as we demonstrated in Chapter Two. Only NPs can bear oblique Case features. Like numerals, stol'ko and skol'ko show adjectival behavior in oblique positions. They cease to require or allow genitive Case on their noun; instead, the noun bears the appropriate oblique Case, and stol'ko/skol'ko agrees with the noun, like any adjective. When an oblique modifier in NP, skol'ko may

still be extracted:

(52)a. xudožnik^v napisal [stol'ko že^v portretov], [skol'kim staruškam]_i^v
 artist drew as many portraits as many old ladies
 (dat plur) (dat plur)

ja xotel, čtoby ty pomog e_i

I wanted that you help

b. xudožnik^v napisal [stol'ko že^v portretov], [skol'kim]_i ja xotel
 artist drew as many portraits as many I wanted
 (dat plur)

čtoby ty pomog [_{NP-e}_i staruškam]^v

that you help old ladies
 (dat plur)

'the artist drew as many portraits as I wanted you to help
 old ladies'

(53)a. [skol'kimi jazykami]_i ty xoeš'^{v v}, čtoby ja vladel?^v
 how many languages you want that I command
 (instr plur) (instr plur)

b. [skol'kimi]_i ty xoeš'^{v v}, čtoby ja vladel [_{NP-e}_i jazykami]?^v
 how many you want that I command languages
 (instr plur) (instr plur)

'how many languages do you want me to command?'

Returning to (48)b, we note that what goes wrong with extraction from within a subject, here as in other languages, is the fact that the path from within the subject lacks the INFL' node present in the INFL'-S' path between INFL and COMP, while the INFL'-S' path lacks the subject NP.

As in French and English, we expect that the addition of another gap lower on the tree will ameliorate the situation, because the union of this path with the offending path from inside the subject will include INFL'.

Such parasitic gap structures with skol'ko movement strike informants as extremely odd. Nonetheless, for some speakers, the addition of a second gap appears to improve Subject Condition violations:

(54)a. *[stol'ko^v ze devu^vsek] ^včitajut knigi, [skol'ko]_i ja xo^vcu, ^včtoby
 as many girls read books as many I want that

[e_i mal'^včikov] smotreli fil'my
 boys watch films
 (gen plur) (plur) (acc plur)

b.??[stol'ko^v ze devu^vsek] ^včitajut knigi, [skol'ko]_i ja xo^vcu, ^včtoby
 as many girls read books as many I want that

[e_i mal'^včikov] smotreli [e_i fil'mov]
 boys watch films
 (gen plur) (plur) (gen plur)

'as many girls read books as I want boys to watch films'

The interrogative equivalent of (54)b does not appear to be better than the interrogative equivalent of (54)a, for mysterious reasons. Also, because all verbs like persuade take an animate object, and because animate genitive plurals have the same form as accusative plurals, we cannot test the full paradigm that we examined in French, to see if the position of the second gap is crucial. Nonetheless, the general parallelism with French suggests that the same forces are at work. Once again, the PCC and KECP offer explanations of the paradigm, with the KECP facing the objections

we have discussed.

2.4 Russian QPs Revisited

We are now in a position to reexamine our Chapter Two analysis of quantificational constructions in Russian. We will show that the PCC accounts for certain subtleties in judgments about these constructions that remained unexplained under an ECP account. A translation of our analysis into the PCC framework is fairly straightforward, but raises certain problems which bear on the distinction between the PCC and the KECP. Unfortunately, we will leave some of these issues unresolved, but they are worth noting nonetheless.

Kayne (1981a) sketches an analysis of the French counterpart to the Russian genitive of negation construction which translates naturally into the KECP framework of Kayne (1982). In essence, he proposes that the subject/object asymmetry characteristic of this construction is the same asymmetry as that found in the combien-extraction construction, with the negation playing the antecedent role. Thus, he equates the following two paradigms:

(55)a. combien_i est-ce que tu penses que Paul aime [e_i de femmes]? (cf. (29)b)

how many do you think that Paul loves of women

b. ??combien_i est-ce que tu penses que [e_i d'hommes] aiment Marie?

how many do you think that of men love Marie

(cf. (30)b)

(56)a. ?personne_i ne pense que Paul aime [e_i de femmes]
 nobody (neg) thinks that Paul loves of women

b. *personne_i ne pense que [e_i d'hommes] aiment Marie
 nobody (neg) thinks that of men love Mary

In a KECP framework, the empty QP in (56)b fails to connect with its antecedent negation (personne) as in any Subject Condition case. We might adopt a similar analysis in a PCC framework, at a level at which personne is in an \bar{A} -position, after the rule of QR has applied to it.¹³ The path from the null AP inside the subject to its binder personne will overlap the path between INFL and COMP in the lower clause -- again, as with any violation of the Subject Condition.

The two paradigms diverge, however. Recall that the addition of a second gap to (55)b can save the structure from PCC (or KECP) effects. The addition of a second gap to (56)b, however, does nothing to improve its status:

(55)c. ?combien_i est-ce que tu penses que [e_i d'hommes] aiment [e_i de femmes]

(56)c. *personne_i ne pense que [e_i d'hommes] aiment [e_i de femmes]

Clearly something more is at stake in (56) than in (55). The relevant paradigms can be fairly well duplicated in Russian:

(57)a. ?devu^vski citajut stol'ko že knig, skol'ko_i ja xoču, čtoby mal'čiki
 girls read as many books as many I want that boys
 (nom plur)

smotreli [e_i fil'mov] (= (47)a)
 watch films
 (gen plur)

b.* . . . skol'ko_i ja xoću, čtoby [e_i mal'čikov] smotreli fil'my
 as many I want that boys watch films
 (gen plur) (acc plur)

c.?? . . . skol'ko_i ja xoću, čtoby [e_i mal'čikov] smotreli [e_i fil'mov]
 as many I want that boys watch films
 (gen plur) (gen plur)

(58)a. nikakie mal'čiki ne_i smotreli [e_i nikakix fil'mov]
 no boys NEG watched no films
 (nom plur) (plur) (gen plur)

b.*[e_i nikakix mal'čikov] ne smotreli nikakie fil'my
 no boys NEG watched no films
 (gen plur) (acc plur)

c.*[e_i nikakix mal'čikov] ne_i smotreli [e_i nikakix fil'mov]
 (gen plur) (gen plur)

Either of two lessons may be drawn from this failure of parallelism between the constructions. We believe that both are probably correct, but only one follows from the discussion below.

The conclusion that does not follow from our future discussion, but which is worth considering nonetheless, is that the contrast between (55)c and (56)c, (57)c and (58)c derives from the fact that only in (55) and (57) (combien and skol'ko extraction) is there an antecedent, an \bar{A} -binder for the empty categories in question. That is, assuming that the negation in (56) and (58) acts like the WH-moved quantifier in (55) and (57) correctly predicts a subject/object asymmetry, but incorrectly predicts the possibility of a parasitic gap cancelling the asymmetry.¹⁴ If we assume, on the other hand, that the empty categories in (56) and (58) are antecedentless, we must find another account for the subject/object

asymmetry observed there, but have some hope of not making a false prediction about (56)c and (58)c. We return to this question shortly.

The account we have in mind is, of course, a version of our account in Chapter Two. Recall that we argued in Chapter Two that the Russian genitive of negation construction involved a constituent headed by a null Q:

(59) nikakie mal'^vciki ne smotreli [_{QP}[_{Q-}^e] nikakix fil'mov] (= (58)a)

From this analysis follows immediately the immunity of these constituents to the Case Filter assumed there, and the inability of these constituents to occur where oblique Case was required.

We reasoned further as follows: verbs like smotret' 'watch' c-select an NP, not a QP. The very existence of these QPs implies that c-selection does not apply at all levels. If it applies at LF, something must happen to replace the QP in structures like (59) with the needed NP. We argued that the rule of QR applying at LF can have this effect: the trace of the moved QP, given the null hypothesis that trace theory follows in toto from other principles, can be the required NP. The QP will be a logical binder of the NP, but will not be a syntactic antecedent of the sort which can govern by coindexation, given Chomsky's (1981a) definition of government and his ECP. It thus followed that a QP can never occur where an NP is required unless its position is lexically governed; government by coindexation will never allow it to meet the terms of the ECP. Specifically, given the following structures --:

(60)a. QP_i [. . . [_{VP} V [_{NP} e]_i]

b. QP_i [[_{NP} e]_i INFL VP]

-- only in (60)a is the NP empty category properly governed. In (60)a it is lexically governed by V; in (60)b it is governed, but not properly governed by INFL, and it is not properly governed by coindexation with the QP. Hence the subject/object asymmetry seen in (58).

Before recasting this analysis in a PCC framework, let us take another look at the failure of additional gaps to ameliorate the French or Russian genitive of negation in subject position. Recall that parasitic gaps are locally \bar{A} -bound (Chomsky 1981b; we return to this in the last section of this chapter). From this it follows that an anti-c-command condition prevails among the various gaps bound by a common binder. If one gap c-commanded a second gap, the second would be locally A-bound by the first. They would form a chain which would violate the θ -criterion, since the gaps bear distinct θ -roles.

Kayne (1981a) assumes, as he must, given more traditional assumptions about c-selection, that the French phrases of the form $[[Q_e] \text{ de } N]$ are NPs, headed by the N. On a theory like Kayne's, one could explain the absence of the parasitic gap structure of (56)c by assuming, with our Chapter Two, that the Q is the head. It would follow that the index of the Q would be the index of the constituent as a whole. The two empty Qs could not then be locally \bar{A} -bound by a common antecedent, because a constituent bearing the index of one would c-command a constituent bearing the index of the other:

(61) *personne_i ne pense que [_S [_{QP} [_{Q⁻} e] d'hommes]_i [_{VP} aime^t [_{QP} [_{Q⁻} e] de femmes]_i]] (= (56)c)

But this analysis would raise the problems with c-selection that our

analysis in Chapter Two raised. If we were right, this analysis leads directly to a rather different account of the subject-object asymmetries in (56) and (58) -- in particular, an analysis like that in Chapter Two. In a sense, however, this account of the failure of (56)c and (58)c will be maintained in our analysis.

In order to accommodate our analysis of these contrasts to a PCC framework, we must change our notion of paths somewhat. In the ECP framework, we captured the ungrammaticality of the NP trace of QR in (60)b by treating it as a syntactically unbound variable. That is, the coindexing with QP in (60)b is not syntactically relevant, due to the difference in category. We could capture this effect in a PCC framework if empty categories like this subject NP generated infinite paths. Recall that the paths we have considered so far are by definition line segments, bounded by two nodes on the tree. We wish to suggest that syntactically unbound empty categories like that in (60)b also create paths, which are not line segments but rays, bounded by only one node. In other words, consider first an embedded instance of a structure like (60)b, abstractly represented:

(62) . . . [_{INFL₁'} [_{VP₁^V} [_{S₂'} [_{S_{2b}^{QP_i}} [_{S_{2a}'} [_{NP₂^e} [_{INFL₂'} [_{VP₂^V} . . .]]]]]]]]]

Paths

(i) Between INFL and COMP of S₂':

{INFL₂', S_{2a}', S_{2b}', S₂'}

(ii) From NP₂:

{S_{2a}', S_{2b}', S₂', VP₁, INFL₁' . . . }

Notice that paths (i) and (ii) overlap, but neither one contains the other. The situation is analogous to a CTP effect in character.

Now consider an embedded instance of a structure like (60)a:

(63) . . . [_{INFL₁'} [_{VP₁} V [_{S₂'} [_{S_{2b}} QP_i [_{S_{2a}} NP [_{INFL₂'} [_{VP₂} V [_{NP₂} e_i]]]]]]]]]]

Paths

(i) Between INFL and COMP of S₂':

{INFL₂', S_{2a}, S_{2b}, S₂'}

(ii) From NP₂:

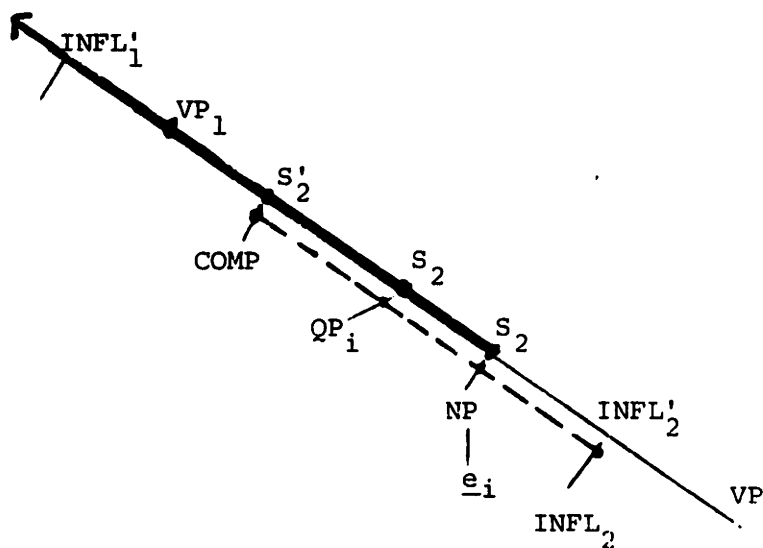
{VP₂, INFL₂', S_{2a}, S_{2b}, S₂', VP₁, INFL₁' . . . }

In (63), paths (i) and (ii) overlap, and path (ii) contains path (i).

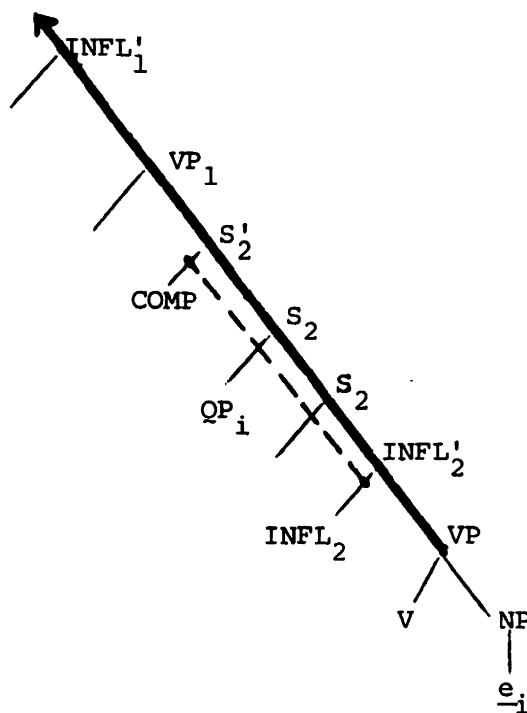
The situation is analogous to long movement of an object.

We can represent the paths in (62) and (63) graphically as follows:

(62)



(63)



Treating the NP empty categories in these constructions as generating infinite paths raises some serious problems, to which we return. In the next chapter we will examine some less problematic cases of infinite paths. For now let us look at some of the advantages of this approach.

The definition of path given in (2) may be revised in one of a number of ways to allow for infinite paths. One possibility is to replace the requirement that \underline{t} be locally \bar{A} -bound with the requirement that it be locally A-free. This allows \underline{t} to be either locally \bar{A} -bound or A- and \bar{A} -free. Certain other adjustments are then required, yielding (64):

(64) Definition of Paths (revised: II)

Consider T such that for some index \underline{i} :

$T = \{ \underline{t} \mid \underline{t} \text{ is an empty category locally A-free and indexed } \underline{i} \}$

- (i) for a set $A = \{ \alpha \mid \exists t \in T: \alpha \text{ is the first maximal projection dominating } \underline{t} \}$

- (ii) for β the first maximal projection dominating \underline{b} , where
 $\forall t \in T: \underline{b}$ locally binds t
- (iii) $P = \{x \mid (x \in A) \vee (x = \beta) \vee (\exists \alpha \in A: x \text{ dom. } \alpha \ \& \ \neg x \text{ dom. } \beta)\}$ ¹⁵

We interpret (64)ii as not implying the existence of \underline{b} . In such a case, members of \underline{t} will generate an infinite path beginning with the maximal projections dominating each member of \underline{t} . Thus, we allow identically indexed free categories to form a united path. An instance of this might arise, for example, as the result of WH-movement of a QP, if the QP logically binds two NP gaps, which it cannot syntactically bind.

Thus, it appears so far that our Chapter Two analysis of the Russian genitive of negation construction, as well as other quantificational constructions, translates fairly straightforwardly into the PCC framework. (We will shortly see some problems with this translation.) Notice now that the translation requires that the PCC, like Chomsky's ECP and Kayne's KECP, apply at LF, after the rule of QR has applied. Let us now ask if the PCC applies only at LF. We have only shown that it must apply at least at LF. There is, in fact, some evidence that the PCC applies at S-structure as well (and cf. 4.4 of this chapter).

We wish to suggest, in fact, that Kayne was right in paying attention to the status of the empty Q with respect to the KECP. Let us consider, first, what predictions the PCC makes about the LF grammaticality of all QPs that occupy NP positions in S-structure. Because of the results of QR, they should be grammatical as direct objects, grammatical as subjects of small clauses governed by a higher V, and ungrammatical as subjects of tensed sentences. This is because the unbound NP trace of QP after QR has essentially the status of a trace of long WH-movement with

respect to the PCC.

Now consider the null Q that occurs inside the genitive of negation construction, but not in other QPs, such as numeral phrases, which have a phonologically realized head. If the PCC applies to this null Q, it will have essentially the status of any empty category contained inside an argument -- for example, the trace of French combien or Russian skol'ko 'how many', discussed above. The PCC predicts that such a Q will be grammatical inside a direct object, ungrammatical inside the subject of a small clause, and ungrammatical as the subject of a tensed sentence. In other words, if the PCC applies both at S-structure and at LF, it has the following consequences for the genitive of negation construction:

(65) Genitive of Negation

	Direct Object	Subject of Small Clause	Subject of Tensed S
S-structure	OK	*	*
Logical Form	OK	OK	*

The PCC will have no effect on other QPs, like numeral phrases, which contain no empty categories, until QR takes place at LF. Thus, the PCC has the following consequences for numeral phrases, po-phrases, and the like:

(66) Numeral Phrases, etc.

	Direct Object	Subject of Small Clause	Subject of Tensed S
S-structure	OK	OK	OK
Logical Form	OK	OK	*

Now recall two problems with our analysis of Russian QPs that we raised briefly in Chapter Two. First, we noted that the subject/object

asymmetries found with the genitive of negation appeared to be much stronger for most speakers than those found with numeral phrases and po-phrases. That is, although for simplicity's sake we assigned stars equally to (67)a and (67)b in Chapter Two, a more accurate reflection of speakers' judgments would be:

(67) a. * $[_{QP} [_{Q^e}] \text{ni odnogo studenta}] \text{ ne ubilo lošad}'$
 not one student NEG killed horse
 (gen sing) (neut sing) (acc)

b. ?? $[_{QP} [_{Q^v} \text{šest' studentov}]] \text{ ubilo lošad}'$
 six students killed horse
 (gen plur) (neut sing) (acc)

c. ?? $[_{QP} [_{Q^po}] \text{studentu}] \text{ ubilo lošad}' \text{ v každoj komnate}$
 po student killed horse in each room

Notice now that the PCC rules (67)a out at two levels -- S-structure and LF -- while (67)b and (67)c are ungrammatical at one level only -- LF. Supposing, as seems reasonable, that the degree of ungrammaticality is proportional to the number of levels at which ungrammaticality is determined, then we explain the contrast very simply.

A second problem was the unexpected reduced acceptability of the genitive of negation as the subject of a small clause:¹⁶ after QR, as we see in (65), an NP trace in this position which is not syntactically bound should be grammatical. The reduced acceptability is also found in the French equivalent of the genitive of negation:

(68)a. ??ja ne scitaju [_{A*} [_{QP} [e] inostrannyx fil'mov] [_{AP} interesnym(i)]]
 I NEG consider foreign films interesting
 (gen plur) (instr sing/plur)

b. ??je ne considère pas [_{A*} [_{QP} [e] de films étrangers] [_{AP} intéressants]]

(68)a-(68)b, while grammatical at LF, are ungrammatical at S-structure: they are, essentially, Subject Condition violations. Hence the weak ungrammaticality. In other words, by applying the PCC to QPs at both S-structure and LF, we explain not only where QPs are ungrammatical, but the degree of ungrammaticality, if we associate two "OKs" with "OK", one "OK" with "??", and two stars with a star.

One final note: (66) predicts that QPs like numeral phrases should be fully grammatical as subject of small clauses. We noted in Chapter Two that there was no outward difference in form between a QP numeral phrase and an NP numeral phrase in a non-oblique Case position. Therefore, all things being equal, we cannot tell whether the numeral phrase in (69)a is an NP or QP. The impossibility of the adjective being singular, seen in (69)b, however, suggests that a QP may still be impossible, for obscure reasons:

(69)a. ja scitaju [s^vest' fil'mov] interesnymi
 I consider six films interesting
 (gen plur) (instr plur)

b. *ja scitaju [s^vest' fil'mov] interesnym
 (instr sing)

Alternatively, this may bear on the question of whether the lack of agreement in tensed sentences is due to an intrinsic QP/NP distinction (e.g. QP does not bear person, number and gender features) or to the

general impossibility of surface QP subjects. Since we predict that a QP is possible in (69)a, it may be that the impossibility of (69)b shows that QPs do bear person, number and gender features. In that case, structures like (67)b would only arise if the QP was actually a direct object at S-structure, suggesting a reanalysis of the verb and its "semantic" direct object (cf. Pollock 1981 for a suggestion along these lines for French il-impersonals.)

Finally, some severe empirical problems that arise in the translation of our Chapter Two analysis into a PCC framework. We raise them here, without offering a solution.

We noted several times that the unbound NP found after LF movement of a QP from a position where an NP is selected will have the properties of a WH-trace bound outside its clause. That is, the strong ungrammaticality of a QP moved from the subject of a tensed sentence derives from the interaction of its path with the path between INFL and COMP in exactly the way the ungrammaticality of a CTP violation derives. Unfortunately, these unbound NPs do not act in all respects like the trace of long-moved WH. If we look at examples from French, where the effects seem clearer, we note that, as predicted, the French "genitive of negation" construction cannot occur inside a subject, but only if the negation is outside the subject:

(70)a. ?personne ne croit qu'il est nécessaire de manger [_{QP}e de pommes]
 nobody (neg) believes that it is necessary to eat of apples

b. *personne ne croit que manger [_{QP}e de pommes] est nécessaire

(71) je crois que ne pas manger e de pommes est nécessaire
 I believe that to not eat of apples is necessary

These examples suggest that the path from the QP, or its NP trace, is stopped in some fashion by the negation, which does act as an antecedent, as in Kayne's (1981a) theory. Only (70)b, then, will show a Subject Condition effect, since only in (70)b does the path from the QP or its trace extend beyond the sentential subject. In (71), the negation internal to the sentential subject appears to limit the path from QP or its trace to within the sentential subject itself, preventing a Subject Condition effect. How this works, however, and how to integrate this effect into our general theory of these constructions, is unclear.

Also, the alleged infinite path from the NP trace of QP does not interact with paths created by WH-movement, as predicted. That is, (72)a does not have the status of (72)b, as it should:

- (72)a. ces livres, que_i je n'ai pas persuadé [_{QP}e d'hommes] de lire e_i
 these books which I have not persuaded of men to read
- b.*ces hommes, que_j je sais quels livres_i persuader e_j de lire e_i
 these men who I know what books to persuade to read

For this construction, these problems may suggest an abandonment of infinite paths, if further research presents a better analysis in the PCC framework. Lest the reader be overly discouraged, however, we note that in Chapter Five we will present good evidence that WH-in-situ constructions involve infinite paths that do interact exactly as predicted with the path between INFL and COMP and with paths created by WH-movement.

In this section, we have shown, first, that extraction of left branches in English, French and Russian, has exactly the properties of extraction of right branches, contrary to what an approach that relies crucially on

left/right asymmetries might lead one to expect. We have shown that left branch extraction in Russian provides some new evidence that apparent QP subjects are actually S-structure direct objects, that our QPs are indeed headed by the quantifier, and perhaps that paths should be allowed that do not terminate with a node dominating an antecedent binder. In the next section we discuss the interaction of Crossing effects with our theory of parasitic gap constructions.

2.5 Parasitic Gaps and Crossing Effects

The PCC claims to unify subject/object asymmetries captured by Chomsky's and Kayne's ECP with Crossing effects. In Chapter Three we showed that this unification is technically possible: we suggested that the dependency between COMP and INFL is an instance of \bar{A} -binding, and constraints on extraction from the subject resulted from Crossing effects with respect to the COMP-INFL dependency. This unification will carry weight, however, if we can show that subject/object asymmetries and Crossing effects do have the same properties: in particular, if properties of grammar that interact with subject/object asymmetries also interact with Crossing effects, and if analyses of subject/object asymmetries make correct predictions about Crossing effects. (A failure on this score was one of the problems with our revision of our Chapter Two analysis of Russian QPs.) In this section we will show that our "forked path" theory of parasitic gap constructions shows the sort of Crossing effects predicted by the PCC.

First of all, consider a straightforward Crossing violation like (73):

(73) *Warlpiri is one language that_i I know who_j to persuade native speakers of e_i to talk to e_j

(73) violates the PCC, whether or not native speakers of is reanalyzed with the verb persuade (assume talk to is reanalyzed, which is equally irrelevant):

(73)a. *one language [_{S'₁} that_i I know [_{S'₂} who_j to [_{VP₂} [_V persuade-native-speakers-of] e_i [_{S'₃} to [_{VP₃} [_V talk to] e_j]]]]]]

b. *one language [_{S'₁} that_i I know [_{S'₂} who_j to [_{VP₂} [_V persuade] [_{NP₂} native speakers [_{PP₂} of e_i]] [_{S'₃} to [_{VP₃} [_V talk to] e_j]]]]]]

Ignoring irrelevant nodes and paths, (73)a and (73)b both violate the PCC:

(73)a. (i) Between e_j and COMP of S'₂:

{VP₃, S'₃, VP₂, S'₂}

(ii) Between e_i and COMP of S'₁:

{VP₂, S'₂, S'₁}

b. (i) Between e_j and COMP of S'₂:

{VP₃, S'₃, VP₂, S'₂}

(ii) Between e_i and COMP of S'₁:

{PP₂, NP₂, VP₂, S'₂, S'₁}

Under either analysis, the two paths overlap, but neither one contains the other. Hence the PCC is violated.

Let us follow the same logic we used at the beginning of this chapter when we considered Subject Condition effects. What operation could we

perform on the paths in (74) that would eliminate the violation of the PCC? In (74)a, if we could add the nodes S'_3 and VP_3 to path (ii), path (ii) would contain path (i), and the PCC would be satisfied. These nodes can be added to path (ii) if another gap bound by that_i is added to VP_3 . As predicted, the result of adding such a gap noticeably improves (73). Compare (73) with (74):

(74) ??Warlpiri is one language that_i I know who_j to persuade native speakers of \underline{e}_i^1 to talk to \underline{e}_j about \underline{e}_i^2

With or without reanalysis of native speakers of, (74) now satisfies the PCC:

(74)a. one language [S'_1 that_i I know [S'_2 who_j to [VP_2 [V persuade native speakers of] \underline{e}_i^1 [S'_3 to [VP_3 [V talk to] \underline{e}_j [PP_3 about \underline{e}_i^2]]]]]]]

b. one language [S'_1 that_i I know [S'_2 who_j to [VP_2 [V persuade]

[NP_2 native speakers [PP_2 of \underline{e}_i^1]] [S'_3 to [VP_3 [V talk to] \underline{e}_j [PP_3 about \underline{e}_i^2]]]]]]]

Paths

(74)a. (i) Between \underline{e}_j and COMP of S'_2 :

{ VP_3 , S'_3 , VP_2 , S'_2 }

(ii) Between \underline{e}_i^1 , \underline{e}_i^1 , and COMP of S'_1 :

{ PP_3 , VP_3 , S'_3 , VP_2 , S'_2 , S'_1 }

b. (i) Between e_j and COMP of S'_2 :

{VP₃, S'₃, VP₂, S'₂}

(ii) Between e_i^1 , e_i^2 , and COMP of S'_1 :

{PP₃, VP₃, S'₃, PP₂, NP₂, VP₂, S'₂, S'₁}

In both (74)a and (74)b, paths (i) and (ii) overlap, and path (ii) contains path (i), as required.

Notice, however, that the anti-c-command condition on parasitic gaps rules out (74)a, in which persuade native speakers of is reanalyzed. In other words, the structure in (74)a is ruled out for the same reasons as (75) is ruled out: e_i^1 c-commands e_i^2 :

(75) *John, who_i I persuaded e_i^1 that I could talk to e_i^2

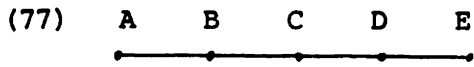
The anti-c-command requirement on parasitic gaps thus implies what we have already argued to be true: reanalysis is not a precondition for preposition stranding in English.

We have noted that the contrast between (73) and (74) is explained exactly the same way as the contrasts discussed at the beginning of this chapter, e.g.:

(76)a. *a person who_i close friends of e_i admire me

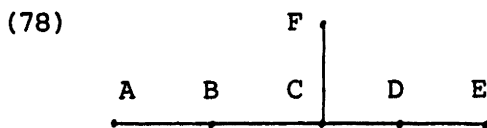
b. ?a person who_i close friends of e_i admire e_i (= (1))

Let us consider the geometry of the contrasts (73)/(74) and (76)a/(76)b somewhat more abstractly than we have so far. For example, consider a line segment \overline{ABCDE} below. \overline{ABCDE} can represent a set of contiguous nodes in a tree -- a path, where each point is a node in the path:

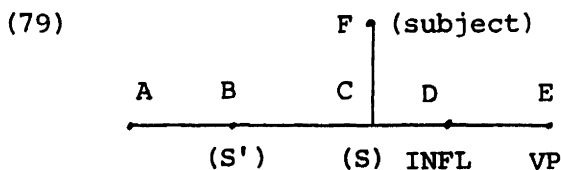


Given \overline{ABCDE} , we can use subsegments of \overline{ABCDE} as abstract illustrations of the paths created by empty categories, according to our definitions. The PCC can be taken to apply equally well to these subsegments, or to the paths made up of the labelled nodes in these subsegments. Suppose we have the paths $\{A, B, C, D, E\}$ and $\{B, C, D\}$ -- that is, the segments \overline{ABCDE} and \overline{BCD} . The first path or segment contains the second, and the PCC is satisfied. Suppose we have the paths $\{A, B, C, D\}$ and $\{B, C, D, E\}$ -- or the corresponding segments. Neither path nor segment contains the other, and the PCC is violated. Add E to the first, or A to the second, however, and the PCC is no longer violated. This "adding" of a point to a path is what happens when a parasitic gap is added to a tree. We speak of "adding" a point to a path as a metaphor: to understand the geometry of paths formed by multiple gap constructions, it is useful to compare the paths with and without the extra gaps.

In (73)-(74), and in (76)a-(76)b, we actually have a path configuration more like (78):



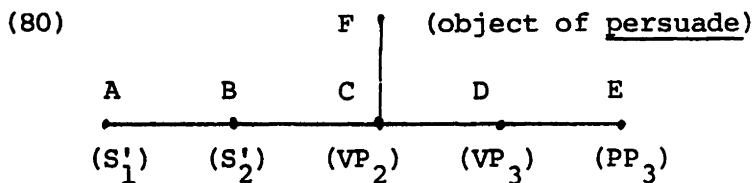
For (76), think of the points in (78) as corresponding to the nodes below:



In (76)a, the path from inside the subject to COMP is {B, C, F}. The path between INFL and COMP is {B, C, D}. The PCC is not satisfied, since the two paths overlap, and neither one contains the other. In (76)b, however, the path from the two gaps to COMP is {B, C, D, E, F} and the PCC is satisfied.

Alternatively, we may think of (76)b as satisfying the PCC in the following way: the PCC is "primarily" satisfied by virtue of the path from the object gap to COMP and the path between INFL and COMP. The path from the object gap is the line segment \overline{BCDE} ; the path between INFL and COMP is the line segment \overline{BCD} . The former contains the latter. Since that is so, it does not matter to the PCC if we add the line segment \overline{CF} to the longer segment, \overline{BCDE} , as a "spur". \overline{BCDE} contains \overline{BCD} with or without the spur of \overline{CF} .

For (73)-(74), think of the points in (78) as corresponding to the following nodes:



In (73), the path from e_j to COMP to S'_2 is {B, C, D}. The path from e_i to COMP of S'_1 is {A, B, C, F}. They overlap, but violate the PCC. In (74), where an additional gap is added, the path from both occurrences of e_i to COMP of S'_1 is {A, B, C, D, E, F}, and the PCC is satisfied

Once again, we may think of (74) as "primarily" satisfying the PCC because the path between the lowest occurrence of e_i and its binder -- \overline{ABCDE} -- and the path between e_j and its binder -- \overline{BCD} -- satisfy the PCC.

Since \overline{ABCDE} already contains \overline{BCD} , it does not matter at all that we have attached a spur, \overline{CF} , to \overline{ABCDE} . Even the union of \overline{CF} and \overline{ABCDE} still contains \overline{BCD} .

The principle involved is, of course, quite simple: given sets P, Q and R, if P is a subset of Q, it is a subset of the union of Q and R. Hence, if paths A and B satisfy the PCC, and if path A contains path B, path A will still contain B even if we add extra nodes to it, as in a multiple gap construction.

If our analyses have been correct, we expect an even more interesting consequence to follow from our approach. In (80), we considered the segments \overline{ABCDE} and \overline{BCD} . We noted that the first continues to contain the second even if we add to it the line segment \overline{CF} . The addition of nodes to the longer of two paths has no effect on containment relations. Suppose, however, that we add the spur \overline{CF} to the shorter of the two paths, yielding paths {A, B, C, D, E} and {B, C, D, F}. Clearly the PCC is violated. In other words, the PCC predicts that, given paths A and B, such that A contains B, we may add nodes not contained in B to A, but we may not add nodes not contained in A to B. We believe that this prediction is correct.

Consider (81):

(81) ?this Volvo is one car that_i I know who_j to persuade [owners of e_i¹]
to [talk to] e_j [about e_i²]

This sentence has the structure of (74), discussed in (80) above. The path from e_i² to that_i contains the path from e_j to who_j. It is thus no problem to add a spur from the NP containing e_i¹ to the larger of the two paths.

But now compare (82), which differs from (81) in the index on e^1 :

(82) *this Volvo is one car that_i I know who_j to persuade [friends of e_j^1]
to [talk to] e_j^2 [about e_i]

As indexed, (82) is completely impossible. It has a good interpretation, of course, in which e_j^1 bears instead the index i , such that the sentence speaks of "friends of the car", but has no interpretation as given. It is worth mentioning that, despite the complexity of both examples, the contrast seems to be exceptionally sharp for all speakers I have asked. This is not surprising, if the sentence is ruled out exactly as other Crossing sentences are ruled out.

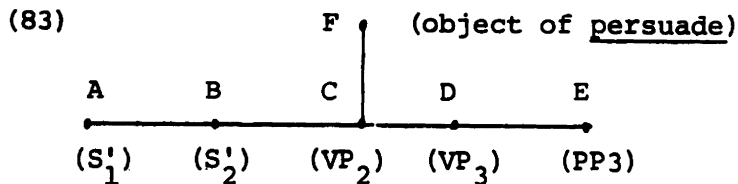
Assuming the structure for (82) given below, the paths are as follows:

(82) *car [_{S₁} that_i I know [_{S₂} who_j to [_{VP₂} persuade [_{NP₂} friends [_{PP₂} of e_j^1]] [_{S₃} to [_{VP₃} [talk to] e_j^2 [_{PP₃} about e_i]]]]]]]

Paths

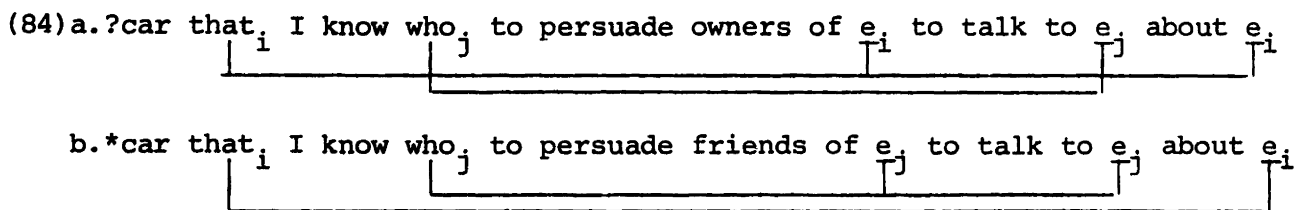
- (i) Between e_j^1 , e_j^2 and COMP of S'_2 :
{VP₃, S₃, PP₂, NP₂, VP₂, S₂, S₁}
- (ii) Between e_i and COMP of S'_1 :
{PP₃, VP₃, S₃, S₂, S₁}

Paths (i) and (ii) overlap, but neither one contains the other, violating the PCC. Returning to our diagram in (78), we have the following picture:

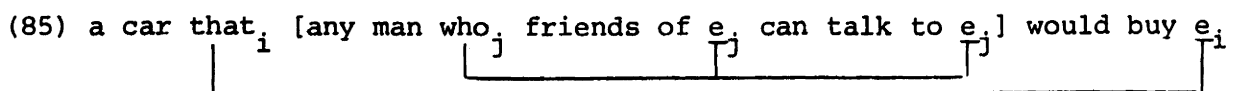


(83) is the same as (80). As in (80), the path from e_i to its binder, $\{A, B, C, D, E\}$, satisfies the PCC, because it overlaps and contains the path $\{B, C, D\}$ between e_j^2 and its binder. The line segment \overline{ABCDE} contains \overline{BCD} . Where (82) goes wrong is in adding the spur \overline{CF} to the smaller line segment instead of the larger. The union of $\{B, C, D\}$ with $\{C, F\}$ neither contains nor is contained by $\{A, B, C, D, E\}$.

As a side remark, we note that the contrast between (81) and (82) is a direct counterexample to simple linear theories of Crossing effects, such as that discussed in Chapter Three. Consider the linear representation of the dependencies in (82)-(83), repeated below:



It is the nested arrangement in (84)b that is bad, while the unnested dependencies in (84)a are acceptable. Lest it be thought that some other linear principle rules out the configuration in (84)b, we note that other examples of this configuration, where the structural relations are different, are acceptable (N. Chomsky, personal communication):



Returning to our discussion of the interaction of Crossing effects with parasitic gaps, notice that our analysis predicts that if a construction contains two paths and a spur, and is not congruent to (81) (with the structure of (74)), it will be ungrammatical: (81) shows the only grammatical outcome. For example, in (86) the lowest traces are reversed:

(86)a.*this Volvo is one car that_i I know who_j to persuade [owners of e_i^1]
to [talk about] e_i^2 [to e_j]

b.*this Volvo is one car that_i I know who_j to persuade [friends of e_j^1]
to [talk about] e_i [to e_j^2]

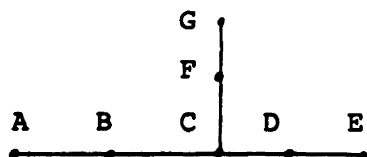
Referring just to the diagram in (83), we note that (86)a contains paths {B, C, D, E} between e_j and its binder, and {A, B, C, D, F} between the two instances of e_i and their binder. Even without the spur \overline{CF} , the two remaining line segments, \overline{BCDE} and \overline{ABCD} , violate the PCC, and the addition of the spur to \overline{ABCD} does not help matters any.

Similarly, (86)b contains paths {B, C, D, E, F} between the two occurrences of e_j and its binder, and {A, B, C, D} between e_i and its binder. The situation is symmetric to that in (86)a. Even without the spur \overline{CF} , the segments \overline{BCDE} and \overline{ABCD} violate the PCC, and adding the spur \overline{BCDE} is no more helpful than adding it to \overline{ABCD} was.

The geometry of the PCC makes some even more complex predictions, which seem to be true, insofar as the relevant sentences can be parsed in the first place. It was somewhat incorrect to state, as we did above, that "given paths A and B, such that A contains B, we may add nodes not contained in B to A, but we may not add nodes not contained in A to B".

In fact, we can add a spur to B, so long as we also add a larger spur to A, such that A's spur contains B's spur. Graphically:

(87)



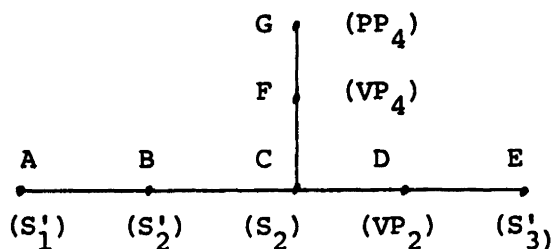
Consider the line segments \overline{ABCDE} and \overline{BCD} . Suppose we add the spur \overline{CF} to \overline{BCD} , and \overline{CFG} to \overline{ABCDE} . The resulting paths, $\{B, C, D, F\}$ and $\{A, B, C, D, E, F, G\}$, will satisfy the PCC. A sentence illustrating this possibility is (88):

(88) (??)this Volvo is one car that_i I know who_j [people that [talk to] e_j [about e_i]] can persuade e_j to buy e_i

Given the structure below for (88), (87) is realized as in (89):

(88) car [_{S'₁} that_i I know [_{S'₂} who_j [_{S₂} [_{NP₄} people that [_{VP₄} [talk to] e_j^1 [_{PP₄} about e_i]]] [_{VP₂} can persuade e_j^2 [_{S'₃} to buy e_i]]]]]]]

(89)



The paths from e_j^1 and e_j^2 to their binder is $\{A, B, C, D, E, F, G\}$. The path from e_i is $\{B, C, D, F\}$. The two paths overlap and satisfy the PCC, since the first contains the second. The "primary" satisfaction of the PCC is due to the two lower traces: the path from e_j^1 to its binder is \overline{BCD} ;

from e_i to its binder is \overline{ABCDE} . To the first path we add the spur \overline{CF} ; to the second, \overline{CFG} . The PCC is satisfied.

If we incorrectly add the spur \overline{CFG} to the smaller "primary" path, and the spur \overline{CF} to the larger primary path, the result, to our ears, is worse than (88):

(90) (*)this Volvo is one car that_i I know who_j [people that [talk about]
 e_j [to e_i]] can persuade e_j to buy e_i

Finally, of course, if the "primary" paths violate the PCC by themselves, addition of spurs will not help them:

(91)a. (*)John is one guy who_j I know what cars_i [people that [talk to]
 e_j [about e_i]] can persuade e_j to buy e_i

b. (*)John is one guy who_j I know what cars_i [people that [talk about]
 e_i [to e_j]] can persuade e_j to buy e_i

(91), without the spurs, has the same status as the familiar Crossing violations:

(92) *John is one guy who_j I know what cars_i Bill persuaded e_j to buy e_i

(Note that (88)-(91) also contain a Subject Condition violation; this is nullified by the lower gaps, equally in all cases, as discussed earlier in this chapter.)¹⁷

In this section we have shown something extremely important for the PCC theory: we have shown that the paths postulated to exist in multiple gap constructions interact with paths created by familiar examples of WH-movement in the same way they interact with the path postulated between

INFL and COMP. Without demonstrations of this sort, we could not be sure that our collapsing of subject/object asymmetries with Crossing effects was a correct move. In the next sections we will show how the PCC accounts very simply for aspects of the Coordinate Structure Constraint, and we will then produce another argument that the relation between INFL and COMP behaves like relations created by movement.

The clarity of some of the judgments in this section also illustrates an important point about Crossing effects. Note that the judgments explained here arise as the result of (1) multiple gaps in contexts where (2) the WH-island condition is violated, resulting in (3) sentences of rather difficult length. Because of (3), all the sentences discussed above are difficult to parse. (2) implies that all these examples are somewhat ungrammatical in the first place. It is highly unlikely that the grammar contains special principles that concern themselves with degrees of ungrammaticality, which are in any case difficult to parse, and unlikely to occur in speech. Chomsky (1981b) notes that the same arguments apply to multiple gap structures in the first place: it is very improbable that Universal Grammar contains principles specific to multiple gap structures. Because of (1), (2) and (3), the phenomena discussed here must be traceable to some more general principle. We feel that we have provided strong evidence that this principle is not a principle of performance, governing the word by word processing of sentences. The principle is therefore a principle of competence. Such a principle we expect to be very general -- not specifically designed for WH-island violations or for multiple gap structures. The PCC meets these requirements.

3.0 Coordination and Multiple Gaps

3.1 The Coordinate Structure Constraint

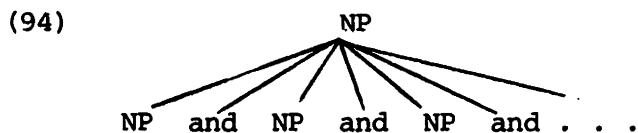
In this section, we show how a single assumption about the properties of coordinating conjunctions can account for exactly that part of Ross's (1967) Coordinate Structure Constraint that does not fall under the A/A principle of Chomsky (1964). In 3.2, we show how the properties of Across-the-Board exception to the Coordinate Structure constraint discussed by Williams (1978) derive from combining the analysis of multiple gaps we have been developing with our analysis of coordination.

It has long been noticed that in a coordinated structure of the form [X C Y C Z ...], where C is a coordinating conjunction and X, Y, Z are elements coordinated by the conjunction no rule may extract X, Y or Z:

(93)a. *who_i did Bill see [Mary and e_i]

b. *Bill, who_i I didn't think that you knew [e_i or Susan]

As Ross (1967) notes, assuming a structure like that in (94), examples like (93) can be easily ruled out by some version of Chomsky's A/A constraint:



Consider the version of the A/A constraint presented by Chomsky (1968; 1973):

(95) If a transformation applies to a structure of the form

[α . . . [_A . . .] . . .]

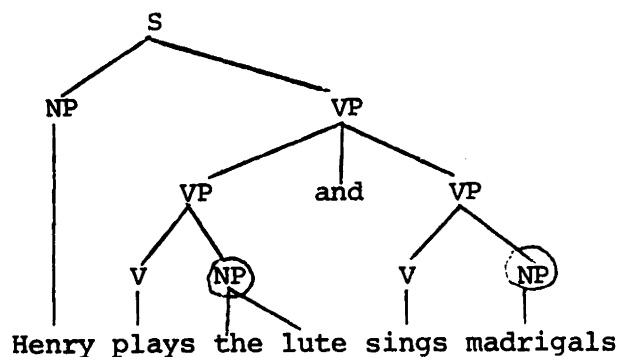
where α is a cyclic node, then it must be so interpreted as to apply to the maximal phrase of the type A.

(Chomsky 1973, (3))

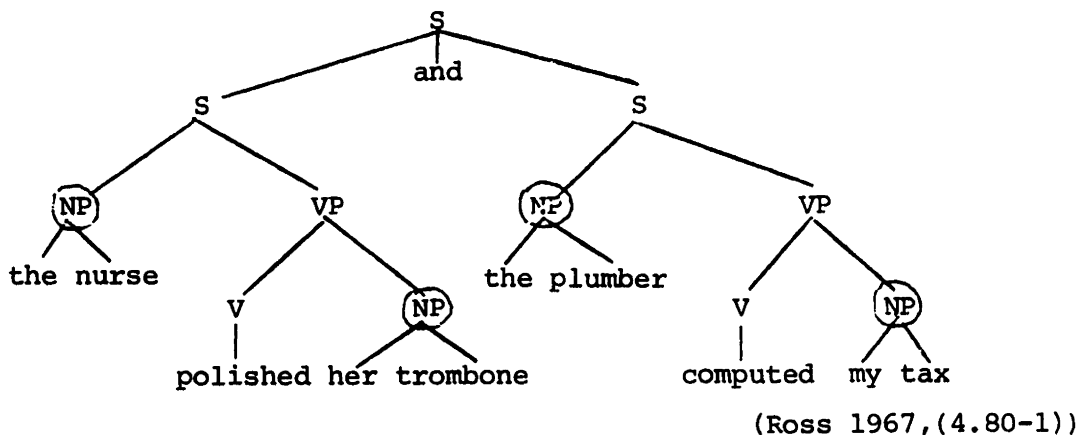
It follows from (95) that no transformational operation can apply to any NP in (94) except the dominating NP (unless the transformation specifies otherwise, which Move α does not!). Hence the ungrammaticality of (93).

Ross goes on to point out that the A/A principle says nothing about extraction of the circled NPs in structures like (96):

(96) a.



b.



Nonetheless, none of the circled nodes may be moved (to an \bar{A} -position):

(97) a. *the lute which_i Henry [_{VP} [_{VP} plays e_i] and [_{VP} sings madrigals]] . . .

b. *the madrigals which_i Henry [_{VP} [_{VP} plays the lute] and [_{VP} sings e_i]] . . .

c. *the nurse who_i [_S [_{S_i} polished her trombone] and [_S the plumber
computed my tax]] . . .

d. *the trombone that_i [_S [_S the nurse polished e_i] and [_S the plumber
computed my tax]] . . .

Ross suggests that the A/A explanation for (93) be abandoned, since it does not extend to (97), and that the grammar include a Coordinate Structure Constraint:

(98) Coordinate Structure Constraint

In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct.

The first clause of (98) covers (93), and the second clause covers (97). Notice, however, that (98) is not an explanatory improvement over the A/A principle. It contains two clauses, the first of which subsumes exactly those facts that are covered by the A/A principle and no others, the second of which describes the facts of (97). Since the Coordinate Structure Constraint does not really unify the cases that fall under A/A with those that do not, there is no argument against continuing to assume the A/A condition, and noting separately the cases that do not fall under it:

(99) In a coordinate structure, no element contained in a conjunct may be moved out of that conjunct.

Grosu (1973) gives a number of arguments that the conflation of cases that do and do not fall under the A/A condition accomplished by Ross's constraint is false. A few of the arguments retain their validity in our framework. One solid argument derives from an observation of Ross's. Ross noted that when a conjunction is asymmetric, as in (100), the Coordinate Structure Constraint may be violated:

- (100)a. I [_{VP} [_{VP} went to the store] and [_{VP} bought some whiskey]]
 b. the whiskey which_i I [_{VP} [_{VP} went to the store] and [_{VP} bought e_i]]
 c. the store that_i I [_{VP} [_{VP} went to e_i] and [_{VP} bought some whiskey]]

Grosu points out, however, that such conjunctions still obey the A/A condition;

- (101)a. John is looking forward to [_{NP}[_{NP}going to the store] and [_{NP}buying some whiskey]]
- b. it is [_{NP}[_{NP}going to the store] and [_{NP}buying some whiskey]] that_i John is looking forward to e_i
- c.*it is [_{NP}going to the store] that_i John is looking forward to [_{NP}[NP_ie] and [_{NP}buying some whiskey]]
- d.*it is [_{NP}buying some whiskey] that_i John is looking forward to [_{NP}[_{NP}going to the store] and [NP_ie]]

We will have more to say about these cases in the Appendix to this chapter.

We have called attention to the apparent incorrectness of Ross's conflation of the A/A principle with (99) because we will be presenting a PCC account of (99) which does not subsume the A/A cases. In the immediately following sections we will show how our proposal works, and use it to present some further evidence for the parallelism between the INFL-COMP relationship and other instance of \bar{A} -binding. We then will try to place our analysis in the wider context of a general theory of multiple gap constructions, the PCC and subjacency. In this last section we will also be able to argue against some counteranalyses.

Many languages contain coordinating conjunctions of various types, and it is conceivable that the notion "coordinating conjunction" is a primitive, like, perhaps, the notion "reciprocal" or "reflexive". Optimally, we expect that the child who is acquiring English has to learn at most the following sorts of propositions:

(102) and is a conjunction meaning \wedge ; or is a conjunction meaning \vee ...

Ignoring thorny questions about the truth-functional properties of natural language coordination, let us try to get a grip on what "X is a conjunction" means to the child. A first approximation may lie in the phrase structure rule given by Sjoblom (1980):

(103) $X \rightarrow (C X)^*$ (where C is some conjunction, and X some constituent)

(103) legitimates constituents like (104):

(104) $*[_{NP} \text{and John and Mary and Sue . . . and Bill}]$

It is a language-specific fact about English that, for conjunctions like and, the first occurrence provided for by (103) must delete:

(105) $[_{NP} \text{John and Mary and Sue . . . and Bill}]$

(For other conjunctions like or, there is suppletion of the first occurrence; we push aside many relevant semantic questions here.) Also, the set of conjunctions that do not immediately precede the final conjunct are preferably deleted en bloc:

(106) $[_{NP} \text{John <and> Mary <and> Sue . . . and Bill}]$

The facts illustrated in (105)-(106) are the sort that a child may very well learn by experience, like other low-level properties of natural language. In Russian and earlier stages of English, for example, (104) is grammatical without deletion of the first conjunction.

Let us now examine (103) somewhat more closely. The fact that certain instances of C may or must be deleted, given the principle of

Recoverability of Deletion, suggests that the separate occurrences of C motivated by (103) are linked, and perhaps constitute a single discontinuous element. This recalls in part the analysis of Lakoff and Peters (1966), who suggested that certain cases of coordination involve generating a single, initial conjunction followed by a string of conjuncts: [C X*]. Later rules copy the conjunction onto the conjuncts, subject to the appropriate deletions.

Thus, we may wish to indicate that there is only one distinct conjunction present in structures like (105). A rule like (103) does not capture this: the various C's are as distinct as the various X's. The rule could, of course, be revised as:

$$(107) \quad X \rightarrow (C^i X)^*$$

-- where superscript i on each occurrence of C will indicate that the occurrences are non-distinct: C is discontinuous.

Let us consider (103) or (107) further. Note that the asterisk in (103) must be interpreted as "obligatory". That is, the string C-X must be repeated at least once, since there are no constituents of the form [_{NP} and NP]. This is a rather odd property for a principle of phrase structure, like (103) or (107), to incorporate.

The general line suggested by these observations is clear enough. The phrase structure rules suggested for coordination do not explain anything about the properties of coordination per se. They furthermore conflate two distinct aspects of coordination: First, the "obligatoriness" of the asterisk in (107) and the very presence of the asterisk recapitulates a basic property of coordination: conjunctions are connectives that must link an element to another element. In standard predicate calculus, such

connectives are taken to link a maximum of two expressions (i.e. $x \vee y, z$ is ill-formed), but we need put no such restriction on natural-language conjunctions. Second, the rule (107) incorporates language-specific information about the position of the conjunction which is in any case modified by later deletion rules.

Let us see how we might go about reducing (107) to simpler principles. We may follow some of the logic outlined in Chomsky (1981b), discussed at the beginning of Chapter Three, for eliminating the content of the base component. First let us note that the fact that X occurs on the right side of (107) should follow from \bar{X} -theory. The only concession we have to make to allow the simplification of (107) is to permit multiply-headed constituents -- a reasonable proposal for coordination. We can simplify (107) to (108):

(108) $X \rightarrow (C^i \text{ Head})^*$

We have just argued that there is essentially one occurrence of C in coordinate structures. The fact that there is more than one conjunct follows from the interpretation of logical forms with such conjunctions. In making certain requirements on constituents it governs, a conjunction is acting very much like a predicate that θ -marks an argument. What is special about a conjunction is that it " θ -marks" more than one argument at a time. This is the syntactic content of the term conjunction. If we take this analogy literally, we can divide (108) into two principles:

(109)a. $X \rightarrow C^i \text{ Head}$

b. Conjunctions θ -mark n constituents simultaneously, $n > 1$

And we add to (109) principles governing the distribution of C. The combination of (109)b and (109)a will guarantee that there is more than one Head to X. It may also be the case, as is implicit in (107) and in Ross's discussion, that the constituent formed by the coordination of occurrences of X^n , where n is some number of bars, is itself X^n -- that is, that structures are as in (97). We will assume this, but nothing important seems to follow from this stipulation.

We do not pretend to have elaborated a complete or insightful theory of coordinate structures here. Our main point was to suggest that (109)b should be factored out from the aspects of coordinate structure that fall under X-bar theory. The fact that conjunctions require more than one "complement" is a fact which should be kept separate from the linear order of conjunctions with respect to the conjuncts and from the \bar{X} properties of conjunction structures.

Suppose we take literally the idea that conjunctions do θ -mark their conjuncts, as suggested in (109)b. We suggested in 4.6 of Chapter Three that an idea of Stowell (1981) be adapted into the PCC framework. Stowell suggested that θ -marking of an argument be regarded as a form of \bar{A} -binding by the " θ -grid" of the θ -marker. We showed that adopting this idea allowed us to explain the Subject Condition effects found in small clauses and S'-deletion infinitivals. Treating θ -marking as a form of \bar{A} -binding, it followed that a path runs between the maximal projection of the lexical category that receives the θ -role and the first maximal projection dominating the θ -marker. Thus, in (110) the nodes {NP, VP} form a path:

(110) John [_{VP} [_V ate] [_{NP} the chocolate ice cream]]

If conjunctions assign a θ -role simultaneously to each conjunct, we may regard the conjunction as containing a θ -grid from which the conjuncts are bound. If the conjuncts are all bound from the same position in the θ -grid, the structure is analogous to a series of parasitic gaps all bound by a common \bar{A} -binder. Following our analysis of multiple-gap constructions, we will assume that a single path runs from the maximal projection dominating the lexical projections of the conjuncts to the maximal projection dominating the conjunction:

- (111) [_{NP₀} [_{NP₁} a picture of Mary] and [_{NP₂} a story about Bill] and
 [_{NP₃} a book]]

Path: Between conjunction and conjuncts:

{NP₀, NP₁, NP₂, NP₃}

- (112) [_{VP₀} [_{VP₁} plays [_{NP₁} the lute]] and [_{VP₂} sings [_{NP₂} madrigals]]]

Path: Between conjunction and conjuncts:

{VP₀, VP₁, VP₂}

If paths such as these exist, then the second part of the Coordinate Structure Constraint, presented in (99), follows immediately. Consider:

- (113) *John, [_S who_i I bought [_{NP₀} [_{NP₁} a picture of Mary] and [_{NP₂} a story about e_i] and [_{NP₃} a book]]]

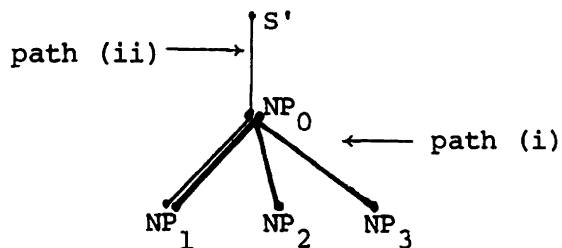
Paths

- (i) Between conjunction and conjuncts:

{NP₁, NP₂, NP₃, NP₀}

- (ii) Between e_i and COMP of S':

{NP₁, NP₀, S'}



We have omitted irrelevant nodes. Path (i) and (ii) overlap, but neither one contains the other. Consider also:

- (114) *the lute [_S which_i Henry [_{VP₀} [_{VP₁} plays e_i] and [_{VP₂} sings madrigals]]]

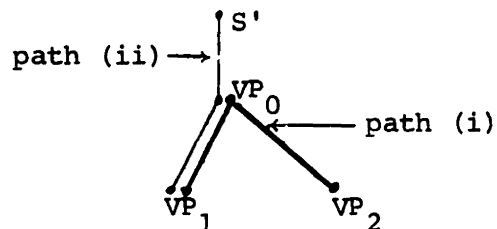
Paths

- (i) Between conjunction and conjuncts:

{VP₁, VP₂, VP₀}

- (ii) Between e_i and COMP of S':

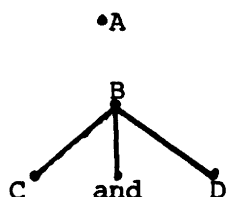
{VP₁, VP₀, S'}



Once again, paths (i) and (ii) overlap, but neither one contains the other.

Quite generally, if there is a path that includes all conjuncts and a node dominating the conjuncts, extraction from any single conjunct will violate the PCC. Consider the abstract situation:

- (115)

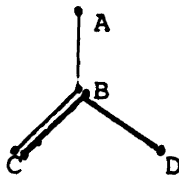


By virtue of the conjunction and, there is a path between the conjuncts: {C, D, B}. Conjunction thus yields a "forked path" much like parasitic gap constructions. Now suppose we extract from within C, such that a path runs from C or lower to A: {C, B, A}. This second path is doomed to violate the PCC, because it includes two members of the path formed by conjunction, but does not include the nodes that dominate the other conjuncts. We thus derive (99): that part of Ross's constraint that does not derive from the A/A condition.¹⁸

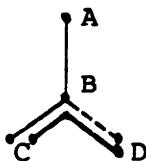
Note, however, that we do not derive what does follow from A/A: the impossibility of extracting a whole conjunct (*who did you see John and t). Suppose we extracted C in (115). Then the empty category left in place of C would create a path that began with B -- the first maximal projection dominating C. The resulting path, {B, A, . . .}, would overlap the path created by conjunction, and would thus not violate the PCC. We therefore retain the A/A condition to prevent extraction of an entire conjunct.

3.2 Across the Board

3.2.1 Consider (115) once more. Consider the path made by conjunction, {C, D, B}, and the path made by extraction from within one conjunct, e.g. {C, B, A}. Let us ask a familiar question: what operation could we perform on the path {C, B, A} that would allow it to satisfy the PCC?



The obvious answer is to add the nodes dominating the other conjuncts -- for example, node D:



How could we add these new nodes to the path {C, B, A} formed by extraction from inside C? The answer lies in our paths theory of multiple gap constructions, adapted from Kayne. The way to add the nodes of each of the conjuncts is to add "parasitic gaps" in each of these conjuncts, such that the gaps in all of the conjuncts share a common \bar{A} -binder. In other words, our "forked path" theory of parasitic gaps predicts that clause (99) of the Coordinate Structure Constraint can be violated if and only if each conjunct contains a gap locally \bar{A} -bound by the same binder. These are well-known "Across-the-Board" (ATB) violations of the Coordinate Structure Constraint:

(116) John, [_S who_i I bought [_{NP₀} [_{NP₁} a picture of e_i¹], [_{NP₂} a story about e_i²] and [_{NP₃} a book by e_i³]]

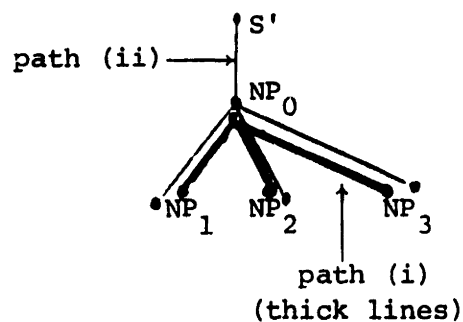
Paths

(i) Between conjunction and conjuncts:

{NP₁, NP₂, NP₃, NP₀}

(ii) Between e_i¹, e_i², e_i³ and COMP of S':

{NP₁, NP₂, NP₃, NP₀, S'}
(cf. (113))



(117) the lute [_S, which_i Henry [_{VP₀} [_{VP₁} bought e_i¹] and [_{VP₂} sold e_i²
to Segovia]]

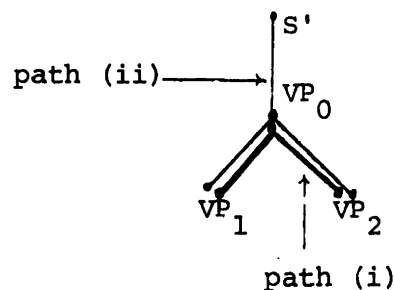
Paths

(i) Between conjunction and conjuncts:

{VP₁, VP₂, VP₀}

(ii) Between e_i¹, e_i² and COMP of S':

{VP₁, VP₂, VP₀, S'}
(cf. (114))



The ATB violations of the Coordinate Structure Constraint exhibited in (116)-(117) are thus saved from the constraint in exactly the same way as the apparent violations of the Subject Condition with which we began the chapter were saved. The addition of extra gaps to a structure can serve to eliminate PCC violations. (N.B. we must assume that NP₀ does not count as bounding for subjacency, as noted by N. Chomsky; cf. 4.5)

3.2.2 We can verify some slightly more complicated predictions of this explanation for ATB violations of the Coordinate Structure Constraint.

For example, when two paths lead out of a conjunct, these same paths must lead out from every conjunct:

(118)a. *a book that_i I know who_j to [[talk to Bill about e_i] and [persuade e_j to buy e_i]]

b. *a book that_i I know who_j to [[talk to e_j about Mary] and [persuade e_j to buy e_i]]

c. a book that_i I know who_j to [[talk to e_j about e_i] and [persuade e_j to buy e_i]]

This is because the paths from \underline{e}_j and the paths from \underline{e}_i each must contain the path linking the conjoined VPs and the VP dominating them. Of course, they must be in the proper containment relation among themselves, in each conjunct:

(119)a. *a book that_i I know who_j to [[talk about \underline{e}_i to \underline{e}_j] and [persuade \underline{e}_j to buy \underline{e}_i]

b. *a man who_j I know what book_i to [[talk about \underline{e}_i to \underline{e}_j] and [persuade \underline{e}_j to buy \underline{e}_i]

WH-movement within one of the conjuncts is not blocked. The path between the WH and its trace does not overlap any path created by conjunction:

(120) a book that I [[knew who_j to talk to \underline{e}_j about \underline{e}_i] and [wanted to buy \underline{e}_i]]

The Coordinate Structure Constraint affects all the constructions Chomsky (1977) characterizes as WH-movement, but not constructions that involve A-binding:

(121)a. *this lute is easy to [[play on \underline{e}] and [improvise on the higher strings]]

a'. this lute is easy to [[play on \underline{e}] and [improvise on \underline{e}]]

b. *this chicken is too cold to [[put \underline{e} on the table] and [eat some fish]]

b'. this chicken is too cold to [[put \underline{e} on the table] and [eat \underline{e}]]

c. *John made more matzohs than I could [[put these napkins in my pocket] or [eat \underline{e}]]

c'. John made more matzohs than I could [[put \underline{e} in my pocket] or [eat \underline{e}]]

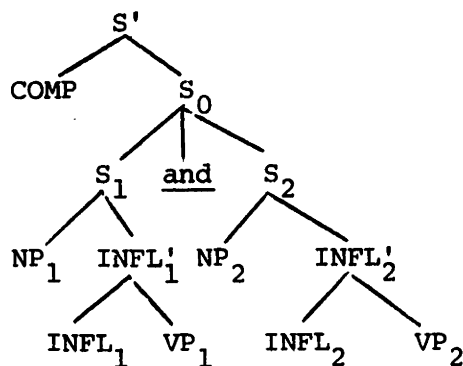
- (122)a. John_i [either [tried PRO_i to phone us] or [is coming here now]]
 b. John_i is [both [glad PRO_i to leave] and [sad that you are staying]]
 c. John_i [either [was registered e_i by the police] or [invented some story]]
 (cf. Williams 1977,(17))
 d. John_i is [neither [likely e_i to win the race] nor [in good shape]]¹⁹

Recall that A-bound empty categories do not create paths, and do not show Crossing effects.

3.2.3 We have looked at structures with conjoined VPs and NPs. Let us now examine the properties of conjoined tensed Ss. We will show that certain apparent "subject/object asymmetries" or "parallelism requirements" discovered by Williams (1978) follow straightforwardly from the PCC. Williams attributed these effects to a condition on analyzability: we discuss Williams' own theory in section 4.4 of this chapter, where we address ourselves to the general question of the etiology of multiple gaps in parasitic and ATB gap constructions.

Consider an abstract structure resulting from conjoining two tensed Ss:

(123)



We know that (123) contains at least the following path:

(123) (i) Between the conjunction and the conjuncts:

$$\{S_1, S_2, S_0\}$$

We also know that there is a path in tensed sentences between INFL and COMP. What sort of path is there between INFL and COMP when two Ss share a COMP, as in (123)? Suppose each INFL inaugurated its own path to COMP. Then we would have the paths $\{INFL'_1, S_1, S_0, S'\}$ and $\{INFL'_2, S_2, S_0, S'\}$. These paths would each overlap the other, and each would overlap path (i), and nothing would contain anything. The result would be an extreme violation of the PCC. We conclude that $INFL_1$ and $INFL_2$ cannot generate separate paths to COMP in (123).

The alternative is to assume that $INFL_1$ and $INFL_2$ share a path to COMP. On this alternative, in effect, the relation between INFL and COMP is established "across the board". COMP is assumed to bind both INFLs simultaneously. We return shortly to a deeper exploration of this move. For the moment let us assume it and see its consequences. We thus assume that (123) contains the following second path:

(123) (ii) Between $INFL_1$, $INFL_2$, and COMP:

$$\{INFL_1, S_1, INFL_2, S_2, S_0, S'\}$$

Paths (ii) and (i) overlap, and (ii) contains (i). The PCC is satisfied.

Now suppose WH-movement applies into the COMP of (123). Path (i) all by itself is enough to require that the movement be ATB -- that there be a gap in both conjuncts bound by WH. Familiar Coordinate Structure Constraint effects obtain:

- (124)a. *John, [_S'who_i [_S₀ [_S₁ Bill [_{INFL}'₁ [_{VP}₁ saw e_i]]] and [_S₂ Mary
 [_{INFL}'₂ [_{VP}₂ likes Tom]]]]]
- b. John, [_S'who_i [_S₀ [_S₁ Bill [_{INFL}'₁ [_{VP}₁ saw e_i¹]]] and [_S₂ Mary
 [_{INFL}'₂ [_{VP}₂ likes e_i²]]]]]

In (124)a, we have the following third path:

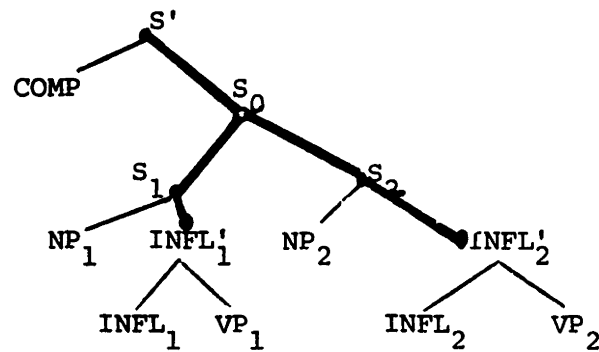
- (124)a. (iii) Between e_i and COMP of S':
 {VP₁, INFL'₁, S₁, S₀, S'}

This path overlaps path (i), but neither contains the other. The same is true of path (ii). In (124)b, on the other hand, we have the following third path:

- (124)b. (iii) Between e_i¹, e_i² and COMP of S':
 {VP₁, INFL'₁, S₁, VP₂, INFL'₂, S₂, S₀, S'}

This path (iii) also overlaps path (i), but this time it contains path (i). Again, the same is true of path (ii): path (iii) overlaps and contains it.

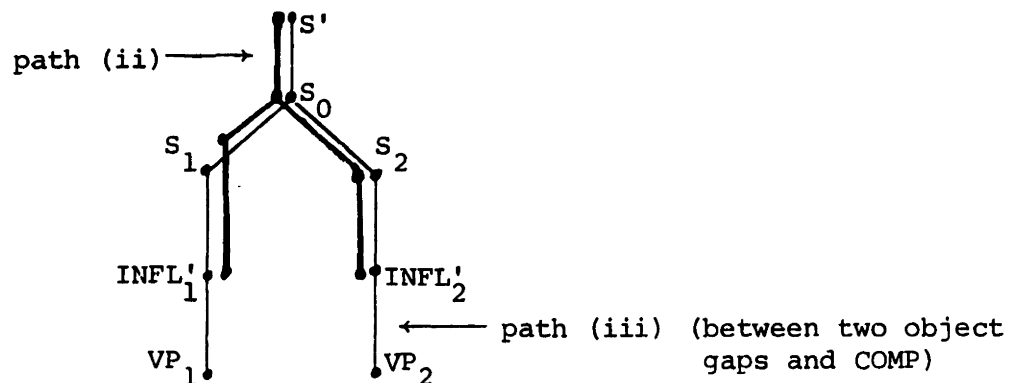
So far the path from the two INFLs to COMP has not played a crucial role. Nonetheless, it has a role to play, in deriving the "subject/object" asymmetries of Williams (1978). Let us begin by looking at it in a tree:

(125) Path between two INFLs and COMP (path (ii))

The upside down \bar{Y} in (125) is path (ii), between the two INFLs and COMP. We have already seen that placing an \bar{A} -bound gap in one of the conjoined Ss in (125) will violate the PCC, with respect to path (i) or path (ii) above. On the other hand, if we have a gap in each conjunct, as in (124)b, the result satisfies the PCC.

Notice that in (124)b the two gaps are both below the lowest points of the upside down \bar{Y} in (125). In other words, placing gaps in the object positions of VP_1 and VP_2 created a path (iii) which took the form of a second upside down \bar{Y} , each prong of which was longer than the prongs of the \bar{Y} in (125). Schematically:

(126)



In other words, our extraction from below the two INFLs yielded a big \bar{Y} that contained the INFL-COMP \bar{Y} .

Now suppose we wish to have gaps in the subject positions of S_1 and S_2 . If we extract only one subject, Coordinate Structure Constraint effects obtain, as expected. These effects are due to path (i), created by the coordination, and not to path (ii) between the INFLs and COMP:

(127) *John, [S , who_i [S_0 [S_1 e_i [$INFL_1$ [VP_1 saw Bill]]]] and [S_2 Mary
 [$INFL_2$ [VP_2 likes Tom]]]

Paths

(i) Between the conjunction and the conjuncts:

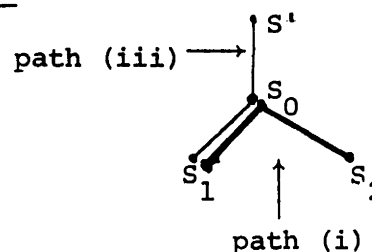
{ S_1 , S_2 , S_0 }

(ii) Between $INFL_1$, $INFL_2$, and COMP:

{ $INFL_1$, S_1 , $INFL_2$, S_2 , S_0 , S' }

(iii) Between e_i and COMP:

{ S_1 , S_0 , S' }



Paths (i) and (iii) overlap, and neither one contains the other. Thus the PCC is violated. Path (ii) overlaps both path (i) and path (iii), but contains each of them; thus path (ii) plays no role in the star on (127).

Suppose now we add a second gap bound by who_i in the subject position of S_2 . The result is, predictably, grammatical. (128) is simply another ATB violation of the Coordinate Structure Constraint:

(128) John, [_S' who_i [_S₀ [_S₁ e_i¹ [_{INFL}'₁ [_{VP}₁ saw Bill]]]] and [_S₂ e_i²
 [_{INFL}' [_{VP} likes Tom]]]]

Paths

(i) Between the conjunction and the conjuncts:

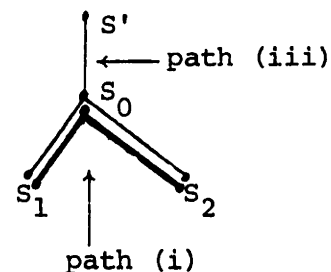
{s₁' s₂' s₀'}

(ii) Between INFL₁' , INFL₂' , and COMP:

{INFL₁' , s₁' , INFL₂' , s₂' , s₀' , s' }

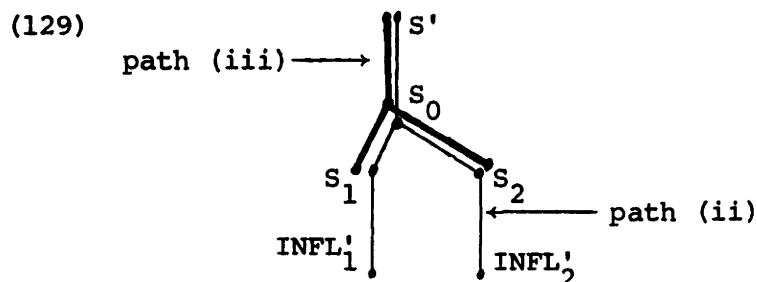
(iii) Between e_i¹ , e_i² and COMP:

{s₁' s₂' s₀' , s' }



Paths (i) and (iii) overlap, and path (iii) now contains path (i). Paths (i) and (ii) overlap, and path (ii) contains path (i). Paths (ii) and (iii) overlap, and path (ii) contains path (iii). The PCC is satisfied.

Let us examine the relationship between paths (i) and (iii) in (128) schematically:

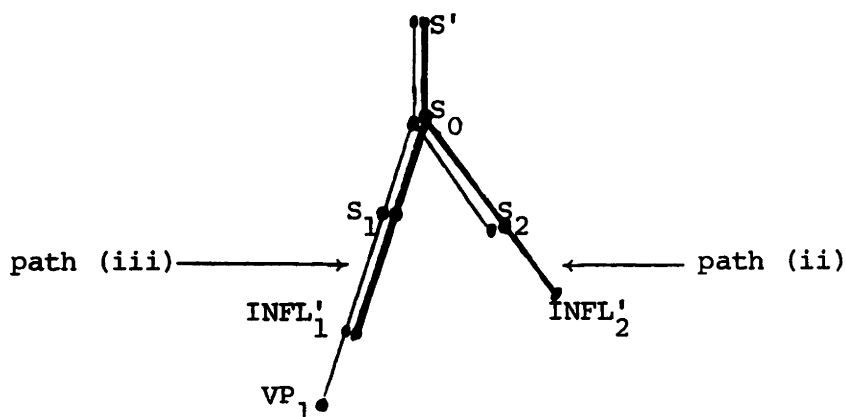


As in (126), the path between the INFLs and COMP is an upside down Y. Recall that in (126) we have gaps in two positions that were below the lower points of the INFL-COMP Y, yielding a second Y that was bigger than the INFL-COMP Y. In (129), we have gaps in two positions that are above

the lower points of the INFL-COMP Y. The result is a second Y that is smaller than the INFL-COMP Y. Because of the PCC, a second Y is grammatical whether uniformly smaller or larger than the first. If smaller, the first Y contains it. If larger, it contains the first Y.

Suppose, however, we have two gaps, such that one of them creates a path that begins above the lower point of one fork of path (ii), and the other creates a path that begins below the lower point of the other fork of path (ii). The situation we have in mind would look like (130):

(130)



On the left fork, path (iii) contains path (ii); but on the right fork, path (ii) contains path (iii). All in all, such a structure would violate the PCC, since path (ii) contains the $INFL'_2$ node, not contained in path (iii), while path (iii) contains the VP_1 node, not contained in path (ii).

What sort of a sentence could yield the situation in (130)? (130) would arise if ATB WH-movement yielded a gap in the highest subject position of one conjunct, and a gap lower than VP of the other conjunct. As Williams (1978) noticed, this situation is indeed ungrammatical:

(131) *John, [_S, who_i [_S₀ [_S₁ Bill [_{INFL}'₁ [_{VP}₁ saw e_i¹]]]] and [_S₂ e_i²
 [_{INFL}'₂ [_{VP}₂ liked Mary]]]] (=Williams 1978, (16))

Compare (131) with (124)b, where both gaps are in VP, and with (128), where both gaps are the subject of the highest S. The paths in (131) are as follows:

(131) (i) Between the conjunction and the conjuncts:

$$\{s_1, \quad s_2, s_0\}$$

(ii) Between INFL₁, INFL₂ and COMP:

$$\{\text{INFL}'_1, s_1, \text{INFL}'_2, s_2, s_0, s'\}$$

(iii) Between e_i¹, e_i² and COMP:

$$\{\text{VP}_1, \text{INFL}'_1, s_1, \quad s_2, s_0, s'\}$$

Paths (ii) and (iii) overlap, but neither contains the other; hence the PCC violation. The situation is exactly that envisioned in diagram (130). (130) contains one upside down Y formed by the two INFLs. The two \bar{A} -gaps create a second Y.

Compare the two Ys. In our previous examples, the Y formed by the gaps has been uniformly longer than the INFL-Y (when both gaps were objects) or uniformly shorter (when both gaps were subjects). In (130)-(131), however, the Y formed by the gaps is longer on one prong and shorter on the other prong. This is the situation that the PCC rules out. This situation arises whenever one gap is a matrix subject, and the other gap is lower.

We thus derive the "subject/object" asymmetry noted by Williams. Notice that this result has followed directly from (1) the PCC; (2) the hypothesis that multiple gaps sharing a common \bar{A} -binder form a single path; and (3) the hypothesis that a path runs between a conjunction and its

conjuncts. (1) has been motivated throughout these last two chapters. (2) is motivated by the parasitic gap exceptions to the Subject Condition, as discussed by Kayne (1982). (3) is the minimum necessary to derive the Coordinate Structure Constraint and its ATB exceptions from (1) and (2). Williams' contrasts thus derive entirely from basic, independent assumptions of the theory.

Let us note some further consequences of our analysis. Williams notes that one might try to rule out (131) by invoking some sort of "parallelism" requirement on multiple gaps in conjunct structures, blocking, for example, ATB constructions where one gap is a subject and the other an object. He shows, however, that parallelism of grammatical relations is not at stake here. Let us consider his example.

On our theory, what is important is that the ATB gaps be either all above the highest INFL, or all below the highest INFL. When all the gaps are matrix subjects, they are all above the highest INFL. When they are objects, they are all below. When one is a matrix subject, and the other an object, one is above and one is below, and the PCC is violated.

Suppose now that one gap is a matrix object, and the other a subject of a subordinate clause. Both gaps will be below the matrix INFLs, and the PCC will be satisfied, although parallelism of grammatical relations (or of Case) is not met. As Williams notes, such structures are grammatical:

significance for the PCC.

The parallelism hypothesis is also disproved by structures in which one gap is a matrix subject, and the other an embedded subject. To our ears, such structures are as unacceptable as (131):

- (134) *John, [_S who_i [_S₀ [_S₁ Bill [_{INFL}₁' [_{VP}₁ thinks [_S₃' e_i^{1a} [_S₃' e_i^{1b} will win]]]]]]
 and [_S₂' e_i² [_{INFL}₂' [_{VP}₂ likes Mary]]]]]

Paths

- (i) Between the conjunction and the conjuncts:

{_S₁', _S₀', _S'}

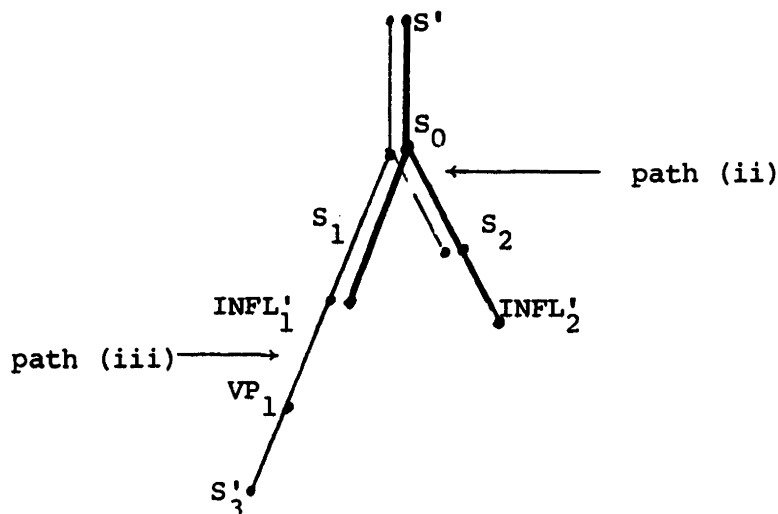
- (ii) Between INFL₁' , INFL₂' , and COMP:

{INFL₁' , _S₁' , INFL₂' , _S₂' , _S₀' , _S'}

- (iii) Between e_i^{1a} , e_i² , and COMP of matrix _S' :

{_S₃' , _{VP}₁' , INFL₁' , _S₁' , _S₂' , _S₀' , _S'}

(135)



Paths (ii) and (iii) overlap, and neither one contains the other, violating the PCC. (135) shows the situation schematically: path (iii) is longer in

the left conjunct, and shorter in the right conjunct, than path (ii). The situation is exactly the same, from the point of view of the PCC, as that in (130), where one gap was a matrix object and the other a matrix subject. Thus, any "parallelism" requirement on ATB structures would be too strong to admit (130), where one gap is an object and the other a subject, and too weak to rule out (131), where both gaps are subjects, if parallelism of grammatical relations were required.

When each of two conjoined Ss contains a gap below INFL', there is no problem with additional gaps being added above INFL' -- as in the structures we considered in the early sections of this chapter. Thus:

(136) John, [_S who_i [_S₀ [_S₁ Bill [_{INFL}'₁ [_{VP}₁ likes e_i¹]]]] and [_S₂ [_{NP}₂ friends of e_i²] [_{INFL}'₂ [_{VP}₂ admire e_i³]]]]]

Paths

(i) Between the conjunction and the conjuncts:

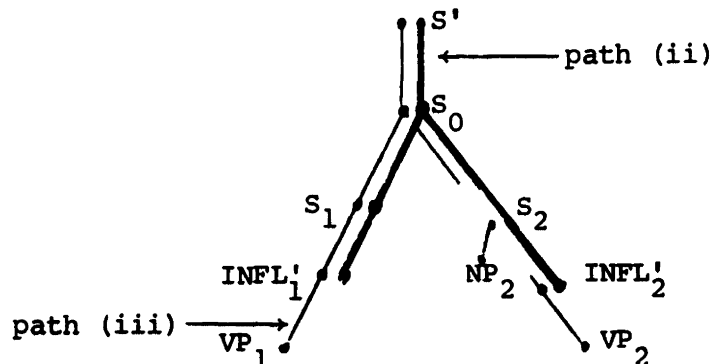
{S₁, S₂, S₀}

(ii) Between INFL'₁, INFL'₂ and COMP:

{INFL'₁, S₁, INFL'₂, S₂, S₀, S'}

(iii) Between e_i¹, e_i², e_i³ and COMP:

{VP₁, INFL'₁, S₁, VP₂, INFL'₂, NP₂, S₂, S₀, S'}



We return to structures like (136) in section 4.4, where we use them as an argument for our approach to ATB and parasitic gap constructions over the theory of Williams (1978). In the next section, we concentrate our attention on the path between INFL and COMP in coordinate and non-coordinate structures, and try to explore its properties and origins in greater depth than we have so far.

3.3 The Path between INFL and COMP

In discussing (123), we showed that the PCC predicts that the relationship between INFL and COMP is subject to Coordinate Structure Condition effects. We reasoned as follows: Consider conjoining several Ss, so that they share a common COMP. Suppose that there was a separate path from the INFL of each S to the common COMP. We demonstrated that these distinct paths would (1) overlap the path between the conjunction and the conjoined Ss, and (2) overlap each other. In each case, the PCC would be violated, since no path would contain any other path. For this reason, we concluded that there must be a single, "forked" path between the COMP and the several INFLs. We then showed that if this forked path does exist, we can explain various apparent subject/object asymmetries in ATB constructions, discovered by Williams. We thus justified our original contention, that the PCC forces a single path to run between the common COMP and all the INFLs.

In this section, we will explore the consequences of this forked path between COMP and INFLs in greater detail. More generally, we will suggest a reason for the connection between INFL and COMP, which makes some surprising and apparently correct predictions with respect to the PCC. Along the way, we will show that our theory derives the effect of our

Chapter Two stipulation that expletive empty categories are not subject to the ECP. A problem will arise with respect to the paths produced by empty categories. We will suggest a solution to this problem which appears to allow us to eliminate the stipulation that only non-A-bound empty categories create paths. This discussion will function as a lemma to our main point: the path between INFL and COMP is real; there is a reason for its existence; it interacts with more than just paths created by WH-movement from subject position; we can show empirically that it displays Coordinate Structure Constraint effects of its own.

Throughout our discussion of the PCC, we have alleged an important parallel between the INFL-COMP relationship in tensed sentences and the relationship between a trace and a local \bar{A} -binder. Since COMP is an \bar{A} -position, it is natural to ask whether we can make this parallel more explicit. In particular, does INFL contain a trace of something that binds it from COMP? If INFL did contain such a trace, then our postulated forked path between COMP and the many INFLs in conjoined Ss would have exactly the same status as the path in other multiple-gap constructions.

We have been assuming that INFL contains two bundles of features: one containing features for person, number and gender, called AGR; and another bundle containing features for tense, called TNS. Of these two, TNS is the one that is an operator, when specified positively for the tense feature, and might reasonable be required to appear in an operator position at LF. Suppose we were to assume the following claims to be true:

- (137)a. TNS appears in INFL at D-structure
- b. [_{TNS}+tense] governs a complete proposition at LF

(138) Complete Proposition (definition)

A category X is a complete proposition iff for all predicates P dominated by X, X contains all members of the chains θ -marked by P.

Notice that government, as a special case of c-command, implies scope.

(137)b says that the constituent in the immediate scope of [+tense] must be a complete proposition. From (137), it follows that a sentence like Mary likes John has the following D-structure and LF representations:

(139)a. D-structure:

$$[_S, [_{COMP}] [_S \text{Mary} [_{INFL}, [_{INFL} [_{TNS} +\text{tense}] \text{AGR}] [_{VP} \text{likes John}]]]]$$
b. LF:

$$[_S, [_{COMP} [_{TNS} +\text{tense}]_i] [_S \text{Mary} [_{INFL}, [_{INFL} [_{TNS} e]_i \text{AGR}] [_{VP} \text{likes John}]]]]$$

Thus, if (137) is correct, INFL in a tensed S does contain a trace bound from COMP: the trace of TNS. Ideas similar to (137) have been suggested by Stowell (1981), and by Den Besten (1978) (who argues for an opposite derivation in which TNS in COMP lowers into S at S-structure).

The second of our two claims, made in (137)b, seems entirely natural. It is common to treat tense as an operator with sentential scope. In our formulation, the tense operator binds a variable internal to its proposition from a position governing the proposition. The claim that tense binds a variable, while not uncontroversial, does not seem to pose any problems of interpretation.

The first of our claims, however, made in (137)a, remains a stipulation. We might take INFL' to bear a grammatical relation defined as [INFL', S], and to be an A-position set aside for grammatical formatives that are

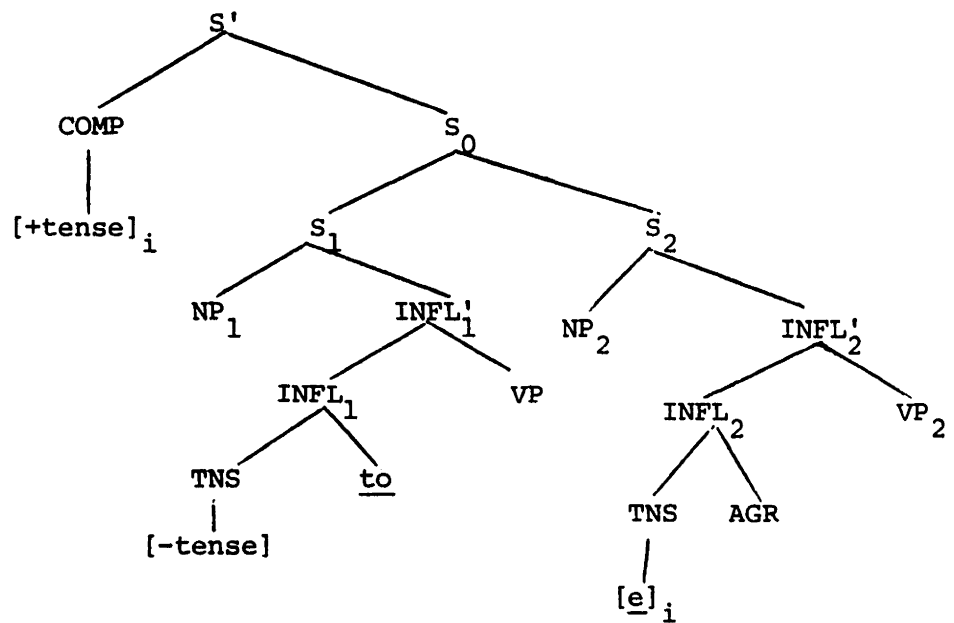
relevant to sentence-level interpretation (cf. Steele et al. 1981). If this were the case, then the D-structure position of TNS in INFL might follow from the definition of D-structure as a pure representation of thematic structure. Perhaps VP selects INFL, in some fashion. In any case, we will continue to stipulate the D-structure position of TNS in INFL.

Given (137), then, we have a reasonable motivation for the path between INFL and COMP. It is formed by the movement of TNS to COMP, where it has scope over a complete proposition. Other principles will surely be necessary -- to prevent tense, for example, from moving too far, into a higher clause. Nonetheless, the general picture seems clear.

Suppose we assume that only operators may occupy \bar{A} -positions at LF, as argued by Higginbotham (forthcoming), based on reconstruction phenomena. Suppose we accept also the traditional assumption that infinitives are [-tense] (contra Stowell (1981, 1982), who presents interesting semantic arguments for the opposite assumption). It follows that the [-tense] TNS node of an infinitive may not move to COMP, and there will be no path between INFL and COMP in infinitives. We will offer a demonstration that this is correct, based on Coordinate Structure Constraint effects.

We have shown the following to be true: if a path runs between a conjunction and all the conjuncts, the presence of an \bar{A} -bound gap in one conjunct, bound outside that conjunct, requires the presence of an identically \bar{A} -bound gap in every other conjunct. This is Ross's Coordinate Structure Constraint (as in (99), less the A/A Condition), with its ATB exceptions. The same conclusion should follow for a "gap" within INFL, if INFL is contained by a conjoined S. Consider the structure in (140):

(140)



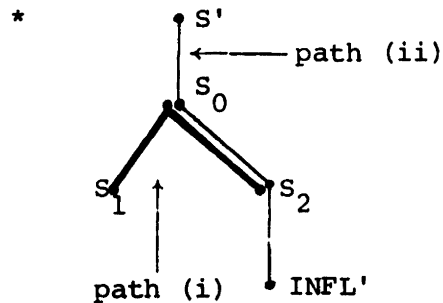
In (140), there is a path between the conjunction and the conjoined Ss. There is also a path between TNS of INFL₂ and COMP, but no path between TNS of INFL₁ and COMP. These paths clearly overlap and violate the PCC:

(140)

(i) Between the conjunction and the conjuncts: $\{S_1, S_2, S_0\}$ (ii) Between TNS of INFL₂ and COMP of S': $\{INFL'_2, S_2, S_0, S'\}$

Schematically:

(141)

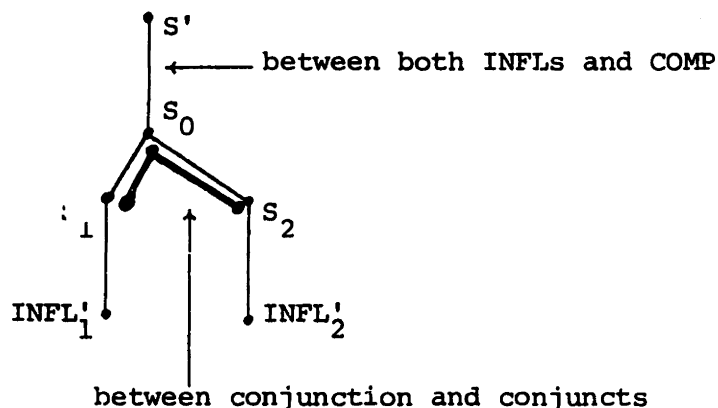


(140) is the abstract representation of a conjunction of a tensed S with a non-tensed S. We saw earlier that there is nothing ungrammatical about the conjunction of two tensed Ss in English, for example:

(142) I wonder [_S,what books_i [_{S₀} [_{S₁} I like e_i] and [_{S₂} you hate e_i]]]

In (142), there is a path from INFL in each conjunct:

(142)



Similarly, the conjunction of infinitival Ss should be grammatical, although indistinguishable from INFL'-conjunction:

(143) I wonder [_S,what books_i [_{S₀} [_{S₁} PRO to like e_i] and [_{S₂} PRO to hate e_i]]]

In (143), there is no path between any INFL and COMP.

When a tensed S is conjoined with an infinitival S, however, the result is ungrammatical -- exactly as predicted by the PCC in (140)-(141):

(144) *I wonder [_S,what books_i [_{S₀} [_{S₁} PRO to like e_i] and [_{S₂} you hate e_i]]]

On our analysis, (144) is a straightforward Coordinate Structure Constraint effect, which derives from the interaction of the path created by TNS movement with the path between the conjunction and the conjuncts.

We have chosen examples with a [+WH] complementizer for two reasons. First, since the [+WH] complementizer does not optionally delete, we can be sure that we are conjoining Ss, and not S's. Second, since the [+WH] complementizer is compatible with both tensed and infinitival Ss, we cannot easily blame the contrast between (142)-(143) and (144) on complementizer conflict. Notice that the actual WH-movement in these sentences is irrelevant to their status under the PCC, and that parallel examples can be constructed in which the WH-phrase is probably in COMP at all levels:

- (145)a. I don't know [_S,whether [_{S₀} [_{S₁} I should be sick] or [_{S₂} I should be well]]]
- b. I don't know [_S,whether [_{S₀} [_{S₁} PRO to be sick] or [_{S₂} PRO to be well]]]
- c. *I don't know [_S,whether [_{S₀} [_{S₁} I should be sick] or [_{S₂} PRO to be well]]]

Certain counteranalyses of this contrast can be disposed of. For example, the difficulty with (144) or (145)c cannot be blamed on a conflict between the "futurate" quality of the infinitive and the tense of the indicative, since there is no evidence that tense or aspect concord is required under conjunction. In the ungrammatical (145)c, the infinitive and tensed S appear to share the relevant tense and aspect properties. In (146), the lack of concord creates no problems:²⁰

- (146)a. I wonder [_S,what books_i [_{S₀} [_{S₁} John bought e_i] and [I should buy e_i]]]
- b. I wonder [_S,what books_i [_{S₀} [_{S₁} John will buy e_i] and [_{S₂} Mary bought e_i]]]
- c. these are the books [_S,for (WH_i) [_{S₀} [_{S₁} Mary to buy e_i] and [_{S₂} you to have bought e_i by next week]]]

Furthermore, the paradigm of (142)-(144) or of (145) cannot be due to a general constraint against conjoining tensed with infinitival clauses, unless this constraint were limited, ad hoc, to conjoined Ss. Conjunction of tensed with infinitival S's, while perhaps stylistically clumsy, is clearly grammatical, unlike such conjunction of Ss:

- (147)a. I hope [_{S₀} [_{S₁} PRO to be there myself] and [_{S₂} that John will be there too]]]
- b. I wonder [_{S₀} [_{S₁} what books_i you are reading e_i] and [_{S₂} what books_i PRO to read e_i myself]]]
- c. I want to learn [_{S₀} [_{S₁} what books_i PRO to read e_i] and [_{S₂} who_j I should ask e_j for advice]]]

There thus appears to be a genuine constraint on conjoining infinitival with tensed S, which is explained without a problem by assuming TNS movement and the PCC theory of conjoined structures.

This argument is important, because it represents the first independent evidence for the existence of a path between INFL and COMP. So far, we have shown that if we assume such a path, the interaction of WH-movement paths with it yields CTP and Subject Condition effects. One might object that the path between INFL and COMP was stipulated ad hoc, to bring these effects together with the Crossing effects. In this section, we have shown that the path between INFL and COMP makes semantic sense, and that it interacts in completely different ways with a separate theory of coordinate structure paths to explain a contrast among conjoined Ss. In effect, we have shown that TNS movement is subject to the Coordinate Structure Constraint, subsumed by the PCC.

Nonetheless, one might still propose that there is a general prohibition against conjoining tensed with infinitival Ss, arising perhaps because of some incompatibility of features, which has nothing to do with the PCC or with TNS movement into COMP. We next present a further argument that the effects observed above arise from the PCC applying to paths created by TNS movement. This argument will involve us in a long detour that is of intrinsic interest. We first present a theory about how expletive subjects should interact with the PCC and the definition of paths, and suggest a parameter distinguishing languages like English from languages like Russian or Italian. Certain problems arise with our theory of expletives, which call for an apparently ad hoc solution. We will try to show that this ad hoc solution actually has other valuable consequences: in particular it may help us to eliminate the stipulation that paths are not formed by A-bound empty categories. We then return to the main thread of discussion, and show how our theory of expletives makes a correct prediction with respect to conjoined Ss with expletive subjects.

In (137), we made two claims concerning the position of TNS. First, we said that TNS is in INFL at D-structure. Second, we claimed that [+tense] TNS must have scope over a complete proposition at LF, where a complete proposition was defined in (138) as a phrase containing all members of all the chains θ -marked by the predicates it dominates.

When the subject of a predicate is indirectly θ -marked by that predicate, in composition with its complements, the minimal phrase dominating the predicate and its θ -chains (if the subject does not undergo NP movement) is the minimal S dominating the predicate and subject. In a tensed sentence, TNS must move to COMP to have scope over this S. The same conclusion follows when the subject is a member of a chain directly

θ -marked by the predicate, as in the passive construction:

(148) [_S, [COMP] [_S John_i [_{INFL}, [_{INFL} TNS AGR] [_{VP} was seen e_i]]]]

TNS in (148) must move to COMP to have scope over the predicate see plus all members of the chain (John_i, e_i) that it θ -marks. Passive sentences do act like tensed sentences with respect to conjunction with an infinitive.

(149)a. I wonder [_S, where [_S₀ [_S₁ Mary_i was tattooed] and [_S₂ Bill_j was shaved e_j]]]

b. I wonder [_S, where [_S₀ [_S₁ PRO_i to be tattooed e_i] and [_S₂ PRO_j to be shaved e_j]]]

c. *I wonder [_S, where [_S₀ PRO_i to be tattooed e_i] and [_S₀ Bill was shaved e_j]]]

Movement of TNS to COMP is also motivated by (137) in raising constructions, which again show the same paradigm under conjunction:

(150)a. I know [_S, what book_i [_S₀ [_S₁ John_j seems [_Se_j to have read e_i]] and [_S₂ Mary_k seems [_Se_k not to have understood e_k]]]]]

b. I know [_S, what book_i [_S₀ [_S₁ PRO_j to seem [_Se_j to have read e_i]] and [_S₂ PRO_k to seem [_Se_k not to have understood e_i]]]]]

c. *I know [_S, what book_i [_S₀ [_S₁ John_j seems [_Se_j to have read e_i]] and [_S₂ PRO_k to seem [_Se_k not to have understood e_i]]]]]

A raising construction has the following S-structure:

(151) [_S, [COMP] [_S₁ John_i [_{INFL}, [_{INFL}^{TNS AGR}] [_{VP}₁ seems [_S₂ e_i [_{INFL}₂ to
 [_{VP}₂ have read this book]]]]]]]

In (151), the first complete proposition is S_1 , since it is the first node that contains all the members of the chains θ -marked by the predicates that it dominates. S_2 , for example, dominates the predicate read, and all members of the singleton chain (this book), which is θ -marked by read. It does not dominate all members of the chain ($John_i, e_i$), indirectly θ -marked by read, however, and is thus not a complete proposition. VP_1 similarly contains all members of all chains θ -marked by seems, since the subject of seems is non- θ , and the sentential complement of seems is not part of any chain including the subject. Nonetheless, VP_1 , like S_2 , dominates read and does not contain all members of the chain ($John_i, e_i$). Therefore VP_1 is not a complete proposition. Since TNS in (151) does not have scope over a complete proposition, it must move to COMP, where it does have scope over the complete proposition S_1 . Since TNS must move to COMP in raising constructions, as in the other constructions we have considered, the paradigm in (150) follows from the PCC (subsuming (99) of the Coordinate Structure Constraint).

Cases do arise, however, in which [+tense] TNS should not have to move in order to have scope over a complete proposition. If the subject position of S does not contain a member of any θ -chain, and if the VP dominated by that S is a complete proposition, (137) allows TNS to remain in INFL. Even without moving to COMP, it has scope over a complete proposition. We thus predict that no path between INFL and COMP should have to exist, when the subject position is expletive. Let us consider this prediction in greater detail.

Suppose a path did have to exist between a tensed INFL and COMP, even when the subject was expletive. Suppose the expletive subject position were empty. Then, by our revised definition of paths that we invoked in (64) to allow for infinite paths, the expletive subject should provoke an infinite path, since it is not A-bound:

- (152) *John [_{VP₁} said [_{S'₂} [_{COMP₂} TNS_i that] [_{S₂} [_e] [_{INFL₂} [_{INFL₂} e_i AGR] [_{VP₂} was necessary [_{S₃} that I buy these books]]]]]]]]

Paths

- (i) Between INFL₂(e_i) and COMP₂(TNS_i):

{INFL'₂, S₂, S'₂}

- (ii) From [e]:

{S₂, S'₂, VP₁, . . . }

If path (i) did exist, despite the expletive subject, the PCC would clearly be violated, just as in a sentence showing a CTP effect. Paths (i) and (ii) overlap, but neither one contains the other.

Since (152) is ungrammatical in English, we must assume, contrary to (137), that TNS always moves to COMP in English, regardless of whether VP is a complete proposition or not. We have two choices: we may replace (137)b with a simple statement that TNS appears in COMP at LF, or we may maintain (137)b as the unmarked case, and suggest that TNS movement in English is "grammaticalized", applying even when not obligatory. The question can be resolved, of course, by looking at other languages.

We stated (137)b as we did, referring to the notion "complete proposition", because we believe (137)b to be the null hypothesis. The

principle that [+tense] TNS must have scope over a predicate and all its arguments (a "saturated predicate") seems to make sense semantically, and as a principle of UG. If [+tense] TNS is required to have scope over elements with null semantic content, such as an expletive subject, the naturalness of the requirement diminishes. Thus, we suggested (137)b as the unmarked principle governing the placement of the tense operator. On the other hand, if a language generally lacks null subjects, as English does, it is reasonable to suppose that this general prohibition should be extended to expletive subjects (though not always, cf. German; Safir 1982), via an extension of the universal principle governing the movement of TNS.

Languages with free pronoun drop, like Russian, or with null pronouns and free inversion, like Italian and Spanish, may be expected to maintain (137)b as the sole principle governing TNS movement. In these languages, no principle forces TNS to move to COMP when it has scope over a complete proposition in INFL. In INFL, TNS has scope over VP. VP is a complete proposition when its subject is expletive.

Recall that the ungrammaticality of (152), an English sentence with a null expletive subject, was due to the overlapping of the path between INFL and COMP and the infinite path from an expletive subject. The situation was thus analogous to a CTP effect arising from long WH-movement of a subject. In the framework of Chomsky (1981b) and our Chapter Two, both CTP effects and the impossibility of null expletives derived from the ECP (cf. Pesetsky 1982 for discussion of CTP effects, null expletives and pronouns in an NIC approach). A subject WH-trace unbound in S' and a null expletive subject are both not properly governed, since they are not locally coindexed with anything, nor are they lexically governed (by a governor ≠ AGR).

Remember now that we discovered in Chapter Two that a null expletive subject was grammatical in Russian, while long movement of a subject was not. We were forced to stipulate that expletive elements were not subject to the ECP in Russian, while they were in English. (We also explored the possibility that expletives were not subject to the ECP in English, but that other requirements prevented them from existing; cf. Safir 1982). It remained a mystery why WH-trace and expletives should differ with respect to the ECP. We are now in a position to derive this distinction. "ECP effects", for us, are the result of the interaction of a path created by TNS movement with other paths. Because of the PCC, the paths that interact with the path created by TNS movement are those that originate in the subject position. If there is no TNS movement, there will be no ECP effects. If TNS is not required to move when the subject is expletive, expletive subjects will not show ECP effects. We thus derive the distinction between expletive and non-expletive subjects with respect to the ECP.

More needs to be said, however, about the status of empty subjects in PRO-drop languages. We do not pretend to offer a cogent theory of the "PRO-drop parameter" here, but certain immediate questions arise that demand an answer.

First, suppose we claim, as we have suggested, that in a language like Italian, Spanish or Russian, there is no path between INFL and COMP when the subject is expletive. Then in the equivalent of example (152), path (i) will not exist. If the expletive subject creates an infinite path, as we argued it does in English, to explain the star on (152), then it will not violate the PCC with respect to any other path in the Italian, Spanish or Russian equivalent of (152). We are thus free for the moment to suppose the following path in the Italian equivalent of (152):

(153) Giovanni [_{VP₁} ha detto [_{S₂} che [_{S₂} [e] [_{INFL₂'} [_{INFL₂} TNS AGR] [_{VP₂} è
necessario [_{S₃} comprare questi libri]]]]]]]]

Path From [e]:

{S₂, S₂', VP₁, . . .}

So far, so good. If the expletive generates an infinite path in Italian, as in English, however, we make the wrong prediction with respect to (154). We give below the paths so far predicted to exist:

(154) non [_{VP₁} so [_{S₂} che libri_i [_{S₂} [e] [_{INFL₂'} [_{INFL₂} TNS AGR] [_{VP₂} è
necessario [_{S₃} comprare e_i]]]]]]]]

Paths

(i) Between e_i and COMP:

{S₃', VP₂, INFL₂', S₂, S₂'}

(ii) From [e]:

{S₂, S₂', VP₁, . . .}

Paths (i) and (ii) overlap, and neither one contains the other. The PCC predicts that (154) is ungrammatical, if these paths exist. The sentence, however, is completely grammatical.

We will tentatively suggest a solution which is for the moment ad hoc, but which, if true, has interesting implications and may resolve certain problems with our paths theory. In particular, it may allow us to get rid of the stipulation that only non-A-bound empty categories create paths. After we discuss this solution, we will pick up again the main thread and show how our theory of expletives interacts with the PCC theory of coordination

in a pro-drop language.

We can eliminate the problem in (154) if we claim that empty categories only generate infinite paths when in the domain of TNS in their clause. Thus, if the subject of S is empty and expletive, and if TNS thus does not have to move to COMP, the expletive subject will not generate a path, since it will not be in the domain of TNS. More specifically, we might assume the following principle:

(155) An empty category e is visible for the definition of paths if:²¹

- (i) for P the minimal category dominating e which is a maximal projection and a complete proposition
- (ii) for T the TNS of P
- (iii) e is in the scope of T

In (154), the minimal complete proposition which is a maximal projector and dominates the expletive [e] is S_2 . Since [e] is expletive, TNS does not have to move to COMP, and [e] does not have to be in the scope of TNS. It follows that [e] may be invisible for the definition of paths, and that path (ii) -- which caused all the problems -- does not have to exist. Note that in the English sentence (152) path (ii) must exist, since TNS always moves to COMP in English, and all subjects of tensed sentences thus count for the definition of paths.

(155) is troubling because it commits us to saying two separate things about expletive subjects and the PCC. Were it not for examples like (154), we could let expletive null subjects in both English and Italian generate infinite paths. The only difference between the two languages would be in the status of [+tense] TNS movement to COMP: in English it always applies, but in Italian it does not have to apply when the subject is expletive.

(154) forces us to add that when TNS movement does apply, the expletive null subject also does not generate an infinite path. (Note that TNS movement will fail to apply only with expletive subjects, so the word "expletive" does not figure in (155).) It is interesting, therefore, to see if (155) has other consequences.

First, (155) forces us to make a distinction between a clause that "lacks the TNS node" and a clause that has a TNS node which is negatively specified for the tense feature. Suppose we were to assume, as we have not, that infinitives lack a TNS node under INFL. Then it would follow from (155) that the object of an infinitival verb is never obliged to create a path, when null, since when P is an infinitival S', a null object in P is not in the scope of any TNS of P. This would be a false conclusion, since an \bar{A} -bound empty category in the object position of an infinitive must be forced to create a path in order to yield familiar Crossing contrasts:

- (156)a. what subject_i do you know [_S,who_j to talk to e_j about e_i]
 b. *who_j do you know [_S,what subject_i to talk to e_j about e_i]

We therefore conclude that the INFL of an infinitive does contain TNS, which is, however, negatively specified for the feature [tense]:

- (157) I wonder [_S,whether [_S^{PRO} [_{INFL}' [_{INFL} [TNS -tense] to] [_{VP}eat cheese]]]]

Since cheese is in the scope of the TNS of S', a non-A-bound empty category in its place will always generate a path.

Recall that (155) was motivated to get us out of postulating an infinite path for an expletive subject when TNS does not move to COMP. Now notice that (155) has produced a further result. (155) singles out two

sorts of empty categories as exempt from obligatory path formation. First, expletive subjects may remain outside the scope of TNS, as we have seen, and may fail to generate a path when null. Second, while objects in infinitival S's are in the scope of [_{TNS}-tense], as we have just seen, subjects of infinitival sentences remain outside the scope of TNS, and thus may also fail to generate a path.

Now recall that PRO, by the binding theory, is effectively limited to the subject position of an infinitival S' (or gerund), where it will lack a governing category. From (155) it follows that PRO will not have to generate a path. So far, we have eliminated PRO from the definition of paths the same way we have eliminated NP-trace, by restricting path formation to empty categories that are not A-bound. This is not really adequate, however, since PRO may be entirely unbound, and be arbitrary in interpretation:

(158) I [_{VP₁} think [_{S'₂} that [_{S'₂} [_{S'₃} PRO to kill tigers] [_{INFL'₂} [_{VP₂} is bad]]]]]

In (158), PRO should generate an infinite path, since it is not A-bound and has no grammatical antecedent. This would give (158), incorrectly, the status of a Subject Condition violation:

(158) Paths

(i) Between INFL'₂ and COMP of S'₂:

{INFL'₂, S'₂, S'₂}

(ii) From PRO:

{S'₃, S'₂, S'₂, VP₁, . . .}

The path from PRO overlaps path (i), but neither contains the other. The PCC is violated, if path (ii) exists. Since (158) is acceptable, we must

assume that the empty category PRO does not create a path at all, whether A-bound or not. This follows as a theorem, it seems, from the Binding Theory, which limits PRO to the subject of an infinitival S', and (155), which allows the empty subject of an infinitival S' to be invisible for the definition of paths. Recall in this context the fact that PRO, like expletive subjects, was considered exempt from the ECP of Chomsky (1981a). Since the ECP is being subsumed under the PCC, it is an advantage if we can derive the special status of PRO from an independently necessary principle.

The stipulation that A-bound empty categories do not create paths was designed to cover NP-trace as well. The facts about NP-trace do not follow from (155), and it is worth considering them briefly. Suppose that NP-trace created a path. Such a path might take one of two forms. Suppose that it were a path which, like the path from WH-trace, terminated at the maximal projection dominating its antecedent. Then simple passives in tensed sentences would be, incorrectly, excluded, since the path between an NP-trace and a subject antecedent would overlap the path between INFL and COMP, and the PCC would be violated:

(159) [_S, [_S John_i [_{INFL}' [_{VP} was seen e_i]

Paths

(i) Between INFL and COMP:

{INFL', S, S'}

(ii) Between e_i and John_i:

{VP, INFL', S}

Alternatively, we might suppose that NP-trace created a path, but that only an \bar{A} -binder could provide an upper bound to a path, so that an A-bound empty category like NP-trace would generate an infinite path. This would solve (159), since path (ii) would now be {VP, INFL', S, S', . . .}. On the other hand, (160) would be incorrectly ruled out:

(160) [_{S₁} I wonder [_{S₂} what books_i [_{S₂} John_j [_{INFL'₂} [_{VP₂} seemed [_{S₃} e_j [_{INFL'₃} to [_{VP₃} read e_i]]]]]]]]]]]]

Paths (i) Between INFL'₂ and COMP of S'₂:

{INFL'₂, S₂, S'₂}

(ii) From e_j (infinite path):

{VP₂, INFL'₂, S₂, S'₂, S'₁}

(iii) Between e_i and what books_i:

{VP₃, INFL'₃, S₃, VP₂, INFL'₂, S₂, S'₂}

Although the infinite path from the NP-trace overlaps the path between INFL and COMP, the PCC is satisfied there. The PCC is violated, however, by the overlapping of the infinite path from the NP-trace and the path created by WH-movement -- paths (ii) and (iii), respectively. They overlap, but neither contains the other.

We thus conclude that NP-trace does not (have to) generate a path at all. Recall that an "NP-trace" is simply an empty category that is locally A-bound by an antecedent in a non- θ -position. If we do not want to stipulate in the definition of paths any difference between A- and \bar{A} -binding, how can we make NP-trace invisible to the definition? One possibility is supplied by the fact that an NP-trace and its antecedent or antecedents

form a chain, which, we have seen, is the unit relevant to the θ -criterion and to the Case filter. An NP-trace, unlike any other empty category will not head its chain. Limiting path formation to empty categories that head chains is, perhaps, more natural than limiting it to empty categories that are not A-bound, since heads of chains appear to figure elsewhere as a natural class, while not A-bound categories do not:

(161) Definition of Paths (revised: III)

Consider T such that for some index i and chain C:

T = {t | t is an empty category heading C and indexed i}

. . . (as in (64))²²

Before returning to the main thread of this discussion, a final remark might be made. By assuming the "visibility" principle (155), we have provided an answer to the problem raised by the non-interaction of expletive subjects and WH-movement in Italian (example (154)) which seems also to explain why PRO does not show effects of the PCC. We noted that this is the PCC version of the question that arises in the framework of Chomsky (1981a): why is PRO not subject to the ECP? In Chapter Two we noted that both PRO and expletive empty categories appear to be exempt from the ECP. We have now derived both these facts. A further question remains, however, which may be related. Why can't an expletive element occur where PRO can? In other words, why can't the ungoverned PRO in (162) act like e in the Italian examples (153)-(154)?

(162) *it is impossible [_S, [_S PRO [_{INFL}, [[_{TNS} -tense] to] [_{VP} be obvious
[_S, that John is a fool]]]]]

We noted that (155) forced us to claim that infinitives have a TNS node, although specified [-tense], as above. Remember that (137)b stated that [_{TNS}+tense] governs a complete proposition at LF. (137)b was, in effect, a conditional: if TNS is [+tense], then it governs a complete proposition at LF. Suppose we strengthen (137)b to a biconditional:

(163) TNS governs a complete proposition at LF iff TNS is [+tense]

It follows from (163), but not from (137)b, that [-tense] TNS will be simply impossible when the subject is expletive. Given (137)a, it will appear in INFL in D-structure, and govern VP. Whether it moves or stays, it will govern a complete proposition, since VP is a complete proposition when the subject is expletive. We thus correctly rule out (162), independent of the question of PRO. On the other hand, we now have a problem with lexical expletive subjects of [-tense] clauses:

(164) John considers [_Sit to be obvious that he is a fool]

Perhaps VP of S is not complete, because the lexical it bears some sort of quasi- θ -role, for the purposes of (163). I leave these problems as a matter for further investigation.

Returning now to our main argument, let us summarize our results so far. TNS movement appears to show Coordinate Structure Constraint effects, which we derive from the PCC, in (142)-(144), (145) and (146). The argument went as follows: while tensed Ss can be conjoined, and infinitival Ss can be conjoined, and while a tensed S' and an infinitival S' (with separate COMPs) can be conjoined, a tensed and an infinitival S cannot be conjoined. Our examples involved tensed Ss with θ -marked subjects, forcing TNS to move to COMP. The movement of TNS to COMP violated the PCC, when it did

not take place from all conjuncts equally; hence the parallel with other Coordinate Structure Constraint and ATB effects.

With the help of principle (155), which appears to solve other problems for us, we can maintain the position that TNS never has to move to COMP when the subject is expletive (except in languages like English, where TNS movement is obligatory). Thus, with respect to TNS movement, in languages not like English, tensed Ss with expletive subjects have the status of infinitives. We claimed that TNS movement is the culprit in the non-conjoinability of tensed with infinitival Ss in the examples we have considered. If this claim is correct, and if TNS does not have to move to COMP when the subject is expletive, we predict that conjunction of a tensed S with an infinitival S will be grammatical if the subject of the tensed S is expletive. This prediction appears to be correct.

First, we must look at a language like Italian or Spanish, in which TNS movement is not obligatory, to test this prediction. In both languages, as in English, tensed Ss may be conjoined with tensed Ss, and infinitival Ss with infinitival Ss. Some speakers find the results somewhat unacceptable, but the contrasts we will describe hold for all speakers I have asked:

(165)a. No sé [_S,que libros [_{S₀} [_{S₁} leyó Maria] ni [_{S₂} Juan compró]]]

I don't know what books read Maria nor Juan bought

b. no sé [_S,que libros [_{S₀} [_{S₁} leer] ni [_{S₂} comprar]]]

I don't know what books to read nor to buy (Spanish)

(166)a. non so [_S,che libri [_{S₀} [_{S₁} Maria legge] o [_{S₂} Giovanni compra]]]

I don't know what books Maria is reading or Giovanni is buying

b. non so [_S,che libri [_{S₀} [_{S₁} leggere] o [_{S₂} comprare]]]

I don't know what books to read or to buy (Italian)

As in English, a tensed S' may be conjoined with an infinitival S', perhaps with some stylistic clumsiness:

(167) no sé [_{S₀}, [_{S₁} que libros leer] ni [_S, que libros compró Juan]]

I don't know what books to read nor what books bought Juan

(168) non so [_{S₀}, [_{S₁} che libri leggere] o [_{S₂} che libri Giovanni ha comprato]]]

I don't know what books to read or what books Giovanni has bought

As we expect, a tensed S may not be conjoined with an infinitival S, when the subject of the tensed S is θ -marked:

(169) *no sé [_S, que libros [_{S₀} [_{S₁} leer] ni [_{S₂} Juan compró]]]

I don't know what books to buy or Juan bought

(170) *non so [_S, che libri [_{S₀} [_{S₁} leggere] o [_{S₂} Giovanni compra]]]

I don't know what books to read or Giovanni is buying

Finally, as our theory predicts, the ungrammaticality of conjoining a tensed S with an infinitival disappears when the subject of the tensed S is expletive. For some speakers, there is some residual unacceptability, perhaps due to a lack of parallelism, but the sharp contrast between (169)-(170) and the following sentences is clear for all speakers:

(171) no sé [_S,que libros [_{S₀} [_{S₁} leer] ni [_{S₂} [e] es necesario [_S,que
Juan compre]]]]

I don't know what books to read nor [e] is necessary that Juan read

(172) ?non so [_S,che libri [_{S₀} [_{S₁} leggere] o [_{S₂} [e] è necessario
[_S,comprare]]]]

I don't know what books to read or [e] is necessary to buy

(169)-(170) are ungrammatical for the same reason (144), discussed in (140), is out (*I wonder what books to like and you hate). The movement of TNS to COMP in only one of two conjuncts violates the PCC by overlapping and not being in any containment relation with the path between the conjunction and conjuncts. In (171)-(172), neither clause contains any movement of TNS to COMP: hence no PCC violation can result. Notice that the WH-phrase in COMP in these sentences binds a gap in each conjunct, so it satisfies the PCC.

We take the contrast just discussed to be strong evidence for our theory of TNS movement.²³ Our theory of TNS movement, in turn, supports our constant hypothesis that the INFL-COMP relationship has the properties of the relationship between an empty category and an \bar{A} -binder. If this hypothesis is true, then our derivation of the basic ECP effects and Subject Condition effects from the same constraint that explains Crossing effects acquires great plausibility.

In addition, this contrast supports the claim that TNS movement is indeed subject to the effects of Ross's Coordinate Structure Constraint. This in turn indirectly supports our contention in the previous section (3.2) that the non-redundant part of this constraint should be reduced to the PCC. In the previous section we showed that if TNS movement is subject

to the Coordinate Structure Constraint, then the PCC derivation of Ross's constraint (but not Ross's constraint itself) explains the subject/object asymmetries in ATB constructions discovered by Williams. In this section we have showed that TNS movement must be subject to the effects of Ross's constraint for independent reasons, thereby supporting our analysis in the previous section, and our derivation of Ross's constraint.

Finally, a reminder that technical problems in our argument led us to postulate principle (155), which held that empty categories not in the domain of tense in a particular domain are not visible for the definition of paths. This principle, it appears, allows us to eliminate the stipulation that only non-A-bound empty categories generate paths -- itself a step forward.

Appendix to 3.3

We have shown that expletive empty subjects in languages like Italian do not have to generate paths because they are not necessarily in the scope of the TNS of their S. In this they contrast with subjects that are \bar{A} -bound by an operator -- for example, the trace at LF of the quantifier nessuno 'noone', discussed by Rizzi (1982). Like French personne, discussed in Chapter Two, nessuno as subject cannot have scope outside its S, a straightforward subject/object asymmetry captured by the PCC, subsuming the ECP.

There is another type of empty subject in Italian and Spanish that our theory does not shed any light on as yet. This is the empty subject that is interpreted as a definite pronoun:

(173) [_S, [_S [_e] [_{INFL}, [_{VP} mangia i macaroni]]]

he/she/it eats the macaroni

The subject in such constructions appears to act as if it bore the θ -role assigned by VP, since sentences like (173) cannot be conjoined with an infinitive. TNS apparently must move to have scope over the subject.

(174)-(175) are as bad as (169)-(170):

(174) *no se [_Sque libros [_{S₀} [_{S₁} leer] ni [_{S₂} [e] compro]]]

I don't know what books to read nor he/she/it bought

(175) *non so [_Sche libri [_{S₀} [_{S₁} leggere] o [_{S₂} [e] compra]]]

I don't know what books to read or he/she/it is buying

We have nothing intelligible to say about such "null pronouns" here. In this study, we will simply stipulate that such empty categories, perhaps bound, as Safir (1982) suggests, by a null clitic, count as lexical pronouns for the definition of paths, and therefore do not create any paths.

In (171)-(172), the complement of the predicate meaning 'necessary' is an S', which, as we have seen, does not require Case. In Spanish and Italian, complements and subjects can occur non-initially which do require Case. It is generally assumed that they receive Case, perhaps directly by transmission from the subject position, via cosuperscripting, as suggested by Chomsky (1981a):

(174) non so che libri [_S[e]ⁱ compra Giovanniⁱ]

I don't know what books buys Giovanni

What is the status of the cosuperscripted empty subject. For the purposes of path construction, it appears to act like a lexical subject. First,

(174) shows that [e]ⁱ must not generate an infinite path, since this would overlap the path created by WH-movement, and violate the PCC (cf. our

discussion of (154)). Second, the evidence from conjunction shows that TNS moves to COMP to have scope over $[e]^i$:

(175) *no sé [_S que libros_j [_{S₀} [_{S₁} leer e_j] ni [_{S₂} [e]ⁱ fueran comprados e_j^i]]

I don't know what books to read nor were bought

(176) *non so [_S che libri_j [_{S₀} [_{S₁} leggere e_j] o [_{S₂} [e]ⁱ sono comprati e_j^i]]

As with null pronouns considered above, we will simply stipulate that these empty subjects do not count as empty categories for path construction. Why and how, we leave for further investigation.

Before leaving this topic, it is worth noting that evidence from parasitic gaps provide some, rather paradoxical, evidence that the subject in (175)-(176) does bear the θ -role of the inverted subject. Such subjects appear to count for the anti-c-command requirement on parasitic gaps:²⁴

(177)a. quanti_i ne avevano catalogati e_i , senza che tu li avessi letti

how many of them have they filed without you having read them

b. quanti_i ne avevano catalogati e_i , senza che tu avessi letto e_i

how many of them have they filed without you having read

(178)a. quanti_i ne erano stati catalogati e_i , senza che tu li avessi letti

how many of them were filed without you having read them

b. *quanti_i ne erano stati catalogati e_i , senza che tu avessi letto e_i

how many of them were filed without you having read

Chomsky (1981a) argues for a cosuperscripting relation in structures like (174)-(176), which links the postverbal subject with the subject position without locally A-binding it and violating principle C of the Binding Theory. On the other hand, (178)b is presumably ruled out, like its English counterpart, because the parasitic gap, the object of letto

is locally A-bound by the subject. We know, from (177)b, that the object of catalogati does not A-bind it.

If the parasitic gap is locally A-bound by the preverbal subject position, it is not clear why the postverbal subject is not also locally A-bound from this position; the parasitic gap should be able to be coindexed with the postverbal subject without bearing any relation relevant for principle C to the subject position. We leave this puzzle, like the others, for future research.^{25,26}

4.0 Multiple Gaps, the PCC and Subjacency

In this section, we develop a general theory of multiple gap constructions, extending the analysis of Chomsky (1981b) in the light of the PCC. We will use the term "multiple gaps" to refer to gaps that share a local binder: thus, in our analysis, both the "parasitic gaps" of Taraldsen and Engdahl and the ATB gaps of Williams are "multiple gaps". This section is devoted to showing that this unification of parasitic and ATB gaps in one theory is correct, and to exploring the theoretical consequences of a general theory of multiple gaps. In this, we differ from Taraldsen, Engdahl and Chomsky, who all assume or argue that parasitic gaps and ATB gaps are essentially different phenomena. We have assumed throughout that UG treats the multiple gaps in both of these constructions exactly the same way. They generate the same kinds of paths, subject to the PCC. The special properties of multiple gaps in conjoined structures arose from one hypothesis only: that a path exists between a conjunction and its conjuncts.

In 4.1, we discuss in more detail the apparent differences between the two constructions that have been noted by Engdahl and Chomsky, and show how the PCC theory explains them. In 4.2, we outline Chomsky's (1981b) analysis of the parasitic gap construction: in particular, how parasitic gaps become \bar{A} -bound, and how they interact with the subjacency condition. In 4.3, we introduce a new (apparent) difference between parasitic gaps and ATB gaps: parasitic gaps appear to violate subjacency freely, but ATB gaps in the normal case seem to obey the subjacency condition. We present an extension of Chomsky's analysis that explains

this difference. We suggest that our explanation provides significant evidence about the levels at which the PCC applies, and about the nature of the subjacency condition and of Move α . We develop this theme in the following sections.

Some evidence for our approach to ATB gaps is provided in 4.4 by comparing our analysis with others that have appeared in the literature. Chomsky and Lasnik (1977), followed by George (1980) and Sjöblom (1980), suggest that there are no ATB exceptions to the Coordinate Structure Constraint: apparent multiple gaps in conjoined structures actually arise by the deletion under identity of \bar{A} -binders and other material in the separate conjuncts. Williams (1978) suggests a special three-dimensional structure for conjoined constituents, and a condition on joint factorization that yields the Coordinate Structure Constraint and the ATB exceptions. We show that the appearance of Coordinate Structure Constraint effects and ATB exceptions in parasitic gap structures argue against both these analyses, and for our unified theory. Our arguments shed further light on the levels at which the PCC and the subjacency constraint apply. In 4.5 we show that our theory links two issues: (1) does Move α allow for the simultaneous factorization of more than one element? (2) is subjacency a condition on Move α or a condition on representations at some level? We argue that if the answer to (1) is no, then subjacency is a condition on representations, but if the answer to (1) is yes, then subjacency may be a condition on Move α . We discuss these alternatives, but leave the issue open.

4.1 A Unified Theory of Multiple Gap Constructions

As has been frequently noted, the study of empty categories is of more than technical interest. Phonologically null elements are not observable in the stream of speech. A child learning his language cannot induce their existence from what he hears. To the extent to which they can be shown to have significant properties, these properties must reflect something of the innate knowledge the child has of grammar, independent of anything he learns about his particular language. As we suggested in Chapter Two, the existence of empty categories and their nature derives largely from the Projection Principle, from the theory of c-selection, and from the binding theory and other subtheories of grammar. The Projection Principle and c-selection contribute to determining the distribution and properties of empty categories by utilizing lexical information that the child might acquire partially from experience. The use to which the grammar puts this information, in determining the empty categories, cannot possibly be learned by any primary experience.

This argument holds with even greater force, if such is possible, for multiple gap constructions. As Chomsky (1981b) notes, in discussing parasitic gap constructions, multiple gaps are always rather marginal. It is impossible to believe that any facts about the properties of parasitic gap constructions are learned by experience, given the unlikelihood of such structures occurring in the primary linguistic data. Surely the properties of these constructions must result from independently specified principles of universal grammar.

Chomsky makes a further point about parasitic gap constructions, which we consider important and wish to elaborate. Even if parasitic gaps

did not exist, the other known phenomena of natural language, because of their surface variety and complexity, lead us to posit a rich system of interacting principles and subtheories, whose independent operation and interaction can yield the pattern of phenomena that we find. Thus, in Chapter Two, the interaction of a Case theory, a government theory for empty categories (replaced now by the PCC), the Projection Principle and a theory of c-selection all contributed to the explanation for the odd distribution of quantificational constructions. Ideally, we expect these principles and subtheories to be extremely simple and natural: a normal scientific assumption. It is possible, for example, that UG will specifically provide for variable binding, as an aspect of the mental representation of quantification. One might expect some principles of UG to refer to variables. This is not a sine qua non, but seems reasonable. Similarly, UG might provide for bound anaphora, and include some principles determining where anaphora is possible and where it is impossible. A child learning English, then, might be expected to figure out from context that himself is a reflexive anaphor. This information immediately indicates what principles of UG are relevant to himself, and the child knows the properties of himself.

On the other hand, as Chomsky notes, it is unlikely, from any point of view, that UG contains special principles governing parasitic gaps. For one thing, parasitic gaps are clearly not a sine qua non of natural language. In most contexts, as Engdahl discusses, they are replaceable by a pronoun:

- (179)a. the article that I filed e without reading e
 b. the article that I filed e without reading it

There is no functional motivation for their existence, to be sure. For these reasons, we may consider a theory that does not have to specify the properties of parasitic gaps to have the advantage over one that does. It is an advantage of a theory incorporating the Θ -criterion and Projection Principle, for example, that the anti-c-command requirement on parasitic gaps, discussed several times above, does not have to be stated. It is an advantage of a theory in which lexical insertion is optional that the existence of parasitic gaps does not have to be stipulated: we expect empty categories to be freely generable. Similarly, it is an advantage of a theory incorporating the PCC (or Kayne's KECP) that we do not have to specify the conditions under which the Subject Condition can be violated in parasitic gap constructions.

A similar, more subtle argument applies to multiple gap constructions in coordinate structures -- so-called ATB gaps. Consider Ross's original discussion of the Coordinate Structure Constraint. Williams (1977) notes that his constraint (cf. our (99)) could be rephrased to take into account the ATB exceptions: "If a rule applies into a coordinate structure, then it must affect all conjuncts of that structure." This is an essentially adequate statement of some of the facts of the matter, ignoring the subject-object asymmetries later discovered by Williams himself (1978), and ignoring the distinction between A- and \bar{A} -binding that seems relevant.

Nonetheless, this Coordinate Structure Constraint is just the sort of principle we do not hope to find in UG. For one thing, in allowing ATB exceptions, it is completely unlike any other "locality principles" (Koster 1978b) that we know of. Compounding a Coordinate Structure Constraint violation makes the structure better; compounding a subjacency

or Binding Theory violation does not generally improve anything.

Furthermore, Ross's constraint has no obvious connection with any intrinsic property of conjunctions. In this, it is unlike, for example, the Binding Theory as it applies to anaphors. Anaphors must be bound, for semantic reasons; it is not surprising, though not inevitable, that UG should guide and limit the possibilities for binding. On the other hand, nothing about conjunction leads us to expect that extraction from a conjoined structure should be limited, and in such a curious way. When a child learns "and is a conjunction", UG tells him what a conjunction is. Other properties of UG may dictate what can be conjoined, and when. It would be odd (though not impossible) if UG also assigned completely wild properties to conjoined structures: why doesn't UG restrict extraction from clauses containing anaphors?

Finally, extraction from a conjoined structure appears to be possible when a multiple-gap configuration of a certain kind is present. We no more expect such ATB multiple gaps to be governed by special principles of UG than we expect parasitic gaps to be. Chomsky's arguments about parasitic gaps apply to ATB gaps as well.

For all these reasons, it is good to assume that the correct theory of UG will not contain special principles concerning extraction from coordinate structures. Rather, the best theory will derive the Coordinate Structure Constraint effects from some intrinsic property of conjunctions, the most general theory of multiple gaps (one which says nothing at all about them), and other general principles. The PCC theory has these properties. By contrast, Ross's constraint, or the special three-dimensional notation and conditions on analyzability of Williams (1979), stipulate something special about conjoined structures, and

are thus less plausible candidates for UG. We shall show later that Williams' theory, though extremely elegant, also has some empirical defects, when extended as needed to handle parasitic gaps.

Let us mention another approach to the problem, which dates at least to Chomsky and Lasnik (1977). The Coordinate Structure Constraint (as revised by Williams (1977)) is, as we have said, unlike other known principles of UG, because it stipulates that extraction from a conjunct is ungrammatical unless the violation is compounded in each conjunct. The constraint would look much more like other locality conditions if there were no ATB exceptions. This approach would leave us with an ad hoc condition for the moment, but would suggest the possibility of collapsing it with other known constraints, like subjacency. Chomsky and Lasnik's proposal will also be considered later. Assume for now that multiple gaps do improve Coordinate Structure Constraint violations -- that ATB constructions exist.

Now let us consider Engdahl's suggestion that the ATB construction falls under a different theory than parasitic gaps. She presents two arguments to support this claim, both of which disappear on a PCC account.

At first sight, the most compelling has to do with optionality of the multiple gaps. We have already seen in (179)b that gaps in parasitic gap constructions may be replaced by a pronoun. While there are sometimes slight differences in acceptability that depend on which gap is replaced by a pronoun, discussed by Engdahl, the result seems generally good, if one remaining gap does not violate subjacency. Similarly, all the gaps may be replaced by pronouns. The result is as acceptable as any other

English sentence with a resumptive pronoun:

(180) the article that I'm sure you filed it without reading it

In a coordinate structure, however, where each conjunct contains a gap, replacing some proper subset of the gaps with resumptive pronouns always results in ungrammaticality. Only if all gaps are replaced with resumptive pronouns is the result fairly acceptable:²⁷

- (181)a. a man who I consider e pleasant and you regard e as annoying
 b. *a man who I consider him pleasant and you regard e as annoying
 c. *a man who I consider e pleasant and you regard him as annoying
 d. a man who I consider him pleasant and you regard him as annoying

Engdahl argues that in the parasitic gap constructions, where only one gap is necessary for grammaticality, it makes sense to call other gaps parasitic. What is more, we can distinguish possible "real gaps", which obey subjacency, from doubtlessly parasitic gaps, which do not. In ATB constructions, however, Engdahl claims that it makes no sense to ask which gaps are parasitic and which are created by movement, because they are all necessary to grammaticality. The status of this argument is unclear. In any case, the important contrast between (179)-(180) and (181) is straightforwardly accounted for by the PCC, on the assumption that the multiple gaps in the two constructions have the same properties.

We must simply assume that resumptive pronouns do not generate paths -- something which follows from our definition of paths. This is unquestionably true in other cases where we have invoked paths and the PCC. Resumptive pronouns do not, for example, show CTP effects or Subject

Condition effects, indicating that they produce no path that interacts with the path between INFL and COMP:

- (182)a. *the man that I think that e likes Mary
 b. the man that I think that he likes Mary
 c. *the man who for John to praise e would surprise Mary
 d. the man who for John to praise him would surprise Mary

Similarly, a resumptive pronoun cannot "save" a sentence from showing Subject Condition effects, unlike an empty category. That is, there is no unified path between e, him and COMP in (183)b, so that (183)b has the same status as (182)c:

- (183)a. ?a man who I think that close friends of e admire e
 b. *a man who I think that close friends of e admire him

A resumptive pronoun shows no Crossing Effects:

- (184)a. *John, who_i I know what book_j to persuade e_i to read e_j
 b. John, who_i know what book_j to persuade him_i to read e_j

Notice that whether a gap is obligatory, optional, or impossible is only a superficial issue. The real question is whether the presence of the path produced by a gap saves a sentence from violating the PCC, is irrelevant to the PCC, or causes a sentence to violate the PCC. In (182), a gap is impossible, because its presence would violate the PCC, In (183), a gap is obligatory, because only its presence prevents the PCC from being violated. In (184), a gap is impossible, or else the PCC is violated. In parasitic gap constructions like (179), a gap is optional, since the PCC is not violated with it or without it. In

coordinate structures like (181), a gap must be present in all conjuncts or no conjuncts, to avoid violating the PCC. The real issue is the PCC, a principle of UG, and not the apparent obligatoriness, impossibility or optionality of a gap.

Furthermore, we have seen nothing so far to support the assumption that ATB constructions are intrinsically different from parasitic gap constructions, except insofar as they contain one additional path created by conjunction. Even the question of the optionality of gaps offers no evidence for this distinction. (183) shows a parasitic gap construction in which a gap is "obligatory"; (179)-(180) shows a parasitic gap construction in which a gap is "optional". The superficial distinction between obligatory and optional gaps cuts across the alleged distinction between parasitic gap and ATB constructions, and derives in any case from the PCC.

Chomsky (1981b, note 40) notes a second specific property of ATB constructions, namely the apparent "parallelism" requirement, or "subject-object asymmetry", discussed above in connection with (124)b, (128), (131) and (132), repeated below:

- (185)a. John, who Bill saw e and Mary likes e
 b. John, who e saw Bill and e likes Tom
 c. *John, who Bill saw e and e likes Mary
 d. John, who Bill likes e and we hope e will win

We have shown, however, that this paradigm follows immediately from the PCC, if the path created by TNS movement between INFL and COMP is subject to the PCC, as it must be. (185) is thus explained, and does not support an intrinsic difference between the constructions.

In general, the PCC appears to contribute to a general theory of multiple gap constructions that makes no distinction between ATB and parasitic gap constructions. By Occam's razor, at least, the PCC account seems preferable to a theory in which special principles account for ATB constructions. We will shortly return to a difference between ATB and parasitic gap constructions that is more difficult to explain away than those we have considered. First, let us consider in greater detail Chomsky's analysis of the parasitic gap construction, which will lead us to a necessary discussion of the subjacency condition.

4.2 Chomsky's Analysis of Parasitic Gaps²⁸

Chomsky's (1981b) analysis of parasitic gaps, like Taraldsen's and Engdahl's, alleges a distinction between "real gaps", which stand in the relationship Move α to their antecedents, and "parasitic gaps", which acquire antecedents as the result of another type of linking. Before considering this distinction further, let us examine the properties of the relationship called Move α in Chomsky's approach.

Recall that the level of D-structure is by definition a pure representation of the relation $GF-\theta$. It is a level at which each θ -position is occupied by an argument and each argument occupies a θ -position. The θ -criterion is always satisfied by arguments at D-structure. At S-structure, arguments may be assigned to non- θ -positions, and θ -positions may be occupied by non-arguments. The θ -criterion may be satisfied at S-structure by chains which contain an argument, where the argument does not necessarily occupy the θ -position.

The mapping between D-structure and S-structure, so described, is characterized by the rule Move α . The rule Move α , in the framework of

Chomsky (1981a), has three characteristic properties:²⁹

- (1) Move α relates two positions.
- (2) The category in the antecedent position does not bear an independent θ -role.
- (3) The subjacency condition is obeyed obligatorily.

Much of our argument will concern property (1), which plays no central role in Chomsky's discussion, but is extremely important to us. Note that property (2) derives from the θ -criterion and the Projection Principle, with the definition of c-command and the Binding Theory giving content to the notion "antecedent position". (One position must c-command the other; the one that does is the antecedent.) Property (3) states that Move α falls under Bounding Theory.

For our purposes, we may assume a fourth property -- that the non-antecedent is an empty (that is, phonologically null) category. Not all relationships between two positions, one of which is empty, fall under Move α . Chomsky demonstrates that the relationship between the ungoverned empty category PRO and its antecedent does not show properties (2) and (3). The same is true for property (1). The binding of PRO falls under Control Theory. Control Theory may relate several positions: PRO may have split antecedents:

(186) John_i told Mary_j that it would be hard PRO_{i,j} to see each other_{i,j}

PRO's antecedents may be non-subjacent and in a θ -position:

(187) Mary_i regrets [_{NP} the fact [_S that it will be impossible PRO_i to find herself_i a job in New York]]

The properties (1)-(3) are diagnostic of Move α , but are not necessarily exclusive to Move α , in Chomsky's theory. Parasitic gaps are a case in point: they show property (2), but not property (1) or property (3). A set of parasitic gaps is a set of gaps that share a local binder. If any gap, parasitic or otherwise, is bound from an A-position bearing a θ -role, it will enter into a chain with its binder. If the gap bears a θ -role, then the θ -criterion will be violated.³⁰ Thus parasitic gaps must show property (2), for the same reason other gaps that are arguments do:

(188) *John_i objected to e_i (in order to punish e_i)

On the other hand, the relation of a binder to a set of parasitic gaps is one-many, violating (1) trivially. Finally, as we have seen, the relation between a parasitic gap and its antecedent may violate subjacency:

(189)a. John, who [_{NP} the man [_S, who liked e]] had to destroy e
 b. the article that I published e without meeting [_{NP} the person [_S, who wrote e]]

Suppose we take property (3), obligatory obedience to subjacency, as being definitive of Move α , unlike (2). (We continue to delay discussion of (1).) It follows then that the relationship between non-subjacent, parasitic gaps and their \bar{A} -binder in sentences like (189)a-! is not Move α . Rather, parasitic gaps become bound by their \bar{A} -binder as the result of some other process that coindexes them.

But now a new question arises: if a coindexing process that is

not Move α and does not obey subadjacency can link an empty category with a binder, why is the subadjacency constraint observable? In other words, suppose that a parasitic gap is coindexed with an \bar{A} -binder by means of a rule R which does not obey subadjacency. Why isn't (190) a grammatical output of rule R?

(190) *an article that_i I met [NP the person [$_{S}$, who wrote \underline{e}_i]]

This is the "licensing" problem raised by Engdahl. Non-subjacent, parasitic gaps can only exist if they are "licensed" by a subjacent gap -- in short, if the \bar{A} -binder of the parasitic gap is in the Move α relationship with some other empty category, as in (189)a-b.

Chomsky solves the problem as follows. With Taraldsen, he assumes that parasitic gaps are null pronominals in D-structure. The existence of such null pronominals is the null hypothesis: if phonological content is optional, we expect null elements to function freely as arguments at D-structure, if they properly satisfy the Binding Conditions and other conditions at S-structure and at LF. These null pronominals become \bar{A} -bound at S-structure, as we have indicated, by an indexing rule. Again the null hypothesis is that coindexing at S-structure is free. To avoid (190), however, Chomsky suggests that coindexing at S-structure only applies to A-positions. In order to generate (190), we must coindex \underline{e} with that; since that is in an \bar{A} -position, this is impossible at S-structure.

A-positions are coindexed with \bar{A} -positions by Move α , however. If Move α applies from an A-position into an \bar{A} -position, as with WH-movement, then the A- and \bar{A} -positions will be coindexed at S-structure.

Now suppose a sentence in which such movement occurs also has a null pronominal. The null pronominal cannot be directly coindexed with the \bar{A} -binder, as we have seen. It can, however, be coindexed with the A-bound trace of the \bar{A} -binder. By transitivity of indexing, it is thus bound by the \bar{A} -binder. In other words, only when a potential \bar{A} -binder is already coindexed with an A-position at S-structure -- and movement is the only procedure that can effect this -- can a null pronominal acquire the index of the \bar{A} -binder, by coindexation with the bound A-position. Thus Chomsky solves the licensing problem.

A number of new problems arise, however, on this solution. First, why must null pronominals eventually be coindexed with an \bar{A} -binder? That is, why can't they pick their own index freely, or appear coindexed with another NP that is not \bar{A} -bound?

(191)a. *John_i bought e_j

b. *John bought the book_i and Mary bought e_i too

Chomsky relates this problem to the general problem of "pro-drop" phenomena. Roughly speaking, an empty category needs to be "identified" (in something like the sense of Jaeggli 1980) in order to function as an argument a_t some particular level. Control Theory performs the necessary identification for PRO. The trace of a clitic is identified by the clitic. A variable is identified by its \bar{A} -binder, which assigns it features and determines its range. In (191), the null pronominals are not identified in English, although they might be identified in languages like Russian or Japanese, where sentences parallel to (191) (or homonymous with their equivalents)

are grammatical. Note that if we identify \underline{e}_i by control theory, and call it PRO, it will be ruled out by the Binding Theory in (191).

Another problem has to do with resumptive pronouns, which may be locally \bar{A} -bound and violate subjacency, but do not require a licensing gap created by movement:³¹

(192) an article_i that_i I met [_{NP} the person [_S who wrote it_i]]

If only A-positions are coindexed at S-structure, how do it and that come to be coindexed in (192)? Chomsky proposes that \bar{A} - as well as A-positions are subject to free coindexation at LF. The identification requirement on empty categories will then have to operate before LF, to prevent a gap in the position of it in (192). Thus:

(193) S-structure: free indexing of A-positions
 identification requirement

 LF: free indexing of all positions (A and \bar{A})

The ordering in (193) also accounts for why movement rules of LF, like the rule that moves a Wh-in-situ into COMP (Aoun, Hornstein and Sportiche 1981; cf. Chapter Five), do not license parasitic gaps. That is, (194)a is ungrammatical even if it has an LF representation like (194)b:

(194)a. * [_S who_i [_S \underline{e}_i filed what article_j without reading \underline{e}_j]]

 b. [_S who_i what article_j [_S \underline{e}_i filed \underline{e}_j without reading \underline{e}_j]]

The S-structure represented in (194)a fails the identification requirement for \underline{e}_j , in the same way the empty categories in (191) fail.

Alternatively, we might propose that LF contains no free indexing. Robert May (personal communication) has pointed out that resumptive pronouns are much more acceptable in relative clauses than in questions, at least in English:³²

- (195)a. this book, that I think Sue left it on the bed
 b. ??what book do you think Sue left it on the bed

If this is the case, we might suppose that resumptive pronouns are coindexed, not with a WH-phrase, but rather with the head of a relative clause (as suggested by Chomsky (1981b, note 14), following suggestions of Williams). The resumptive pronoun ends up coindexed with the WH-phrase because the Wh-phrase undergoes a later predication rule (cf. Chomsky 1977) that coindexes it with the head of the relative. This view is supported by the fact that resumptive pronouns are impossible in cases where the predication rule applies to only a piece of the \bar{A} -binder in COMP, as with pied-piping of NPs:

- (196)a. John, who_i I'm sure that he_i will come
 b. *John, whose mother_i I'm sure that she_i will come
 c. *John, pictures of whom_i I'm sure that they_i are on sale cheap

If there is no coindexing of \bar{A} -positions, but only coindexing with the head of a relative, perhaps at S-structure, we can explain the impossibility of (196)b-c, since she cannot be coindexed with John and yield the correct interpretation. We must assume that coindexation by the head of a relative clause does not identify an empty category. This would yield a simpler picture than (193), although it is not clear which picture is the null hypothesis, in default of powerful arguments for one or the

other:

(197) S-structure: free indexing of A-positions
 identification requirement

LF: ----- (for present purposes)

In any case, limiting free indexing to A-positions at S-structure has the desired result in parasitic gap constructions. A parasitic gap, violating subjacency, can only appear if its \bar{A} -binder also binds a gap created by Move α , which must obey subjacency. As a consequence, among each set of gaps that share an \bar{A} -binder, one gap at least must obey subjacency. This is true, not only in examples like (190), but also in examples like:

- (198)a. *the article that John filed [_{NP} the review [_S, that I wrote of e]] without meeting [_{NP} the person [_S, who edited e]]
- b. *John, who [_{NP} people [_S, who like e]] admire [_{NP} the things [_S, that I say about e]]³³

Notice that nothing in Chomsky's analysis of parasitic gaps ever forces more than one gap to obey subjacency. This will be important in the next section, where we discuss a problem for our PCC and the general theory of multiple gap constructions.

4.3 Subjacency and the PCC

We have argued that nothing except the path created by conjunction distinguishes parasitic gap constructions from ATB constructions. At first sight, however, there is an important difference between the two constructions. In the ATB construction, all the gaps must obey

subjacency:

- (199)a. a book that John [read e] and [disapproved of e]
 b. *a book that John [read e] and [disapproved of [_{NP} the man
 [_S, who reviewed e]]]
 c. *a book that John [read [_{NP} the advertisement [_S, that I wrote
 for e]] and [disapproved of e]
 d. *a book that John [read [_{NP} the advertisement [_S, that I wrote
 for e]] and [disapproved of [_{NP} the man [_S, who reviewed e]]]

Nothing so far accounts for this fact. We expect (199)d, where all gaps violate subjacency, to be ungrammatical. Nothing yet explains why (199)b-c, where one gap obeys subjacency, should also be ungrammatical. One gap should derive by Move α , and the other gap by free coindexation at S-structure, if ATB gaps are no different from parasitic gaps. Similarly, the PCC makes no distinction among the sentences of (199), as things stand. In each sentence, a path should run between the \bar{A} -binder in COMP and all the gaps in all the conjuncts, which should contain the path created by conjunction.

This problem is not specific to ATB constructions, however. Consider the examples discussed in section 2 (and (183) above), in which a gap inside a subject is "saved" from the PCC by a lower gap in object position. To our ears, only the gap inside the subject can violate subjacency. The "saving" gap cannot:

- (200)a. John, who_i [_S, for you to visit e_i] would encourage e_i
- b. John, who_i [_{NP} people [_S, who know e_i]] like e_i
- c. *John, who_i [_S, for you to visit e_i] would encourage [_{NP} people
[_S, who support e_i]]
- d. *John, who_i [_{NP} people [_S, who know e_i]] like [_{NP} articles
[_S, that I write about e_i]]

The crucial sentence is (200)c. In (200)a, neither gap violates subjacency. In (200)b, the gap inside the subject violates subjacency, but the saving gap does not. In (200)c, the gap inside the subject does not violate subjacency (unless S' is dominated by NP), but the object gap does. In (200)d, both gaps violate subjacency. We can explain all the judgments except (200)c: since one gap obeys subjacency, why it is ungrammatical?

The problem noted in (199) thus seems to cut across the distinction between parasitic gaps and ATB gaps, and suggests an explanation that has nothing to do with this distinction. Instead, let us ask what the constructions in (199) and in (200) have in common.

We have already had occasion to link these two constructions in section 4.1 (examples (181) and (183)). There we noted that in each of these constructions one of the gaps is obligatory, if another gap is present. In the ATB construction, the PCC tells us that if one gap is found free in one conjunct, all the other conjuncts must contain gaps as well. In the construction in (200), the PCC tells us that if the subject contains a gap, there must be another gap further down the tree to save it.

In the unexpectedly bad ATB sentences (199)b-c, and in the unexpectedly bad parasitic gap sentence (200)c, the gap that violates

subjacency is a gap which is crucial to grammaticality under the PCC. In all the other cases in which a non-subjacent gap is possible, as in (200)b, that gap is not crucial to the PCC. We proved this earlier by noting what happens if these gaps are replaced by resumptive pronouns. In (199)b-c and (200)c, the sentence becomes ungrammatical. In (200)b, the sentence remains acceptable.

Suppose we decide that this is what is at stake. In other words, suppose that the following is a fact about UG:

(201) An empty category E must be subjacently bound in a sentence S if S would violate the PCC without E.

This captures the facts of (199)-(200) at the descriptive level. We might equally well state that "if a gap is in a position where it cannot be replaced by a resumptive pronoun, it must obey subjacency".

Thus, in the ATB constructions of (199), no proper subset of the gaps may be replaced by resumptive pronouns (or other lexical NPs). Therefore, these gaps must obey subjacency. In other words, (202)a (= (199)c) is ungrammatical because (202)b is ungrammatical:³⁴

- (202)a. *a book that John [read [_{NP} the advertisement [_S that I wrote for e¹]]] and [disapproved of e²]
- b. *a book that John [read [_{NP} the advertisement [_S that I wrote for it]]] and [disapproved of e]

Similarly, in the parasitic gap construction of (200), the gap inside the subject position, but not the saving object gap, may be replaced by a resumptive pronoun (cf. (183)). It follows that the gap

inside the subject may be non-subjacent, but the saving gap must be subjacent. In other words, (203)a-b are grammatical for the same reason, and (204)a is ungrammatical because (204)b is ungrammatical:

(203)a. John, who [_{NP} people [_S, who know e]] like e

b. John, who [_{NP} people [_S, who know him]] like e

(204)a. *John, who [_S, for you to visit e] would encourage [_{NP} people [_S, who support e]]

b. *John, who [_S, for you to visit e] would encourage [_{NP} people [_S, who support him]]

How is a non-subjacent empty category like a resumptive pronoun? To answer this riddle, recall why the resumptive pronouns in (202)b and (204)b result in ungrammaticality. We claimed in 4.1 that resumptive pronouns, like all lexical categories, do not generate paths. As a result, (202)b contains a path from a gap in one conjunct to its \bar{A} -binder that overlaps the path between the conjunction and the conjuncts; it violates the PCC. Similarly, the path from inside the subject in (204)b overlaps the path between INFL and COMP, also violating the PCC. (202)b is a normal Coordinate Structure Constraint effect, and (204)b is a normal Subject Condition effect.

The data we have been examining show that non-subjacent empty categories behave like resumptive pronouns, which do not create paths. We thus conclude that non-subjacent empty categories also do not create paths, at one level of representation, at least. The rest of this chapter is devoted to fleshing out this proposal and to exploring its consequences

for the general theory.

Let us speculate, then, that there is a level at which the PCC applies, at which an empty category is an element that stands in the relation Move α to its antecedent. The defining property of Move α is obligatory obedience to subjacency. Thus, the PCC will, at this level, see only paths from subjacent empty categories. At this level, the PCC will rule out (202)a as a Coordinate Structure Constraint effect, and (200)c as a Subject Condition effect, exactly as desired.

This solution has many consequences for the status of Move α , paths and the PCC:

First, it leads us to ask more generally at what level or levels the PCC applies. We have tacitly assumed that it applies at LF, because the ECP, which it largely subsumes, applies at LF (as discussed in Chapter Two). We have just argued that the PCC applies at a level at which only subjacent gaps count. We will shortly argue that it also applies at a level at which both subjacent and non-subjacent gaps count. Can we identify the former level with S-structure, and the latter with LF?

Second, if there is a level at which multiple gaps are all subject to subjacency, and if subjacency is a defining property of Move α , can we conclude that Move α applies simultaneously to more than a pair of positions? If so, this would deny property (1) of Move α , discussed in the preceding section. It would also follow that not all parasitic gaps must be derived by coindexation at S-structure.

In the next section, we deal with the first of these issues. We will present our argument that the PCC applies at two levels as a result

of considering some alternative analyses of ATB phenomena. In 4.5, we turn to the second of these issues, and discuss the status of property (1) of Move α .

4.4 The PCC at LF; Counteranalyses of ATB Phenomena

In the previous section we noted that in ATB constructions deriving from WH-movement all gaps must obey subjacency. This raised a problem for Chomsky's analysis of parasitic gaps, which only predicts that one of a set of multiple gaps must obey subjacency. We linked this to a parallel problem with parasitic gap constructions that violate the Subject Condition. One might object to this linkage, since some cases of parasitic gap violations of the Subject Condition seem acceptable, even when the saving gap or both gaps violate subjacency. Chomsky cites (205) (his (79)) as acceptable:

(205) a man whom [_{NP} everyone [_S, who meets e]] knows [_{NP} someone
[_S, who likes e]]

We consider (205) acceptable, but probably ungrammatical, since when the sentence is "unbalanced" slightly the result is much worse, to our ears (as noted in fn. 5):

(206) *a man whom [_{NP} everyone [_S, who meets e]] knows [_{NP} someone
[_S, who writes novels about e]]

As Chomsky notes, there is no reason, given his analysis, to expect (205) to be grammatical. (He speculates that it is not a parasitic gap construction, citing some Italian data suggested by Rizzi.) This is true

independently of our arguments in the preceding section.

Nonetheless, one might decide that examples like (205) indicate that we are wrong to link such parasitic gap constructions to ATB constructions, particularly since alternative analyses of ATB constructions have appeared in the literature that automatically explain the fact that all the ATB gaps obey subjacency, without special assumptions about levels and empty categories. In this section, we examine two such proposals. Among other arguments against these proposals, we will find that they fail to extend to coordinated structures that contain parasitic, non-subjacent gaps. We will find that these gaps also show Coordinate Structure Constraint effects and ATB phenomena, whose properties the PCC theory, but not the others, can explain. This will lead us to posit that the PCC applies both at S-structure and at LF.

4.4.1 Consider again a typical Coordinate Structure Constraint violation with its ATB counterpart:

- (207)a. *I wonder who [_S Mary called e an idiot] and [_S June called him a cretin]
- b. I wonder who [_S Mary called e an idiot] and [_S June called e a cretin]

If the second e in (207)b were a parasitic gap, we could not explain why it must obey subjacency, unless we take the position we took in the previous section. Suppose, however, that the second e is not and cannot be a parasitic gap. In other words, suppose that the Coordinate Structure Constraint were absolute, with no ATB exceptions. Chomsky and Lasnik (1977) point out that (207)b and examples like it admit of another analysis, in which (207)b contains conjoined S's and not conjoined Ss, and in which "the WH-word that appears derives from the first clause, while some sort of deletion applies in the second."³⁵ On this analysis, (207)b would have the following derivation:

(208) ... [_S, who Mary called e an idiot] and [_S, who June called e
 a cretin] ↓
∅

Since the second gap derives by ordinary WH-movement within its clause, it must obey subjacency: there is no problem on that score.

(207)a would be ruled out by the Coordinate Structure Constraint on the analysis given in which Ss are conjoined, as would (207)b, with the structure given. On the other hand, something special would have to be said about structures like (207)b analyzed as having conjoined S's:

(209)a. *a man who Mary called him an idiot and June called e a cretin
 b. a man [_S, who Mary called him an idiot] and [_S, who June called e a cretin]

(209) does not have the status of a sentence with a resumptive pronoun, but rather seems to be a Coordinate Structure Constraint violation. On this analysis, one might take a very strict view of the identity required for deletion. Perhaps a WH-word that binds an empty category is distinct from a WH-word that binds a resumptive pronoun, preventing the deletion. We will shortly see that this does not work.

More lengthy deletions could account for apparent ATB phenomena that seem to involve VP conjunctions and other non-S'-conjunction:

(210)a. I wonder who June considered e an idiot and called e a cretin
 b. I wonder [_S, who June considered e an idiot] and [_S, who June
 called e a cretin] ↓ ↓
∅ ∅

Having sketched an outline of an analysis in the spirit of Chomsky and Lasnik's proposal, let us consider why it might be attractive. We remarked earlier that Ross's Coordinate Structure Constraint was like no other locality principle, because its effects can be annulled when the violation is compounded. If Chomsky and Lasnik's proposal were viable, this odd feature of Ross's constraint would disappear. All extraction from

conjoined structures would be forbidden, and the Coordinate Structure Constraint would look more like other principles. We might then try to derive Ross's condition from other locality principles. Chomsky and Lasnik's proposal thus merits serious examination, since it does suggest interesting avenues of research.

Suppose that UG contained an inviolable Coordinate Structure Constraint. If this constraint were not derived from the PCC, it might take one or two general forms: (1) a condition on Move α , or (2) a condition on the distribution of certain empty categories. If (1) were correct, we would expect the Coordinate Structure Constraint to operate in tandem with the subjacency constraint. That is, a relation should observe the Coordinate Structure Constraint if and only if it also observes subjacency. At first blush, this seems to be true. Control of PRO, for example, obeys neither constraint. On the other hand, parasitic gaps seem to disprove (1). As we have seen, they do not obey subjacency. They do, however, obey the Coordinate Structure Constraint:

(211)a. *an article that I filed e after [contacting the man who wrote it] and [meeting the woman who edited e]

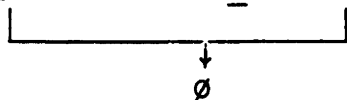
b. *John, who [men who like Mary] and [women who know e] admire e

We conclude that the Coordinate Structure Constraint is not a condition on Move α . (We return to the implications of this for the PCC theory below.) If the Coordinate Structure Constraint is absolute, and is a condition on gaps ("no conjunct may contain an empty category (of the appropriate type) free in that conjunct"), then we can rule out (211)a-b with the structures given above. But how could we rule out a

derivation by deletion, as in (212)?

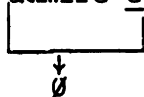
(212)a. an article [that I filed e after contacting the man who wrote it]

and [that I filed e after meeting the woman who edited e]



b. John, [who men like Mary admire e] and [who women who know e

admire e]



In each conjunct, the WH delendum binds a gap.

Compounding the problem, Coordinate Structure Constraint violations like (211)a-b, with non-subjacent parasitic gaps, also admit ATB exceptions:

(213)a. an article that I filed e after [contacting the man who wrote e]

and [meeting the woman who edited e]

b. John, who [men who like e] and [women who know e] admire e

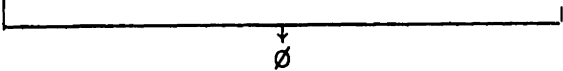

If the Coordinate Structure Constraint does allow ATB exceptions (as in a theory that derives the constraint from the PCC), then the contrast between (211) and (213) is straightforward. If apparent ATB exceptions arise by means of S'-conjunction and deletion, then it seems impossible to rule out the structures in (212) and admit their counterparts for the sentences of (213). Rather, this contrast seems to argue that ATB exceptions do exist, and, furthermore, that no deletion rule of the type envisaged exists -- since if such deletions existed, on any theory, we could not rule out the strings in (211) and allow (213).

Furthermore, we can construct examples which cannot be derived as an S'-conjunction, but which nonetheless show both Coordinate Structure Constraint effects and ATB exceptions:

(214)a. *John, who [the man who likes him] and [the woman who hates e]
often talk to each other about e

b. John, who [the man who likes e] and [the woman who hates e] often
talk to each other about e

Even supposing that we can account for the contrast between (211) and (213) by some condition on deletion across conjoined S's, (214) shows that we must still allow for ATB exceptions to the Coordinate Structure Constraint. (214)b cannot be derived from (215), since the anaphor each other requires a plural antecedent:

(215) *John, [who the man who likes e often talks to each other about e] and

 [who the woman who hates e often talks to each other about e]


(214)a-b must be genuine cases of conjoined NPs, and the contrast must be accounted for by allowing ATB exceptions to the Coordinate Structure Constraint.³⁶

If this is so, however, the arguments for Chomsky and Lasnik's proposal unravel. If we must allow for ATB exceptions to the Coordinate Structure Constraint in any case, then the deletion proposal does not solve anything. Furthermore, if such deletions under identity are allowed, as we have seen, they will require ad hoc conditions in order to properly distinguish constructions with gaps from those with empty categories, and -- more seriously -- constructions with one gap from constructions with two (examples (212)a-b). Furthermore, this proposal, modified as necessary, turns out to leave us in the same difficulties with subjacency

as before: if ATB exceptions are allowed, why must the multiple gaps be uniformly subjacent in some cases?

4.4.2 Let us examine briefly how the problems raised in this section might be solved if the Coordinate Structure Constraint derives from the PCC. At first sight, we have a paradox:

We considered first (in 4.3) sentences with only a set of ATB gaps. The relation between each of these gaps and its local \bar{A} -binder must obey subjacency. We proposed that there was a level at which the PCC applied only to those paths that arise from subjacent gaps. At this level, a non-subjacent gap in one conjunct would simply not contribute to any path. For the purposes of the PCC, it would not be there, and such a structure would be ruled out by the PCC like any Coordinate Structure Constraint violation. In other words, an example like (199)b, repeated below, would have the relevant paths indicated:

(216) *a book [_{S₁} that_i [_S John [_{INFL'} [_{VP₀} [_{VP₁} read e_i¹] and [_{VP₂} disapproved of [_{NP₂} the man [_{S₂} who reviewed e_i²]]]]]]]]

Paths (subjacent)

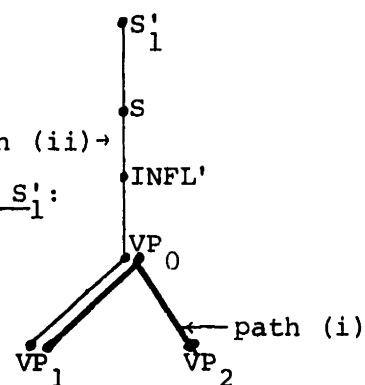
(i) Between the conjunction and conjuncts:

{VP₁, VP₂, VP₀}

path (ii) →

(ii) Between e_i¹ (the subjacent gap) and COMP of S₁:

{VP₁, VP₀, INFL', S, S₁'}



Paths (i) and (ii) overlap, but they violate the PCC because path (ii) is missing the crucial VP_2 node that would be supplied by e_i^2 , if e_i^2 satisfied subjacency.

In 4.4.1, we considered sentences which contained a single, real gap, and conjoined constituents which themselves could contain parasitic gaps. We showed that such parasitic gaps, which could violate subjacency, showed the same Coordinate Structure Constraint effects and ATB exceptions as subjacent gaps. If these phenomena result from the PCC, and the PCC sees only paths from subjacent gaps, how can we explain these same phenomena in parasitic, non-subjacent gaps?

We propose that the PCC applies at two levels. At the first level, the PCC sees only subjacent gaps, and (216) is ruled out as indicated. At this level, the PCC will say nothing about the non-subjacent parasitic gaps that we have discussed. At the second level, the PCC sees the full set of path-creating gaps, subjacent and non-subjacent. At this level, (216) will have paths that satisfy the PCC, since path (ii) will include the crucial VP_2 node contributed by the second subjacent gap:

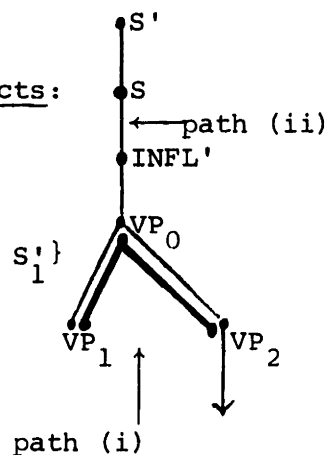
(216) Paths (non-subjacent and subjacent)

(i) Between the conjunction and conjuncts:

{ VP_1 , VP_2 , VP_0 }

(ii) Between e_i^1 , e_i^2 , and COMP of S'_1 :

{ VP_1 , S'_2 , NP_2 , VP_2 , VP_0 , $INFL'$, S , S'_1 }



On the other hand, a sentence like (211)b, repeated as (217), would have the following paths at the two levels (we ignore the path between INFL and COMP):

(217) *John, [_S, who_i [_S [_{NP₀} [_{NP₁} men [_{S₁} who like Mary]]] and [_{NP₂} women [_{S₂} who know e_i¹]]] [_{INFL'} [_{VP} admire e_i²]]] }

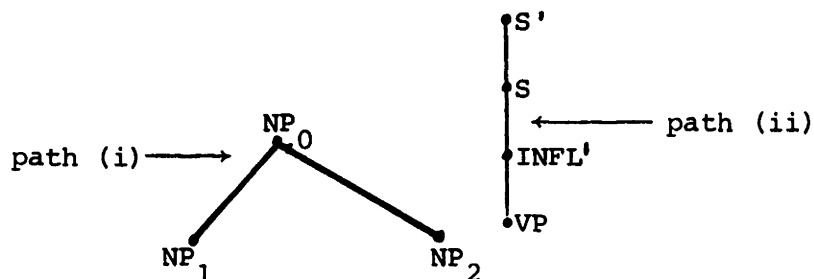
Level I Paths (subjacent):

(i) Between the conjunction and conjuncts:

{NP₁, NP₂, NP₀}

(ii) Between e_i² and COMP of S':

{VP, INFL', S, S'}



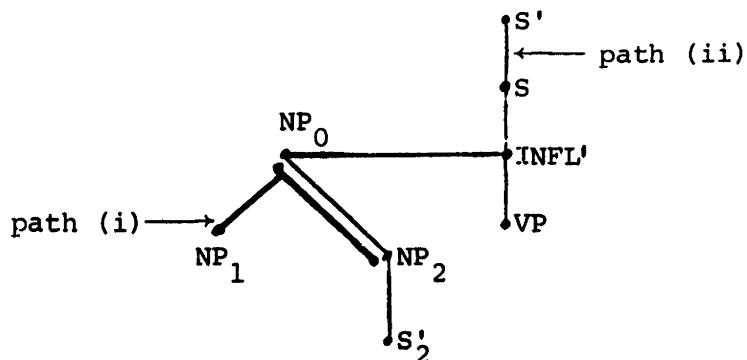
Level II Paths (Non-Subjacent and Subjacent):

(i) Between the conjunction and conjuncts:

{NP₁, NP₂, NP₀}

(ii) Between e_i¹, e_i² and COMP of S':

{VP, INFL', S₂, NP₂, NP₀, S, S'}



At Level I, where only subjacent gaps contribute to paths, path (i) and path (ii) do not overlap. At Level II, where all appropriate gaps count, paths (i) and (ii) overlap, and violate the PCC. Hence (217) is ruled out at level II.

Now recall that we have argued that the PCC must apply at least at LF. Certainly, at LF all gaps (particularly if interpreted as variables) must be visible. Therefore, it is plausible to identify Level II with LF. On the other hand, can we identify Level I with S-structure? Given Chomsky's analysis of parasitic gaps, which we have assumed, we know that all gaps, subjacent and non-subjacent, are visible to the free indexing of A-positions at S-structure. Non-subjacent gaps could not wait until LF to receive an index, because the identification requirement on empty categories applies before LF. (It was this assumption that explained why LF movement and resumptive pronouns do not license parasitic gaps.) The identification requirement, of course, makes no distinction between subjacent and non-subjacent gaps.

It cannot be the case, then, that non-subjacent empty categories are generally invisible at S-structure. We are left, then, with an important unsolved question. Why does the PCC ever apply to only subjacent gaps? This question will not be completely answered in this study, but we will return with some speculations in 4.5.

4.4.3 Let us review our discussion so far in section 4. We noted first that there were cases in which ATB gaps all had to obey subjacency. We suggested that this implied the existence of a level at which Coordinate

Structure Constraint effects, derived from the PCC, only applied to subjacent gaps. We then examined a counteranalysis of ATB phenomena, based on a proposal of Chomsky and Lasnik, which, at first sight, seemed to explain the subjacency properties of ATB constructions. This counteranalysis seemed to fail for principled reasons in accounting for ATB phenomena in non-subjacent gaps. On the other hand, we have claimed that the PCC can handle these phenomena if it reapplies at a level at which both non-subjacent and subjacent gaps were considered.

Here too, however, there is a counteranalysis of ATB phenomena, that of Williams (1978), which could be suitably extended to handle ATB facts that arise with all kinds of gaps. We will present the main features of Williams' analysis, which does not take parasitic gaps into account. We will then propose a simple extension of his analysis to handle parasitic gaps and show how this extension fails, again for principled reasons. Since Williams himself did not consider parasitic gap constructions, the proposal we will consider is a straw man, and not Williams' own. Nonetheless, following the logic whereby the best theory of multiple gap constructions is one that says nothing special about them whatsoever, our extension of Williams' proposal appears to be natural and optimal. Its difficulties are of interest for this reason.

Williams proposes a special notation for conjoined structures, in which "conjuncts in a coordinate structure [are] written on top of each other, and ... factor lines that split coordinate structures [are] drawn so as to split all conjuncts of that structure." The factorization Williams has in mind is the factorization imposed by movement. Thus,

the phrase (218) would result from the pre-WH-movement factorization like (218)b, where the lines that split the conjuncts represent the simultaneous factorization of the conjuncts:

(218)a. Mary, who [John saw e] and [Bill hit e]

b. COMP | [[S John saw | who] | and]
 | [[S Bill hit | who] |]
 | S

1 2 3 4 (Williams (1978), (6))

The terms affected by WH-movement are 1 and 3. Williams formalizes variants of the definitions of "well-formed labelled bracketing" and "factor" that allow for structures like (218)b. Quite crucially, Williams requires that "if one conjunct is split by a factor line, all must be split."

To capture the subject/object asymmetries discussed in 3.2.3 of this chapter, which Williams discovered, he stipulates that "if the conjuncts are split [by a factor line], then the left conjunct brackets must all belong to the same factor". We will not repeat Williams' demonstration of this condition here; we simply recall that these asymmetries followed directly from the PCC applying to the path between INFL and COMP. In fact, given the existence of such a path, the facts could not be otherwise.

Important for our purposes are Williams' extensions of the principle of Recoverability of Deletions and of the relation "is a" in the structural description of transformations (= Williams' (31), (33)):

(219)a. Recoverability of Deletions (ROD)

If T is a term moved or deleted by a transformation, and T consists of simultaneous factors $F_1 \dots F_n$, then it must be the case that $F_1 = \dots = F_n$.

b. "is a"

If F is a factor consisting of simultaneous factors $F_1 \dots F_n$, then F "is a" X if F_1 "is a" X and ... , ... and F_n "is a" X.

(219)a straightforwardly prevents, for example, the simultaneous movement of who and what or of who and Bill. The two constraints, Williams shows, derive the Coordinate Structure Constraint of Ross, plus the ATB exceptions. For example, the sentence in (220)a might be factored, before WH-movement, either as (220)b or as (220)c (= Williams' (34)-(36)):

(220)a. *who did John see e and Bill hit Mary

b. COMP	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 5px;">[_S John saw</td> <td style="padding: 5px;">who]</td> <td rowspan="2" style="padding: 5px;">and</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">[_S Bill hit</td> <td style="padding: 5px;">Mary]</td> </tr> </table>	[_S John saw	who]	and	[_S Bill hit	Mary]
[_S John saw	who]	and				
[_S Bill hit	Mary]					
1	2	3	4			

c. COMP	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 5px;">[_S John saw</td> <td style="padding: 5px;">who]</td> <td rowspan="2" style="padding: 5px;">and</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">[_S Bill hit</td> <td style="padding: 5px;">Mary]</td> </tr> </table>	[_S John saw	who]	and	[_S Bill hit	Mary]
[_S John saw	who]	and				
[_S Bill hit	Mary]					
1	2	3	4			

In neither of the two factorizations is there a term which is subject to WH-movement and which does not violate ROD. In (220)b, term 3 "isn't a" WH-term, since the second simultaneous factor is not. Therefore, movement of who alone cannot apply. (Williams writes before Chomsky's (1976) proposal to subsume WH-movement under Move α .) Movement of both who and Mary violates ROD. Similar considerations block any movement in (220)c, where term 3 contains who and the identity element, which is assumed to be

distinct from who for both "is a" and ROD. These two principles, plus the "ATB format" developed by Williams, thus derive the content of Ross's constraint, plus the ATB exceptions, for gaps derived by movement.

On Williams' account, there is no mystery about why the gaps in sentences like (218)a obey subjacency. They are derived by a simultaneous application of WH-movement in each conjunct, not by coindexing A-positions. It would be surprising if ATB gaps in such sentences did not obey subjacency. Similarly, once Williams' account is extended, it will come as no surprise that ATB gaps do not obey subjacency.

Let us therefore consider how Williams' theory might account for the now-familiar contrast in (221):

- (221)a. *John, who [men who like Mary] and [women who know e] admire e
 b. John, who [men who like e] and [women who know e] admire e

The gaps in subject position of (221)a-b do not arise by movement, since they are non-subjacent. Suppose, however, that we rewrite Williams' ROD condition as a condition on representations with empty categories (of the appropriate sort -- here, \bar{A} -bound):

(222) Recoverability of Deletions - Revised (ROD')

If T is the minimal term containing an empty category E (qualified as in text), and T consists of simultaneous factors $F_1 \dots F_n$, then if E is free in the conjunct containing E, it must be the case that $F_1 = \dots = F_n$.

Notice that the factorizations in (220)b-c both fail the revised ROD', if who is replaced by e bound from COMP. ROD' and "is a" also correctly rule out (221)a, while allowing (221)b:

movement in this extension of Williams' theory, which makes it interesting.

Structures of the following type, however, are incorrectly analyzed by ROD':

(225) I know who [close friends of e admire e] and [Mary hates e]

There are a number of possible factorizations of (225), all of which violate ROD', e.g. (226):

$$(226) \quad \begin{array}{c} [S, \text{who}_i \\ 1 \end{array} \left| \begin{array}{c} \left[\begin{array}{c} [S \text{ [close friends of} \\ [S \text{ [Mary]} \end{array} \right. \\ 2 \end{array} \right. \left| \begin{array}{c} e_i^1 \\ \\ \end{array} \right. \left| \begin{array}{c} [\text{admire} \\ [\text{hates} \\ \end{array} \left| \begin{array}{c} e_i^2 \\ e_i^3 \\ \end{array} \right. \left. \right] \text{ and} \right. \left. \right] \end{array} \quad \left. \right] \end{array}$$

Terms 3 and 5 are the minimal terms containing instances of E, an empty category. In 5, both simultaneous factors contain E, and satisfy ROD'. In 3, only one does, and ROD' is violated. ROD' thus predicts, contrary to fact, that (226) should be ungrammatical. (226) is, however, completely acceptable.

How does (226) differ from (223)-(224), where ROD' made the correct predictions? Remember that we exchanged ROD for ROD' because the ATB phenomenon and the Coordinate Structure Constraint are found with parasitic gaps as well as with gaps created by movement. A number of alternatives to ROD' are easily imaginable. For example, one might take parasitic gaps to result from a pronoun deletion rule, instead of from a free indexing of empty categories. Then Williams' original ROD, which covers deletions, would work correctly in (223)-(224), but would still stumble at (226). The essential difference between (223)-(224)

would hold, it seems, however ROD is reformulated: namely, in (223)-(224) only one term subject to ROD(') is found in the conjoined structure, while in (226) two terms subject to ROD(') are found in the conjoined structure.

In order not to rule out (226), ROD' or its equivalent would have to apply only once per conjoined structure. In (223)-(224), ROD' applies in term 3. In (226), if ROD applies in term 3, the sentence is incorrectly ruled out, but if ROD applies in term 5 and does not reapply, the sentence is correctly allowed. We would thus have to add a rider to ROD':

(227) In a conjoined structure C , for a term T_i the minimal term containing E an empty category, T_i may violate ROD' if C contains another term T_j , $i \neq j$, such that T_j is the minimal term containing another instance of E , and T_j does not violate ROD'.

(227) appears to be completely adequate from a descriptive point of view. On the other hand, it stipulates a fact that the PCC theory explains. ROD and ROD' impose a kind of parallelism requirement on conjuncts, but the parallelism requirement they impose appears to be too strong -- it is a parallelism of factorization. The PCC, on the other hand, imposes a different kind of requirement altogether. Conjunction creates a path from each conjunct. If there is another path caused by a gap free in one conjunct it will inevitably overlap the path created by conjunction. This path must therefore be big enough to contain the conjunction path. This is accomplished only if there is a gap in each conjunct.

Once we have a gap in each conjunct, and a path from these gaps that contains the conjunction path, we can add as many "spurs" to this path as we like, without changing the situation. Thus (225) is grammatical

because (228) is:

(228) I know [_S, who_i [_S₀ [_S₁ Bill [_{VP}₁ admires e_i¹]]] and [_S₂ Mary
 [_{VP}₂ hates e_i²]]]

Paths

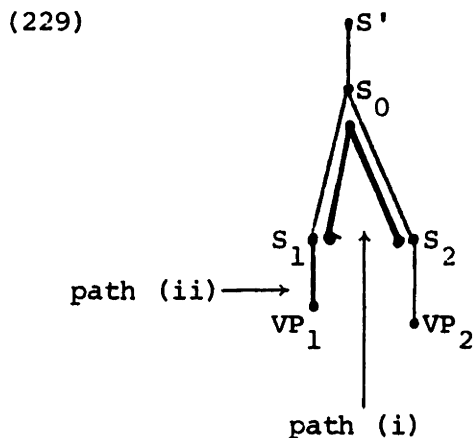
(i) Between conjunction and conjunct:

{S₁, S₂, S₀}

(ii) Between e_i¹, e_i² and COMP:

{VP₁, S₁, VP₂, S₂, S₀, S'}

These paths can be represented schematically as follows:



Path (ii) clearly contains path (i), and the PCC is not violated.

(225) is just like (228), except that path (ii) contains more nodes. As we showed in section 2.5 of this chapter, if one path contains another, adding more nodes to the larger path can't possibly change the containment relation. Thus, consider again (225) and associated paths:

- (225) I know [_{S'} who_i [_{S₀} [_{S₁} [_{NP₁} close friends of e_i³] [_{VP₁} admire e_i¹]]]]
 and [_{S₂} Mary [_{VP₂} hates e_i²]]]]

Paths

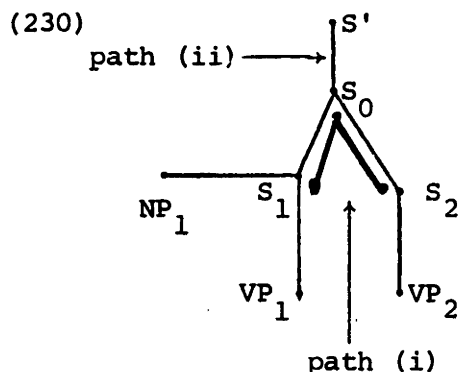
- (i) Between conjunction and conjuncts:

{S₁, S₂, S₀}

- (ii) Between e_i¹, e_i², e_i³ and COMP:

{VP₁, NP₁, S₁, VP₂, S₂, S₀, S'}

Schematically, the paths for (225) are simply the paths for (229) plus a "spur":



Path (ii) still contains path (i); the spur makes no difference.

Thus, although Williams' theory, with the ROD extended as the ROD', appears to solve our problems with subjacency, it runs into trouble in cases where parallelism of factorization is not met, but where the Coordinate Structure Constraint appears not to apply. These cases, on the other hand, show exactly the properties expected of them on a PCC approach; nothing special need be said. We will therefore pursue further the implications of the PCC theory for subjacency, rather than assume

Williams' notation, which seems not to extend to cases of parasitic gaps without problems and stipulations.³⁷

4.5 Move α , Subjacency and the PCC

In the preceding sections, we have linked a rather technical question of analysis to a deeper issue. In this section, we link both to a still deeper issue: the nature of the relationship Move α .

The technical issue we considered was the apparent distinction between parasitic gaps and ATB gaps. We argued that there was no difference in either source or properties between the gaps in the two constructions. Both constructions contain a set of empty categories that share a local binder. These empty categories generate paths of the same sort, subject to the same PCC, in both constructions. The apparent differences arise from two facts:

First, coordinate structures contain a path linking the conjunction and the conjuncts. Once this is assumed, the various phenomena that fall under Ross's Coordinate Structure Constraint (cf. (99)), the ATB exceptions, and the alleged "parallelism requirements" or "subject/object asymmetries" all follow automatically from the PCC.

Second, coordinate structures show apparently special properties with respect to subjacency because the PCC applies once at a level where only subjacent gaps contribute to paths. This set must meet the PCC on its own. We showed that this assumption explained certain facts about parasitic gaps: it is not specific either to ATB gaps or to parasitic gaps.

In the course of discussion, we showed that parasitic gaps were subject to Coordinate Structure Constraint effects and to ATB exceptions

(cf. (214)a-b). Similarly, conjuncts in coordinate structures may contain parasitic, as well as "real" gaps (cf. (225)). In our framework, it makes no sense to ask whether a gap arises as an "ATB gap" or as a "parasitic gap": rather the grammar allows for multiple gap constructions because of free indexing and the optionality of phonetic content, and these gaps have certain properties in certain contexts.

This technical discussion, as we have noted, raised a deeper issue: where does the PCC apply, and why does it differ at the various levels? In the course of discussing Chomsky and Lasnik's and Williams' proposals, we concluded that the PCC must apply at two levels. At one of them, only subjacent gaps contribute to the paths it sees. At the other, all appropriate gaps contribute. It is reasonable to identify the second level with LF, for reasons discussed above, but what is the first level?

This question leads us to the even deeper question. First, does the distinction between two levels indicate that the definition of paths is different at each of them, or is there a constant definition of paths, dependent on differing properties of the levels themselves? A priori, the second view seems most reasonable. We know that subjacency is a property of the relation Move α . We know that the relation Move α holds at S-structure between a binder in a non- θ position and a bound empty category, before free indexing of A-positions. This suggests that the first level at which the PCC applies is this level, and that the definition of paths at this level counts as "empty" only categories which bear the relation Move α to their antecedent. But an interesting problem now arises. In ATB sentences like (231) (= (199)a-b) a non-singleton set of gaps must contribute to a path at the first level at

which the PCC applies; all the gaps must obey subjacency. Following our line of reasoning, each member of this set must bear the relationship Move α to the common binder:

(231)a. a book that John [read e] and [disapproved of e]

b. *a book that John [read e] and [disapproved of the man who wrote e]

Can more than one element stand in the relationship Move α to a single category in a single position? So far we have assumed not. Let us examine the issue more closely.

4.5.1 We have been tacitly integrating our claims about the PCC into Chomsky's analysis of parasitic gap phenomena. Chomsky, in turn, tacitly assumes that Move α cannot simultaneously relate more than one category to a single position -- i.e. that the two gaps in (231)a could not both arise by movement to COMP. Let us call Move α with such a restriction non-multiple Move α . To be more precise about non-multiple Move α , we can employ the notation of Lasnik and Kupin (1977) for representing transformational substitutions. If X/Y represents the substitution of X for Y, we may generally prohibit derivations that contain both X/Y and Z/Y, where X \neq Y (unless, perhaps, the derivation also contains Y/Z, a situation generally excluded by the Strict Cycle). If one gap in (231) is X and the other Y, then a derivation of (231) by Move α would involve X/COMP and Y/COMP, excluded under non-multiple Move α .

If Move α is non-multiple, then we are envisioning an extension of Chomsky's analysis of parasitic gaps (cf. (193)) along the lines of (232). Let us call this Theory A:

(232) Theory A

- (1) Move α (non-multiple): subject to subjacency
- S-structure: (2) free indexing of A-positions
- (3) PCC: all and only subjacent empty categories count
- (4) Identification requirement for empty categories (all)
- (5) LF movement
- LF: (6) PCC: all appropriate empty categories count

We might ask why UG contains each of the steps in (232), but the most pressing question for us concerns (3). Both (1) and (3) are subject to the subjacency constraint. (1) must be subject to subjacency in this model to ensure that at least one empty category in a multiple gap construction obeys subjacency. Recall that limiting (2) to A-positions requires all non-subjacent gaps to be licensed by a subjacent gap derived by (1), yielding the desired result.

If Move α is non-multiple, then (3) must be ordered after (2), to allow ATB gaps to satisfy the PCC. Only through (2) can gaps in more than one conjunct share an \bar{A} -binder and index; it is this multiplicity of gaps that is crucial when conjoined structures face the PCC in (3). On the other hand, the ordering of (3) with respect to (4) or (5) is not crucial, although the ordering of (4) and (5) is, as we saw in 4.2 (cf. Chomsky 1981b).

Under Theory A, since the ordering (1)-(2)-(3) is critical, we cannot say that they all characterize a single level, where processes and conditions obey subjacency. Free indexing is obviously not subject to subjacency. Theory A thus contains at least an inelegance: it is completely unclear why (1) and (3) alone are subject to subjacency. Since it is often profitable to try to eliminate such inelegances, let us

consider another theory.

Let us consider a Theory B, under which Move α is not subject to the restrictions described above. In Theory B, we will allow multiple movement to a single position, subject to some reasonable Recoverability condition. Thus, suppose that a derivation may contain X/Y and Z/Y, $X \neq Z$, where X and Z are identical in the sense relevant for Recoverability. Under this theory, the output of such multiple Move α may include a non-singleton set of empty categories sharing a common binder, as in (131)a, so long as the relation in each case obeys subjacency.

Under Theory B, the nodes that contribute to paths for the first application of the PCC can be easily characterized: they are the nodes that are relevant to Move α , and satisfy the definition of paths. The PCC becomes, as it were, a filter on the output of Move α . We must stipulate, however, that only empty categories related to their binders by Move α count; recall the evidence in the first section of this chapter that we cannot allow each gap in a multiple gap construction to generate its own path. Theory B thus might be represented as in (233):

(233) Theory B

- (1) Move α (multiple): subject to subjacency
- (2) PCC: all and only paths resulting from Move α count (N.B. paths due to Θ -marking, such as the path created by conjunction, count as resulting from Move α (cf. Stowell 1981).)

- S-structure:
- (3) free indexing of A-positions
 - (4) identification requirement for empty categories
 - (5) PCC: all appropriate empty categories generate paths

Theory B reduces the inelegance of Theory A, because it links the

restrictions on the first application of the PCC to Move α . Note that under Theory A, empty categories bound via Move α and empty categories bound by free indexing are indistinguishable when the PCC first applies. On Theory B, if we assume that unique indices are assigned to empty categories that do not acquire indices by movement, empty categories that are not bound by application of Move α will be unbound. The PCC could then limit itself to bound categories at this level. (There is some conflict here with our analysis of Russian QP constructions in the PCC framework; crucial in our discussion of (65) was the assumption that the PCC saw the unbound empty QP at S-structure, but see the discussion of WH-in-situ in the next chapter.)

4.5.2 There is also a subtle empirical difference between the two theories. At first sight, this difference favors Theory B, showing Theory A to be untenable. On the other hand, an appropriate revision of Theory A to accommodate these counterexamples has interesting general implications for the theory of Move α , suggesting a simplification of Chomsky's analysis of parasitic gaps, and a rapprochement between Theory B and our revision of Theory A. Our discussion will be speculative; we consider a number of options, but leave the issue open.

Under Theory B, all empty categories derived by Move α obey subjacency; it does not follow that all empty categories that obey subjacency are derived by Move α . Under Theory B, it is entirely possible that a subjacent gap might owe its index and binder to free indexing and not to Move α , as long as the sentence contains some other empty category from which it gets its index, which is derived by Move α .

A subjacent gap that gets its index from free indexing will not be visible to the first application of the PCC in (233).(2): it will not contribute to any paths.

Theory A, however, makes a different prediction. Under Theory A, all subjacent empty categories should count for the first application of the PCC. Under Theory A, as under theory B, a subjacent gap can get its index by free indexing, as opposed to Move α . On Theory A, however, the PCC has no way of telling the difference: one subjacent gap looks like the next.

To test these different predictions, we must find an example of a multiple gap configuration in which at least two empty categories obey subjacency. The example must have the following further properties: if the PCC is forced to look at all the empty categories, it will rule the sentence out; but if the PCC can ignore all but one of the empty categories, it will allow the sentence. The first case is that predicted by theory A; such a sentence should be ungrammatical. The second is that predicted by theory B: the sentence should be grammatical. Clear, uncontroversial examples are hard to find. In the examples that are constructable, the facts seem to support theory B. Below, we consider two examples:

Although adverbial clauses are, in some cases, inaccessible to movement (cf. Engdahl and Chomsky's discussion of this point), they appear to be subjacent domains. Thus (233)b is grammatical, implying that the second gap in (233)c is subjacent to its binder:

- (233)a. an article that I filed e after leaving the room
 b. an article that I left the room after putting e in the computer
 c. an article that I filed e after putting e in the computer

On the other hand, the second gap in (234)b is clearly not subjacent, as (234)a shows:

- (234)a. *an article that I talked to the man who wrote e
 b. an article that I filed e after talking to the man who wrote e

As predicted by both theories, (235) is ungrammatical, since e² must derive by Move α , in order for e¹ to receive its index by free indexation; e² cannot derive by Move α , as (234)a shows:

- (235) *an article that I [talked to the man who wrote e¹] and [put e²
 into the computer]

Now consider (236), which is clearly better than (235), if somewhat worse than (233)-(234):

- (236) ?an article that I filed e¹ after [talking to the man who wrote e²]
 and [putting e³ in the computer]

(236) distinguishes Theories A and B. Theory A requires that the PCC look, on its first application, at both subjacent gaps: e¹ and e³. The path from e³ would violate the PCC with respect to the path created by conjunction: it would be a Coordinate Structure Constraint violation. In other words, (236) would be as bad as (235), which it is not.

Since we know from (235) that the PCC on its first application does not look at the contribution of non-subjacent gaps like e² in (236), the PCC must also be overlooking e³ -- a subjacent gap -- in order not to rule (236) out. This possibility is expected on Theory B. Under Theory B, we are free to assume that only e¹ is indexed by Move α , and is thus the

only empty category relevant for the first application of the PCC.

The following examples show the same paradigm, where the subjacent status of e^2 is more certain, although the apparent violation of the anti-c-command condition is surprising. (The S'-complement might be moved by Heavy NP Shift, in which case the subjacency issue becomes murky again.)

- (237)a. a man who I can easily persuade e that Mary likes Bill very much
 b. a man who I can easily persuade you that Mary likes e very much
 c. a man who I can easily persuade e that Mary likes e very much
 (cf. (233))
- (238)a. *a man who Mary would kill anyone who insulted
 b. ?a man who I can easily persuade e that Mary would kill anyone
 who insulted e (cf. (234))
- (239) *a man who Mary likes e very much and would kill anyone who
 insulted e (cf. (235))
- (240) ??a man who I can easily persuade e^1 that Mary likes e^2 very much
 and would kill anyone who insulted e^3

Again (240) does not have the status of (239), it seems, although the complexity of the examples makes secure judgments difficult. Assuming the judgments as indicated, it follows that the first application of the PCC must be able to ignore the subjacent empty category e^2 as well as the non-subjacent e^3 . Once again, this possibility arises naturally in Theory B, where we may take only e^1 to bear the relation Move α to its binder, but not in Theory A, where the PCC applies to all subjacent empty categories.

The preceding discussion thus supports multiple Move α and Theory B over a restricted Move α relationship and Theory A. Recall that we listed

three properties that might characterize Move α in section 4.2:

- (1) Move α relates two positions
- (2) The antecedent does not bear an independent Θ -role
- (3) The subadjacency condition is obeyed

By (1), we meant that movement was expressible in Lasnik and Kupin's notation (A/B). We have not abandoned this position here. Nonetheless, non-multiple Move α does not preserve the spirit of (1). By allowing more than one position to be related to a single position in different applications of Move α , we are in effect allowing Move α to be a many-one relation, and not a one-one relation. (1) is thus no longer a crucial defining property, if Theory B is correct.³⁸

Nonetheless, suppose we retain the spirit of property (1), and continue to limit Move α to non-multiple applications (in the sense introduced above). How might we modify Theory A, so that it would correctly distinguish structures like (235) from (236), and (239) from (240)?

Theory A incorrectly rules out (236) and (240), because it has the PCC look at all and only subjacent gaps. An obvious modification of Theory A does not work: eliminating the all. If we allowed the PCC to overlook subjacent gaps freely, we would not be able to rule out (235) and (239). In fact, if the PCC could freely overlook any or all empty categories on its first application, it would have no effect at all. We would lose our ability to explain the one non-trivial difference between ATB constructions and parasitic gap constructions: their different status with respect to subadjacency.

Instead, we must consider what the real difference is between (235)

and (236), and between (239) and (240). In (235) and (239), we find only a set of ATB gaps, and no other relevant empty categories. These ATB gaps must all be subjacent. In (239) and (240), there is a set of ATB gaps and another, subjacent gap (e^1). There is now no subjacency restriction on the ATB gaps.

The difference here is, of course, reminiscent of the basic distinction between "real" gaps and "parasitic" gaps, as noted above. In the familiar cases, a gap may violate subjacency only if it is parasitic. A parasitic (non-subjacent) gap is allowed only if there is also a "real" (subjacent) gap that licences it. In our sentences, a set of ATB gaps can violate subjacency only if this set is licensed by a real (subjacent) gap outside the conjoined structure.

Let us outline first a mechanical solution to the problem. We might stipulate, in a revised Theory A -- call it Theory AA -- that the PCC sees paths to which only subjacent gaps and at least one subjacent gap has contributed. This stipulation would prevent the PCC from ignoring the one subjacent gap in (235) and (239), as desired, while allowing the PCC to ignore the corresponding subjacent gaps in (236) (e^3) and (240) (e^2), where another subjacent gap (e^1) is present. In other words, theory AA would allow the PCC to ignore a subjacent gap if it is possibly parasitic: i.e. if there is another, real gap in the sentence.

This is not enough, however: in order to rule out (235) and (239), the PCC must see not only the path created by the one subjacent gap, but also the path created by the conjunction -- or else there will be no Coordinate Structure Constraint effects. So what Theory AA will have to stipulate about the first application of the PCC is:

(241) Theory AA: Condition on the First Application of the PCC

- a. Let $E = \{e_i^1, \dots, e_i^n\}$ be a set of empty categories locally bound by B_i . At the first application of the PCC, the definition of paths counts as an empty category at least one member of set E which is subjacent to B_i , and no members of set E not subjacent to B_i .
- b. Let L be an element Θ -marked by T . At the first application of the PCC, the definition of paths counts L as an empty category bound by T .

Recall that we claimed that the relationship between a conjunction and its conjuncts was a kind of Θ -marking. Note also that the definition of paths will ensure that only the proper sort of binding is relevant.

Theory AA, like theory A, will have the following structure (cf.

(232)):

Theory AA

- (1) Move α (non-multiple): subject to subjacency
- S-structure: (2) free Indexing of A-positions
- (3) PCC: subject to (241)
- (4) identification requirement for empty categories (all)
- (5) LF movement
- LF: (6) PCC; all appropriate empty categories count

Recall again that (3) must follow (2) because multiple gaps arise only by free indexing, and because the PCC must see multiple gap sets in (3), to avoid ruling out simple ATB structures like:

I wonder who [you saw e] and [Bill liked e]

The fact that both (1) and (3) refer to subjacency was the inelegance that we sought to eliminate by admitting multiple Move α in Theory B. The

inelegance is still there: if anything it is greater, if Theory AA assumes (241).

Notice something rather odd about (241). It requires that there must be a path corresponding to each assignment of a θ -role. It further requires that there must be at least one path between every (\bar{A}) binder and a subjacent bindee. The fact that each θ -role must be assigned in the first place is separately enforced by the θ -criterion (here, at S-structure). The fact that each binder (that lacks an independent θ -role, see below) must actually have one subjacent bindee is separately enforced by the subjacency condition on Move α in tandem with the limitation of free indexing to A-positions. What is odd about (241) is this: it requires the existence of a path when exactly those two relationships exist that independent principles require to exist. In other words:

(1) The θ -criterion requires that every θ -role be assigned to some chain; (241) requires that if a θ -role is assigned, this relationship must contribute to a path at S-structure.

(2) If subjacency is a condition of Move α and if free indexing is limited to A-positions, each binder (of the appropriate sort) must have at least one subjacent bindee; (241) requires that if such a binder has a subjacent bindee, this relationship (at least) must contribute to a path.

There is thus a curious and striking redundancy between (241) -- our patch-up of Theory A --, one clause of the θ -criterion, and what we can all the movement/free indexing account of subjacency phenomena. The obligatoriness of θ -assignment and the obligatoriness of subjacency for one binder-bindee pair are vested in the θ -criterion and in the

movement/free indexing system; exactly these two relations appear again as those that must create paths on the first application of the PCC. This redundancy is also related to the inelegance we first found in Theory A and in Theory AA: subjacency characterizes both path formation and movement, but cannot be reduced to the property of a single level.

Suppose we try to eliminate this redundancy by vesting the obligatoriness of θ -role assignment and the obligatoriness of subjacency in the theory of paths itself. In other words, suppose we eliminate (at least at S-structure) the clause of the θ -criterion that requires each θ -role to be assigned and suppose we also eliminate subjacency as a property of Move α . In the latter case, we can stipulate that whenever there is a set of empty categories sharing a common binder at the level where the PCC first applies, there must be a path between the binder and a subjacent bindee, if the empty category is the sort that generates a path. This captures (241)a, and -- by requiring the path envisioned in (241)a to exist -- explains why each \bar{A} -binder must have at least one subjacent bindee, without restricting free indexing to A-positions. Most importantly, it also suggests removing obedience to subjacency from the list of properties of Move α .

Similarly, let us stipulate that for any θ -marker L (or, more precisely, for any position in the θ -grid of L), there must be a path between L and some position θ -marked by L. Since this path is generable only by θ -marking, the assignment of this θ -role is obligatory. This suggests eliminating from the θ -criterion the obligatory assignment of θ -roles. Since the path created by this assignment must exist, the assignment must exist. The remaining content of the θ -criterion would consist, in effect, of Freidin's (1978) principles of Functional

Uniqueness and Functional Relatedness, which would require every chain to bear one and only one Θ -role.

In other words, we might revise (241)a as (242);

(242) Let $E = \{e_i^1, \dots, e_i^n\}$ be a set of empty categories locally bound by b_i , such that the members of E would count for the general definition of paths. At the first application of the PCC in (3) of (242):

- (i) the definition of paths counts as an empty category no members of E not subjacent to b_i
- (ii) the definition of paths optionally counts as an empty category members of E subjacent to b_i
- (iii) for some $e_i^k \in E$, a path runs between e_i^k and b_i

(242) prevents non-subjacent gaps from contributing to paths, as required, and ensures that there will be at least one subjacent gap, by requiring a path to exist between some gap and its binder. Note that nothing prevents several subjacent gaps from contributing to a path at this level, as in simple ATB structures.

Similarly, we might revise (241)b as (243):

(243) Let T be a Θ -marker. At the first application of the PCC, in (3) of (242), there exists a non-null set of positions L such that a path runs between T and all members of L .

If T is actually a position in a Θ -grid, the only way a path can run between T and some other position is by Θ -marking. Hence, stipulating the existence of such a path makes Θ -marking obligatory. In most cases, set L will be singleton. In the case of conjunction, L will be the set of all conjuncts linked by the conjunction.

Let us now concentrate our attention on (242), and ask whether (242) allows us to derive the subjacency condition in full as a condition on paths. Let us look at the various configurations in which we find bound empty categories and ask (1) whether the empty category counts for the general definition of paths, and (2) whether the binding relation in question obeys subjacency. If (242) derives the subjacency condition, all and only those empty categories which count for the definition of paths should obey subjacency.

The important issue is the following: on a standard theory, subjacency is a defining property of Move α . Another property of Move α is that it is a relation between a position and an antecedent that does not have an independent θ -role. This last is almost a defining property of Move α , except for the case of parasitic gaps. On the standard theory, we expect all non-parasitic empty categories whose antecedent lacks an independent θ -role to obey subjacency. On a theory in which (242) subsumes subjacency, we expect only empty categories that produce paths to be subject to the subjacency condition. The two theories will differ empirically if there are empty categories whose antecedents lack an independent θ -role but which do not produce paths. There are such categories, and we shall discuss the evidence. Let us examine the spectrum of empty categories.

First, consider empty categories whose antecedent does bear an independent θ -role: the pronominal anaphor PRO. If subjacency is a property of Move α , then we predict that the relationship between PRO and its antecedent should not obligatorily obey subjacency, which is true, as we saw. But PRO also does not create a path, for reasons we have

discussed in 3.3. Recall that PRO shows no CTP effects, nor Crossing effects, nor Coordinate Structure constraint effects. Thus, since PRO does not count for the general definition of paths, it does not fall under (242). Thus, (242) makes the same predictions about PRO and subjacency as does linking subjacency to Move α .

Next, consider empty categories whose antecedent does not bear an independent θ -role. The antecedent in such a case may be in an \bar{A} -position or in an A-position. Let us consider first \bar{A} -bound empty categories.

Since \bar{A} -bound empty categories may arise by Move α , we expect them to obey subjacency under the conditions determined by Move α and free indexing of A-positions, if subjacency is a condition on movement. Indeed, all our illustrations of subjacency phenomena have involved \bar{A} -bound empty categories. Since \bar{A} -bound empty categories head chains, they fall under the general definition of empty categories, and hence under (242). Hence, (242) makes the same predictions for \bar{A} -bound empty categories as a theory in which subjacency is a property of Move α .

Consider now empty categories A-bound by an antecedent that does not bear an independent θ -role -- in other words, NP-trace. Here (242) and a Move α theory of subjacency make different claims. A Move α theory predicts that NP-trace is subject to subjacency. Recall that NP-trace, since it does not head a chain, does not fall under the general definition of paths, and hence does not fall under (242). Thus, there should be no subjacency requirement on NP-trace, if (242) subsumes the subjacency condition.

Are NP-traces subject to subjacency? NP-traces are treated as anaphors by the Binding Theory. Since the Binding Theory generally puts

much stronger locality conditions on NP-trace than subjacency could, it is almost impossible to tell. There is one case, discussed by Chomsky (1981a, pp. 58, 306), and credited to G. Longobardi, in which NP-trace has been claimed to show subjacency effects:

Normally, the binding theory requires an anaphor to be bound in the minimal category containing a "SUBJECT" -- a subject or AGR:

- (244)a. *they think he AGR considers pictures of each other to be hanging on the wall
- b. *Mary seems that he AGR considers t to be happy

In (244)a, each other must be bound by they, but they is outside the domain of the S containing the SUBJECTs he and AGR.

On the other hand, under certain conditions, SUBJECTs do not seem to count for lexical anaphor -- for example, when the nearest subject is an expletive coindexed with the clause containing the anaphor:

- (245) they think that it_i AGR_i is certain [_{S₁} that pictures of each other are hanging on the wall]

Ignoring the details of Chomsky's analysis, which involves the notion "accessible SUBJECT", it is clear that the Binding Theory for anaphors is relaxed in such configurations. Despite this relaxation of the Binding Theory, however, NP-trace cannot be bound in a configuration like (246):

- (246) *Mary_j seems [_S that [_S it_i AGR_i is certain [_{S_i} t_j to be happy]]]

Since the Binding Theory cannot rule out (246), Chomsky suggests

that (246) is ruled out by subjacency. If S is a bounding node, then the relation between t and its A-binder Mary crosses two bounding nodes, and violates subjacency. As Chomsky notes, this raises problems for the analogous bad sentences in languages like Italian, where S' may be bounding instead of S.

We, however, have another way to exclude (246). We argued in 3.3 that a [+tense] TNS must take scope over a complete proposition, defined as a constituent such that for all predicates that it contains it also contains all members of the chains they θ -mark. In (246), the [+tense] TNS of the S it is certain takes scope over VP, which contains the predicate be happy. TNS cannot take scope over both members of the chain θ -marked by be happy, (Mary, t). It thus cannot have scope over a complete proposition as required, and (246) is ungrammatical.

On this account, too, various questions arise, concerning, for example, clauses that are [-tense] but contain AGR, like subjunctives.³⁹ Nonetheless, it appears that an account of (246) is available that does not rely on subjacency. Consequently, since in all other cases the Binding Theory would make the effects of subjacency on NP-trace undetectable, we have no solid evidence for or against the proposition that NP-trace is subject to subjacency, and thus no evidence against subjacency as a condition on paths, as in (242).⁴⁰

Therefore, it looks plausible to make subjacency a condition on paths at S-structure, as in (242), and not a condition on Move α . At the same time, this allows us to eliminate the restriction of free indexing to A positions, since (242) will prevent an A position from being indexed with a non-subjacent \bar{A} -position when no licensing subjacent gap is

present. If (243) is correct, we might also wish to derive the obligatoriness of Θ -assignment from paths theory, but we leave this matter to further investigation.

If (242) is correct, we may revise Theory AA to Theory AAA:

(247) Theory AAA

- (1) Move α (non multiple)
- S-structure: (2) free indexing of all positions
- (3) PCC, subject to (242) and (243)
- (4) LF movement
- LF: (5) PCC: all appropriate empty categories count

The defining properties of Move α are now:

- (1) Move α relates two positions (of which the non-antecedent is an empty category)
- (2) The antecedent does not bear an independent Θ -role

I suspect that (247) can be further simplified. Property (1) now has very little content, since relations that satisfy property (2) can also be established by free indexing. The only substantive difference between relations derived by Move α and relations derived by free indexing lies in the fact that the latter type of relation might include an antecedent that does bear an independent Θ -role. The issue of "multiplicity" of movement, that led us to distinguish Theory B from Theory A and its successors, similarly seems to fade, since relations established by Move α may be a subset of those established by free indexing.

These considerations suggest the elimination of the principled distinction between Move α and free indexing, as suggested by Koster (1978b) and others. As often noted, this collapsing requires that the subjacency condition be defined on representations, which we have done

in (242). It remains unclear, if subjacency is subsumed by (242), what to say about rightward movement, which also appears to be subject to Bounding Theory and the subjacency condition. A paths account of rightward movement remains to be given. Other issues arise, concerning, for example, the different status of doubly-filled COMPs for the PCC and the subjacency condition, that also invite caution in assuming (242).

These remarks have been speculative, among other reasons because they take as their starting point a very narrow set of observations concerning subjacency and ATB constructions, where judgments are less than sharp and the analytical issues complicated. I have pursued this discussion for two reasons, one technical and one general. First, it is of technical interest that by subsuming the properties of ATB constructions under the PCC we seem led to choose between a theory like B in which Move α can apply multiply and subjacency is a condition on movement, and a theory like AAA in which Move α cannot apply multiply, and subjacency is a condition on paths.

More generally, ATB constructions may provide one of the rare cases in which one can distinguish a theory containing movement from a theory that replaces movement with an indexing or construal schema. As Chomsky (1981a, 91) notes, there is no a priori advantage in replacing Move α by another system of rules with exactly its properties. Rather, the two proposals must be distinguished empirically, for the effort to be interesting -- an extremely difficult task. We have claimed that in a framework incorporating the PCC a redundancy between an extremely odd condition on path formation and the subjacency condition on movement can be eliminated if subjacency is made to be a condition on paths. We argued

that if this step is taken, we remove from Move α the one property that distinguishes it from the independently necessary free indexing convention, since free indexing and movement can both relate a category to an antecedent lacking an independent θ -role. If this argument is sound, there may be an empirical reason for eliminating movement, since we do not have to add any new subsystem with movement's properties.⁴¹

On the other hand, since this argument disappears if we accept Theory B, and may disappear under some further adjustment of the definition of paths, we consider our argument most tentative. Still, since genuine empirical arguments that address the question of movement's existence are rare, we have cautiously offered this one of our own.

5.0 Conclusion

In this chapter, we have developed the outlines of a PCC approach to multiple gap constructions. We have combined Kayne's (1982) approach to parasitic gaps, in which they form a single path leading to their common binder, with the PCC as developed in the previous chapter. It turned out that almost no auxiliary assumptions were necessary to explain a wide variety of properties of these constructions. In particular, the PCC, like Kayne's own analysis explained the patterns of Subject Condition violations discovered by Kayne -- a phenomenon which we showed to extend to left branch extractions as well as more typical extractions of arguments. Our previous analysis of Russian QPs was reformulated (with some problems), and we discovered new evidence for some of our proposals in Chapter Two. Finally, the PCC and Kayne's "forked path" approach to parasitic gap phenomena correctly predict that parasitic gaps will show Crossing effects

of a very particular sort.

We further extended this analysis of parasitic gaps to coordinate structures. We derived that part of Ross's Coordinate Structure Constraint that does not follow from A/A from the PCC, the general path theory of multiple gap constructions, and the hypothesis that conjunctions assign a sort of θ -role to their conjuncts. The same theory also predicts the existence of ATB exceptions to Ross's constraint. The existence of the path between INFL and COMP was confirmed in a series of arguments that showed this path to be subject to our derivation of Ross's Coordinate Structure Constraint. We derived the apparent "subject/object asymmetries" or "parallelism requirement" discovered by Williams from predicted properties of the path from INFL in conjoined Ss. We further accounted for constraints on coordination of tensed with infinitival clauses, and exceptions to these constraints in languages like Italian.

Along the way, we ran into difficulties with expletive subjects, which led us to formulate a "visibility condition" on the definition of paths. This auxiliary, seemingly ad hoc proposal actually turned out to have interesting consequences for the general theory, allowing us to remove an ad hoc restriction on the definition of paths and perhaps explaining why PRO cannot be expletive in infinitives. These difficulties arose in the course of motivating the existence of the path between INFL and COMP: we argued that the path is formed by TNS, which moves to take scope over a complete proposition.

Finally, we compared our proposal to other accounts of coordinate structure phenomena. In this discussion, we justified our unified treatment of ATB and parasitic multiple gap constructions. We were also led to propose that the PCC applies twice, at S-structure and at LF. A strict

visibility requirement, referring to subjacency, seems to hold on the S-structure application of the PCC. In the final section we speculated about the source of this restriction, following a line of reasoning that suggested that the subjacency condition and part of the Θ -criterion might be reformulated in terms of paths.

In general, this chapter represents an exploration of extensions and consequences of the definition of paths and the PCC introduced in Chapter Three. The PCC places powerful constraints on the links that may obtain between constituents at S-structure and LF. In this chapter, we have used the PCC in an extended analysis of the complete range of multiple gap configurations -- structures whose properties must surely derive from independent principles of UG. Since the PCC seems to explain many of these properties, it is likely that something like the PCC and the theory of paths forms a subsystem of UG. In Chapter Three we demonstrated that the PCC applies to non-multiple gap configurations, and in following sections we will explore certain other places where the PCC seems relevant.

Appendix: Slash Categories and Coordination

In this appendix, we briefly compare the PCC account of ATB phenomena with a recent proposal of Gazdar (1981) (cf. Williams 1981 for a reply defending Williams 1978), and present two arguments against Gazdar's theory. Gazdar suggests that UG does not incorporate a transformational component. Since Gazdar (1981) offers no account of subjacency, it is not clear what content this claim has. Nonetheless, even on the technical level, we may find arguments against Gazdar's approach.

Gazdar suggests that categorial labels contain information about the presence of a gap somewhere in dominated material. This information is encoded by means of a special "slash" notation. Thus, an S, which contains no gaps, is taken to be categorially different from an S/NP -- an S containing an NP gap.⁴²

Gazdar argues that the slash notation allows the Coordinate Structure Constraint to be reduced to a basic principle of coordination. Following Williams (1981), we can call this the Law of the Coordination of Likes. It is well-known that a constituent labeled X can be conjoined only with other constituents labeled X. Now suppose that a constituent containing a gap is, as Gazdar claims, categorially distinct from any constituent that does not contain a gap. For example, S is distinct from S/NP. It follows that a category with a gap (e.g. S/NP) cannot be conjoined with a category without a gap (e.g. S) without violating the Law of the Coordination of Likes. Thus Ross's condition is derived. On the other hand, a constituent with no gap, another S, and a constituent with gap, like S/NP, can be conjoined with another constituent of the same type with the same gap, another S/NP. This latter case is constituted by the ATB exceptions to Ross's constraint. Thus:

- (1)a. ?the man who [_S Mary loves him] and [_S Sally hates him] (S & S)
 b. *the man who [_S Mary loves him] and [_{S/NP} Sally hates e] (S & S/NP)
 c. the man who [_{S/NP} Mary loves e] and [_{S/NP} Sally hates e] (S/NP & S/NP)

As Chomsky (1981b) notes, the slash notation can be objected to a priori. It allows a rather wild proliferation of categorial labels, without explaining the basic properties of Move α . We might also add that slash categories do not behave like other categories. No verb subcategorizes for NP/NP or for PP/NP or for any other slash category, as we would expect, if they existed as categories.

We offer two arguments against the slash category account of ATB phenomena. First, we can show that Gazdar must allow double-slash categories, if he is to account for Coordinate Structure Constraint and ATB effects with double extraction from S, thus further proliferating the inventory of categories. Worse yet, the distinction between a single slash category and a double slash category must be sensitive to the indices on the gaps. Thus, if a category contains two gaps with different indices, they must both be represented in the slash notation, to capture the facts we noted in 3.2.2, in examples (118)a-c, repeated below:

- (2)a. *a book that_i I know who_j to [_{VP/NP} talk to Bill about e_i] and
 [_{VP/NP-NP} persuade e_j to buy e_i] (VP/NP & VP/NP-NP)
 b. *a book that_i I know who_j to [_{VP/NP} talk to e_j about Mary] and
 [_{VP/NP-NP} persuade e_j to buy e_i] (VP/NP & VP/NP-NP)
 c. a book that_i I know who_j to [_{VP/NP-NP} talk to e_j about e_i] and
 [_{VP/NP-NP} persuade e_j to read e_i] (VP/NP-NP & VP/NP-NP)

If a category contains two gaps with the same index, however, the result cannot be a double slash category, or else we would not rule out (3):

- (3) *a book that_i I know who_j to [_{VP/NP} talk to the author of e_i about e_i] and [_{VP/NP-NP} persuade e_j to buy e_i] (VP/NP & VP/NP-NP)

That category labels should distinguish referential indices is troubling, but worse is yet to come. The labels must actually bear these indices, in order to avoid allowing (4):

- (4) *a book that_i I know who_j to [_{VP/NP_j} talk to e_j] and [_{VP/NP_i} buy e_i]

To correctly star (4), VP/NP_j must be distinct from VP/NP_i for the Law of Coordination of Likes. But this law must distinguish indices only on the right side of the slash, or else it will eliminate (5):

- (5) John read [_{NP_i} a book] and [_{NP_j} a novel]

So Gazdar's proposal, to work properly, must introduce double slash categories, with indices on the components of the category, but distinguish these indices only to the right of the slash. Clearly something is being missed by viewing Coordinate Structure Constraint and ATB effects in this way. The apparent simplicity of the slash category solution fades when more complex cases are considered. By contrast, all the examples considered above follow automatically from the PCC.⁴³

A second argument comes from examples discussed by Ross (1967) and by Grosu (1973), which we considered briefly in 3.1 (100). These are cases in which a part of the Coordinate Structure Constraint seems not to hold, where the conjunction and is interpreted as consecutive:

(6)a. the whiskey [_S, which_i [_S I [_{VP₀} [_{VP₁} went to the store] and
 [_{VP₂} bought e_i]]]]

b. the store [_S, that_i [_S I [_{VP₀} [_{VP₁} went to e_i] and [_{VP₂} bought
 whiskey]]]]

Something special needs to be said about these cases on any analysis. Let us see what we could say on the PCC approach.

We argued that ordinary conjunction created a single path between the conjunction and the conjuncts. We proposed that this path arose as the result of the conjunction assigning simultaneously a single θ -role to a number of complements. This proposal seemed to provide motivation for this common path, since we showed elsewhere that θ -marking produces paths. In sentences like (6), however, the conjuncts appear to be of different status: the right-hand conjunct is interpreted as a consequent of the left-hand conjunct. Suppose we claim that in such structures, the conjunction assigns different θ -roles to the two conjuncts. Then it would follow that there would be a separate path from each conjunct to the conjunction, and we would, correctly, not expect the Coordinate Structure Constraint to obtain. Thus, the paths for (6)a would be as in (7):

- (7) (i) Between the conjunction and VP₁:
 {VP₁, VP₀}
- (ii) Between the conjunction and VP₂:
 {VP₂, VP₀}
- (iii) Between e_i and COMP of S':
 {VP₁, VP₀, S, S'}

The only two paths that overlap are (i) and (iii). Since (iii) contains (i), the PCC is satisfied.

We propose then, that consecutive conjunctions differ from others in assigning distinct roles to each of their conjuncts. On the other hand, these conjunctions show all the other properties of normal conjunctions. They require more than one complement. They obey the A/A constraint⁴⁴ (as Grosu shows; cf. (101)). Most importantly, they obey the Law of the Coordination of the Likes, with respect to the traditional, non-slash categories:

(8) *the store that_i I [_{VP} went to e_i] and [_{NP} yesterday's newspaper]

If they obey this law, why isn't (6b) ruled out, on the analysis in (9):

(9) the store that_i I [_{VP/NP} went to e_i] and [_{VP} bought whiskey]

This problem does not seem to have a natural solution in the slash-category framework. The only available stipulation is to prevent the Law of the Coordination of Likes from looking at the material after the slash when the conjunction involved is consecutive, but this is entirely ad hoc.

FOOTNOTES: CHAPTER FOUR

1. N. Chomsky suggests that the contrast might simply result from some constraint along the lines of Kuno's (1973) "Clause Nonfinal Incomplete Constituent Constraint". Such a constraint might, for example, block clause-internal preposition stranding, although it is unclear how it could account for the contrast in (1)a-b (as Taraldsen 1981 notes). There might be justification for such a constraint if, as N. Chomsky suggests, (i) is ungrammatical:

(i) who did you

persuade	[friends of t] [that I liked Mary]
order	
promise	

All speakers I have asked find (i) fully acceptable, and certainly not on a par with sentences like (4)a or (1)a. In any case, if (i) should be ungrammatical, we might account for it by assuming, with Kayne (1981)b, that the indirect and direct objects of verbs like persuade, order, and promise form a "small clause":

(ii) who did you [_{VP} persuade [_? NP S']]

If (i) is ungrammatical, it would be a case of Subject Condition effects in the subject of a small clause, discussed in the previous chapter. As N. Chomsky points out, this analysis is semantically implausible, since the verb appears to select and θ -mark both objects (cf. Chapter Three, fn. 1). Since (i) appears to be acceptable for most speakers, we will assume that the problem does not arise.

There is a problem that may be related to (i) that arises on any

current analysis of parasitic gaps, including ours and Kayne's. Here the contrasts seem clear for all speakers:

- (iii) a. ?John, who [people who believe [friends of e] to be nice] admire e
 b. *John, who [people who believe [e] to be nice] admire e
- (iv) a. ?Mary, who [people who consider [long articles about e] boring] usually ignore e
 b. *Mary, who [people who consider [e] boring] usually ignore e
- (v) a. ?the man that [the woman who persuaded [close friends of e] to leave] later denounced e
 b. *the man that [the woman who persuaded [e] to leave] later denounced e

I know of no account of these contrasts. What is interesting is that the first object of persuade patterns here with the subject of a small clause and S'-deletion infinitival. (Example (v)a-b was pointed out to me by J.-R. Vergnaud.)

2. If VP were the maximal projection of V, then S_2 would immediately dominate the subject (W) and VP_2 (Z), a maximal projection of V_2 (X). Since W precedes Z, Z is in the canonical government configuration for English. Thus S_2 would be a g-projection of VP_2 , even if it were not a projection of V_2 .

If S' is a projection of COMP, then S will have to be a projection of V to allow S' to be a g-projection of S, unless INFL is a structural governor. But INFL cannot be a structural governor in Kayne's framework, as we discuss in connection with (18) and (19) below.

3. Kayne formalizes these requirements as follows:

(i) A g -projection set G_β of a category β , where γ governs β :

(a) $\forall \pi, \pi =$ a g -projection of $\gamma \rightarrow \pi \in G_\beta$

(b) $\beta \in G_\beta$ and (b') δ dominates β and δ does not dominate

$\gamma \rightarrow \delta \in G_\beta$

(ii) KECP (revised)

Let $\beta_1 \dots \beta_j, \beta_{j+1} \dots \beta_n$ be a maximal set of empty categories in a tree T such that \exists a c -commanding α , $\forall j, \beta_j$ is locally bound by α . Then

(a) $\bigcup_{1-j-n} G_{\beta_j}$ must constitute a subtree of T

and (b) there must exist a ρ such that $\rho \in \bigcup_{1-j-n} G_{\beta_j}$
and ρ dominates α .

(Kayne's (22)-(23).) Kayne remarks that (a) is intended to read "the union of the g -projection set of all the empty categories in question must constitute a subtree".

4. Kayne apparently assumes that the antecedent is the head of the relative in his diagrams. This issue is irrelevant here, since the same effects can be duplicated with questions.

5. Chomsky (1981b, (79)) shows that there are acceptable sentences that violate this generalization:

(i) a man who [everyone who meets e] knows someone who likes e]

As Chomsky notes, however, sentences like (i) may be acceptable by virtue of meeting some "strong parallelism requirement". Examples without parallelism, like (26)b, seem as bad as any subadjacency

violation.

6. The contrast in (28) holds regardless of any intervening specificity or definiteness conditions on variables in NP, as can be seen by replacing the first chapter of with a chapter of, or by replacing which chapter of with what chapter of.

7. Kayne (1982, fn. 27) briefly discusses examples of the sort we have considered here, and suggests that some speakers might accept examples like our (25), which we consider a subjacency violation. His suggestion, which would extend to (28)b, is to claim that the higher verb might govern into the COMP of the embedded S', thus beginning the g-projections with the higher VP and bypassing the left-branch effect. We consider this move dubious, even within Kayne's own approach, since it opens the door to wholesale violations of the Subject Condition, particularly in small clauses, where the higher V should be able to govern into the subject of the embedded clause, and also in the subject position of a tensed S, where a binder in COMP should be able to govern inside the subject position by coindexation.

8. (30)b probably actually has a bi-clausal structure, as the que/qui alternation in (29)a/(30)a indicates (cf. Obenauer 1981a, for discussion of est-ce que). The contrast repeats itself with clear long extraction, but the judgments are much weaker for some speakers (Obenauer 1976; Kayne 1981a, 117-8):

(i) [combien]_i crois-tu que j'ai invité [_{NP}[_{QP}^e]_i d'hommes]
 how many do you think that I've invited of men

(ii)??[combien]_i crois-tu que [_{NP}[_{QP}^e]_i d'invités] viendront
 how many do you think that of guests will come

The corresponding sentences in Russian show a much sharper contrast (see below in text).

9. Kayne (personal communication) indicates that he would rather maintain the analysis of Kayne (1981a), in which the subject/object asymmetries derive from government across NP into the specifier. Thus the contrasts in the text would be exactly parallel in analysis to basic CTP contrasts, with caveats as in Kayne (1981a). Given the KECP of Kayne (1982), however, the contrasts also derive as discussed in the text; the solution of Kayne (1981a) requires abandonment of the definition of government we are assuming (from Aoun and Sportiche 1982). Thus the KECP seems committed either to a different definition of government, in which maximal projections may be transparent, or to maintaining the asymmetry in the KECP as is. As indicated in the text, our argument is in any case less than overwhelming, but worth presenting because it bears directly on the issue of the KECP vs. the PCC.

10. Barry Schein (personal communication) reports some contrasts that are relevant here:

- (i) ??more women are sexy than [e men] are ___
- (ii) more women are buxom than [e men] are bearded

He suggests that (ii) might be acceptable because the predicate bearded can be focused. Treating focus as a movement rule at LF, we might suppose that the comparative operator undergoes absorption (Higginbotham and May 1981; cf. Aoun, Hornstein and Sportiche 1981) with the focus

operator, creating a binary quantifier:

(iii) ... [_{COMP} _{O_i}-FOCUS(_{bearded})_j] [_{e_i} men] are _{e_j}

If _{e_i} can act like a parasitic gap to _{e_j}, so that they form a united path, then _{e_j} can help _{e_i} escape the PCC in the familiar way. This recalls Kayne's discussion of WH-in-situ, where "ECP effects" can be averted if the offending WH undergoes absorption with a lower WH (we return to these cases in Chapter Five):

(iv) *I wonder what book who bought

(v) I wonder what book who bought for whom

In (i), the VP deletion removes the possibility of focus, accounting for the reduced acceptability.

11. Actually the ergative-passive vs. transitive contrast discussed in Chapter Two is very weak with stol'ko and skol'ko phrases. The effects are already weaker with numeral phrases, as noted in Chapter Two, and explained below. The extra weakness with stol'ko and skol'ko is unexplained.

12. The particle ze is optional, but preferred in these clauses. It adds the notion of "preciseness" to stol'ko: 'exactly as many...'. It is irrelevant to our points here.

13. See Chapter Five, for some discussion of personne.

14. On the other hand, difficulties we will shortly encounter do suggest that negation plays some antecedent-like role.

(ii) (*) John, who_j I wonder what books_i I can talk about e_i with e_j,
 [in order to impress e_j.]

Ignoring the purpose clause, the path between e_j and its binder is longer than the path between e_i and its binder. By the logic discussed in the text, only the longer path should permit the addition of a "spur" in the purpose clause; hence, only (i) should be acceptable. I can persuade myself that there is some contrast between (i) and (ii), but it is certainly not sharp.

18. A problem arises if we allow non-maximal projections to be conjoined:

(i) I consider [_{A₀}* [_{A₁}* John [_{AP} angry at Mary]] and [_{A₂}* Bill [_{AP} sad]]]

The Coordinate Structure Constraint appears to hold when the subject of a small clause (or S'-deletion infinitival) is extracted, which is not predicted by our formulations. Note that the first maximal projection dominating both the conjunction and e_i is VP, and that the first maximal projections dominating the head of the predicate phrases are the APs:

(ii) *John [_S, who_i I [_{VP} consider [_{A₀}* [_{A_i}* e_i [_{AP₁} angry at Mary]] and
 [_{A₂}* Bill [_{AP₂} sad]]]]]

Paths (a) Between the conjunction and conjuncts:

{AP₁, A₁*, AP₂, A₂*, A₀*, VP}

(b) Between e_i and COMP of S':

{VP, S'}

Paths (a) and (b) do not overlap, and the PCC cannot rule (ii) out.

Note that the PCC properly rules out (iii), where the Coordinate Structure Constraint is violated in the predicate phrase AP:

(iii)*John [_S who_i I [_{VP} consider [<sub>A₀^{*} [<sub>A₁^{*} Sue [_{AP} angry at e_i]]] and
 [_{A₂^{*} Bill [_{AP₂} sad]]]]]}</sub></sub>

Paths (a) Between the conjunction and conjuncts:

{AP₁^{*}, A₁^{*}, AP₂^{*}, A₂^{*}, A₀^{*}, VP}

(b) Between e_i and COMP of S':

{AP₁^{*}, A₁^{*}, A₀^{*}, VP, S'}

Paths (a) and (b) overlap and violate the PCC.

I can offer two suggestions for dealing with (ii). First, if we adopt the convention that Xⁿ is always maximal in the environment when dominated by Xⁿ, it will follow that A₁^{*} and A₂^{*} are maximal in (i). Paths (a) and (b) will then be:

(ii) (a) Between the conjunction and conjuncts:

{A₁^{*}, A₂^{*}, A₀^{*}, VP}

(b) Between e_i and COMP of S':

{A₁^{*}, VP, S'}

Paths (a) and (b) now overlap and violate the PCC, as desired. The problem with this solution is that it seems to prevent the V consider from governing and assigning Case to the subjects of the small clauses in (i).

Alternatively, we might suppose that there is no conjunction of non-maximal projections. Rather, (i) derives from (iv) below, with deletion of the second consider under identity:

(iv) I [_{VP₀} [_{VP₁} consider [_{A*} John angry at Mary]] and [_{VP₂} consider
 [_{A*} Bill sad]]] |
 ∅

The ungrammaticality of (ii) will now result from the (straightforward) PCC violation in (v), with deletion of the second consider:

(v)*John [_{S, who_i} I [_{VP₀} [_{VP₁} consider [_{A*₁} e_i angry at Mary]] and
 [_{VP₂} consider [_{A*₂} Bill sad]]]] |
 ∅

In 4.4 of this chapter we give arguments against such deletions under identity, but our arguments actually concern deletion of \bar{A} -binders. If deletions of V are allowed, and certain other deletions, we can account for (ii) by prohibiting conjunction of non-maximal projections (with Chomsky 1965, 212). This prohibition might be supported by the (slight?) unacceptability of sentences like (vi), where an analysis as VP conjunction is implausible:

(vi)a.??I consider neither John angry nor Bill sad

b.??I believe both Bill to be a fool and Mary to be quite smart

c.??John proved either Mary guilty or Peter innocent

The problem with (vi) cannot be a violation of the adjacency condition on Case marking, because (vii) is fine:

(vii)a. I considered neither John nor Bill

b. I believe both Bill and Mary

c. John proved either the theorem or the lemma

19. Non-trivial modifications of the θ -criterion are necessary, if (122)c-d can have the structures assigned to them in our examples. In (122)c, the argument John seems to bear two θ -roles, one assigned to it by the VP told some story, and one assigned to it in the chain (John, e) by the verb registered. I believe that the solution has two parts:

First, we must define the grammatical relation subject in terms of sisterhood to a predicate phrase (as argued in section 1 of Chapter Three), rather than in terms of daughterhood to a clause (as in Aspects). We must also modify the definition of sister so that in the configuration:

(i) [A [_{B₀} B₁ and B₂...]]

A is a sister both of B₁ and B₂. In other words, B₀ is overlooked.

Thus, in the structure:

(ii) [_S NP [_{VP₀} VP₁ and VP₂]]

The position [NP,S] is the value for the functions [subject of VP₁] and [subject of VP₂].

We then assume that θ -roles are assigned to grammatical functions (GFs) rather than to positions directly. Chains will be sequences of GFs, much as in Chomsky (1981a). Crucially, we will relativize the θ -criterion to GFs. Let us say that an argument "fills" a GF α if it occupies the position associated with α . We assume further the convention that a chain C bears θ -role R if and only if one of its members is assigned R. We state a Relativized θ -criterion in (iii):

(iii) Relativized θ -criterion

- a. Every θ -role is assigned to exactly one chain.
- b. Every chain is assigned at most one θ -role.
- c. For α a GF in chain C, α is assigned a θ -role if and only if C contains exactly one GF filled by an argument.

The crucial feature of (iii) is that it allows an argument to be a member of more than one chain, under narrowly restricted conditions.

Consider (iv):

(iv) John [_{VP₀} [_{VP₁} ate meat] and [_{VP₂} drank beer]]

Example (iv) contains two chains:

- (a) ([subject of VP₁])
- (b) ([subject of VP₂])

Both VP₁ and VP₂ assign θ -roles. VP₁ assigns it to the sole member of chain (a); VP₂, to the sole member of chain (b). The θ -roles may be assigned under clause (c) of the Relativized θ -criterion, because each chain contains a GF filled by an argument. Each chain bears only one θ -role, satisfying clause (b). Note that the argument John fills a GF in both chains, which is allowed under the Relativized θ -criterion.

Consider now (122)c, repeated below:

(v) John_i [_{VP₀} either [_{VP₁} was [_{V₁} registered] e_i by the police] or
 [_{VP₂} invented some story]]

Example (v) contains two chains:

- (a) ([subject of VP₁], [object of V₁])
- (b) ([subject of VP₂])

The verb registered assigns its θ -role to [object of V_1], and hence to chain (a). This is allowed under clause (c) of the Relativized θ -criterion, because chain (a) contains a GF filled by the argument John. The VP invented some story indirectly θ -marks the [subject of VP_2], and thus chain (b). The GF [subject of VP_2] may receive this θ -role under clause (c), because it contains the argument John. Again, it is irrelevant that John fills GFs in two chains.

Notice also that (v) calls for a revision of terminology. The position occupied by John is a " θ -position" as the subject of VP_2 , but is a "non- θ -position" as the subject of VP_1 . Clearly, what is relevant is the notion " θ -GF" (in fact, as in Chomsky 1981a), and not the notion " θ -position", which appears to have no status.

The Relativized θ -criterion continues to capture the other effects of the θ -criterion. Consider, for example, the typical θ -criterion violation:

(vi) * $John_i$ [VP [V praised] e_i]

Because the subject locally A-binds the object, (vi) contains the chain:

([subject of VP], [object of V])

Both GFs in this chain are assigned θ -roles. The chain thus violates (b) of the revised θ -criterion. Failure to assign the θ -role of praise would violate clause (a).

Consider also (vii):

(vii) * $[_S$ it [VP_0 [VP_1 was claimed that Mary lied] and [VP_2 shocked John]]

Example (vii) is ungrammatical with the interpretation indicated by our

brackets. The ungrammaticality derives from the θ -criterion, since it must be both an argument and a non-argument:

it fills two GFs: [subject of VP_1] and [subject of VP_2]. VP_2 must assign its θ -role to [subject of VP_2]. Suppose it is an argument. Then VP_2 can assign its θ -role to it. Suppose it is not an argument. Then VP_2 cannot assign its θ -role, by clause (c) of the Relativized θ -criterion. But it must assign its θ -role, by clause (a); hence it must be an argument.

On the other hand, if it is an argument, then the chain ([subject of VP_1]) must contain a GF assigned a θ -role. But VP_1 does not assign a θ -role to [subject of VP_1]; if VP_2 assigned its θ -role, then that θ -role would be assigned to two chains, and violate clause (a). Hence it cannot be an argument. Since it cannot be an argument and a non-argument, the structure is ungrammatical.

This discussion has two hidden assumptions. First, that expletive it does not form a chain with an extraposed S' . This is argued independently by Safir (1982). Second, note that a chain does not have to bear a θ -role -- though, if it does, the chain must contain an argument and the θ -role must be unique. This allows for chains headed by an expletive.

This "relativization" of the θ -criterion to GFs has other consequences, with respect to small clauses and other constructions, which lie outside of this study (cf. Schein 1982 for discussion of a similar proposal). We merely wish to show that structures like (122)c-d can be made to fall under a natural modification of the θ -criterion, to reinforce our claim that A-bound empty categories are not subject to Structure Constraint, any more than they show Crossing

20. This suggests that what moves to COMP is the feature [+tense] and not [\pm past], as pointed out to me by J.-R. Vergnaud, since the element in COMP binds all the "variables" in the various INFLs, which may be variously [+past] and [-past]. I leave the semantic questions raised by this problem open.

21. Empty categories must be only optionally "visible" when they don't fall under (155); see 4.4 of this chapter for more discussion. The notion "visible" is somewhat imprecise: what we really mean is to append (155) to the definition of paths: "Where \underline{t} is an empty category satisfying (155), there is a path such that... and where \underline{t} is an empty category that does not satisfy (155), there may be a path such that ...". It is clear that simpler formulations should be available, and our discussion in this section is largely speculative.

22. Or as in (2). The question being considered here is logically independent on the question of infinite paths.

23. Our theory makes another prediction, which seems to be fulfilled, although the effects are disappointingly weak. If tensed Ss containing expletive subjects are like infinitives with respect to conjunction, as we claim, we predict that they should not be conjoinable with tensed Ss that have thematic subjects. Thus we expect the contrast indicated in parentheses:

(i) non so [_Sche libri [_{S₀} [_{S₁} Maria legge] o [_{S₂} Giovanni compra]]

(=(166)a)

(ii) (*) non so [_Sche libri [_{S₀} [_{S₁} e necessario leggere] o [_{S₂} Giovanni compra]]

Some contrast exists, but it can apparently be overridden by providing a suitable context, more plausible examples, etc.

24. Note that the ne cliticization from the trace of quanti shows that its trace is indeed postverbal -- in fact, in object position (cf. Belletti and Rizzi 1981 for a demonstration).

25. M.-R. Manzini suggests that (178)b is excluded because AGR is coindexed with the postverbal subject. Being an \bar{A} -binder, AGR will not violate principle C by being coindexed with the postverbal subject. Suppose that AGR, like a clitic in the system of Chomsky (1981)b, bears a θ -role and enters into chains. AGR and the postverbal subject will thus form a chain. AGR also c-commands and binds the parasitic gap, and will also form a chain with it. Since the parasitic gap bears an independent θ -role, the θ -criterion will be violated. The case is analogous to the non-licensing of a parasitic gap by a clitic, discussed by Chomsky. This approach is consistent with the analysis of pro-drop phenomena in Safir (1982), as far as I can tell.

26. The paradigm of (177)-(178) shows up also with inverted subjects in the French il construction, arguing that the postverbal subject is coindexed with the subject position (or with AGR; cf. note 25), contra Pollock (1981):

(i)a. quel article_i as-tu classé e_i sans le lire?
 what article did you file without reading it

b. ?quel article_i as-tu classé e_i sans lire e_i?
 what article did you file without reading

(ii)a. ?quel article_i a-t-il été classé e_i sans le lire?
 what article was there filed without reading it

b.*quel article_i a-t-il été classé e_i sans lire e_i?
 what article was there filed without reading

27. We will treat the resumptive pronoun strategy in English as fully grammatical, as it is for many speakers, in colloquial speech. Readers who object may add one question mark's worth (or more) of unacceptability to our diacritics. The contrasts we discuss remain clear.

28. We simplify Chomsky's analysis on some points and elaborate on others, to bring out the issues that are important for our discussion.

29. Chomsky (1981a,b) also considers the ECP to be a telling property of Move α , since traces of movement must be properly governed, while PRO need not (and may not). Recall that we have not dealt with the ECP for NP-trace.

30. Other complications might arise when the gap does not bear a θ -role. Many of these cases are independently ruled out by the Binding Theory. See also Bouchard (1982) for some discussion of these issues from a different point of view.

31. Taraldsen's solution to the "licensing problem" is to say that WH-phrases cannot be in COMP at D-structure, and must move there by Move α . It is unclear how this solution can allow for non-subjacent resumptive pronouns.

32. I. Haik suggests that a definiteness-clash might be at stake here, since interrogative words are indefinite (possibly excepting which, as N. Chomsky points out), but resumptive pronouns like it are definite. She notes the analogy with clitic-doubling, which is often sensitive to indefiniteness.

33. On exceptions to this generalization noted by Chomsky, cf. note 5.

34. Non-subjacent empty categories have another property in common with resumptive pronouns. Coordinate Structure Constraint violations in which one or more would-be ATB gaps is replaced with a resumptive pronoun are better when the resumptive pronouns are right-peripheral.

In other words, the following contrast seems to hold:

(i) *a man who [John considers him foolish] and [Mary considers e smart]

(ii)?*a man who [John considers e foolish] and [Mary considers him smart]

Subjacency violations seem to act like resumptive pronouns in this respect:

(iii)*a man who [John considers the woman who likes e to be foolish] and
[Mary considers e to be smart]

(iv)?*a man who [John considers e to be foolish] and [Mary considers the
woman who likes e to be smart]

This contrast recalls various "crossover" contrasts discussed by Engdahl:

(v)a.*which student_i did your attempt to talk to him_i scare e_i to death

b. which student_i did your attempt to talk to e_i scare e_i to death

(judgments Engdahl's.)

If the analogy is correct (and, as Chomsky notes, there is no adequate account of Engdahl's contrasts), then our approach to subjacency in coordinate structure may need revision. If we regard (ii) and (iv) as grammatical, and (i) and (iii) as out by some crossover condition, then we must allow resumptive pronouns to create paths, creating problems for our treatment of (183)b.

A relevant question is whether the contrast between (i)/(ii) and

(ii)/(iv) really relates to the question of resumptive pronouns, or to a more general left/right asymmetry in Coordinate Structure Constraint effects. At issue is whether there is any contrast between (vi) and (vii) comparable to (i)/(ii) and (iii)/(iv), following the lines indicated in parentheses:

(vi) (*)a man who [John considers Bill foolish] and [Mary considers e smart]

(vii) (?*)a man who [John considers e foolish] and [Mary considers Bill smart]

My feeling is that there is, but the judgments throughout are shrouded in murk. In any case, the general parallelism of non-subjacent gaps and resumptive pronouns (and lexical NPs?) is interesting, in view of our discussion in text. I am grateful to Barry Schein for discussion of these questions.

35. Sjoberg (1980) and George (1980) adopt Chomsky and Lasnik's proposal, but do not deal with the problems raised in our discussion.

36. This argument is based on the assumption that the deletions envisioned must take place in PF, on the left side of the grammar. One might suppose an analysis under which deletion applied at S-structure, followed by some sort of restructuring or pruning so that (215) had the LF structure of (214)b. Since the Coordinate Structure Constraint and ATB phenomena apply to parasitic gaps as well as gaps deriving from movement, however, this analysis would only lead us back to the conclusion that ATB phenomena exist at LF, rendering the deletion rule almost otiose.

37. A number of other cases have figured in the recent discussion of ATB phenomena initiated by Williams (Williams 1978; Gazdar 1981; Williams 1982), which we may briefly list and consider.

Example (i) is ruled out by specific features of Williams' ATB conventions, but not by the PCC, in any straightforward way.

(i) *John, [who] and [whose friends] I saw e

Gazdar (crediting G. Pullum) notes that interrogatives corresponding to (i) are much better:

(ii) I wonder who and whose friends he handed over to the FBI

Williams (1982) discusses a possible derivation for (ii), which involves WH-movement of the S in both conjuncts, followed by Right-Node Raising across the board, which he blocks in (i) for phonological reasons. We omit discussion of this proposal here. We can only note that, if there were no contrast between (i) and (ii), we might relate the badness of (i) to a structure after reconstruction of friends:

(ii) *John, [[who]_i and [who]_j e] I saw [e]_i and [e_j's friends]

Perhaps the empty category in [who e] creates an infinite path, which violates the Coordinate Structure Constraint (PCC). Other than this, we have nothing to offer for these examples, and consider the matter undecided.

Williams (1978) also considers examples like (iii) relevant:

(iii) *John has more cows [_S than [_S₀ [_S₁ Bill has [e₁ dogs]]] or [_S₂ Pete wants to have e₂]]]

As he notes, his approach would require the QP occupying the \underline{e}_1 slot and the NP occupying the \underline{e}_2 slot before the comparative rule (for him, deletion) to be simultaneously factored, in violation of ROD. (iii) presents no problems for us, if both Sub-C and ordinary comparatives involve movement. The empty categories \underline{e}_1 and \underline{e}_2 share a common binder in COMP. \underline{e}_1 expects this binder to be an NP, while \underline{e}_2 expects a QP. This conflict causes the ungrammaticality. Note that when S's are conjoined, so that each empty category has its own binder, sentences corresponding to (iv) become acceptable:

- (iv) John has more cows [S'_0 [S'_1 than Bill has [\underline{e}_1 dogs]] or
 [S'_2 than Pete wants to have \underline{e}_2]

Finally, Williams (1982) suggests that Gapping must refer to his ATB format. We discuss Gapping briefly in Chapter Five, but leave this matter open. On Gazdar's general theory, see the Appendix to this chapter.

38. One argument against allowing multiple Move α is presented in Chomsky (1981b; a less strong version appears in Taraldsen 1981), based on CTP effects. Recall that long WH-movement of the subject of a tensed S is possible only if the subject position can be bound from the nearest COMP:

- (i) a man who_i I said [S , t_i [S t_i came]]
 (ii) *a man who_i I said [S' that [S t_i came]]

(i) is possible because movement is possible "COMP-to-COMP". On the other hand, coindexation is not possible "COMP-to-COMP", particularly

if limited to A-positions at S-structure. Therefore, nothing can save the subject of a tensed S if it is a parasitic gap:

(iii)*the student that_i everyone thinks \underline{e}_i^1 is intelligent because John said (that) \underline{e}_i^2 was intelligent

(iv) ?the student that_i everyone thinks \underline{e}_i^1 is intelligent because John said (that) Mary liked \underline{e}_i^2

(iii) is ruled out by the ECP in Chomsky's framework, or by the PCC for us, since there can be no binder in the COMP nearest to \underline{e}_i^2 . The object gap \underline{e}_i^2 in (iv) has no such requirement to meet; hence (iv) is grammatical. We therefore might conclude that multiple Move α cannot exist, if we are to rule out (iii).

This conclusion might be too strong. First, notice that because clauses of the type seen in (iii)-(iv) are relatively accessible to movement in some cases (as discussed in Chomsky):

(v) ?the student that Susan left because John said (that) Mary liked \underline{e}

Given (v), where \underline{e} must be derived by movement according to all the systems we have been considering, how do we know which of the gaps in (iii)-(iv) are derived by movement, and which are derived by coindexation? If \underline{e} in (v) can be derived by movement, then we expect that \underline{e}_i^2 in (iii) and (iv) can be derived by movement, in which case the explanation for the contrast is lost.

More important, even if a gap in a because clause is clearly derived by movement, the subject/object asymmetry of (iii)-(iv) reappears, for absolutely mysterious reasons. Thus (vi) contrasts clearly with (v), in the same way that (iii) contrasts with (iv):

(vi) *the student that Susan left because John said (that) e was intelligent

Clearly something prevents e in a because-clause from being locally \bar{A} -bound and satisfying the ECP or PCC, independent of the question of parasitic gaps. This may be related to the mysterious prohibition against pied-piping of PPs from such clauses, noted by Chomsky (1981b), who credits A. Belletti:

(vii)a. ?John, who Susan left because Peter spoke to [_{NP} e]
 b. *John, to whom Susan left because Peter spoke [_{PP} e]

Similar things happen when the WH-island condition is violated:

(viii)a. ?the student that Bill knows why I said (that) Mary hates e
 b. *the student that Bill knows why I said (that) e hates Mary

(ix)a. ?a man who Bill knows when to talk to [_{NP} e]
 b. ??a man to whom Bill knows when to talk [_{PP} e]

An ECP account of (ix) is given by Huang (1982). If these sentences form a paradigm, however, the reason for these effects is presently unknown on any approach to the phenomena. In any case, the argument against multiple Move α is weakened.

Also note that both (iii) and (vi) become still worse when that is not deleted.

39. In Chapter Five, we present an analysis of subjunctives in which the higher predicate on which they are (often) dependent plays, in effect, the role of TNS in indicative clauses, solving the problems of sentences like (i):

in a Case-marked position, so it must get Case from some other member of its chain. After free indexing, it can get Case from there, but e_i is not in a Case-marked position. To rule (iii) out, there must not be coindexed with three men when the Case filter applies.

Suppose NP-movement does not obey subjacency, but we assume our convention that [+tense] TNS has scope over a complete proposition. Suppose there moves to its S-structure position from the position of e_i . We will still violate the convention about [+tense] TNS, since the TNS of is certain cannot have scope over the chain (there, e_i , three men) without moving out of its clause. Thus, the Case filter must apply at S-structure before free indexing. This does not exclude the possibility that it also applies at LF, but we have argued in Chapter Two that it does not.

Nothing prevents non-subjacent binding of an empty category by an expletive that is not crucial to Case assignment, as far as I can see:

(vi) it_j seems^a that it is certain e_j to be claimed that Bill is sick

See also Burzio (1981) for other locality requirements on A-chains.

41. Chomsky restricts free indexing to A-positions, for reasons we have seen. Actually, the same result is achievable if free indexing is restricted to θ -positions, since expletive A-positions never licence parasitic gaps, not being operator-bindable. If free indexing is restricted to θ -positions, the redundancy with Move α disappears for NP-

movement (along with our argument in note 40). If free indexing is restricted to θ -positions, we might also account for the impossibility or moderate unacceptability of PP parasitic gaps, discussed in Chomsky (1981b):

- (i) ??a table on which you set a book [_{PP}e] without putting a record [_{PP}e]
 (ii) a table which you set a book on [_{NP}e] without putting a record
 on [_{NP}e]

Questions arise with respect to NP-traces in conjoined structures, discussed in footnote 19.

42. Gazdar seems to assume that NP-trace and PRO do not count as "gaps", in order to permit sentences like those discussed in footnote 19, where a passive and an active VP are conjoined. This is independent from the claim that "there is no passive transformation" which Gazdar explicitly makes. If one wishes to integrate Gazdar's analysis into a quasi-GB framework, one might suppose that only Case-marked traces count as "gaps".

43. Gazdar (note 6), refers to a forthcoming literature on these questions, to which I have had no access.

44. Actually, the A/A constraint needs to be generalized to parasitic gaps, in our theory, to rule out:

- (1) an article that I filed e without reading a book and e

-- given the fact that our account of the Coordinate Structure Constraint does not extend to these cases.

CHAPTER FIVE: FURTHER PCC EFFECTS

1.0 In Chapters Three and Four, we have argued that a subtheory of grammar, "Path Theory", whose basic principle is the PCC, subsumes the "Government Theory" of Chomsky (1981a), whose basic principle is the ECP. We have shown that "Path Theory" can account, not only for the ECP effects, but for the Subject Condition effects and Coordinate Structure Constraint effects. Augmented with a theory of multiple gaps adapted from Kayne (1982), Path Theory also accounts for apparent exceptions to the Subject Condition and the Coordinate Structure Constraint.

In this chapter, we explore some further areas in which the PCC seems to be relevant. In section 2, we consider a cluster of properties characteristic of WH words that do not move to COMP in syntax, some of which have been accounted for by the Superiority Condition of Chomsky (1973). We suggest that the PCC provides a unified account of these effects, given one assumption about the paths created by such "WH-in-situ". We suggest further that the PCC accounts for a very similar group of properties observed in the "Gapping" construction, and suggest a new analysis of Gapping that fits into our explanation. Our remarks about Gapping shall be speculative.

In section 3, we consider a different topic. Working in a framework including the ECP, Piccolo (1982) has presented an argument suggesting a crucial difference between indicative and subjunctive INFL with respect to proper government. We discuss a possible reanalysis of her result in a PCC framework.

In section 4, we consider briefly some possible problems with Path Theory, as we have developed it. In 4.1, we consider some problems raised for Path Theory by May's (1977) rule of QR. In 4.2, we discuss the phenomenon of preposition stranding, which has played a role in discussions of the ECP. In 4.3, we deal with the properties of NP-trace that have been derived from the ECP.

Finally, in section 5, we summarize the results of Path Theory, and discuss some prospects for future research. Among the topics we consider is the definition of government: we show that replacing the ECP for \bar{A} -bound empty categories by the PCC allows a drastic simplification of the definition of government (and of c-command) developed by Chomsky (1981a). In fact, the PCC allows us to adopt in full the conceptually optimal definition proposed by Aoun and Sportiche (1982).

2.0 Infinite Paths: WH-in-situ and Gapping

2.1 WH-in-situ

We adopt the term "WH-in-situ" from Aoun, Hornstein and Sportiche (1982):

- (1) A WH-in-situ is a WH-phrase which has not been subject to WH-movement (in the sense of not having been the target of this rule).

Alternatively (ignoring questions about adverbials like where, when, in situ), we may define WH-in-situ as in (2):

- (2) A WH-in-situ is a WH-phrase in an A-position.

Examples of WH-in-situ are the underlined phrases in (3):

In French, Russian (cf. Sinicyn 1982a), and other dialects of English, matrix WH-questions need not show WH-movement. (8)a, like (8)b, is an acceptable S-structure in French:

- (8)a. $[_S, [_{COMP}] [_S \text{ Jean a vu qui }]]$ 'John saw who?'
 $[_S, [_{COMP} \text{ qui}_i] [_S \text{ Jean a vu } \underline{e}_i]]$ (= Aoun, et al. (12)-(13))

In general, where not required by (5), WH-movement to COMP is optional in syntax.³ As is well-known, there are cases in which it is impossible -- for example, when COMP is already filled. (9)b is not a grammatical S-structure:

- (9)a. $[_S, [_{COMP} \text{ what}_i] [_S \text{ did Mary give } \underline{e}_i \text{ to whom }]]$
 $[_S, [_{COMP} \text{ to whom}_j \text{ what}_i] [_S \text{ did Mary give } \underline{e}_i \underline{e}_j]]$

Early studies of WH-movement (cf. Chomsky 1964) focused on the conditions under which WH-words can and must undergo WH-movement. Since Baker (1970), much attention has been devoted to the properties of WH-words themselves when they are in situ. Much of this discussion has concerned their status at LF, and has touched on the ECP. Since we are deriving the ECP from a more general condition -- the PCC -- we should examine what light the PCC sheds on the properties of WH-in-situ.

First let us consider the LF representation of WH-in-situ. Three types of arguments may be advanced, which tend to suggest to support the following generalization:

- (10) WH may not be in situ at LF⁴

Because of (10), WH-phrases that have not undergone movement at S-structure

must undergo movement to COMP at LF.⁵ Thus, for example, an S-structure like (9)a will have an LF representation somewhat like (9)b. Similarly, an S-structure like (11)a will have an LF representation like (11)b (we complicate this picture somewhat a bit later):

- (11)a. $[_S [_{COMP} \text{who}_i] [_S \underline{e}_i \text{ saw what }]]$
 b. $[_S [_{COMP} \text{what}_j \text{ who}_i] [_S \underline{e}_i \text{ saw } \underline{e}_j]]$

If (10) is true, it is probably an instance of a broader generalization, whose implications are explored by Higginbotham (forthcoming):

- (11) At LF, α occupies an operator position iff α is a logical operator

(10) is a special case of (11) if WH-words are logical operators, and if only \bar{A} -positions are operator positions.

In the case of WH-in-situ, the first argument for (10) that we will consider shows that an S-structure WH-in-situ may be not in situ at LF. The second and third arguments suggest that it must not be in situ at LF.

The first argument derives from Baker (1970). WH-in-situ displays the phenomena of scope. In particular, there are cases where a sentence containing WH-in-situ shows scopal ambiguities with respect to WH-phrases in COMP. Consider the following example from Baker (his (69)-(70)):

- (12) $[_{S_1} \text{who remembers } [_{S_2} \text{where we bought } \underline{\text{which book}}]]$

Sentence (12) has two readings. On one reading, which book appears to be "associated" with where, taking narrow scope with respect to who. On this reading, Baker notes, (12) may be answered as in (13):

(13) John and Martha remember where we bought which book (narrow scope)

On the second reading, which book is "associated" with who, so that a sentence like (14) is a natural answer:

(14) John remembers where we bought the physics book, and
 Martha remembers where we bought The Wizard of Oz, and ...
 (wide scope)

Following Chomsky (1973), we may derive this ambiguity at LF by allowing WH-in-situ to move to COMP in LF, freely choosing whether it ends up in the COMP of S'_2 (in which case (13) is a natural answer) or in the COMP of S'_1 (in which case (14) is a natural answer). The two logical forms that may result are illustrated approximately in (15)a-b:

- (15)a. [S'_1 [$_{COMP}$ who_i $which\ book_j$] e_i remembers [S'_2 [$_{COMP}$ $where_k$]
 we bought e_j e_k]]]
- b. [S'_1 [$_{COMP}$ who_i] e_i remembers [S'_2 [$_{COMP}$ $where_k$ $which\ book_j$]
 we bought e_j e_k]]]

What is the meaning of the notion "associated", which we used above to characterize the relationship between the several WH-phrases in COMP in the logical forms of WH-in-situ sentences? Clearly there is a special link of some sort between them. In the structures of (15), where two WHs appear in COMP, what seems to be involved is the establishment of ordered pairs of members of two sets. The pair of WH-words in COMP binds two variables; whether a particular value for one variable makes an answer true depends on the value chosen for the other variable. In Chomsky's (1973) terms, the interpretation of a

set of WH-phrases in COMP is "uniform in [the] COMP node".

This "uniform interpretation" of multiple WH-words in COMP has been studied in detail by Higginbotham and May (1981). They propose that the WH-operators in such constructions undergo a rule of absorption. This rule takes a set of n operators in COMP and forms from them a single n -ary operator. We omit the details of their rule here, and simplify its properties somewhat.⁶ The crucial point is that a cluster of WH-phrases in COMP behaves as if it were a single operator, which binds a number of different variables.

For reasons that will become clear later, let us represent the operation of absorption syntactically by means of coindexation -- specifically, cosuperscripting. As a first approximation:

$$(16) \quad [{}_{\text{COMP}} \text{WH}_1, \dots, \text{WH}_n] \longrightarrow [{}_{\text{COMP}} \text{WH}_1^i, \dots, \text{WH}_n^i], \text{ for some index } i$$

We may suppose that (16) is obligatory. Finally, as a notational matter, we may interpret the members of the set $\{\text{WH}_1^i, \dots, \text{WH}_n^i\}$ as constituting a single n -ary quantifier, as defined by Higginbotham and May (1981).

One more piece must be added to the analysis of WH-in-situ constructions that we are outlining. Consider a sentence like (17):

$$(17) \quad [S_1' \text{ who}_i [e_i \text{ said } [S_2' \text{ that Mary liked } \underline{\text{what}}]]]]$$

By principle (10), we know that what must move to COMP at LF. We also know that if it moves to the COMP of S_1' it will undergo absorption with who, and an interpretation will be available. We do not, however, as yet prevent what from moving to the COMP of S_2' , where it would yield an interpretation like (18):

(18) who said what Mary liked

We might rule this out by the convention of Chomsky (1973, (249)):

(19) Assign a WH-phrase not in COMP to some higher structure $\begin{bmatrix} \text{COMP} \dots \\ [+WH] \end{bmatrix}$.

By (19), what could move to the lower COMP in (17) only if it were [+WH]. If it were [+WH], on the other hand, filter (5) would force it to contain a WH-phrase at S-structure. Since it does not, the WH-in-situ cannot move there, as desired.

More generally, we might assume a convention like (20), which subsumes Chomsky's convention (19) and filter (5):

(20) COMP contains an interrogative WH-phrases iff it is [+WH]⁷

Assume that (20) applies both at S-structure and at LF. If a verb selects a question and has a clausal complement, its COMP must be [+WH]. If it is [+WH], then it must contain an interrogative WH-phrase at S-structure -- subsuming filter (5) -- and at LF. Suppose we assign what to the lower COMP in (17). By (20), this COMP is [+WH]. But by (20) again, it must contain an interrogative WH-phrase at S-structure. Since it does not, what cannot be moved there at LF -- subsuming principle (19).

Thus, filter (5) and convention (19), subsumed by (20), allow WH-in-situ to move only to a [+WH] COMP. There it will undergo absorption.

The scopal properties of WH-in-situ suggest that it may appear in COMP at LF. The interpretation of WH-in-situ sentences strongly

suggests that it must appear in COMP at LF. A further argument to that effect comes from the phenomenon of "weak crossover":

- (21)a. his_i brother visited John_i
 b. *who_i did his_i brother visit x_i

In (21)a, coreference between his and the name John is possible. In (21)b, however, binding of the pronoun his is impossible. A number of different accounts of this phenomena have been proposed (e.g. Higginbotham 1980; Koopman and Sportiche 1982; Haik 1982). For our purposes, the condition proposed by Chomsky (1976) will suffice:

- (22) A variable cannot be the antecedent of a pronoun to its left

That weak crossover effects really are diagnostic of the presence of a variable can be seen in (23). No weak crossover effects are observed with NP-trace:

- (23)a. John_i seemed to his_i brother [_S e_i to be crazy]
 b. *who_i did it seem to his_i brother [_S, x_i was vrazy]

Chomsky (1976) notes (following Wasow) that quantifiers that are "in situ" at S-structure show weak crossover effects:

- (24) *his_i brother visited someone_i

Chomsky argues that (24) can be explained by principle (22) if at LF (24) shows a quantifier-variable structure. In other words, (24) argues for QR at LF, yielding the LF structure (25), which falls under principle (22):

- (25) [someone_i [_S his_i brother visited x_i]]

Now notice that weak crossover effects appear with WH-in-situ as well:

- (26)a. [_S, which woman_i [_S e_i claims that his_j brother visited John_j]]
 b. *[[_S, which woman_i [_S e_i claims that his_j brother visited which man_j]]]

As before, we can explain the contrast between (26)a-b with principle (22) if (26)b obligatorily shows a quantifier-variable structure at LF:

- (27) [_S, which woman_i^k which man_j^k [_S e_i claims that his_j brother visited x_j]]

The variable x_j cannot be the antecedent of the pronoun his_j, which is to its left, by (22). (27) is thus parallel to (21)b at LF, if we assume LF movement of the WH-in-situ to COMP. Note the cosuperscripting resulting from absorption in COMP.

We take these two arguments to establish the case for LF movement of WH-in-situ. There is a third argument, however, to which we turn our attention in the next section. This argument derives from Kayne (1981e), who argues that the trace of WH-in-situ obeys the ECP at LF.⁸

We will argue that Kayne's analysis, with the ECP replaced by the PCC, is both wrong and right. We will suggest that the effects which Kayne discusses do derive from the PCC, but not at LF. On the other hand, we will present evidence that the trace of WH-in-situ does indeed obey the PCC at LF, but that the relevant cases are not the ones Kayne considered.

2.2 ECP, Superiority, and the PCC

2.2.1 Pure ECP Effects

Kayne (1981e) presents an analysis of subject/object asymmetries like the following (discovered by Hankamer 1974), which argues for WH-movement at LF:

(28)a. [_{S₁'} who_i [_{e_i} said [_{S₂'} (that) [John tripped over what]]]]

b.*[_{S₁'} who_i [_{e_i} said [_{S₂'} (that) [what fell on John's head]]]]

(29)a.?I know perfectly well [_{S₁'} who_i [_{e_i} thinks [_{S₂'} (that) [he's in love with who]]]]

b.*I know perfectly well [_{S₁'} who_i [_{e_i} thinks [_{S₂'} (that) [who is in love with him]]]]

(30)a. I don't know [_{S₁'} who_i [_{e_i} expects [_S who to marry Mary]]]]

b.*I don't know [_{S₁'} who_i [_{e_i} expects [_{S₂'} (that) [who will marry Mary]]]]

(Examples (29)a-b from Kayne (1981e, (12)-(13); (30)a-b from Kuno and Robinson (1972), (2-4).)

Kayne suggests that the ungrammaticality of the (b) sentences above derives from the ECP applying to the variables left after WH-movement of the WH-in-situ in subject position. (28)a-b, for example, have the LF representations in (31), when that is not deleted:

(31)a. [_{S₁'} who_i^k what_j^k [_{e_i} said [_{S₂'} that [John tripped over e_j]]]]

b. [_{S₁'} who_i^k what_j^k [_{e_i} said [_{S₂'} that [e_j fell on John's head]]]]

(31)b is a case of long movement of the subject of a tensed sentence,

and violates the ECP quite straightforwardly. The violation is maintained if we replace the ECP by the PCC. In other words, this analysis implies that the contrast in (31) is essentially the same as the contrast in (32):

- (32)a. [_{S₁'} what_i [_S did Mary say [_{S₂'} that [John tripped over e_j]]]]]
 b.*[_{S₁'} what_i [_S did Mary say [_{S₂'} that [e_j fell on John's head]]]]

A problem now arises with respect to the omission of that. We have seen in Chapters Two and Three that the absence of the complementizer in a sentence like (32)b permits the subject trace to be bound by a trace of WH in COMP, and to escape the ECP or PCC:

- (33) [_{S₁'} what_i [_S did Mary say [_{S₂'} t_i [e_j fell on John's head]]]]

The presence or absence of the complementizer in a WH-in-situ sentence like (30)b, however, has little if any effect on the ungrammaticality, as indicated by the parentheses around that.⁹ This is unexpected, since there is no a priori reason why (29)b, with the complementizer deleted, cannot have the LF structure in (34), which should not violate the ECP or PCC:

- (34) [_{S₁'} who_i^k what_j^k [e_i said [_{S₂'} t_j [e_j fell on John's head]]]]

This problem, of course, admits a variety of technical solutions, given Kayne's analysis. For example, we might assume, with Aoun et al., that WH-movement at LF cannot target a WH in an \bar{A} -position. As a result, WH-movement cannot apply successive cyclically at LF, moving a WH from one COMP to another. Thus, the structure of (33) cannot arise.

Alternatively, we might assume, also with Aoun et al., that a subject empty category is properly governed from COMP (or bound for the PCC) only if the COMP itself bears the index of its contents, by a rule like (35) (Aoun et al. (57)):

$$(35) \quad [{}_{\text{COMP}} X_i'' \dots] \longrightarrow [{}_{\text{COMP}_i} X_i'' \dots] \text{ iff COMP dominates only } i\text{-indexed elements.}$$

Aoun et al. suggest that (35) applies only at S-structure. Consequently, a subject can be governed from COMP only if WH-movement takes place in the syntax, as in (33), and not at LF, as in (28)b. Other solutions to this problem are sketched in Chomsky (1981a, 236 ff.).

We will be using Aoun et al.'s rule as a plausible counterproposal to the analysis we will present in 2.2.4. We wish to emphasize here that if the effects discussed above are accounted for at LF, there is indeed a problem, albeit one that admits a solution like (35). On our later proposal, however, this proposal does not arise; furthermore, we will present some evidence in 2.4 that movement at LF is perfectly free to deposit traces in COMP, which can act like any other trace in COMP to bind (or "properly govern", in ECP terms) the subject. We return to this topic below.

To summarize: The requirement that a WH-in-situ must move to a [+WH] COMP in LF predicts a subject/object asymmetry when the movement is "long". Long movement of an object violates neither the ECP nor our PCC. Long movement of the subject of a tensed S, on the other hand, violates both the ECP and the PCC. Thus, if the theory in this section is right, there is an almost exact analogy between the properties of

long WH-movement in syntax, and the properties of long WH-movement at LF. We might thus have a third argument for WH-movement at LF. We later suggest that the argument is correct, but that the phenomena considered here are not the right data on which to base this argument.

2.2.2 ECP/Superiority Effects

On the analysis just sketched, ECP/PCC effects arise with WH-in-situ when the conventions governing interrogatives force a subject WH-in-situ to "long move" to a COMP in a higher clause. All things being equal, we might expect no problems when the subject WH-in-situ can "short move" to the COMP in its own clause. This prediction is false, however, as is well-known. If anything, the subject/object asymmetry is sharper in such cases:

(36)a. I wonder [_S, where_i [John bought what e_i]]

b. *I wonder [_S, where_i [who bought this record e_i]]

This is unexpected, if the LF representations of these sentences are as in (37):

(37)a. I wonder [_S, where_i^k what_j^k [John bought e_j e_i]]

b. I wonder [_S, where_i^k what_j^k [e_j bought this record e_i]]

Under normal assumptions, e_j is locally bound (and properly governed, under the ECP) by who in COMP in (37)b. The ungrammaticality of such "short movement" is unexplained. (We ignore Aoun et al's rule (35) for the moment.)

We cannot blame the ungrammaticality of (37)b on the fact that the COMP is doubly filled, because of the grammaticality of (38)a, which has the LF representation in (38)b:

- (38)a. I wonder [_S, who_i [e_i bought what]]
 b. I wonder [_S, who_i^k what_j^k [e_i bought e_j]]

If we are to rule (36)b out by means of the ECP or PCC applying at LF, following the thrust of Kayne's analysis of the cases considered earlier, we must find some reason why e_j violates the ECP or PCC in (37)b. Again, a number of different assumptions might solve this problem. Aoun et al. appeal once more to rule (35) applying at S-structure. If a subject is properly bound (or properly governed) from COMP when the COMP itself bears the subject's index, and if COMP can only acquire an index at S-structure, then only WH-movement before S-structure can leave a subject trace properly bound. In other words, the ungrammatical (37)b is distinguished from the grammatical (38)b at LF because COMP binds the subject in the latter, but not the former case:

- (37)b' I wonder [_S, [_{COMP} where_i^k who_j^k] [e_j bought this record e_i]]
 (38)b' I wonder [_S, [_{COMP}_i who_i^k what_j^k] [e_i bought e_j]]

Other solutions are possible. For example, one might refine the definition of c-command so that only the original occupant of COMP can c-command material outside of COMP (and cf. Chomsky 1981a, 279 fn. 2). What is clear is that the ECP or PCC does not account straightforwardly for the subject/object asymmetries that show up even with short WH-movement in LF.

Notice that there is another way to look at all the subject/object asymmetries with WH-in-situ, which indirectly recalls Aoun et al.'s approach. In all our cases, WH-in-situ appears to have the ECP/PCC properties of an empty category bound by a WH in the uppermost clause.

In the cases considered in the last section, this picture is an accurate reflection of the situation at LF. In the cases considered in this section, this picture does not reflect the situation at LF: at LF the position of the WH-in-situ is occupied by an empty category bound within its clause. Nevertheless, it acts like as if it was bound from some higher clause. Looked at in this light, the contrast in (36)a-b is parallel to the contrast in (39), which follows straightforwardly from the ECP/PCC:

- (39)a.?_{S₁} what_j [do you wonder [_{S₂} where_i [John bought e_j e_i]]]]
 b.*_{S₁} who_j [do you wonder [_{S₂} where_i [e_j bought this record e_i]]]]

If all WH-in-situ act as if they were empty categories bound maximally far away, then the correspondence between this property and the real LF representation of the sentences in the previous section might be a red herring. This is what we will argue. In the next section we present specific arguments to this effect. First, however, let us consider another account of the data in this section.

Chomsky (1973) was the first to notice the contrasts in (36). He proposed that they could be explained by a condition on the factorization of transformations, which he called the "Superiority Condition"

(Chomsky's (73)):

(40) Superiority Condition

No rule can involve X, Y in the structure

...X... [_α...Z... -WYV ...]...

where the rule applies ambiguously to Z and Y, and Z is superior to Y.

A is superior to B if A c-commands B (on our definition of c-command) and B does not c-command A.

Let us see how (40) works. Consider the structure in (41):

(41) $[_S, [_{COMP}] [_S \text{ who } [_{VP} \text{ saw what}]]]$

WH-movement applies ambiguously to what and who, which are both WH-phrases. Because who is superior to what, however, the rule cannot apply to what. Thus only (42)b is an appropriate outcome of WH-movement:

(42)a. $*[_S, [_{COMP} \text{ what}_i] [_S \text{ who } [_{VP} \text{ saw } \underline{e}_i]]]$

b. $[_S, [_{COMP} \text{ who}_i] [_S \underline{e}_i [_{VP} \text{ saw what}]]]$

On the other hand, consider (43):

(43) $[_S, [_{COMP}] [_S \text{ you } [_{VP} \text{ give what books [to whom}]]]]]$

Both (44)a and (44) b are possible outcomes of WH-movement:

(44)a. $[_S, [_{COMP} \text{ what books}_i] [_S \text{ did you [give } \underline{e}_i \text{ [to whom}]]]]]$

b. $[_S, [_{COMP} \text{ [to whom]}_i] [_S \text{ did you give what books } \underline{e}_i]]]$

Neither one of the WH-phrases what books or to whom is superior to the other, since each c-commands the other. The Superiority Condition is thus not invoked.

Supposing for the moment that both the Superiority Condition and the ECP or PCC are principles of UG, we have seen that they overlap in their coverage of the facts discussed in this section. The facts originally presented by Kayne, which we discussed in the previous section, fall only under the ECP/PCC, and not under the Superiority Condition. For this reason, we may call them "pure ECP effects". The facts considered in this section fall under the Superiority Condition, but not

the ECP/PCC at LF. We will call these "pure Superiority effects". We will show that these effects cannot follow from the PCC at LF. They also cannot follow from the ECP at any level. On the other hand, all the facts we have been considering can be made to follow naturally from the PCC at S-structure. This analysis, then, will show that the PCC subsumes both the ECP and the Superiority Condition, further demonstrating its explanatory power.

2.2.3 Pure Superiority Effects

Consider a pre-WH-movement structure like (45):

(45) I know [_S'₁ [_{COMP}] [PRO to persuade who [_S'₂ PRO to read what books]]]

WH-movement applies ambiguously in (45) to either of the two underlined WH words. Since who is superior to what books, the Superiority Condition dictates that only who can be the actual target of WH-movement:

(46)a. I know [_S'₁ [_{COMP} who_i] [_S PRO to persuade e_i [_S'₂ PRO to read what books]]]

b. *I know [_S'₁ [_{COMP} what books_i] [_S PRO to persuade who [_S'₂ PRO to read e_i]]]

Although some informants find the contrast weaker than indicated, it is clear for almost all speakers.

The same contrast appears in (48)a-b, derived from a pre-WH-movement structure like (47). Again, Superiority makes the correct predictions:

(47) [_S'₁ [_{COMP}] [_S Mary expects [_S who [_{VP} to buy what]]]]]

- (48)a. [_S' [_{COMP} who_i] [_S does Mary expect [_S e_i [_{VP} to buy what]]]]
 b.*[_S' [_{COMP} what_i] [_S does Mary expect [_S who [_{VP} to buy e_i]]]]

Superiority also explains (49)-(50) (from Fiengo 1980, 123):

- (49) [_S' [_{COMP}] [_S you [_{VP} told who [_{PP} about what topic]]]]
 (50)a. [_S' [_{COMP} who_i] [_S did you [_{VP} tell e_i [_{PP} about what topic]]]]
 b.*[_S' [_{COMP} what topic_i] [_S did you [_{VP} tell who [_{PP} about e_i]]]]¹⁰

Similar effects hold in other languages, including French, Spanish, and Modern Hebrew (cf. Hendrick and Rochemont 1982, citing H. Borer). These are contrasts that follow from Superiority, but do not follow from the ECP in any way. Neither do they follow from the PCC at LF. Consider, for example, the LF we might associate with (46)b:

- (51) I know [_S'₁ [_{COMP} who_j^k what books_i^k] [_S PRO to persuade e_j]
 [_S'₂ PRO to read e_i]]

Paths (i) Between e_i and COMP of S'₁:

{S'₂, S, S'₁}

(ii) Between e_j and COMP of S'₁:

{S, S'₁}

Path (i) contains path (ii), and nothing is violated.

On the other hand, the contrasts in (46), (48) and (50) are reminiscent of the Crossing effects we discussed in Chapter Three.¹¹ Suppose, contrary to apparent LF fact, that the WH-in-situ has the properties of a variable bound by an operator with maximal scope: in particular, with wider scope than the WH in COMP with which we expect

it to be absorbed at LF. Recall that the ECP/Superiority effects considered in the previous section also seemed to be explainable if the WH-in-situ was taken to have the properties of a variable bound maximally high in the tree. On such a hypothesis, (46), (48) and (50) are analogous to (52), (53) and (54) -- familiar Crossing contrasts:

(52)a. ?what books_i do you know [who_j to persuade e_j [to read e_i]]

b. *who_j do you know [what books_i to persuade e_j [to read e_i]]
(cf. (46))

(53)a. ?what_i do you know [who_j Mary expects [e_j [to buy e_i]]]

b. *who_j do you know [what_i Mary expects [e_j [to buy e_i]]] (cf. (48))

(54)a. ?what topic_i do you know [who_j to tell e_j [about e_i]]

b. *who_j do you know [what topic_i to tell e_j [about e_i]]

To recapitulate: we have seen three sets of phenomena. The first fell naturally under the ECP at LF, with minor stipulations about traces in COMP. The second fell under the ECP, with rather major stipulations, and also fell under Chomsky's Superiority Condition. The third set, presented in this section, falls naturally under the Superiority Condition, but not under the ECP, in any version. Alternatively, we have suggested that the properties of WH-in-situ in all cases considered fall immediately under the PCC, if we can treat the WH-in-situ as having the properties of an empty category bound by an operator with the widest possible scope.

To put this last theory in relief, we may repeat the analogies we have suggested. We suggest that each (a) sentence below is

ungrammatical for the same reason the familiar (b) sentences are:

(55)a. *you know [who_j [e_j said [that what fell on John's head]]]

b. *what_i do you know [who_j [e_j said [that e_i fell on John's head]]]

(56)a. *he wonders [where_j [who bought this record]]

b. *who_i does he wonder [where_j [e_i bought this record]]

(57)a. *you know [what books_j to persuade who [to read e_j]]

b. *who_i do you know [what books_j to persuade e_i [to read e_j]]

If we can find some rationale for treating WH-in-situ in this way, we will have a unified account of the restrictions on WH-in-situ that are presently accounted for, some by the ECP, some by Superiority, some by both conditions. What rationale can we find? In the next section, we will suggest that the notion of "infinite paths", introduced tentatively in the last chapter, can help us toward an answer.

2.2.4 Infinite Paths and WH-in-situ

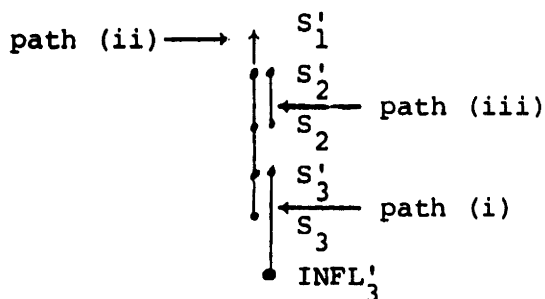
Clearly it makes no sense, either at S-structure or at LF, to treat the examples of WH-in-situ in the (a) sentences of (55)-(57) as actually taking scope over the root S', as in the (b) sentences. When they move to take scope at LF, they move to an internal, [+WH] COMP instead, where they can undergo absorption.¹² What we thus need is an explanation of why WH-in-situ acts as if it had widest scope which does not commit us to claiming that it actually does take widest scope at any level.¹³

A natural suggestion in the PCC framework we have been developing is that the position of the WH-in-situ acts at some level like a

completely unbound empty category. In section 2.4 of Chapter Four, we suggested that unbound (specifically, non A-bound) empty categories generate paths that have no natural upper bound. These we called infinite paths. The paths from such positions would have the properties, with respect to the PCC, of the paths of variables bound by a widest-scope operator. Thus, the (a) sentences of (55)-(57) would all violate the PCC, if the WH-in-situ's position creates an infinite path:

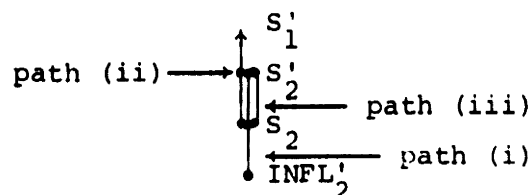
(55)a. * $[_{S_1}$ you know $[_{S_2}$ who $_j$ $[_{S_2}$ e said $[_{S_3}$ that $[_{S_3}$ what
 $[_{INFL'_3}$ $[_{VP}$ fell on John's head]]]]]

- (i) Between $INFL'_3$ and COMP of S'_3 :
 $\{INFL'_3, S_3, S'_3\}$
- (ii) From what (infinite path):
 $\{S_3, S'_3, S_2, S'_2, S'_1\}$
- (iii) Between e_j and COMP of S'_2 :
 $\{S_2, S'_2\}$



(56)a. * $[_{S_1}$ he wonders $[_{S_2}$ where $_j$ $[_{S_2}$ who $_i$ $[_{INFL'_2}$ $[_{VP}$ bought the
 record] e_j]]]¹⁴

- (i) Between $INFL'_2$ and COMP of S'_2 :
 $\{INFL'_2, S_2, S'_2\}$
- (ii) From who (infinite path):
 $\{S_2, S'_2, S'_1\}$
- (iii) Between e_j and COMP of S'_2 :
 $\{S_2, S'_2\}$



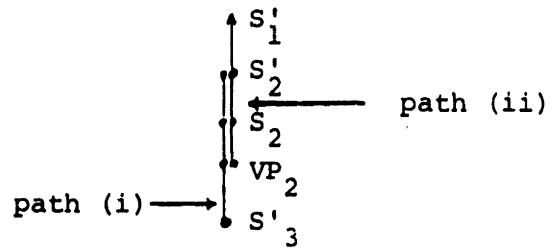
(57) a. * $[_{S'_1}$ you know $[_{S'_2}$ what books $_j$ $[_{S'_2}$ to $[_{VP_2}$ persuade who $[_{S'_3}$ to read \underline{e}_j]]]]]

(i) Between e_j and COMP of S'_2 :

$\{S'_3, VP_2, S_2, S'_2\}$

(ii) From who (infinite path):

$\{VP_2, S_2, S'_2, S'_1\}$



Why should the position of a WH-in-situ generate an infinite path?

We have two proposals which work at this point, both somewhat ad hoc.

We shall be able to distinguish them in section 2.4.

The first proposal, which we shall reject in 2.4, is to adopt into the PCC approach the COMP indexing rule (35) of Aoun et al. Recall that they suggested that an element in COMP can only bind an element outside of COMP (for the ECP) if COMP inherits the index of its content. They proposed that COMP can inherit this index by a special rule, which only applies at S-structure. Hence, only if WH is in COMP at S-structure can it bind its trace and -- in an ECP framework -- properly govern it.

If we adopt this rule, changing "properly govern" to "function as a binder for the definition of paths", we achieve the desired result at LF. The trace of WH-in-situ will always have the status of a free empty category, with respect to the definition of paths. The path of such a trace will begin with the first maximal projection dominating it, and will have no natural upper bound, exactly as desired. Note that this solution relies entirely on Aoun et al.'s claim that elements in COMP cannot bind elements outside of COMP (in some special sense) unless COMP itself is indexed. Most crucially, it relies on the stipulation that

this indexing applies only at S-structure.

The second solution, which we shall argue to be more nearly correct in 2.4, is to assume that the effects discussed above do not derive from any properties of WH-in-situ at LF. Instead, suppose we make the following assumption:

(58) WH-in-situ has the properties of an (unbound) empty category.

Recall that LF contains no WH-in-situ (except in echo questions, which we disregard). Therefore, principle (58) is relevant only at D-structure and at S-structure. Neither the PCC nor Binding Theory seem to apply at D-structure. (58) is thus relevant only at S-structure, because of the structure of the grammar.

(58) claims that at S-structure a WH-in-situ will meet the definition of "empty category" relevant to the definition of paths.¹⁵ Such a WH-phrase, since it heads its chain and is not \bar{A} -bound by anything, will, like any unbound empty category, generate an infinite path. The paths of the (a) sentences in (55)-(5-) thus follow from this proposal at S-structure, just as they follow from Aoun et al's COMP indexing rule at LF.

(58) is, of course, a stipulation. Why should a WH-in-situ act like an empty category? We can suggest some rationalization of (58) -- that is, ways to make (58) seem more or less natural -- but we have no really satisfactory answer to our question. Suppose, for example, that the definition of paths does not really concern itself with "empty categories", but with "categories that cannot refer (independently or dependently)". The idea would be that names, pronouns and anaphors refer, even if their reference is linked to the reference of some other

element (as in the case of anaphors), but that variables and operators do not refer. Still, the link between "not referring" and creating a path is obscure. We shall argue briefly in the final section that focused NPs can be said to "not refer". For now, since no real insight is afforded us by such rationalizations, let us continue to base our analysis on stipulation (58), in the hope that a deeper explanation will be forthcoming.

Notice that under our second proposal, assuming (58), it is no surprise that deletion of that does not save a WH-in-situ from pure ECP effects. Since what is wrong with a WH-in-situ subject of a tensed S has nothing to do with movement, and since the WH violates the PCC before such movement, there is no way for a trace to bind this WH from any COMP. This conclusion also follows from Aoun et al.'s COMP indexing rule, since a trace deposited in COMP by LF movement cannot bind a trace in subject position. Both theories therefore solve the problem raised by that deletion in (30)b.

The two proposals are thus extremely similar. Both could be adopted in either an ECP or PCC framework, though the ECP does not explain the pure Superiority (Crossing) effects with either proposal. They differ in one respect, however. Aoun et al.'s proposal entails that the PCC will be relevant to WH-in-situ constructions at only one level -- LF. Our second proposal derives Aoun et al.'s effects at S-structure, and leaves open the possibility that further effects of the PCC might be observed at LF. At LF, (if we do not assume Aoun et al.'s convention on COMP indexing), the positions of S-structure WH-in-situ are filled by real empty categories that are bound from COMP.

Our discussion in the next two sections will provide us with an extremely clear test of this distinction between the two theories. Transposing results of Kayne (1982) once more into a PCC framework, and assuming principle (58), we will have a way of nullifying the effects of the PCC on WH-in-situ at S-structure. This nullification will leave exposed whatever effects the PCC might have on the trace of WH-in-situ at LF. If LF movement does allow proper binding from COMP, we expect such effects to exist. Since these facts do exist, we will have an argument in favor of proper binding from COMP established by LF movement, in favor of principle (58) at S-structure, against our first proposal, which include Aoun et al's COMP indexing rule.

Before turning to cases in which the effects of the PCC are nullified, we need to consider some issues that arise in our discussion so far.

First, if we wish to rule out WH-in-situ in the subject position of a tensed S, as in (55)a and (56)a, by the PCC and principle (58), we must assume that the path between INFL and COMP exists at S-structure. In Chapter Four, we argued that this path exists as a result of TNS movement, forced by a filter that requires TNS to have scope over a complete proposition. It is natural to suppose that this filter applies at LF. To make principle (58) work, however, we must assume that the filter also applies at S-structure. While this is less obviously natural, it does not seem to lead to any contradiction. Note, however, that this assumption is not required on our adaptation of Aoun et al's theory.

Second, if the PCC affects WH-in-situ in the way indicated,

there are at least two other domains in which we expect the effects to show up:

A. If Subject Condition effects are derived from the PCC, we expect WH-in-situ to be ungrammatical inside a subject. Judgments differ on whether this is the case. Huang (1982, 496) assigns to (59)a a single question mark; in our judgment it is unacceptable (except as an echo question).¹⁶ (59)c seems particularly ungrammatical.

(59)a.*who said that [for Bill to marry who] would be a surprise

b.*who believes that [pictures of what building] are appropriate here

c.*I wonder where [books by whom] are stored

If the judgments are as indicated, then WH-in-situ shows the predicted Subject Condition effects.

B. If Coordinate Structure Condition effects are derived from the PCC, we expect WH-in-situ to show these effects. Certainly those Coordinate Structure Condition effects that we continue to attribute to the A/A Condition appear with WH-in-situ, as noted, for example, by Baker (1970). The following example is from Chomsky (1981a, 279, fn. 8):

(60)*I wonder who wrote which textbook and that novel

Cases that derive from the PCC appear to be ungrammatical as well, although the judgments may be weak:

(61)a.*I don't know who tried to read which book or to play that sonata

b.*which article proves your theorem and defends what theory

In the next section, we will see some "ATB exceptions" to the Coordinate Structure Constraint with WH-in-situ.

Third, we must solve a technical problem that arises when we consider matrix questions. In our discussion of the PCC, we have cunningly chosen only examples in which the interrogative clauses are embedded. The same effects appear, of course, with unembedded questions. Compare (62)a-b with the (a) sentences of (56)-(57):

(62)a. *where did who buy this record

b. *what books did you persuade who to read

If the highest S' is also the root node, we cannot rule these sentences out by the PCC. The path from the trace of WH to its binder at S-structure will end at the highest S', as will the theoretically infinite path from the WH-in-situ. Since the two paths end at the same place, the longer will contain the shorter, and the PCC will be satisfied, contrary to our desires.

It is probably possible to solve this problem by refining the formal definition of paths in some way. For our present purposes, we may make things work out correctly by stipulating that the root of all trees is an initial symbol Σ , which will dominate the highest S' in the structures of (62). Thus, the relevant paths for (62)b will be as follows:

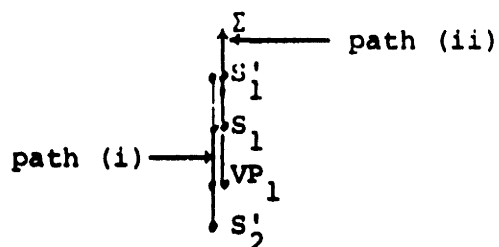
(62)b [Σ [S'_1 what books_i [S_1 did you [VP_1 persuade who [S'_2 to read e_i]]]]]]

Paths: (i) Between e_i and COMP of S'₁:

{S'₂, VP₁, S₁, S'₁}

(ii) From who (infinite path):

{VP₁, S₁, S'₁, Σ }



Finally, it is well to steer clear, in the discussion, of the issues regarding subjacency raised in the last section of the preceding chapter. On a technical level, something must be said to force the definition of paths to count all WH-in-situ as empty categories. Recall that we suggested a number of theories that could account for the fact that non-subjacent empty categories and some adjacent empty categories can be disregarded by the PCC at S-structure. Perhaps we wish to make the predicate "is subjacent" relevant only to categories that have binders, and stipulate that all empty categories that lack binders are visible at all levels to the PCC.

The basic point is clear: the definition of paths must not be allowed to ignore WH-in-situ at S-structure. The details of the theory that will ensure this result remain to be worked out, particularly since our discussion of subjacency was speculative, and did not choose between a number of options. We will therefore act as if no problems existed in this domain, although of course they do.

2.3 Multiple WH-in-situ and Absorption as Coindexation

Kayne (1982) notes, following in part an observation of N. Chomsky (cf. also Kayne 1981e, fn. 7; Chomsky 1981a, 238), that pure ECP effects and ECP/Superiority effects with WH-in-situ are often nullified if an additional WH-in-situ is added to the tree:¹⁷

(63)a. *who said [that what fell on John's head]

b. ?who said [that what fell on which man]

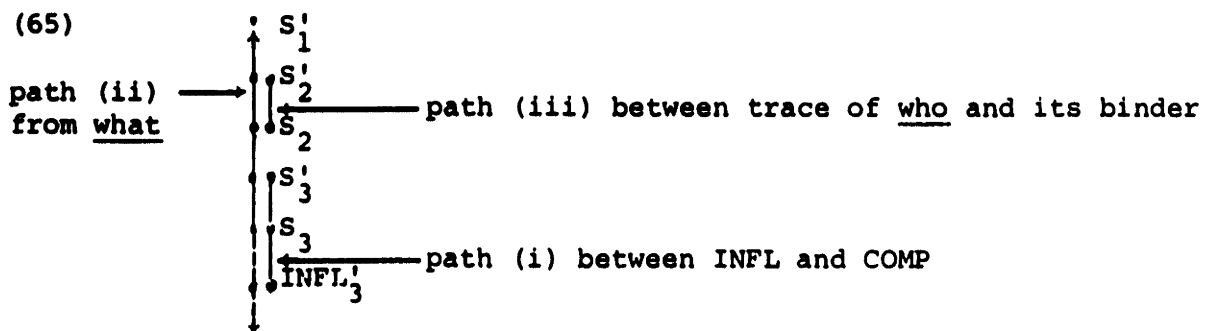
(64)a.*he wonders [where [who [bought this record]]]

b.?he wonders [where [who [bought what]]]

Interestingly, Kayne shows that the additional lower WH-in-situ saves the higher WH-in-situ in (63)b and (64)b in essentially the same way additional lower gaps can save a Subject Condition violation, and provides KECP accounts of the two phenomena that are parallel. We may once more adapt his insight to the PCC framework.

Consider once more the paths in (63)a, which we presented above in (55)₁. Notice what goes wrong with these paths. The infinite path from the in situ what overlaps a path between INFL and COMP. The two paths violate the PCC for the following reason: on the one hand, the path from what contains a number of nodes that the path between INFL and COMP does not; on the other hand, the path from what does not contain the INFL' node that the path between INFL and COMP does.

Let us ask a familiar question. What operation can we perform on the infinite path from what that will allow it to contain the path between INFL and COMP? The obvious answer is to add at least the missing INFL node: compare (65) with our diagram of (55)a:



Only with the addition of the dotted portion of the path, extending

down to INFL, will path (ii) contain path (i).

Let us hypothesize, then, that the addition of the lower WH-word what to (63)a serves to add this INFL' node, and possible others, to the path from who, allowing this path to contain the path between INFL and COMP. Let us ask why.

Continuing to adapt Kayne's analysis, let us claim that who and what, rather than generating two separate infinite paths, unite to form a common path, just like the multiple gaps considered in the previous chapter. Supposing that there is this common path, which is the union of what would be the separate infinite path from who and from what, we can demonstrate that the result indeed does not violate the PCC. Compare (63)b and its paths with (55)a (= (63)a) and its paths:

(63)b. [Σ [S'_2 who_j [S'_2 e_j said [S'_3 that [S'_3 what [$INFL'_3$ [VP_3 fell on which man]]]]]]]]

(i) Between INFL and COMP of S'_3 :

{ $INFL'_3$, S_3 , S'_3 }

(ii) From which man and from what:

{ VP_3 , $INFL'_3$, S_3 , S'_3 , S_2 , S'_2 , Σ }

(iii) Between e_j and COMP of S'_2 :

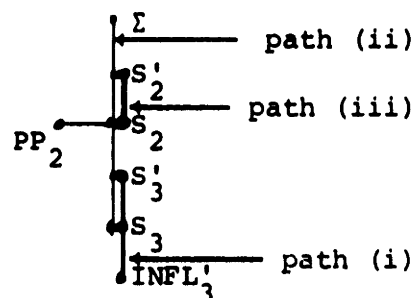
{ S_2 , S'_2 }

The underlined nodes in path (ii) are those contributed exclusively by which man. Notice that path (ii) now contains all the other paths, as in the diagram in (65) above.

The reader may verify that uniting the paths of who and what in (64)b will allow this path to contain the path between INFL and COMP in the same way. Notice, of course, that if we add a second WH-word whose path does not include INFL, the PCC will continue to be violated. Such is the case in (63)c, which we saw above:

(63)c.*[[_{S₂'} who_j [_{S₂'} e_j said [_{PP₂} to whom [_{S₃'} that [_{S₃'} what [_{INFL₃'} [_{VP₃} fell on John's head]]]]]]]]]

- (i) Between INFL and COMP of S₃':
 {INFL₃', S₃, S₃'}
- (ii) From what and from whom:
 {S₃, S₃', PP₂, S₂, S₂', Σ}
- (iii) Between e_j and S₂':
 {S₂, S₂'}



As Kayne notes further, the second gap must not initiate a new violation of (for us) the PCC:

(66) *[[_Σ [_{S₂'} who_j [_{S₂'} e_j said [_{S₃'} that [_{S₃'} who believes [_{S₄'} that [_{S₄'} what fell on John]]]]]]]]]

The lowest WH-in-situ, what, may save the who in S₃ from the PCC, but it runs afoul of the path between INFL and COMP in its own clause, and violates the PCC. Of course, it may be saved by a third WH-in-situ:

(67) [_Σ [_{S₂'} who_j [_{S₂'} e_j said [_{S₃'} that [_{S₃'} who believes [_{S₄'} that [_{S₄'} what fell [_{PP₄} on whom]]]]]]]]]

Now let us add to Kayne's observations some of our own. Exactly the same paradigm obtains with pure Superiority effects, which we have put together with Crossing effects:

(68)a. *what books_j [did you persuade who [to give e_j [to Bill]]]]

b. ?what books_j [did you persuade who [to give e_j [to whom]]]]

(69)a. *what_i [does Mary expect [who [to buy e_i [for Bill]]]]

b. ?what_i [does Mary expect [who [to buy e_i [for whom]]]]

(68)a is essentially the same as (57)a. Recall that the infinite path from who, being infinite, contained high nodes that the path between e_i and its binder did not. At the same time, the path between e_i and its binder contains low nodes that the infinite path does not. If adding a second, lower WH-in-situ adds these low nodes to the infinite path, then the PCC violation will be eliminated. This is what happens in (68)b. Compare the paths of (68)b with (57)a, diagrammed above:

(68)b. [_Σ [_{S'₂} what books_j [_S did you [_{VP₂} persuade who [_{S'₃} to give e_j [_{PP₃} to whom]]]]]]]]

(i) Between e_j and COMP of S'₂:

{S'₃, VP₂, S₂, S'₂}

(ii) From who and from whom:

{PP₃, S'₃, VP₂, S₂, S'₂, Σ}

The underlined nodes are the contribution of whom to path (ii). As can be seen, they are crucial to satisfying the PCC. Predictably, addition of a second WH that does not supply these crucial low nodes leaves the structure in violation of the PCC:

(70)*what books_j [did you tell what man [to persuade who [to give e_j to me]]]

It seems that the same paradigm recurs with "Subject Condition effects" with WH-in-situ, of the type discussed at (59) in the previous section. (71)b shows how a second WH-in-situ that adds the INFL' node to the infinite path saves the structure of (71)a from the PCC. (71)c shows that a second WH-in-situ that does not add the crucial INFL' node does not improve things:

(71)a.*who said [that [for Bill to marry who] would be a surprise for me]

b.?who said [that [for Bill to marry who] would be a surprise for whom]

c.*who said [to whom] [that [for Bill to marry who] would be a surprise for Mary]

Most interestingly, a similar paradigm occurs in conjoined structures. Consider the paths for (72), which we briefly mentioned above, as (61)b:

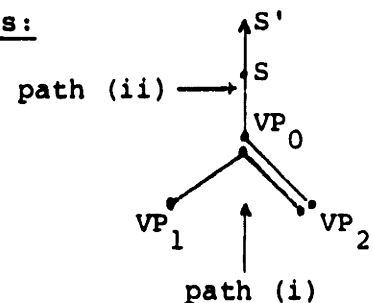
(72)*[_S which article_i [_S e_i [_{VP₀} [_{VP₁} proves your theorem] and [_{VP₂} defends which theory]]]]

(i) Between the conjunction and the conjuncts:

{VP₁, VP₂, VP₀}

(ii) From which theory:

{VP₂, VP₀, S, S'}



Now consider (73), in which a WH-in-situ is added to the other conjunct:

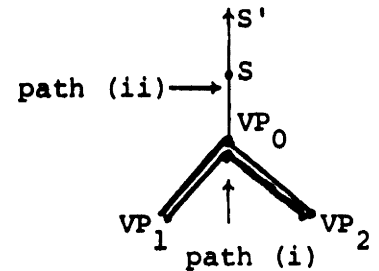
(73) [_S, which article_i [_S e_i [_{VP₀} [_{VP₁} proves which theorem] and
 [_{VP₂} defends which theory]]]]

(i) Between the conjunction and the conjuncts:

{VP₁, VP₂, VP₀}

(ii) From which theorem and from which theory:

{VP₁, VP₂, VP₀, S, S'}



Addition of the extra WH allows the path from the WH-in-situ to contain the path between the conjunction and the conjuncts. In other words, (73) is exactly analogous to the "ATB exceptions of the Coordinate Structure Constraint" discussed in the previous chapter. (72)-(73) are precisely parallel to (74)a-b:¹⁸

(74)a. *which theory_i [does this article [_{VP₀} [_{VP₁} prove your theorem] and
 [_{VP₂} defend e_i]]]

b. which theory_i [does this article [_{VP₀} [_{VP₁} prove e_i] and
 [_{VP₂} defend e_i]]]

Thus, our theory both of Coordinate Structure Constraint effects and of the ATB exceptions has something correct to say about WH-in-situ constructions.

We have just demonstrated the mechanics of a theory of WH-in-situ constructions. Assuming (for the moment without argument) that WH-in-situ generates an infinite path at S-structure, we have accounted for

a variety of effects by means of the PCC. Assuming that a group of WHs-in-situ form a single united path, just like multiple gaps in the constructions of Chapter Four, we explain the exceptions to these effects. We have not so far answered a very important question: why should a group of WHs-in-situ combine their paths? With malice aforethought, we have chosen a notation to express absorption that will help us to answer this very question.

Recall the beginning of our last definition of paths, (64) in Chapter Four:

(75) "Consider a set T such that for some index i :

$T = \{ \underline{t} \mid \underline{t} \text{ is an empty category locally } A\text{-free and indexed } \underline{i} \} \dots$ "

A unified path is begun by all the members of T , in accordance with the rest of the definition. By referring to an index \underline{i} , we guaranteed that all and only gaps that share a common binder will unite their paths in one.

The motivation for uniting the paths of multiple gaps was that we could allow certain gaps to "save" other gaps from the PCC. Subject Condition effects and Coordinate Structure Constraint effects were two cases in point. At the same time, we noticed that certain types of "salvation" were impossible. For example, additional gaps never save a CTP violation:

(76)a. *who _{i} did Bill say [_{S} , that [_{e_i} liked Mary]]

b. *who _{i} did Bill say [_{S} , that [_{e_i} liked e_i]]

As we pointed out, we are in fact free to assume that the lower gap in

(76)b does save the higher gap from the PCC. The sentence is independently ruled out by the requirement that neither gap c-commands the other, a requirement imposed by the θ -criterion and the Binding Theory. If the lower gap and the higher gap form a chain, the θ -criterion is violated; if the lower gap and the higher gap do not form a chain, the lower gap is PRO, and, being governed, violates the Binding Theory. As we noted, this condition is independently needed in cases where the PCC is irrelevant:

- (77)a. who_i does Bill consider [_S e_i to be angry at Mary]
 b. * who_i does Bill consider [_S e_i to be angry at e_i]

On the other hand, "CTP effects" (pure "CTP effects") with WH-in-situ may be saved by adding a lower WH-in-situ:

- (78)a. *who said [_S, that [who liked Mary]]
 b. ?who said [_S, that [who liked who]]

The obvious reason, as Kayne (1982) observes, is that the two WHs-in-situ do not share a referential index: neither the Binding Theory nor the θ -criterion blocks the structure. Therefore the "saving" effects of adding nodes to the path from the subject are observable:¹⁹

- (79)?who said [_S, that [who_j like who_k]]

On the other hand, if the two WHs-in-situ do not share a referential index, but are treated by the definition of paths as empty categories, how do their paths come to be united? On Kayne's approach, path are united, not when they share an index, but when they share a binder. He argues that, in WH-in-situ sentences, the WH-in-COMP with

which the WHs are to be absorbed acts as a binder for all the WHs. (ECP/Superiority effects are captured under the assumption that the subject position is completely ungoverned, so that a subject WH-in-situ can be bound by a WH in COMP only if a lower WH "completes the circuit" to COMP. The WH in COMP cannot, apparently, govern a subject WH-in-situ, although it can govern a subject trace.) While the exact process is not spelled out in Kayne (1982), the suggestion that absorption is what allows multiple WHs-in-situ to combine their paths is worth pursuing. We shall suggest a different approach to the problem, which nonetheless makes use of Kayne's fundamental insight.

Following Higginbotham and May (1981) (also Aoun et al. 1982), we stipulated that the rule of absorption applies to a set of WH-phrases in a particular COMP. In a departure from Higginbotham and May, we suggested that the syntactic rule of absorption is to coindex -- specifically, cosuperscript -- the set of WH-phrases. This set of cosuperscripted WH-phrases is then interpreted as an n -ary operator, in Higginbotham and May's sense. We repeat our rule (= (16)) below:

$$(80) \quad [{}_{\text{COMP}} \text{WH}_1, \dots, \text{WH}_n] \longrightarrow [{}_{\text{COMP}} \text{WH}_1^i, \dots, \text{WH}_n^i], \text{ for some index } \underline{i}.$$

Of necessity, this rule applies at LF, since only at LF are there multiple WH-phrases in COMP. Notice now that specifying COMP in (80) is almost, but not quite redundant at LF, since all WH-phrases that are operators are independently required to be in COMP at LF by principle (10) (or (11)). Suppose we were to eliminate the COMP brackets from (80), and state it more generally as in (81) (yielding a convention similar to Haik's (1982) "indirect binding" among quantifiers in A-positions):

(81) Absorption Cosuperscripting (AC)

Let W be the set of WH-phrases in a phrase marker P . For some superscript \underline{i} , assign \underline{i} to each member of W .

(81), optimally, is optional and applies freely. There are two cases in which (81) overgenerates at LF. First, it does not require all the WH-phrases in a given COMP to be cosuperscripted with each other. That is, it allows some WH-phrases to bear a superscript distinct from others, or to lack a superscript altogether:

$$(82) \quad [{}_{\text{COMP}} \text{WH}_i^m \text{WH}_j^m \text{WH}_k^m \text{WH}_l^m]$$

Given the semantic facts of multiple interrogation, we probably wish to rule this case out. We also want to prevent material in one COMP from being cosuperscripted with material in another COMP:

$$(83) \quad [S_1' [{}_{\text{COMP}} \text{WH}_i^m] [S \dots [S_2' [{}_{\text{COMP}} \text{WH}_j^m] [S \dots]]]]$$

If (83) were possible, we could interpret who asked what to do identically to who asked to do what, with who and what undergoing absorption.

Both these cases can be ruled out by the following convention:

(84) WH-Criterion

Let W be the maximal set of phrases in some COMP C , and W' be the maximal set of phrases in some COMP C' , $C \neq C'$. Then for some superscript \underline{i} ,

- (i) $\forall x \in W, x \text{ bears } \underline{i}$
- (ii) $\forall x \in W', \neg x \text{ bears } \underline{i}$

From (84), it follows that in each COMP all the WH-phrases share a superscript, and that this superscript is unique to that COMP. The first part

probably derives from considerations of interpretation; the second part is probably related in some way to the uniqueness clause of the θ -criterion -- hence our name "WH-criterion".

The rule AC and the WH-criterion essentially divide our previous cosuperscripting rule into two parts: the AC assigns the superscript, while the WH-criterion governs its output. Unlike our previous cosuperscripting rule, the AC and the WH-criterion do not rule out in principle any cosuperscripting between WH-phrases not in COMP. It just so happens that all WH-phrases will be in COMP at LF.

The AC, however, opens up a possibility that our earlier rule, and Higginbotham and May's, does not -- namely, the possibility of applying the rule at S-structure, to WH-phrases in situ.²⁰ Suppose AC is allowed to apply at S-structure. We then expect to allow S-structures like (85):

(85) who said [_S that [_S who_i^k like what_j^k]]

Now let us ask what cosuperscriptings are. Cosuperscripts, for us, as well as for Chomsky (1981a), who uses them in a different context, are means of coindexation that are not relevant to the Binding Theory. In other words, who_i^k and what_j^k in (85) are coindexed, and each has the properties of an empty category (if we are correct), but because the coindexation is not relevant for the Binding Theory neither functions as an antecedent for the other, nor do we have to worry about the domains in which the cosuperscripting occurs. In fact, as we have seen, the subscripts on these WHs, which do count for the Binding Theory,

must be distinct, since who c-commands what in (85).

Consider again our definition of paths -- in particular, the excerpt in (75) above. For a given index i , all empty categories bearing that index form a united path. Taking our formulation quite literally, we note that it makes no distinction between subscript indexing and superscript indexing. In other words, we expect categories that have distinct subscripts but identical superscripts -- like who_i^k and what_j^k in (85) -- to form a single path. Since who and what are both indexed k , they are both members of a set T which create a single path under (75). Thus, if the definition of paths does not distinguish superscripting from subscripting, and if absorption results from superscripting, we explain why multiple WHs-in-situ may form a common path.

This theory makes one clear prediction, which we shall see confirmed in the next section: if there is no "re-cosuperscripting" at LF, we expect that cosuperscripting forced by the PCC at S-structure will have consequences for logical interpretation. In particular, certain potential scope ambiguities should be eliminated. Where we might expect WH_1 to be able to take wide scope, while WH_2 takes narrow scope, they will have to take identical scope whenever the PCC forces them to be cosuperscripted at S-structure.

One problem raised by Kayne (1982) can be resolved in a number of ways on our approach. He asks why an S-structure WH-trace cannot save a WH-in-situ from the PCC -- i.e. why the paths from who and e_i are not united in sentences like (86)a, the way the paths from who and what are united in (86)b:

(86)a. *who said [_S, what_i [_S who liked e_i]]

b. who said [_S, that [_S who liked what]]

For us, this question has three parts. First, it is clear that who and e_i cannot be cosubscripted. Can they be cosuperscripted? Not if we limit cosuperscripting to operators, as we have. This seems reasonable: not only can a trace not save a WH-in-situ, but a WH-in-situ cannot save a trace. Although the path from what in (87) does include the INFL' node missing from the path between e_i and who, the paths obviously do not unite:

(87) *who_i did you say [_S, that [_S e_i [_{INFL'} [_{VP} liked what]]]]

Also, a trace cannot save another trace, unless they are cosubscripted, as in a multiple gap construction. The paths from e_j and from e_i obviously do not unite, or else e_i would not violate the PCC with respect to the path between INFL and COMP:

(88) *who do you know [_S, what_j [_S e_i [_{INFL'} [_{VP} did e_j]]]]

Second, we might ask why we cannot cosuperscript who in situ with what in COMP in (86)a, so that what might count as a "cosuperscript binder" of who and stop its path at the embedded S'. In other words, why can't a structure like (89)a exist, where what is to who as who_i is to e_i in (89)b?

(89)a. who said [_S, what_j^k [_S who_i^k like e_j]]

b. who said [_S, who_i [_S e_i liked Mary]]

One answer might be that only subscripts are relevant for the binding

spoken of in the definition of paths.

Alternatively, and more interestingly, we might regard cosuperscripting as simply a variety of free indexing, applying, like all free indexing, at S-structure and LF. In our discussion of Chomsky's analysis of parasitic gaps in Chapter Four, we noted that free indexing at S-structure must be limited to A-positions. Although we explored some alternatives to this view, suppose it is correct. If cosuperscripting is simply one type of free indexing, it too should be limited to A-positions at S-structure, though not necessarily at LF.

In such a case, there would be no problem with (86)a: what cannot "cosuperscript-bind" who at S-structure, even if we don't rule such binding out in principle, because \bar{A} -positions can't be coindexed at S-structure. At LF, an operator in COMP would be free to cosuperscript-bind an operator in an A-position -- except that there are no operators in A-positions at LF. We thus conclude that cosuperscripting may very well be simply a special case of free indexing, and fall under the laws free indexing is subject to, if we maintain Chomsky's analysis of parasitic gaps. In this case, we have a general explanation of (80). If we do not maintain Chomsky's analysis, we may, as indicated, simply stipulate that "cosuperscript binding" does not count in the definition of paths.

Finally, we might raise a question that is somewhat tangential to this discussion, but becomes important later. Does a trace inherit the superscript of its (subscript-)binder? If we do not maintain Chomsky's analysis of parasitic gaps, we might allow cosuperscripting of \bar{A} -positions at S-structure. If we did, then what and who in (87) might be

cosuperscripted as in (89)a. We have seen that we might prevent what from acting as a "cosuperscript-binder" of who and stopping its path from extending higher than S'. Recall that the problem with this path is that it does not include INFL'. Suppose e_j were allowed to inherit the cosuperscript of its binder in (89)a:

(90) who said [_S, what_j^k [_S who_i^k like e_j]]

Strictly speaking, the cosuperscripting rule AC coindexes only the two operators, so we cannot appeal to AC.

It seems that we do not want to allow for "inheritance" of co-superscripts in this way. Notice that here the problem only arises on some very particular choices of indexing conventions. As we saw, (86)a can be ruled out in a number of other ways, depending on what conventions we assume. Nonetheless, this particular assumption -- that superscripts are not inherited from (subscripts) binders -- will be crucial in the next section, regardless of what theory of indexing we adopt here.

Before proceeding to the next section, let us note that there is already a certain argument for our treatment of pure ECP, ECP/Superiority and pure Superiority effects at S-structure rather than at LF. Remember that we began this discussion by considering two very similar hypotheses. Under one hypothesis, adapted to the PCC framework from a proposal of Aoun et al's, these effects are due to the trace of WH-in-situ at LF: by a special convention, such traces will always act as if they are unbound. Under our other hypothesis, which we have assumed in our discussion, WH-in-situ acts like an empty category at S-structure, and might very well bind its trace at LF.

If our adaptation of Aoun et al's proposal were correct, we

would have to find some reason to unite the paths from the two traces of WH-movement in the logical forms of sentences like (86)b. The reason cannot be cosubscripting, since one trace would c-command the other. Similarly, the reason cannot be cosuperscripting: if we allowed cosuperscripting of traces, we could not distinguish (86)b from (86)a. We could invent some other diacritic: perhaps we could not combine the path of a trace that is bound, like the trace of syntactic WH-movement, with the path of a trace that is not bound, like the trace of LF WH-movement, on this hypothesis. The point is that there is no natural reason to distinguish the two cases at LF.

In the next section, however, we will present the more substantive argument for handling pure ECP, ECP/Superiority, and pure Superiority effects at S-structure. As discussed earlier, on an S-structure theory of these phenomena, we expect additional effects to surface at LF, if the conditions are right. On an LF theory, we do not expect any additional effects. We will show that such additional effects so indeed exist.

2.4 WH-in-situ and Logical Form

In this section, we will show that not only is WH-in-situ subject to the PCC at S-structure, but its trace is subject to the PCC at LF. We will first give a difficult demonstration of this fact, based on certain scope ambiguities that fail to appear. Afterwards, we will demonstrate how this phenomenon can be made to be a question of grammaticality, providing an easier and more powerful demonstration of the PCC at LF. This order is necessary, because the second demonstration depends on the first.

In section 2.1, we showed, following Baker (1970), that WH-in-

situ exhibits scope ambiguities, which we can attribute to its LF representation. Let us briefly review the evidence. We considered a sentence like (91) below (similar to (12)), which, we claimed, could have either the LF representation in (92)a or that in (92)b:

(91) $[S_1, [\text{which man}]_i [e_i \text{ remembers } [S_2, \text{where}_j [\text{we bought } \underline{\text{what}} e_j]]]]]$

(92)a. $[S_1, [\text{which man}]_i^m [e_i \text{ remembers } [S_2, [\text{where}_j^n \text{what}_k^n] [\text{we bought } e_k e_j]]]]]$

b. $[S_1, [\text{which man}]_i^m \text{what}_k^m [e_i \text{ remembers } [S_2, [\text{where}_j^n] [\text{we bought } e_k e_j]]]]]$

In (92)a, what has narrow scope. The question can be answered by naming a single man. What that man remembers is a list of ordered pairs of things bought and places they were bought in:

(92)a'. John remembers where we bought what: he remembers that we bought A in GUM, B in Gostinnyj Dvor, and C on Nevsky Prospect...

In (92)b, what has wide scope. The question can be answered by a set of ordered pairs of men who remember and things bought. What each man remembers is a place where something was bought:

(92)b'. John remembers where we bought A (in GUM); Mary remembers where we bought B (in Gostinnyj Dvor); Sue remembers where we bought C (on Nevsky Prospect)...

Consider now (93):

(93) $[S_1, [\text{who}_i] [e_i \text{ knows } [S_2, [\text{what books}_j] [\text{to persuade } \underline{\text{which man}} [\text{to give } e_j [\text{to } \underline{\text{which woman}}]]]]]]]$

(93) is essentially the same as (68)b, which we considered in the context of apparent violations of pure Superiority effects. As (68)a showed, which woman cannot be replaced by a name (e.g. that woman) to yield a well-formed multiple interrogation.²¹ The infinite paths from which woman and which man are united, enabling which man to not violate the ECP with respect to the path between e_j and what books_j.

In the last section, we argued that these two paths can be united only if which man and which woman are cosuperscripted. When two operators are cosuperscripted, they are interpreted as absorbed. It should follow, as we remarked, that they will have the same scope. This appears to be true, although the judgments are, as always in these matters, extraordinarily difficult to determine. Thus, (93) seems to be two-ways ambiguous. The pair (which man, which woman) may take wide scope or narrow scope, but the members of the pair may not take different scopes. Thus, the only available LFs are:

- (94)a. [_{S₁}, [who_i^m] [e_i knows [_{S₂}' [what books_jⁿ which man_kⁿ which woman₁ⁿ] [to persuade e_k [to give e_j [to e_1]]]]]]]
- b. [_{S₁}, [who_i^m which man_k^m which woman₁^m] e_i knows [_{S₂}' [what books_jⁿ] [to persuade e_k [to give e_j [to e_1]]]]]]]

(94)a, with narrow scope of which man and which woman, allows an answer of the following sort:

- (94)a'. John knows what books to persuade which man to give to which woman: he knows that he should persuade Bob to give book A to Mary, that he should persuade Bill to give book B to Susan,...

(94)b, with wide scope of which man and which woman, allows an answer like (94)b':

(94)b'. John knows what books to persuade Bob to give to Mary; Tim knows what books to persuade Barry to give to Osvaldo; Jean-Roger knows what books to persuade Isabelle to give to Yvon,...

Missing are the "mixed scope" readings in (95):

(95)a. * $[_{S_1} \text{who}_i^m \text{ which man}_k^m] [_{e_i} \text{ knows } [_{S_2} \text{ [what books}_j^n \text{ which woman}_1^n] \text{ [to persuade } \underline{e}_k \text{ [to give } \underline{e}_j \text{ [to } \underline{e}_1 \text{]]]}]]]]]$

b. * $[_{S_1} \text{ [who}_i^m \text{ which woman}_1^m] [_{e_i} \text{ knows } [_{S_2} \text{ [what books}_j^n \text{ which man}_1^n] \text{ [to persuade } \underline{e}_k \text{ [to give } \underline{e}_j \text{ [to } \underline{e}_1 \text{]]]}]]]]]$

(95)a invites an answer like (95)a'. I believe that (95)a' is not an appropriate answer to the S-structure question (93):

(95)a'. *John knows what books to persuade Bill to give to which woman (namely, book A to Mary, book B to Sue...); Ken knows what books to persuade Morris to give to which woman (namely, book F to Dorothy, book G to Sally...); Rita knows what books to persuade Mario to give to which woman (namely, ...) ...

The same is true for (95)b'. The most natural answer that would go with (95)b' is not grammatical (John knows what books to persuade which man to give to Mary, ...: this is worse if who replaces which man). In any case, the invited answer does not seem possible in any form.

If these judgments, difficult to sort out at best, are correct, they support the claim that absorption in the form of cosuperscripting takes place at S-structure, and that this cosuperscripting allows an escape from the PCC at S-structure. The absence of the readings in (95)a-b is an

immediate consequence of this cosuperscripting.

Now let us turn to the main example we will consider in this section. Consider a familiar ECP/Superiority effect:

(96) * $[_{S_1} [\text{who}_i] [\underline{e}_i \text{ knows } [_{S_2} [\text{where}_j] [\underline{\text{who}} \text{ bought this record } \underline{e}_j]]]]]$

Recall that who violates the PCC at S-structure, because its infinite path overlaps the path between INFL and COMP of S_2' , without either containing the other. Recall that adding a lower, cosuperscripted WH-in-situ saves (95), by adding the crucial INFL' node to the infinite path:

(96)' $[_{S_1} [\text{who}_i] [\underline{e}_i \text{ knows } [_{S_2} [\text{where}_j] [\underline{\text{who}}_k^m \text{ bought } \underline{\text{what}}_1^m \underline{e}_j]]]]]$

Having added what₁^m, (96) no longer violates the PCC at S-structure.

Not only is this phenomenon interesting in itself, as we have seen, but it is also an extremely useful tool for probing into the real properties of the trace of WH-in-situ at LF -- if we are correct in assuming that the effects we have discussed derive from S-structure. Any additional PCC effects we find must be attributable to properties of the trace of WH-in-situ at LF, and not to properties of WH-in-situ at S-structure.

Given the discussion above, we expect that who_k^m and what₁^m will have the same scope in (96)'. Supposing, as we have, that all WHs-in-situ move to COMP at LF, we might expect two possible outcomes, here as in the previous set of examples. The pair (who_k^m, what₁^m) might either move to the higher COMP and cosuperscript with who_i, or move to the lower COMP, and cosuperscript with where_j:

(97)a. $[_{S_1} [\text{who}_i^n] [\underline{e}_i \text{ knows } [_{S_2} [\text{where}_j^m \text{ who}_k^m \text{ what}_1^m] [\underline{e}_x \text{ bought } \underline{e}_1 \underline{e}_j]]]]]$

b. $[_{S_1} [\text{who}_i^m \text{ who}_k^m \text{ what}_1^m] [\underline{e}_i \text{ knows } [_{S_2} [\text{where}_j^n] [\underline{e}_x \text{ bought } \underline{e}_1 \underline{e}_j]]]]]$

The reading of (97)a is certainly present. The most natural answer corresponding to this reading would be a name of someone who knows. What this person knows is a set of ordered triples of buyers, things bought and places where the buying takes place:

(97)a'. Rick knows where who what: he knows that Bill bought A in New York, that Dagmar bought B in Munich, that Fred bought C in Boston,...

What about the reading in (97)b? It seems that it is not present. This reading calls for a set of ordered triples of people who know, buyers and things bought. What is known by each person who knows is a place:

(97)b'. John knows where Bill bought A (namely, in New York); Mary knows where Dagmar bought B (namely, in Munich); Sue knows where Fred bought C (namely, in Boston),...

The absence of (97)b is certainly a subtle judgment, but seems real. Let us account for it first, and then show how to make the judgment much stronger.

Our explanation of why (97)b is missing rests crucially on the assumption discussed in the last section that traces do not inherit superscripts from their (subscript-) binders. Although $\underline{\text{who}}_k^m$ and $\underline{\text{what}}_1^m$ are cosuperscripted in situ at S-structure, their traces at LF, \underline{e}_k and \underline{e}_1 are not cosuperscripted. As a result, while $\underline{\text{what}}_1^m$ shelters $\underline{\text{who}}_k^m$ from the PCC at S-structure, nothing shelters the trace of $\underline{\text{who}}_k^m$ from the PCC at LF.

At LF, then, the trace of $\underline{\text{who}}_k^m$ should act like any trace of WH in the subject position of a tensed sentence. The basic properties of such a trace are by now quite familiar. Because the path from the subject of

of a tensed sentence will overlap the path between COMP and INFL, the PCC is satisfied only if the subject trace is bound from the nearest COMP. This is the origin of the basic CTP effects discussed in Chapter Three. In (97)a, the trace of who_k^m , e_k , is bound from the local COMP and satisfies the PCC. In (97)b, e_k is not bound from the local COMP, and the PCC is violated. The absence of (97)b as a logical form for (96) thus follows straightforwardly from the PCC applying at LF to the trace of WH-in-situ.

In order to make this explanation work, we need to make two assumptions, both of which are independently motivated. First, we must assume, contrary to Aoun et al., that movement to COMP at LF does allow for the sort of binding relevant to the PCC (here, = proper government under the ECP). This, of course, is the null hypothesis. Second, however, we do have to assume one difference between S-structure and LF. At S-structure, multiply filled COMPs are not tolerated: in particular, a binder in a multiply filled COMP cannot bind its trace in the sense relevant for the PCC. This is why a subject trace cannot be bound from a COMP containing that or any other undeleted complementizer (cf. our discussion in Chapter Three, section 3.2, based on Pesetsky 1982). At LF, however, if (97)a is an acceptable representation, and does not violate the PCC, a binder in a multiply filled COMP may bind its trace in the sense relevant for the PCC. Thus, "multiply filled COMP phenomena" do not exist at LF. We stipulate this, without attempting an explanation.

If this is so, however, we correctly predict that (98) will be

grammatical, and not violate the PCC:

(98) [_{S₁}, [who_i^m] [e_i said [_{S₂}, that] [who_j^m bought what_k^m]]]]

At S-structure, who_j^m is saved from the PCC by what_k^m. At LF, however, if (98) has the following representation, we expect it to violate the PCC, which it does not:

(99) [_{S₁}, [who_i^m who_j^m what_k^m] [e_i said [_{S₂}, [that] [e_j bought e_k]]]]]

We must assume that (98) can have the representation in (100), where a trace in COMP binds the e_j in subject position:

(100) [_{S₁}, [who_i^m who_j^m what_k^m] [e_i said [_{S₂}, [e_j that] [e_j bought e_k]]]]]

Because multiply filled COMPs are not special at LF, there is no problem with the representation in (100), and the PCC is not violated.

Now we must ask why a similar trace in COMP cannot save (97)b.

That is, what is wrong with a representation of such a wide scope reading for (96)' like (101) below:

(101) [_S, [who_i^m who_k^m what₁^m] [e_i knows [_S, [where_jⁿ e_kⁿ] [e_b bought e₁ e_j]]]]]

Since e_k is bound from the COMP of (101), it does not violate the PCC.

(101) goes wrong as a result of the WH-criterion, which forces e_k in the COMP of S₂' to be cosuperscripted with where. The pair (where_jⁿ, e_kⁿ) will

be interpreted as a binary quantifier, which ranges over two variables.

But e does not assign a range at all. Furthermore, if e_kⁿ is forced to act like a quantifier, then who_k^m will quantify vacuously, violating well-

motivated restrictions on vacuous quantification discussed by Chomsky

(1981b). In any case, it is clear that there is some simple way to

exclude a trace in a [+WH] COMP. As a consequence, WH-movement must either land in a [+WH] COMP, or skip over it, not leaving any trace.

With this in mind, we may proceed to our demonstration that PCC effects at LF can be a matter of grammaticality in WH-in-situ sentences. We rely for this demonstration on the existence of certain WH-words which do not undergo the rule of absorption. The WH-word whether is the most convincing example of such a word, although Aoun et al. argue that why also belongs in this class.²²

That whether does not undergo the rule of absorption can be seen in (102), from Chomsky (1973, (254)). As Chomsky notes, (102) is only interpretable as an echo question, and is ungrammatical as a multiple interrogation:

(102)*I wonder whether Bill saw what

Let us say that what is special about whether is that it must have a unique superscript, triggering a violation of the WH-criterion if any other WH-phrase tries to move into a COMP containing whether. This story is for convenience only; the true story is probably that whether is free to be cosuperscripted with another WH in COMP, but that the resulting n-ary quantifier is not interpretable for semantic reasons.

As a consequence, (103) is unambiguous. The what in situ has only a wide scope interpretation, in which it is absorbed with who in COMP, shown in (104):

(103) [_{S₁}[who_i][e_i wonders [_{S₂}[whether] Bill saw what]]]

(104) [_{S₁}[who_i^m what_j^m][e_i wonders [_{S₂}[whetherⁿ] Bill saw e_j]]]

Some speakers find examples like (103) difficult to interpret. (Hankamer (1974, (55)) stars a similar example.) Nonetheless, I believe they are grammatical, and invite an answer like (105):

(105) John's been wondering whether Bill saw A, Sue's been wondering whether Bill saw B, Josie's been wondering whether Bill saw C,...

What is clear, however, is that there is no (non-echo) logical form for (103) besides (104). In particular, there is no narrow scope reading.

Since whether forces a wide scope reading, we now can test whether the PCC applies to the LF trace of WH-in-situ, over and above whatever PCC effects exist at S-structure. First, we consider a WH-in-situ in the subject position of a tensed S. Second, we neutralize the PCC effects that obtain at S-structure by adding a second WH lower in the tree. Finally, we force the subject WH-in-situ to undergo long movement in LF by placing whether in the nearest COMP. If the PCC applies to the trace of WH-in-situ at LF, we expect the result to be ungrammatical, as indeed it is:

(106) *who knows whether who bought what

The ungrammaticality of (106) is particularly striking when it is compared to (96), which we repeat:

(96) ?who knows where who bought what

(106) is ungrammatical because it can only have the LF representation in (107):

(107) [_S₁ [_{who}_i^m _{who}_j^m _{what}_k^m] [_e_i knows [_S [_{whether}] [_e_j [_{INFL} [_{VP} bought _e_k]]]]]]]

(107) violates the PCC at LF in the same way (108) violates it at S-structure:

(108)*[_{S₁} who_i [does John know [_{S₂} whether [_{e_i} [_{INFL}' [_{VP} bought this book]]]]]]

Notice that (106) cannot have an LF representation like (109), where WH-movement at LF has left a trace in the lower COMP. (109) does not violate the PCC, but does violate the WH-criterion. Here the special properties of whether are also violated: the WH-criterion forces the cosuperscripting that whether does not allow:

(109) [_{S₁} [who_i^m who_j^m what_k^m] [_{e_i} knows [_{S₂} [_{e_j}ⁿ whetherⁿ] [_{e_j} bought e_k]]]]]]

There is thus no way for (106) to satisfy the PCC at LF.

We consider our main point proven with example (106). What is important is that (106) does satisfy the PCC at S-structure: at S-structure there is no difference between (106) and (96). The lower WH-in-situ saves the subject WH-in-situ from the PCC in both cases. A theory in which all the crucial properties of WH-in-situ are derived at one level is at a loss to distinguish (106) from (96), as far as we can tell. Instead, this contrast shows clearly that we wish to distinguish the properties of WH-in-situ at two levels of representation. Consequently, we must assume that WH-in-situ has special properties at some level other than LF. The logical candidate is S-structure. We thus conclude that WH-in-situ has special properties at S-structure: in particular, the property of acting like an unbound empty category.^{23, 24}

2.5 Focus and Gapping

In this section we present some evidence suggesting a connection between some properties of the so-called "Gapping construction" that strongly recall WH-in-situ. The content of this section is speculative; an analysis of Gapping is briefly sketched, but no attempt is made to explain the many properties that have been observed in the literature. First, we begin by considering focused NPs in the context of the PCC.

2.5.1 Focus and the PCC

Besides QR and the LF movement of WH-in-situ to COMP, certain other constructions have been analyzed to involve movement in LF. In particular, Chomsky (1976) discusses the contrast in (110), where capitalization indicates stress:

- (110)a. the woman he_i loved BETRAYED John_i
 b. *the woman he_i loved betrayed JOHN_i

Chomsky notes that (110)b appears to be a case of the weak crossover phenomenon that we have already discussed in connection with WH-in-situ. In other words, the badness of (110)b, with stress and intended coreference as indicated, should be accounted for in the same manner as (111):

- (111) who_i [did the woman he_i loved betray e_i]

This is achievable if we posit an LF rule of FOCUS, which derives from (110)b a structure like (112):

- (112) for x_i = John [the woman he_i loved betrayed x_i]

Let us suppose, tentatively, that the rule of FOCUS moves the focused NP to COMP, so that it is formally identical to WH-movement. (Hendrick

and Rochemont 1982, followed by May (forthcoming), argue for adjunction to S'.)

A problem with this proposal that has been raised by Jaeggli (1980b) and by Chomsky (1981a, 282) concerns the absence of ECP effects with focused NPs. Supposing that (113)a has the LF representation of (113)b, the ECP, even when subsumed by the PCC, should rule the structure out:

(113)a. I wonder [how [JOHN solved the problem]]

b. [_S, [_{COMP} John_i] [I wonder [how [e_i solved the problem]]]]

Contrary to predictions, (113)a is grammatical.²⁵ As Chomsky notes, however, (following a suggestion of D. Sportiche), there is no reason to suppose that the focused NP in (113)a takes wide scope. If we relax our strictures governing [+WH] in COMPs, we might allow (113)a to have the LF representation in (114):

(114) I wonder [_S, [_{COMP} how John_i] [_S e_i solved the problem]]

Something else follows from the grammaticality of (113)a, given our discussion in the previous section: we know that the focused NP John cannot be treated as an unbound empty category, like WH-in-situ, at S-structure. If the focused NP had the properties of WH-in-situ, then it should be grammatical at S-structure, regardless of what happens at LF. This is expected, given that the focused NP appears to "refer". If the ability to "refer" is, as we speculated, what determines whether a path is formed or not, then we expect focused NPs, perhaps, not to create a path at S-structure.

On the other hand, there does not appear to be a focus construction in which a focused NP does not have the properties of an empty category at S-structure, showing "pure ECP", "ECP/Superiority" and -- perhaps -- "pure Superiority" effects. These are the focused NPs in the constructions usually derived by a rule of "Gapping". What appears to distinguish these focused NPs from the ones we have just considered is that these focused NPs undergo the rule of absorption -- our cosuperscripting rule AC. This suggests, in turn, that undergoing absorption might also be a prerequisite for a lexical category to generate a path at S-structure. We discuss this construction in the next section.

2.5.2 Gapping

A rule of Gapping is ususally assumed to account for sentences like (115):

(115) [John bought the book], and [Mary, the record]

Let us call the left-most conjunct in the English Gapping construction the full conjunct. The right-hand conjunct in (115) we may call the Gapping conjunct. There may be an unlimited number of gapping conjuncts:

(116) [John bought the book], [Mary, the record], [Tim, the plant],
and [Barry, the coffee grinder]

Gapping has been widely studied, and we will be accounting for only a few of its properties, some of which appear not to have been previously noticed. For reviews of the literature, as well as two of the more recent analyses, see Sag (1976) and Sjoblom (1980).

All the published proposals with which I am familiar proceed from

a basic intuition that the gapping conjuncts are full conjuncts with something missing. That is, these analyses propose that there is some level at which (115) is represented as in (117):

(117) [John bought the book], and [Mary bought the record]

In Ross (1970), Sag (1976), and elsewhere, structures like (118) are assumed to appear in the syntax, with the gapped structure of (115) generated by a deletion rule, the exact formulation of which has been a matter of some controversy. Williams (1977b) presents an argument for syntactic structures like (118):

(118) [John [_{VP} [_{V₁} bought] the book]] and [Mary [_{VP} [_{V₂} e] the record]]

At LF, a rule will copy the contents of V_1 into V_2 , yielding an LF representation like (117).

What both of these proposals have in common is a focus on what is "missing" from the gapping construction. For example, in a sentence like (115), both analyses take as their main task to define a correspondence between the phonologically real verb in the full conjunct and the phonologically absent verb in the gapping conjunct. These analyses thus have two properties: (1) the crucial correspondence between the full and gapping conjuncts is held to exist between the apparently missing material in the gapping conjunct and the present material in the full conjunct, and (2) the important "action" in the gapping construction takes place in the gapping conjuncts, where either deletion or copying takes place.

We wish to suggest a slight shift of emphasis. First, we wish to pay attention to what is present as well as to what is missing. In

particular, we consider the crucial correspondence between the full and gapping conjuncts to exist between the material that is phonologically real in the gapping conjunct, and its counterpart in the full conjunct. To introduce some terminology, let us call the underlined NPs in (119) the "correspondent", because they are the counterparts of the material in the gapping conjunct:

- (119)a. [John bought the book], and [Mary, the record]
 b. [John considers Mary a genius] and [Bill, Sue]
 c. [We will pay a brief visit to Masha], and [you, to Pasha]

We will claim that the correspondents in a gapping construction share properties with the WH-in-situ discussed in the previous sections.

Second, we will argue that the important "action" in the gapping construction occurs in both conjuncts. Specifically, we argue that gapping is a two-step process, with one step applying in each conjunct.

Let us consider some data first. It is probably the case, as argued by Kuno (1976), that the correspondents in a gapping construction are most acceptable when they are clause-mates:²⁶

- (120)a. [this man eats spaghetti], and [that man, macaroni]
 b.?[this man knows why you eat spaghetti], and [that man, macaroni]

The desired reading in (120)b can be forced more easily if the underlined correspondents in the full conjunct are given strong stress -- a significant fact, to which we return. Despite the contrast between (120)a and (120)b, we believe that (120)b is probably grammatical, although somewhat unacceptable. What is important to us is the subject/object asymmetry

between (120)b and (121), which we judge completely unacceptable:

- (121) *[this man knows why spaghetti makes you sick], and [that man, macaroni]

Similar contrasts can be seen in the following pairs:

- (122)a.?[this doctor thinks that I should buy tunafish], and [that doctor, salmon]

- b.*[this doctor thinks that tunafish will harm me], and [that doctor, salmon]

- (123)a.?[your sister prefers that the committee pick John], and [your brother, Bill]

- b.*[your sister prefers that John pick the committee], and [your brother, Bill]²⁷

The same facts obtain in French:

- (124)a.?[celui-ci sait pourquoi j'aime Marie], et [celui-là, Suzanne]
this one knows why I love Marie, and that one, Suzanne

- b.*[celui-ci sait pourquoi Marie t'aime], et [celui-là, Suzanne]
this one knows why Marie loves you, and that one, Suzanne

The contrasts in (120)b-(124) are immediately reminiscent of the "pure ECP effects" found with WH-in-situ:

- (125)a. who knows that you eat what

- b.*who knows that what makes you sick

We wish to suggest the same analysis. Specifically, suppose that Gapping correspondents, like WH-in-situ, have the properties of unbound empty categories at S-structure. (121) and the (b) sentences of (122)-(124) can then be explained as violations of the PCC, since the infinite path from the subject correspondent will overlap the path from INFL to COMP

in its clause, without either path containing the other. We will flesh this analysis out in discussion below.

The analogy with WH-in-situ extends further, when we consider the following contrasts. Correspondents in gapping structures may be WH-phrases in COMP:

(126) Bill asked [which books I gave to Mary], and [which records, to John]

The other correspondent may not be the subject of a tensed sentence:

(127) *Bill asked [which books Mary likes], and [which records, John]

Further examples:

(128)a. I know [which boy prefers spaghetti], and [which girl, macaroni]

b.*I know [which boy spaghetti appeals to], and [which girl, macaroni]

(129)a. he asked [where I bought the macaroni], and [where, the spaghetti]

b.*he asked [where the macaroni was sold], and [where, the spaghetti]

(130)a. je sais [quels livres tu as donnés à Paul], et [quelles thèses, à Marie]

'I know what books you gave to Paul, and what theses, to Marie'

b.*je sais [quels livres Paul a lu], et [quelles thèses, Marie]

'I know what books Paul read, and what theses, Marie'

In traditional analyses of gapping, these contrasts are often accounted for by stipulating that a non-null variable must intervene between the two correspondents. For us, however, these contrasts immediately recall the ECP/Superiority effects that we have discussed earlier:

- (131)a. Bill asked which records I gave to whom
 b. *Bill asked which records who bought

If gapping correspondents, like Wh-in-situ, generate infinite paths at S-structure, then (127) and the (b) sentences of (128)-(130) may be explained by the PCC, just as (131)b is.²⁸

There may even be some instances of gapping contrasts parallel to the "pure Superiority" effects discussed above, which we in turn reduced to Crossing effects. The contrast in (132) is not sharp, but may exist:

- (132)a. ?I want to know [who_i [e_i convinced John to visit me]] and
 [who, Mary]
 b. ?*I want to know [who_i you convinced John to visit e_i], and
 [who, Mary]

In (132)b, if John provokes an infinite path, this path will overlap the path between e_i and who, but neither will contain the other. The contrast in (132)a-b would thus be parallel to familiar contrasts like:

- (133)a. I want to know [who_i [e_i convinced which man to visit me]]
 b. *I want to know [who_i [you convinced which man to visit e_i]]

For the actual paths involved, the reader may refer to our earlier discussion of WH-in-situ, at (51).

What does a gapping correspondent have in common with a WH-in-situ? First, putting aside for a moment our concern with S-structure, let us look at the gapping construction at LF. The parallel between a reasonable LF representation of gapping and the LF representation of multiple interrogations has been pointed out by Sag (1976). To transpose Sag's suggestion into terms that we have discussed: he proposes that at LF all conjuncts of a gapping sentence are full conjuncts. He argues, in essence, that the correspondents in each conjunct are extracted and undergo absorption. For our purposes, we may ignore his arguments that the extraction operation involves set abstraction and treat the extraction as simple movement to COMP. Thus, a sentence like (134)a has an LF like (134)b, just as a sentence like (135)a has an LF like (135)b:

(134)a. [John bought the book], and Mary, the record

b. [_S, [_{COMP} John^k_i, book^k_j] [e_i bought e_j]] and [_S, [_{COMP} Maryⁿ₁, recordⁿ_m]
[e₁ bought e_m]]

(135)a. [who bought what]

b. [_S, [_{COMP} who^k_i, what^k_j] [e_i bought e_j]]

We use our cosuperscript notation to indicate absorption.

Our argument will proceed as follows. We have suggested that gapping correspondents resemble WH-in-situ in their S-structure properties. If gapping sentences like (134)a have the LF representations that (134)b implies, they resemble WH-in-situ at LF as well. First we will present some suggestive arguments in favor of an LF representation like (134)b, involving movement and absorption. Then we will sketch an analysis that can derive LF representations like (134)b from S-structures like (134)a.

We return to S-structure, and show some further respects in which gapping correspondents resemble WH-in-situ. Finally, in summarizing our analysis, we will point out some problems with gapping that we do not solve.

The suggestion that gapping involves variable binding and absorption at LF is plausible; one has the sense that gapping constructions involve a series of pairs which are "uniformly fitted" into positions in an open sentence in much the way the answers to multiple interrogations come in pairs.

One argument that the correspondents in gapping sentences undergo movement in logical form can be derived from the fact, which Sag notes, that they bear focal stress. We have already suggested, following Chomsky (1976), that focused NPs undergo movement at LF. Unless we suppose that this principle is somehow voided in gapping constructions, we must suppose that the correspondents in gapping sentences similarly undergo movement at LF. We may also demonstrate the existence of weak crossover effects in gapping sentences, somewhat below, where we discuss briefly "sloppy identity" phenomena.

If the correspondents do undergo movement in logical form, then by the WH-criterion they must undergo absorption, as implied in our example (134)b. An argument for absorption can be found in the constructions of (126)-(130), where one correspondent is a WH-phrase in COMP. Recall that some WH-phrases, particularly whether, do not undergo absorption:

(136)a. he asked [where [I bought what]]

b. *he asked [whether [I bought what]] (except as an echo question)

The same contrast reappears in the gapping construction:

- (137) a. he asked [where I bought the macaroni], and [where, the spaghetti]
b. *he asked [whether I bought the macaroni], and [whether, the spaghetti]²⁹

We may account for the contrast in (137) if we force gapping correspondents to move to COMP at LF and undergo absorption. The contrast thus argues for movement as well as for absorption, since there is no way to force absorption unless movement to COMP is forced.

Now let us sketch an analysis for gapping constructions. First, let us discuss S-structure. We have already argued that the correspondents in the full conjunct have the properties we attributed to WH-in-situ. They generate infinite paths at S-structure, and move to COMP at LF. We thus account for the pure ECP, ECP/Superiority, and Superiority -type effects that we have discussed. In the case of the pure ECP and ECP/Superiority effects, we assume crucially that the correspondent subject of a tensed sentence generates an infinite path, which overlaps the path between INFL and COMP and violates the PCC. With this in mind, it may seem that we incorrectly rule out simple gapping structures like (138):

- (138) I know that [_SJohn [_{INFL}, [_{VP} bought the book]]] and [Mary, the record]

If John generates an infinite path all by itself, it will violate the PCC in the manner we have just described. But John does not have to generate an infinite path all by itself. Recall that the cosuperscripting rule AC can apply (at least) to A-positions at S-structure. Thus, we may cosuperscript John and the book:

- (139) I know that [_SJohn_i^k [_{INFL}, [_{VP} bought the book_i^k]]] . . .

Having cosuperscripted the two correspondents, they will now form a united path, which will contain the path between INFL and COMP. In other words, the correspondent the book in (139) saves John from the PCC in the same way what saves who in a familiar WH-in-situ sentence like (140):

(140) which article proves that [_S^{who^k_i [_{INFL}, [_{VP} bought what^k]_j]]}

As for the gapping conjuncts, if we do not adopt a deletion rule in PF (which we argue against below), it is clear that their contents are "fragmentary" at S-structure. We propose that they consist of a bare COMP, which already contains at S-structure a pair of focused NPs. These bare COMPs might be related to the apparent bare COMPs found in the "sluicing" construction in English (Ross 1969; Van Riemsdijk 1978; Levin, forthcoming):

(141) somebody arrived, but I don't know [_S, [_{COMP} who]]

-- and by constructions like (142):

(142) I wonder [which books Mary stole] and [_S, [_{COMP} which records]]

Variants of the "bare COMP" proposal for sluicing are suggested by Van Riemsdijk and Levin, although other alternatives are clearly possible.³⁰ In particular, we leave open whether "bare COMP S's" also contain a null S at S-structure. (Such a null S might fail the identification requirement on empty categories discussed in Chapter Four.) For some reason, focused NPs can occur in such "bare COMPs" only when paired with another focused NP, as, we claim, in the gapping construction:

(143)a. *I know that [John bought the book] and [_S, [_{COMP} Mary]]

b. I know that [John bought the book] and [_S, [_{COMP} Mary, the record]]³¹

This, we claim, is where things stand at S-structure. The "action" at S-structure takes place in the full conjunct, where the correspondents generate infinite paths and may be cosuperscripted by AC. The gapping conjuncts are reduced structures, possibly containing bare COMPs. Notice that the PCC at S-structure already accounts for certain restrictions on gapping that otherwise must be built into the deletion or interpretive rule of gapping in the form of constraints on variables.

At LF, the action is in both conjuncts. Supposing that the focused NPs, which have undergone cosperscripting, have the properties of operators, they must move to an operator position, by Higginbotham's general principle (11), discussed above. Thus, from an S-structure like (144)a we derive a preliminary LF representation like (144)b:

- (144)a. I know [_S, that [_S John^k_i bought the book^k_j]] and [_S, [_{COMP} Mary, the record]]
- b. I know [_S, [_{COMP} that John^k_i the book^k_j] [_S e_i bought e_j]] and [_S, [_{COMP} Mary, the record]]

So far, our analysis has needed no rule or principle which is unique to gapping constructions, except, perhaps, for whatever allows the rather unusual pair of NPs in COMP. At this point, we introduce one rule that has not come up in earlier discussion. It is, in fact, a simple copying rule, similar, perhaps, to the "VP Rule" that Williams (1977b) motivates for VP deletion. Our suggestion for gapping is in part inspired by Williams' treatment of VP deletion. We propose that the S present in the full conjunct be copied under the S' in each gapping conjunct. This will yield (145):

(145) I know [_S, [_{COMP} that John_i^k the book_j^k] [_Se_i bought e_j]] and
 [_S, [_{COMP} Mary, the record] [_Se_i bought e_j]]

If (145) is the final LF, the empty categories in the gapping conjunct will violate the identification requirement, and the contents of COMP in the gapping conjunct will violate the WH Criterion. We therefore apply free indexing to yield the final LF representation in (146):

(146) I know [_S, [_{COMP} that John_i^k the book_j^k] [_Se_i bought e_j]] and
 [_S, [_{COMP} Mary_i¹ the record_i¹] [_Se_i bought e_j]]

(146) is, of course, an alphabetic variant of Sag's LF representation in (134)b.

Notice that the copying operation that derives (145) from (144) copies, as it must, the indices found on the empty categories in the full conjunct. This is no problem in (146), since neither S' c-commands the other, but may well explain the impossibility of gapping when one clause does c-command the other:

(147)a. *John bought the book [before Mary, the record]
 b. [_S, [_{COMP} John_i^k the book_j^k] [_Se_i bought e_j [_S, [before Mary_i¹
 record_j¹] [_Se_i bought e_j]]]]

If the Principle C of the Binding Theory is a true principle, preventing variables from being A-bound, then (147)b violates Principle C, since at least the first occurrence of e_i probably binds the second occurrence. Alternatively, we might appeal to the principle that prohibits vacuous quantification (Chomsky 1981b): John_i^k and book_j^k vacuously quantify over

the second occurrences of \underline{e}_i and \underline{e}_j .

Various phenomena of sloppy identity are explained and expected under this analysis of gapping, as they are in Williams' analysis of VP deletion, particularly if we suppose that pronouns and anaphors are replaced by variables at LF:

- (148)a. John_i gave a book to his_i father, and Mary, a record
 (= '. . . Mary_i gave a record to her_i father')
- b. [_S, [John_i^k a book_j^k] [\underline{e}_i gave \underline{e}_j to e_i 's father]] and
 [_S, [Mary_i^k a record_j^k] [\underline{e}_i gave \underline{e}_j to e_i 's father]]

At this point, we may demonstrate weak crossover phenomena in gapping constructions, a promised argument for movement at LF. Consider the following example:

- (149) [John gave her book back to Mary], and [Susan, to Bill]

(149) is a good sentence, but her cannot refer to Mary. That is, the following LF, of necessity showing "sloppy identity", if (bound) pronouns are replaced by variables, is impossible:

- (150) * [_S, [John_i^k to Mary_j^k] [\underline{e}_i gave her₁ book back \underline{e}_j]] and
 [_S, [Susan_i¹ to Bill_j¹] [\underline{e}_i gave her₁ book back \underline{e}_j]]

The weak crossover phenomenon thus affords more evidence that the LF of gapping sentences does involve a quantifier-variable structure.

Our analysis has some other consequences. First notice that one correspondent in a gapping sentence may be a WH-in-situ, where the other is a WH-in-COMP:

(152) I wonder [_S, when [John bought which book]] and [when, which record]

Movement to COMP of the WH-in-situ, followed by S-copying, proceeds normally, yielding the LF in (153):

(153) I wonder [_S, [when_i^k which book_j^k] [John bought e_j e_i]]

Non-correspondents, however, may not be WHs-in-situ when one correspondent is a WH in COMP. Compare (154)a with (154)b (only correspondents are underlined):

(154)a. I wonder [when John will give this book to Mary], and [when, to Sue]

b. *I wonder [when John will give which book to Mary], and [when, to Sue]

This is predicted if S, and not S', is copied in gapping constructions. In (154)b, both to Mary, a correspondent, and which book, a non-correspondent, will have to move to COMP. S-copying will yield (155):

(155) I wonder [_S, [_{COMP} when_i¹ which book_j¹ to Mary_k¹] [John will give e_j e_k
e_i]] and [_S, [_{COMP} when_i^m to Sue_k^m] [John will give e_j e_k e_i]]

In the second conjunct, e_j is unbound, and violates the identification requirement on empty categories. Thus, (154)b, with the LF in (155), is ruled out. Of course (154)b is acceptable as an echo question, in which case which book does not move to COMP.

We further predict, correctly, that WH-in-situ may be a non-correspondent if its scope is wider than the COMP of the conjuncts:

(156) who said [that John will give a book to whom] and [Bill, a map]

(156) is somewhat difficult to interpret, but is, I believe grammatical.

Its LF representation is:

(156) [who_i^k to whom_j^k] [\underline{e}_1 said [S , that John_1^n a book_m^n] [\underline{e}_1 will give \underline{e}_m \underline{e}_j]
and [S , [Bill_1^o a map_m^o] [\underline{e}_1 will give \underline{e}_m \underline{e}_j]]]

In (156), to whom binds a trace in both conjuncts at LF, violating no constraints. An appropriate answer to (156) (which may make the interpretation clearer) might take the form of (157):³²

(157) A said that John will give a book to B and Bill will give a map to B;
C said that John will give a book to D and Bill will give a map to D;
E said . . .

To summarize: we have assumed that the correspondents in the full conjunct of gapping constructions have the S-structure and LF properties that we attributed to WH-in-situ. The only specific principle we had to assume to handle gapping was a rule that copied the S of the full conjunct onto the gapping conjuncts. It appeared that this copying rule accounted for sloppy identity phenomena in gapping sentences, an anti-c-command restriction on gapping, and the distribution of WH-in-situ in the full and gapping conjuncts.

We present this analysis as a speculation because, while it accounts for certain aspects of gapping that have not previously been considered in this context, it fails to account for a number of other properties of gapping, and remains vague on a number of important points. Two obvious ones:

First, if the correspondents of the full conjunct generate infinite paths, and if the contents of the other conjuncts do not, or are not

cosuperscripted with the correspondents, why is there no violation of the PCC in the form of a Coordinate Structure Constraint? We may propose any of a number of ad hoc solutions, but the problem is unresolved.

Second, if the gapping conjuncts contain focused NPs that are in COMP at S-structure, why can't the rule of FOCUS apply freely in other contexts in the syntax. For example, why is (158) ill-formed as an S-structure?

(158) [_S, [John_i the book_j] e_i bought e_j]

Again, we may make some stipulation here, restating the facts, but no insightful answer is available.

Finally, various other properties of gapping remain unexplained, although they resist insightful treatments on other analyses as well. We choose from the list supplied by Sag (1976):

There appears to be a restriction, in the usual case, limiting the gapping conjuncts to two constituents apiece:

(159) *John persuaded Bill to see a movie, and [Harry, Mary, a TV show]

Sag notes a number of exceptions to this restriction, but the generalization does appear to hold in the majority of cases. Note that this cannot be a restriction on absorption, since, for example, more than one WH may occur in situ.³³

We do not explain why, on our analysis, the copying rule is "clause bounded":

(160) *John believes that Bill likes Mary, and Harry believes that [Joe, Sue]

Koster (1978b) speculates that (160) and similar examples might be ruled out by subjacency; but this seems doubtful, if gapping is derived by an LF copying rule. Other LF rules, such as WH-movement, appear to violate subjacency rather freely.

We do not explain why "preposition stranding" is impossible under gapping:

(161)a. [John sent the books to Susan] and [Harry, to Bill]

b.*[John sent the books to Susan] and [Harry, Bill]

We might argue that pied piping is obligatory at LF, although there are other considerations that suggest the opposite.

We also do not explain something that is very simple on a deletion analysis. The order of constituents in a gapping conjunct is the same as the order of correspondents in the full conjunct. Thus, in the sentence John kicked Bill, and Mary, Sue, we know that Mary, and not Sue, did the kicking in the gapping conjunct. This might follow from perceptual factors, or from some assumptions about c-command inside COMP, but does not find a straightforward explanation, in any case, on our analysis.

Finally, we do not explain, though we may certainly stipulate, the left/right asymmetry in the gapping construction:

(162)a. [John bought the books], and [Mary, the records]

b.*[Mary, the records], and [John bought the books]

Ross (1970) links this asymmetry to the Head First parameter, since in many SOV languages the paradigm of (162) is reversed (cf. Sjoblom 1980 for discussion). We have no account of this linkage, which, if true, is surely significant.

Despite these points on which our analysis is so far silent or vague, we believe that the ability of the PCC to explain certain aspects of the gapping phenomenon is significant. We have presented this analysis as a demonstration of an area in which the PCC may contribute to an explanation of an otherwise baffling complex set of phenomena. Ellipsis constructions like gapping have been unusually resistant to insightful explanation, the work of Williams (1978b) and Sag (1976) being exceptions.³⁴ These constructions appear to require elaborate rules, with stipulated non-null variables, ad hoc rule conditions and other non desiderata. Perhaps the approach to gapping outlined in this section will not meet all of these difficulties.

Gapping is also interesting, if we are correct, for the light it sheds on path theory. The contrast between gapping and simple focus constructions suggests a link between absorption and the phenomenon of infinite paths at S-structure. While we believe that path theory and the PCC offer much in the way of explanation, much remains unclear. If there is a link between absorption and infinite paths, it is worth investigating. What the link may derive from is, for now, unknown.

3.0 Subjunctive and W-Verbs

In this section we turn to a new topic, which has also attracted attention in the context of the ECP. Recent work by Picallo (1982) suggests that indicative and (at least some) subjunctive clauses differ with respect to some of the subject/object asymmetries we have been discussing. We will discuss some of her evidence, and the conclusions she draws within the context of the ECP. Her conclusions are interesting because, if true, they do not translate naturally into the PCC approach we have been developing.

Nonetheless, we will suggest that her observations can be interpreted in a slightly different way, which does suggest a PCC analysis of some explanatory power. Extending our analysis somewhat, we can account for some facts about NP movement from the complements of verbs like want that are not adequately explained in current theories.

In Chapter Two, we presented Kayne's evidence that Chomsky's (1981a) ECP and other similar conditions like the KECP applies at LF. The evidence came from the rule QR, which assigns scope to quantifiers by adjoining them to S at LF. In particular, Kayne noted that the particle ne acts as a scope marker for negative quantifiers like personne 'nobody' in French. Thus, (163)a has the LF seen in (163)b, and (164)a has the LF of (164)b:

- (163)a. je n'ai exigé [_S,qu'ils arrêtent personne]
I ne have required that they arrest nobody
- b. [_Spersonne_i [_Sje n'ai exigé [_S,qu'ils arrêtent e_i]]
- (164)a. j'ai exigé [_S,qu'ils n'arrêtent personne]
I have required that they ne arrest nobody
- b. j'ai exigé [_S,que [_Spersonne_i [_Sils n'arrêtent e_i]]]]

With this in mind, Kayne notes that the absence of a sentence like (163)a with personne in subject position can be blamed on the ECP (NIC, for him) applying at LF:

- (165)a. *je n'ai exigé [_S,que personne soit arrêté]
I ne have required that nobody be arrested
- b. [_Spersonne_i [_Sje n'ai exigé [_S,que e_i soit arrêté]]]]

(166)a. j'ai exigé [_S,que personne ne soit arrêté]

I have required that nobody ne be arrested

b. j'ai exigé [_S,que [_Spersonne_i [_Se_i ne soit arrêté]]]

(165)b is taken to violate the ECP in the same way a familiar CTP violation does:

(167) *l'homme [_S,que_i [_Sj'ai exigé [_S,que [_Se_i soit arrêté]]]]

Rizzi (1982) has shown that the same paradigm obtains in Italian, but in a slightly more complicated fashion. The Italian equivalent of personne, nessuno, also requires a scope marker, the negative particle non. When nessuno is a direct object, this scope marker may be in a higher clause, as in the French example (163)a. When nessuno is a preverbal subject, the scope marker may not be in a higher clause: this is Italian's equivalent of (165)a. On the other hand, when nessuno is a postverbal subject, the scope marker may once more be in a higher clause.³⁵

Picallo (1982) demonstrates the same paradigm in Catalan, for the Catalan equivalent of nessuno and personne: ningú. We draw our examples from her paper; the (b) examples are LF representations:

(167)a. no vull [_S,que tu parlis amb ningú]

NEG I want that you talk to nobody

b. [_S, [_Sningú_i [_Svull [_S,que tu parlis amb e_i]]]]

(168)a. no vull [_S,que ningú vingui]

NEG I want that nobody come

b. * [_S, [_Sningú_i [_Svull [_S,que e_i vingui]]]]

(169)a. no vull [_S,que vingui ningú]

NEG I want that come nobody

b. [_S, [_Sningú_i [_Svull [_S,que vingui e_i]]]]

This paradigm was taken by Rizzi to shed light on why languages like Italian (and Catalan) allow apparent violations of the CTP effect (the *that-trace filter) with syntactic WH-movement. Picallo gives the following Catalan example:

(170) qui creus que vindra

who do you think that will come

In a framework including the ECP, consideration of (170) alone leads naturally to the conclusion that in pro-drop languages like Catalan, unlike English, AGR is a proper governor, just like V. This would allow (170) to have the S-structure in (171):

(171) [_S,qui_i [creus [_S,que [_Se_i [_{INFL}, [_{VP}vindra]]]]]]

If AGR in INFL is a proper governor, then e_i does not violate the ECP. This conclusion is plausible particularly because the trace of WH at S-structure is not pronounced: we cannot tell whether the trace is in subject position. On the other hand, the position occupied by a quantifier subject to LF movement does contain phonologically realized material: in such a case we can tell for sure where the LF trace is. The paradigm in (167)-(168) shows that the trace of long movement cannot in fact be in the preverbal subject position in a tensed sentence, contrary to what the ECP predicts, if AGR in Catalan is a proper governor.

Rizzi thus concludes that AGR is not a proper governor, in a pro-drop

language as in English. Rather, apparent cases of long movement of a syntactic subject actually come from a post-verbal position, inside or adjoined to VP (following the distinction between ergative and non-ergative intransitives discussed in Chapter Two). In (169), we see that long movement of such a post-verbal subject is indeed possible at LF. We thus conclude that the S-structure for (170) is not that seen in (171), but rather (172) below:

(172) [_S,qui_i [_{creus} [_S,que [_S-_i [_{INFL}, [_{VP} vindra e_i]]]]]]

Picallo notes, however, that Rizzi's conclusion is actually somewhat stronger than the facts indicate. Significantly, all the examples that tend to show that AGR is not a proper governor involve complements in the subjunctive mood. Thus, what really follows Kayne's, Rizzi's and Picallo's paradigms in (163)-(169), in an ECP framework, is the conclusion that a subjunctive INFL does not contain a proper governor. In order to justify Rizzi's strong conclusion, we must ask whether indicative INFL does or does not contain a proper governor for the subject.

Although relevant evidence is somewhat difficult to obtain, Picallo suggests that indicative AGR does indeed contain a proper governor for the subject. First, however, we note that it is not possible to reconstruct the sentences of (167)-(169), where a scope-marking no is in a higher clause, with indicative complements. As Picallo notes, in indicative clauses, the "doubling" no obligatorily appears in the minimal clause that contains the negative word it doubles:

(173) *no crec [_S,que ve ningú]
 NEG I believe that comes nobody
 (IND)

Thus, any possible subject/object asymmetry in an indicative parallel to (167)-(169) will be overshadowed by the general impossibility of doubling a negation in an indicative clause with a no in a higher clause.

Picallo points out, however, that a wide-scope interpretation of ningú is marginally possible even when the "scope-marker" no that doubles it is in the same lower clause. Crucially, this marginal wide-scope reading is only available in indicative clauses, when ningú is subject:

(174)a. en Pere diu [_S,que [_Sningú no l'estima]
Peter says that nobody NEG loves him/her

b.??[_S, [_Sningú_i [_Sen Pere diu [_S,que [_Se_i l'estima]]]]]

(175)a. en Pere vol [_S,que [_Sningú no l'estimi]]
Peter wants that nobody NEG love him/her

b.*[_S, [_Sningú_i [_Sen Pere vol [_S,que [_Se_i l'estimi]]]]]

No such contrast is found when ningú is an object; post-verbal subjects share the properties of objects (Picallo, personal communication):

(176)a. en Pere diu [_S,que [la noia no estima ningú]]
Peter says that the girl NEG loves nobody
(IND)

b.??[_S, [_Sningú_i [_Sen Pere diu [_S,que [la noia estima e_i]]]]]

(177)a. en Pere vol [_S,que [la noia no estimi ningú]]
Peter wants that the girl NEG love nobody
(SUB)

b.??[_S, [_Sningú_i [_Sen Pere vol [_S,que [la noia estimi e_i]]]]]

In an ECP framework, the contrast between (174)b and (175)b suggests that INFL contains a proper governor for the subject if and only if INFL is indicative. Picallo also presents certain other contrasts involving

scope of quantification and focus, which tend to support the same conclusion. She relates this contrast to the feature specification of the constituents of INFL: [\pm tense] and [\pm agreement]. Specifically, she assumes the following specifications for the INFL of various clause-types:

- (178)a. INDICATIVE: [+tense] [+agreement]
- b. SUBJUNCTIVE: [-tense] [+agreement]
- c. INFINITIVE: [-tense] [-agreement]

Picallo argues that it is the value for the features [tense] that determines whether or not INFL can properly govern the subject: INFL is a proper governor if and only if [+tense].

If this argument is correct, it follows that we need no longer assume that all apparent violations of the CTP effect derive from extraction from a post-verbal position in or adjoined to VP. In indicative clauses, at least, long WH-movement from the preverbal subject position should be possible: (171) should be a possible S-structure for (170) in Catalan.

If Picallo's conclusions are correct, it is hard indeed to imagine a suitable treatment of her facts in a framework that replaces the ECP by the PCC. In the PCC framework, we have assumed that long movement from the subject position of an indicative sentence is ruled out due to the interaction of the path from the subject trace with the path created by TNS movement to COMP. If anything, we expect movement from the subject of a subjunctive to be freer, not more restricted than movement from the subject of an indicative clause, since no movement to COMP is indicated when TNS is [-tense]. Let us therefore consider some further facts about subjunctive and indicative clauses that suggest a different view of the phenomena that

Picallo has discussed.

First notice that the hypothesis that indicative INFL is a proper governor in Catalan implies that it is not a proper governor in English, since English show CTP effects in indicative as well as subjunctive clauses. Since English lacks free subject inversion, we can be fairly sure that the S-structures are those presented below:

(179)a. *who_i [did you say [_S, that [_Se_{-i} [_{INFL}, [_{VP} came at 6:00]]]]]

b. *who_i [do you desire [_S, that [_Se_{-i} [_{INFL}, [_{VP} come at 6:00]]]]]

If INFL is not a proper governor in English, we do not expect any contrast between the subject position of indicative and subjunctive clauses with respect to quantifier scope. Contrary to this prediction, such a contrast may well exist. The facts are very far from clear (as in Catalan, apparently), but appear to point to a contrast between indicative and subjunctive clauses of the type observed in Catalan. The contrast seems much clearer with only than with other quantifiers. In (180), the phrase only Bill is ambiguous in scope:

(180) I particularly desire [_S, that you visit [only Bill]]

(181)a. I particularly desire [_S, that [_Sonly Bill_i [_Syou visit e_{-i}]]]³⁶

(I particularly desire that Bill be the only person that you visit)

b. [_Sonly Bill_i [_SI particularly desire [_S, that [_Syou visit e_{-i}]]]]

(Bill is the only person about whom I think "you should visit him"; about your other visits I'm more indifferent)

This ambiguity appears to be missing in (182):

(182) I particularly desire [_S,that [_Sonly Bill visit you]]

(183)a. I particularly desire [_S,that [_Sonly Bill_i [_Se_i visit you]]

(I particularly desire that Bill be the only person to visit you)

b.*[_Sonly Bill_i [_SI particularly desire [_S,that [_Se_i visit you]]]]

(Bill is the only person about whom I think "he should visit you")

On the other hand, consider (184), which also shows a scope ambiguity:

(184) I said [_S,that you visited [only Bill]]

a. I said [_S,that [_Sonly Bill_i [_Syou visited e_i]]]

(I said that Bill was the only person who visited you)

b. [_Sonly Bill_i [_SI said [_S,that [_Syou visited e_i]]]]

(Bill is the only person about whom I said "you visited him";
if you made any other visits, I said nothing about them)

In our judgment, (185) probably shows the same ambiguity. The wide scope interpretation is perhaps less likely here than in (184), but seems more accessible than in the subjunctive example (183)b:

(185) I said [_S,that [only Bill] v'isited you]

a. I said [_S,that [_Sonly Bill_i [_Se_i visited you]]]

(=I said that Bill was the only person who visited you)

b.?[_Sonly Bill_i [_SI said [_S,that [_Se_i visited you]]]]

(Bill is the only person about whom I said "he visited you";
if anyone else visit you, I said nothing about it)

I do not propose to base a theory on these judgments, but I take them as suggestive.³⁷ Supposing that they are correct, they suggest that, even on an ECP theory, what distinguishes indicative from subjunctive clauses is not a question of whether the subject position is properly governed or not.

CTP effects clearly show that the subject of an indicative or subjunctive sentence in English is not properly governed from INFL. If this is so, we must ask anew, for English and for Catalan, two questions: (1) why is wide-scope quantification of the subject allowed in indicatives? (2) why is wide-scope quantification of the subject not allowed in subjunctives?

A glimmer of an answer may be found in French. It is unclear whether French shows the contrast between quantification into indicative and subjunctive clauses that we have claimed to exist in English. If there is any subject/object asymmetry in French, of the sort described by Kayne, it seems equally present in indicatives and in subjunctives (as predicted by Picallo). On the other hand, judgments seem to be highly insecure. For the sake of the argument, let us suppose that French also makes the distinction found in Catalan and English. We make this conjecture because many speakers make another contrast between indicatives and subjunctives that may be significant. Recall that French, like English, shows CTP effects. These surface in both indicative and subjunctive clauses:

(186)a. *la femme que_i je crois [_S, que [_Se_i est venu]]

the woman that I think that came

b. *la femme que_i j'exige [_S, que [_Se_i vienne]]

the woman that I require that come

Now recall that French allows an "escape" from CTP effects. If the complementizer que is replaced by qui, which we have hypothesized to bear a referential index (Pesetsky 1982; Chapter Three), the CTP violation in (186)a is nullified. Interestingly, the result is rather bad in subjunctive complements like that in (186)b:

(187)a. la femme que_i je crois [_S,qui_i [_Se_{-i} est venue]]

b.??la femme que_i j'exige [_S,qui_i [_Se_{-i} vienne]]

This difference is not explained by the hypothesis that subjunctive INFL cannot properly govern the subject trace in (187)b, since the indexed qui should be able to act as a proper governor. There might be ways to bring the paradigm of (186)-(187) together with the paradigms discovered by Picallo (perhaps qui is a proper governor only when [+tense] appears in COMP) in an ECP theory, but we shall suggest another approach.

Consider again (185)b, where a wide-scope reading is represented for a quantifier subject of an indicative clause. (185)b was given as an approximation at a logical form for (185). Suppose that the representation we gave there (and for its Catalan equivalent in (174)b) is not quite accurate. In particular, suppose that the subject is actually locally bound by a trace of the wide-scope quantifier in the nearest COMP. Recall our argument in section 2 that there is no prohibition against doubly-filled COMPs at LF:

(188) [_Sonly Bill_i [_SI said [_Se_{-i} that [_Se_{-i} visited you]]]]

If (188) is the actual LF for the wide-scope reading of (185), then neither the ECP nor the PCC is violated, assuming that the trace in COMP binds the subject trace.

Now suppose that the subject of the subjunctive sentence must also be bound from the nearest COMP, but cannot be, for reasons we have yet to explain. (189) is thus a failed LF for the wide-scope reading of (182)b:

(189) *[[_Sonly Bill_i [_SI particularly desire [_Se_{-i} that [_Se_{-i} visit you]]]]]

(191)a. $\{S_1 \text{ only Bill}_i [S_1 I [VP_1 \text{ desire } [S_2 \text{ that } [S_2 e_i [INFL_2' [VP_2 \text{ visit you}]]]]]]\}$
 (SUB)

b. $\{S_1 \text{ que}_i [S_1 j' [VP_1 \text{ exige } [S_2 \text{ que } [S_2 e_i [INFL_2' [VP_2 \text{ vienne}]]]]]]\}$
 (SUB)

Paths: (i) Between e_i and its binder in (191)a-b:

$\{S_2, S_2', VP_1, S_1, S_1'\}$

(ii) Between $INFL_2$ and V_1 (subjunctive):

$\{INFL_2', S_2, S_2', VP_1\}$

The two paths overlap, and violate the PCC. Now let us see that things get no better if e_i is locally bound by a trace in the COMP of S_2' . The path formed between these two traces satisfies the PCC, but the path from COMP of S_2' to the WH in (191)b or the quantifier in (191)a still violates the PCC:

(192)a. $\{S_1 \text{ only Bill}_i [S_1 I [VP_1 \text{ desire } [S_2 e_i^1 \text{ that } [S_2 e_i^2 [INFL_2' [VP_2 \text{ visit you}]]]]]]\}$
 (SUB)

b. $\{S_1 \text{ que}_i [S_1 j' [VP_1 \text{ exige } [S_2 \text{ qui}_i [S_2 e_i^2 [INFL_2' [VP_2 \text{ vienne}]]]]]]\}$
 (SUB)

Paths: (i) Between e_i^2 and e_i^1 or qui_i :

$\{S_2, S_2'\}$

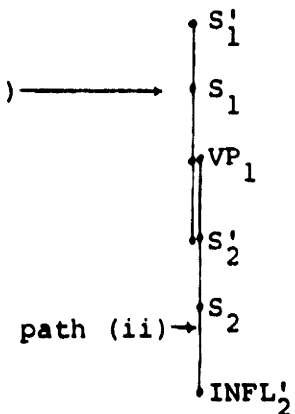
path (i) →

(ii) Between e_i^1 or qui_i and its binder in S_1' :

$\{S_2', VP_1, S_1, S_1'\}$

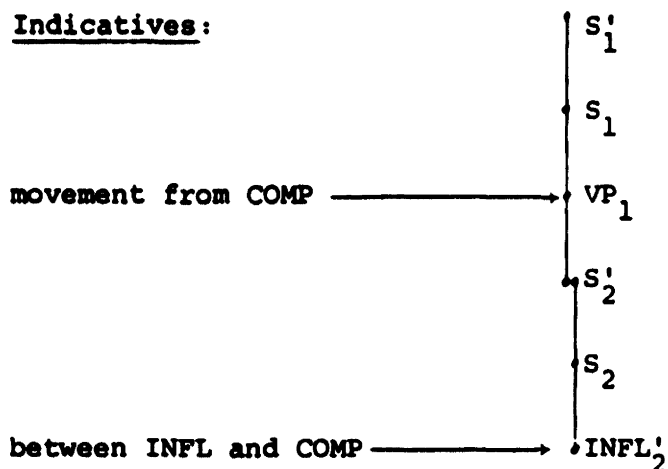
(iii) Between $INFL_2$ and V_1 (subjunctive):

$\{INFL_2', S_2, S_2', VP_1\}$



The two-member path (i) does not, of course, play any role in the ungrammaticality of (192)a-b. What violates the PCC is the interaction between the subjunctive path between INFL and VP and the path from COMP. As the schema to the right of the paths shows, no movement should ever be possible from the COMP of a subjunctive complement, since there is no eligible landing site in VP. Compare this to the situation that exists in indicatives, where the path from INFL ends with the S' node. The definition of overlapping allows movement from COMP in this case:

(193) Indicatives:



The PCC theory of subjunctives thus claims that the true generalization about movement from subjunctive and indicative clauses is different from the generalization drawn by Picallo. Picallo tacitly assumes that QR cannot apply from COMP. The contrast between QR from the subject of indicative and subjunctive clauses is thus seen as a direct contrast between permissible one-step long movement (indicatives) and impermissible one-step long movements (subjunctives). This analysis does not extend to the que/qui facts in French, where movement to COMP should not distinguish indicatives from subjunctives, on Picallo's analysis. On our analysis, QR may apply from COMP. Neither indicatives nor subjunctives allow one-step long

movement from subject position. Indicatives, however, allow this movement to take place in two steps: into the nearest COMP and then out. Subjunctive complements do not allow movement out of the nearest COMP: there is thus no way to move a subject out of a subjunctive clause. The one new assumption of this analysis is the path from subjunctive INFL to a higher verb.

This new assumption raises a number of questions and possibilities, one of which I will consider here, though not in detail. We might ask why there is a path between a subjunctive INFL and the verb that selects the subjunctive. One possibility is that such a verb assigns its θ -role specifically to the INFL of its complement, since it does select the tense features of that complement. This proposal is reasonable, although it raises questions about the locality of θ -role assignment, since this path crosses S', presumably a maximal projection.

Another possibility is that the higher verb acts in some way like the tense of the lower clause, fulfilling the role that TNS in COMP fulfills in an indicative clause. The tense concord that Pica!lo observes in subjunctive complements (discussed in detail by Salamanca 1982) certainly suggests some relation of this sort. Suppose that this is so. One might ask whether the verbs that take subjunctive complements have this property when they take other types of complements. In English, there are two classes of verbs that take subjunctive complements more or less productively: the verbs that Postal (1974) calls W-verbs (want, prefer, desire, . . .), and verbs of requesting (ask, beg, recommend, . . .). Let us look at the former class, which takes a range of complement structures that are of interest to us.

In a manner that demands an explanation (which we cannot offer), the W-verbs share a "family" of complements. There is a good deal of

variation among speakers on exactly which of these complements each W-verb may take. The crucial point is that a speaker of English recognizes the complements we will discuss as possible complements for all the W-verbs, and can distinguish "possible but not actual" selection from impossible selection. Thus, some speakers of English do not accept a for-infinitive immediately adjacent to want, or a subjunctive complement to want:

(194)a. *I want [for you to come earlier]

b. *I want [that John be elected president]

But it is probable that all speakers would recognize the superiority of (194)a-b to (195)a-b, where believe is followed by the same complements:

(195)a. *I believed [for you to come earlier]

b. *I believed [that John be elected president]

Thus, consider the following partial family of possible W-verb complements:

(196)a. Subjunctive Complements

I (prefer, desire, wish, %want, %yearn, %would like, %need, %mean, . . .) [that John be elected president]

b. for-infinitival Complements

I (prefer, %desire, wish, %want, yearn, would like, %need, mean, . . .) [for John to be elected president]

c. PRO-infinitival Complements

I (prefer, desire, wish, want, yearn, would like, need, intend, . . .) [PRO to be elected president]

d. Small Clause Complements

I (prefer, %desire, wish, want, *yearn, %would like, %need, %interd, . . .) [it clear from the start that we will not elect John]

We have omitted from our list one frequent type of W-verb complement, to which we return shortly.

Suppose it were the case that there is always a path between a W-verb and, perhaps, the first maximal projection internal to its complement. In an untensed, subjunctive S', this projection will be INFL', as above. If such a path also ran into the for-infinitival complements of such verbs, we would expect to find a subject/object asymmetry. This seems to be true:

(197) I strongly desire [_S,for you to elect no one]

a. I [_{VP}strongly desire [_S,for [_Sno one_i [_Syou to elect e_{-i}]]]]

b. [_Sno one_i [_SI [_{VP}strongly desire [_S,for [_Syou to elect e_{-i}]]]]]

(198) I strongly desire [_S,for no one to be elected]

a. I [_{VP}strongly desire [_S,for [_Sno one_i [_Se_{-i} [_{INFL}'to be elected]]]]]

b.*[_S₁no one_i [_S₂I [_{VP}₂strongly desire [_S₃,for [_S₃e_{-i} [_{INFL}'₃to be elected]]]]]]]

(198) contains the following paths:

(198)b.(i) Between INFL'₃ and V₂ (W-verb):

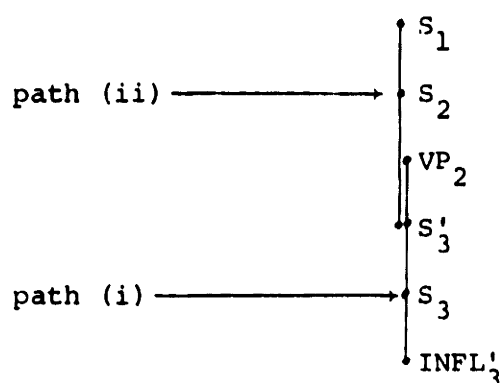
{INFL'₃, S₃, S'₃, VP₂}

path (ii) →

(ii) Between e_{-i} and no one_i:

{S'₃, VP₂, S₂, S₁}

path (i) →



A similar subject/object asymmetry holds under WH-movement (at least in non-"Ozark" dialects of English; cf. Chomsky and Lasnik (1977)):

- (199)a. John, who_i I strongly desire [for you to elect e_i]
 b. *John, who_i I strongly desire [for e_i to be elected]

Many of the W-verbs also allow complements that appear to be normal S'-deletion infinitivals:

- (200) I (prefer, %desire, wish, want, %yearn, would like, need, %intend, . . .) [John to be elected]

If these complements are normal S'-deletion infinitivals, at S-structure and at LF, we can explain the apparent absence of a subject/object asymmetry with wide scope quantification.³⁸ We find both (201) and (202) perfectly ambiguous:

- (201) I strongly desire [you to elect no one]
 (202) I strongly desire [no one to be elected]

We can allow the wide-scope reading for (202) if it has the structure seen in (203). Crucially, the PCC will not be violated, even if there is a path from the higher W-verb, if the complement is dominated by a non-maximal projection. The path from the subject will begin with VP, the first maximal projection that dominates it:

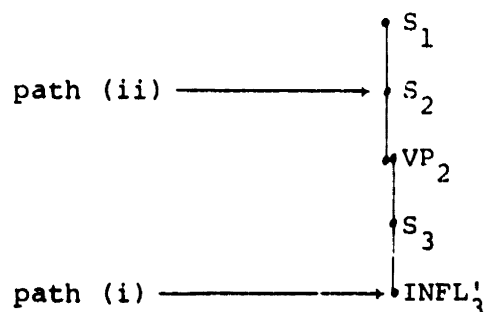
- (203) [_{S₁} no one_i [_{S₂} I [_{VP₂} strongly desire [_{S₃} e_i [_{INFL'₃} to be elected]]]]]]

Paths: (i) Between INFL'₃ and VP₂ (W-verb):

{INFL'₃, S₃, VP₂}

(ii) Between e_i and no one_i:

{VP₂, S₂, S₁}



In like fashion, there is no subject/object asymmetry with WH-movement, just as with ordinary S'-deletion infinitivals:

- (204)a. John, who_i I strongly desire [you to elect e_i]
 b. John, who_i I strongly desire [e_i to vote for Sally]

In view of this evidence, why not assume that some W-verbs may take an S'-deletion infinitival as a complement? One reason, provided by Chomsky (1981a), concerns passivization.³⁹ The subject of W-verb complements, unlike the subject of the complement to verbs like consider or believe, cannot be extracted from its clause by NP-movement:

- (205)a. there_i were believed [e_i to have arrived three men_i]
 b. *there_i were preferred [e_i to arrive three men_i]
 (206)a. Mary_i was considered [e_i to be less competent than Sue]
 b. *Mary_i was desired [e_i to be less competent than Sue]

Chomsky suggests that the apparent S'-deletion infinitival complements to W-verbs are actually for-infinitivals in which the complementizer for is optionally deleted in PF. The S-structures for (205)b and (206)b, according to this analysis, are actually as in (207):

- (207)a. *there_i were preferred [_S, for e_i to arrive three men_i]
 b. *Mary_i was desired [_S, for e_i to be less competent than Sue]

On this analysis, (205)b-(206)b would be ruled out at S-structure or LF by whatever principle rules out (207)a-b. For Chomsky (1981a), this principle might be the ECP, which applies to NP-trace, if the complementizer for is not a proper governor.

The for-deletion analysis has a number of other consequences that might be problematic, besides the wide-scope quantification and WH-movement data presented above. First, if there is a PF rule of for deletion, it must feed the Case filter, and must in any case feed a PF rule of S'-deletion. This can be seen in the following examples:

- (208)a. I prefer [(for) John to win]
 b. I prefer with all my heart [*(for) John to win]
 c. it is preferred [*(for) John to win]
 d. it is preferable [*(for) John to win]
 e. my preference [*(for) John to win]

Some of this paradigm is discussed by Lasnik and Freidin (1981). (208) shows that for may delete only when the subject that it governs before deletion is adjacent to a Case assigner after deletion. This is the case only in (208)a. In order to account for this generalization in a non-arbitrary way, we must assume that the Case filter and a rule of S'-deletion follow for deletion. Only if S'-deletion takes place can there be any case-related interaction between the W-predicate and the embedded subject in (208) (since only S'-deletion allows government into the clause). Only if the Case filter follows deletion can the Case filter then distinguish (208)a from (208)b-e. The paradigm of (208) is then brought together with (209):

- (209)a. I believe [that John has won]/[John to have won]
 b. I believe with all my heart [that John has won]/*[John to have won]
 c. it is believed [that John has won]/*[John to have won]
 d. it is believable [that John has won]/*[John to have won]
 e. my belief [that John has won]/*[John to have won]

The assumption that the Case filter applies at PF is somewhat problematical if it is derived from the θ -criterion, as suggested by Chomsky (1981a) (see Chapter Three, section 1). It seems unlikely that the θ -criterion applies at PF, if PF does not contain the empty categories that can satisfy the θ -criterion. Also, if the grammar contains deletion rules other than for deletion, it is doubtful that the Case filter should follow these rules. We have given an analysis for gapping that does not involve deletion. If we are wrong, however, simple gapping sentences show that the Case filter must apply before verb deletion:

(210) John bought the book, and [Mary -- the record]

There is no V to assign Case to the record, on standard analyses of gapping.

A more serious problem for the for-deletion analysis of W-verbs comes from small clause complements to these verbs. They show exactly the same restriction on passivization that we saw in infinitival complements:

(211)a. I desire [_P*you [_{PP}home by midnight]]

b. *you_i are desired [_P*e_{-i} [_{PP}home by midnight]]

(212)a. we_i need [_A*this car [_{AP}fully overhauled]]

b. *this car_i is needed [_A*e_{-i} [_{AP}fully overhauled]]

(213)a. I prefer [_A*it [_{AP}clear from the start that we will not elect John]

b. *it_i is preferred [_A*e_{-i} [_{AP}clear from the start that we will not elect John]]

This restriction does not, of course, obtain with the small clause complements of other verbs:

(214)a. I consider [_{A*}this car [_{AP}fully overhauled]]

b. this car_i is considered [_{A*}e_{-i} [_{AP}fully overhauled]]

Can we suppose that the small clause complements of W-verbs also have a for-complementizer? We cannot give a definite no as an answer. If these small clauses take a for-complementizer at S-structure, this complementizer is obligatorily deleted in all contexts. since no small clauses ever surface with for:

(215) *we need [for [this car fully overhauled]]

The need to make for-deletion obligatory argues against a for-deletion analysis of small clause complements to W-verbs. On the other hand, it is a general mystery in current theory why small clauses never take (overt) complementizers. The obligatoriness of for-deletion here might be tied to this general mystery.

Note, however, that the subjects of these small clause complements, like the subjects of the infinitival complements to W-verbs, show no restriction against wide-scope quantification or against long WH-movement. (216)a may have a wide-scope reading, and (216)b is grammatical:

(216)a. I desire [_{p*}nobody [_{pp}home by midnight]]

b. who_i do you desire [_{p*}e_{-i} [_{pp}home by midnight]]

Here too, these facts argue against an ECP-type analysis of the passivization facts in (211)-(213), just as wide scope and WH-movement from infinitivals argues against an ECP explanation of the passivization facts in (205)b-(206)b.

Assuming, then, that an S-structure for is not to blame for the restrictions on passivization from W-verb complements, what is to blame? We

speculate that the parallel between W-verbs and the TNS that appears in COMP of indicative clauses might go beyond the question of paths. Recall that we drew an analogy between the INFL/COMP relationship in indicatives and the INFL/W-verb relationship in subjunctives and for-infinitives. We suggested that the W-verb might function to provide the tense specification of its complement. Remember now that [+tense] TNS in COMP was subject to a general requirement, which we stated in Chapter Four ((137)-(138)):

(217)a. TNS appears in INFL at D-structure

b. [_{TNS} +tense] governs a complete proposition at LF

(218) Complete Proposition (definition)

A category X is a complete proposition iff for all predicates P dominated by X, X contains all members of the chains θ -marked by P.

Suppose that the following is true:

(219) A W-verb governs a complete proposition at LF

(219) prevents passivization out of the complement of a W-verb, wherever the element to be passivized is a member of a θ -marked chain. If we assume that the expletives it and there in (205)a-b and (213)a-b are linked in a θ -chain with the postverbal arguments in those sentences, then principle (218) covers all the cases of blocked NP-movement that we have considered.⁴⁰

Suppose we assume that W-verbs bear the TNS features for their complements and that this feature is [+tense], at LF. Suppose we extend (217)a to include small clauses, as in (220):

(220) TNS appears in the predicate phrase at D-structure

-- where "predicate phrase" includes INFL' and the XP of small clauses. Then the properties of W-verbs --both the path from W-verbs to a lower predicate phrase and the constraints on passivization -- follow from (220) and (217)b. (219) is a special case of (217)b.

This proposal has some plausibility. In many languages the notion of volition or intension characteristic of the W-verbs is represented by a special tense affix of some kind. One might suppose that the tense features in such a case would be subject to (217), just like any other tense features, and would move into COMP. In English, volition is lexicalized in the form of a verb. The theory we have outlined suggests that these volitional verbs, like the affixes in some other languages, are also subject to the generalization captured by (217).

Since this is a chapter of speculations, we will leave the story here. Additional work, we hope, will flesh out our suggestions for the treatment of W-verbs. In particular, it is important to discover whether non-W-verbs that take subjunctive complements, in English and in other languages, show any of the properties we have discussed in this section. If so, can these properties be derived in a similar way for these verbs? What we hope to have shown here is another area in which the PCC can capture an insight not capturable by the ECP.^{41,42}

4.0 Problems for the PCC

In the previous two sections, we have considered some domains into which the PCC appears to cast some light. In this section, we will briefly turn the lights off, and consider some domains in which the PCC either meets with problems or fails to capture a generalization that has been subsumed under Chomsky's or Kayne's ECP. In 4.1, we discuss the rule of QR, for

which the PCC creates some problems. These problems are resolvable, but the extent to which our resolutions create new problems or provide new insights awaits further work. In 4.2, we briefly consider preposition stranding, and discuss some facts of Kayne's which suggest that constraints on preposition stranding follow reasonably naturally from the ECP. The PCC, by contrast, seems to have nothing to say about this subject. Finally, in 4.3, we briefly discuss the cases in which NP-trace has been taken to be subject to the ECP. A number of these cases have been discussed in this study in different connections. We make some suggestions for dealing with the cases that we have not explained.

4.1 The PCC and QR

As we discussed in Chapter Two, May (1977) proposes that quantified expressions adjoin to S at LF by the rule QR. This proposal raises certain questions with respect to the PCC, briefly noted in footnote 36. Consider the S-structure in (221)b. Note that QR here moves a subject:

(221)a. [_S, [_S every good boy [_{INFL}, [_{VP} does fine]]]]

b. [_S, [_{S₁} every good boy_i [_{S₂} e_{-i} [_{INFL}, [_{VP} does fine]]]]]

We have argued very crucially that tensed S is a maximal projection. Thus, if every good boy in (221)a were replaced by something that created a path, this path would begin at S. A question arises in (221)b, however: of S₂ and S₁, which are maximal projections? This question is not significant in (221), where the subject undergoes QR. In (221)b, there is a path between INFL and COMP. Suppose S₁ and S₂ are maximal:

(221)b'. (i) Between INFL and COMP:

{INFL', S₂, S₁, S'}

(ii) Between e_i and every good boy_i:

{S₂, S₁}

Path (ii) runs from the first maximal projection dominating e_i to the first maximal projection dominating every good boy. The PCC is not violated.

Neither is the PCC violated if only S₂ is maximal:

(221)b''. (ii) Between e_i and every good boy_i:

{S₂, S₁, S'}

Nor if only S₁ is maximal:

(221)b'''. (ii) Between e_i and every good boy_i:

{S₁}

On the other hand, consider the result of adjoining a direct object to a tensed S:

(222)a. [S, [S_S Mary [INFL, [VP loves every good boy]]]]

b. [S, [S₁ every good boy_i [S₂ Mary [INFL, [VP loves e_i]]]]]

Suppose S₁ and S₂ are maximal, or only S₁ is maximal:

(222)' (i) Between INFL and COMP:

{INFL', S₁, S₂, S'}

(ii) Between e_i and every good boy_i:

{VP, INFL', S₂, S₁}

The PCC is violated. If only S_2 is maximal, then the PCC is satisfied:

- (222)'' (ii) Between e_i and every good boy $_i$:
 {VP, INFL', S_2 , S_1 , S'}

We conclude that if a node created by adjunction to a maximal projection is itself maximal, as in (220)', we must revise May's rule of QR. On the other hand, if the adjunction-created node is non-maximal, May's rule of QR so far works smoothly. Problems arise in subjunctives, as we will shortly see.

First, let us consider how May's rule might be altered if adjunction to S created a new maximal projection. Following Fiengo and Higginbotham (1981), we might adopt a generalization of QR, under which QR can adjoin a quantified constituent to any category, placing it as a sister to the specifier of that category. If SPEC of S is COMP, they note, then May's adjunction to S follows as a special case of their generalized QR. In (222), then, we would not need to assume that every good boy adjoins to S. Instead, it could adjoin to VP, and satisfy the PCC:

- (223) [_S [_S Mary [_{INFL} [_{VP₁} every good boy $_i$ [_{VP₂} loves e_i]]]]]

Paths: (i) Between INFL and COMP:

{INFL', S, S'}

(ii) Between e_i and every good boy $_i$:

{VP₂, VP₁}

This would lead to the prediction that a quantifier binding the subject of a tensed sentence should take scope over a quantifier binding the object. As is well-known, a quantifier binding a subject does prefer

to take wide scope, but narrow scope is also possible. For such cases, we might assume that a quantified object could adjoin to S'. If adjunction to S' is marked in some way, we could account for the preference of the subject to take wide scope.

Now suppose that adjunction to S does not create a new maximal projection. As we saw in (221)b'', this assumption raises no problems when an object is adjoined to an indicative S. We need not assume adjunction to VP or S'. On the other hand, if our analysis of subjunctives, presented in the previous section, is correct, very familiar problems arise when we try to adjoin the direct object to a subjunctive S. We assume that S' is always a maximal projection, although if it were not, we would solve the problem:

(224)a. [_{S₁} I [_{VP₁} prefer [_{S₂} that [_S Mary [_{INFL₂} [_{VP} marry no one]]]]]]

b. [_{S₁} I [_{VP₁} prefer [_{S₂} that [_{S_{2a}} no one_i [_{S_{2b}} Mary [_{INFL₂} [_{VP₂} marry e_i]]]]]]]]

Paths: (i) Between INFL₂ and V₁ (subjunctive):

{INFL₂[!], S_{2b}, S_{2a}, S₂[!], VP₁}

(ii) Between e_i and no one_i:

{VP₂, INFL₂[!], S_{2b}, S_{2a}, S₂[!]}

Obviously, the PCC is violated in (224)b, regardless of whether S_{2a} is a maximal projection. Perhaps the narrow scope reading of (224)a involves adjunction to VP; perhaps some other solution is available.

As things stand, there are a number of different ways to make the PCC compatible with the rule of QR, each one of which makes a number of predictions about possible scope relations. We will not try to sort out

the different possibilities here. We merely note the problem and the direction some solutions might take.

4.2 The PCC and Preposition Stranding

In many languages -- perhaps most, as claimed by Van Riemsdijk (1978)a -- prepositions may not be stranded by movement rules. Thus, the English sentence (225)a has no French equivalent (225)b:

- (225)a. the man that_i I spoke of e_i
 b. *l'homme que_i j'ai parlé de e_i

Kayne (1981a) has suggested that this fact be related to the ECP, by removing P, like AGR, from the roster of proper governors. Chomsky (1981a, 252) suggests that proper governors more generally are those categories that are either [+N] or [+V], thus eliminating P and possibly AGR.

On such a theory, how can we allow for preposition stranding in English? Kayne adopts a version of Hornstein and Weinberg's (1981) suggestion that preposition stranding is due to reanalysis of the stranded preposition with V. Since V is a proper governor, the ECP will not rule the structure out. This proposal was discussed in some detail in Chapter Three, section 4.1. We noted there that sentences like (225) show that reanalysis cannot be a precondition for preposition stranding -- at least when movement is to an \bar{A} -position:

- (226) what subject_i do you know who_j to [_{VP} talk to e_j about e_i]

The PCC did force reanalysis of talk and to, because of the presence of two gaps. Nonetheless, we argued that we could not also assume reanalysis of talk and about. The structure that would result from such a reanalysis

would violate anaphoric island conditions that are operative in other reanalyses of this sort.

If reanalysis is not a precondition for preposition stranding, then the difference between English and other languages must lie in the preposition itself. We might assume, for example, that English P is a proper governor, while French P is not, in a theory with the ECP. If this were true, however, we could not appeal to the ECP to rule out some cases of preposition stranding in English:

(227) *what war_i did you [during e_i] receive a wound

Recall that the ungrammaticality of (227) did seem to follow from the PCC.

In view of these problems with an ECP account of preposition stranding, it is perhaps not very disturbing that the PCC does not deal with the phenomenon. On the other hand, one still must ask what does explain the difference between languages like English and languages like French in this domain.

Particularly interesting is the fact, noted by Kayne, that left branch extractions are impossible from the object of a preposition in French; (228)a is as impossible as (228)b:

(228)a. *[combien de spectateurs]_i a-t-elle été applaudie par e_i
 how many of spectators was she applauded by

b. *combien a-t-elle été applaudie par [e de spectateurs]

Recall from Chapter Four, section 2.3, that such left branch extractions are possible from the object of a verb.

We might speculate that the relevant constraint is phonological, and not specifically syntactic. Suppose that French prepositions are

clitics, while English prepositions are not. We might rule out both (228)a and (228)b by the well-known constraint that rules out cliticization before an extraction site (King 1970), as in the following English examples:

(229)a. John has more pigs [O_i than [I have e_i]]

b. *John has more pigs [O_i than [I've e_j]]

(230)a. John has more pigs [O_i than [I have [e_i ducks]]]

b. ??John has more pigs [O_i than [I've [e_i ducks]]]

(231)a. the table is longer [O_i than [the door is [e_i wide]]]

b. *the table is longer [O_i than [the door's [e_i wide]]]

(Example (231)b is from Bresnan 1973, except for our (here) irrelevant assumption of an empty operator O.)

(230)-(231) show that left branch extractions, as well as extractions of major constituents, block cliticization processes in English. If French prepositions (and prepositions in all languages with the properties of French) are obligatorily subject to cliticization, then the general constraint exemplified in (230)-(231) might account for the restrictions on stranding. We have no explanation for this general constraint, of course.

Notice that English left branch extractions in Sub-C constructions may occur in PP:

(232) John talked to more policemen [O_i than [I talked to [e_i firemen]]]

On the other hand, Kayne (1981e) cites some contrasts involving extraction from gerunds that recall the French facts in (228)b. We pick examples that seem slightly sharper than Kayne's:

- (233)a. close tabs, which_i Bill remembers [Mary keeping e_i on you]
 b. close tabs, which_i Bill remembers [e_i being kept on you]
- (234)a. ?close tabs, which_i Bill talked to me [about [Mary keeping e_i on you]]
 b. *close tabs, which_i Bill talked to me [about [e_i being kept on you]]
- (235) Mary keeping close tabs on you, which_i Bill talked to me [about e_i]

If English prepositions are not clitics, then the contrast between (234)a and (234)b does not admit our phonological explanation. It is thus possible that there is some syntactic explanation for the differences between English and French, possibly involving the PCC, possibly involving some ECP-like constraint.^{43,44}

4.3 The PCC and NP Trace

At various points in this study, we have noted that NP-trace seems not to be subject to the PCC -- in particular, NP-trace does not generate a path. NP-trace shows neither Crossing effects nor Coordinate Structure Constraint effects. In this respect, the PCC does not subsume the ECP, which does govern the behavior of NP-trace. Let us examine briefly some of the instances in which the ECP has been held to govern NP-trace.

First, let us reiterate what we mean by the term "NP-trace". An NP-trace is an empty category with two properties: (1) it is locally A-bound, and (2) its antecedent does not bear an independent θ -role. An NP-trace is thus, by and large, the trace left by passive or raising operations, although indexing could yield an NP-trace, as discussed in Chapter Four, footnote (40). Thus, when we ask whether NP-trace is subject to the ECP or PCC, or to some other condition, we are asking what the conditions are on the occurrence of an empty category with properties (1)

and (2).

We know that NP-trace, being an anaphor under the Binding Theory, must be bound in some local domain -- in most cases, the domain of the nearest AGR or subject. The condition we have imposed on [+tense] TNS, that it govern a complete proposition, further restricts the domain in which NP-trace is bound to the domain of the nearest [+tense] TNS. In addition, the extension of this constraint to the complement of W-verbs further limits the distribution of NP-trace.

The ECP might play a crucial role, then, in exactly one case: when NP trace is bound within the domain of the nearest AGR, subject, TNS of W-verb. In English, there is one configuration in which NP-trace meets these criteria but is nonetheless ruled out: this is when NP-trace is the subject of an infinitive that does not undergo S'-deletion:

- (236)a. *John_i is impossible [_S, e_i to come]
 b. *John_i is impossible [_S, for e_i to come]
 c. *John_i is debatable [_S, when e_i to come]

We know that John is in a non- θ -position in each of these cases, since

(237)a-c are possible:

- (237)a. it is impossible [_S, PRO to come]
 b. it is impossible [_S, for Bill to come]
 c. it is debatable [_S, when PRO to come]

The empty categories in (236)a-c meet the definition of NP-trace. No principle we have discussed, except the ECP, rules out (236)a-c, and accounts for the contrast with (238):

- (238)a. John_i is likely [_Se_i to come]
 b. John_i is sure [_Se_i to come]
 c. John_i is believed [_Se_i to have come]

How can we rule (236)a-c out? A number of options are available, some of which have been discussed in the recent literature. One obvious approach is to maintain a residue of the ECP in the form of (239):

(239) RES(ECP)

NP-trace must be properly governed.

This RES(ECP) is clearly a stop-gap, however. A clear desideratum is the elimination of the distinction between government and proper government, which the PCC accomplishes for \bar{A} -bound empty categories, but not for NP-trace. Notice that we could reduce (239) to (240), were it not for (236)b:

(240) RES(ECP)'

NP-trace must be governed.

If (236)b is ruled out in an independent way, perhaps by Case Theory, then we can maintain (240). (240) has some plausibility; if NP-trace must be bound in its governing category, it might be the case that it must have a governing category in the first place.

Alternatively, we might adopt a suggestion of Aoun's (1982). Aoun suggests the following principle:

(241) S' breaks a chain.

Aoun's principle tells us that we cannot construct the chain (John, e) in

the sentences of (236), although we can in (238). An alternative version of Aoun's constraint might be (242), which strongly recalls our principles governing [+tense] TNS at LF;

(242) COMP must govern a complete proposition.

Principle (242) subsumes (241), since S', but not S, contains COMP.

Finally, a different alternative is proposed by Zubizarreta (1982). She suggests, for reasons developed in her work, that the D-structure of sentences like (236)a is as in (243), with the non- θ subject position and the complement cosuperscripted:

(243) \underline{e}^j is impossible [_S, John to come]^j

When John moves into the subject position, it inherits the superscript and transmits it to its trace:

(244) John_i^j is impossible [_S, e_i^j to come]^j

This structure is then ruled out by a constraint on referential circularity of Chomsky (1981a, 212), which plays an independent role in the Binding Theory:⁴⁵

(245) i-within-i Condition

*[_γ. . . δ . . .], where γ and δ bear the same index.

Since both S' and the empty category it contains bear the index j in (244), the i-within-i Condition rules the structure out.

(238)a has a similar D-structure:

(246) \underline{e}^j is likely [_S, John to come]^j

Since the superscript j is taken to adhere to the S' , deletion of the S' deletes the superscript, yielding the S -structure in (247):

(247) John_i^j is likely [$_{S} e_i$ to come]

(247) does not violate the i -within- i condition.

Any, all, or none of these proposals might be on the right track. Our goal in this brief discussion has been to present that portion of the ECP which the PCC does not subsume, and to suggest other accounts of the phenomena in question. Zubizarreta's proposal is particularly welcome, since it does not propose a special condition governing these cases. Aoun's proposal is rather natural, and overlaps in an interesting way with our proposals about TNS, suggesting a generalization that has not quite been captured. Once more, we will leave matters in an uncertain state.

5.0 Path Theory and the PCC

In the last three chapters, we have tried to justify a new subtheory of grammar: Path Theory. Path Theory consists of a definition of objects called paths, and a constraint on their interaction, the PCC. In addition, we have relied on certain auxiliary principles in the course of our discussion: a visibility condition on the definition of paths, needed to predict the distribution of expletive empty categories in languages like Italian; a condition on the scope of TNS; a "WH-criterion" governing absorption in COMP. Of these auxiliary principles, only the visibility condition is likely to form a part of Path Theory proper. The others presumably reflect principles of other subsystems, perhaps those governing relations at LF.

Path Theory's virtues fall into three categories. First, its

empirical coverage is rather wide, including much that falls under the ECP of Chomsky (1981a), and the KECP of Kayne (1981a, 1982). Second, it simplifies some of the other subtheories of grammar. Third, Path Theory is conceptually extremely simple: the PCC requires nothing more than a containment relation between overlapping paths; the definition of paths is also quite natural, although we have not resolved the question of what counts for the definition in an entirely satisfactory manner. In the preceding chapters, we have focused our attention largely on the first of the virtues: empirical coverage. In this section we wish to focus briefly on the second: the simplification of other subtheories afforded by the introduction of Path Theory. In our conclusions, we will survey once more the empirical coverage of Path Theory, and discuss in slightly more detail the third virtue.

5.1 Path Theory and the Definition of Government

We concentrate in this section on the influence of Path Theory on the definition of government. The definition of government, even if the ECP is eliminated as an independent principle of grammar, is relevant for a number of other subtheories. In particular, θ -theory, Case theory and Binding Theory may invoke the relation of government. In θ -theory, direct θ -marking takes place when the θ -marker governs the recipient of the θ -role. In Case theory, the same relation holds: A Case marker must generally govern the recipient of the Case features. Finally, one of the versions of the Binding Theory considered by Chomsky (1981a) relies crucially on the notion "governing category": the governing category for X is the minimal category that contains a governor for X and a subject or AGR (subject to certain refinements discussed earlier). Government thus

plays a central role.

Quite a number of definitions of government have appeared in the literature. As noted by Aoun and Sportiche (1982), who survey the definitions, they agree only on one "core" configuration. All definitions of government have as a result that Y governs NP in the configuration:



In their paper, Aoun and Sportiche defend the following definition of government (we adopt the slightly different presentation of Chomsky 1981a, 164):

(249) Consider the structure:

$[_{\beta} \dots \gamma \dots \alpha \dots \gamma \dots]$, where

(i) $\alpha = X^0$

(ii) where ϕ is a maximal projection, ϕ dominates α if and only if ϕ dominates γ .

In this structure, α governs γ .

The definition of Aoun and Sportiche is extremely simple and has, as they note, a number of correct empirical consequences. For example, if gerunds are dominated by NP, as their distribution seems to predict, and if PRO must be ungoverned, as argued by Chomsky (1981a), then Aoun and Sportiche's definition of government correctly predicts that PRO may be the subject of a gerund:

(250) I like $[_{NP} PRO [_{VP} [_{V} \text{reading}] \text{books}]]$

PRO is not governed by like because the maximal projection NP dominates PRO but not like. Similarly, PRO is not governed by the V reading, because this V is dominated by the maximal projection VP, which does not dominate PRO. PRO is thus ungoverned, as required.

On the other hand, PRO is governed in (251):

(251) *I like [_{NP} PRO [_N, [_N book]]]

The first maximal projection dominating PRO is also the first maximal projection dominating book. Thus, there is no maximal projection that dominates one and not the other. Hence, the N book governs PRO, and the sentence is ruled out.

As Aoun and Sportiche note, their definition of government is a symmetrized version of an equally natural definition of c-command:

(252) Consider the structure:

[_β . . . γ . . . α . . . γ . . .], where

(i) ϕ is a maximal projection, if ϕ dominates α , then ϕ dominates γ .

(ii) $\alpha \neq \gamma$

In this structure, α c-commands γ .

The only substantive difference between (249) and (252) lies in the strengthening of (252)'s conditional in (i) to (249)'s biconditional in (ii).

Now consider an alternative definition of government provided by Chomsky (1981a, 165):

(253) Consider the structure:

[β . . . γ . . . α . . . γ . . .], where

(i) $\alpha = X^0$

(ii) ϕ is a maximal projection, if ϕ dominates γ then
 ϕ dominates α .

(iii) α c-commands γ .

In this structure, α governs γ .

If we adopt Aoun and Sportiche's definition of c-command in (252), Chomsky's definition of government is equivalent to their definition in (249), since clause (iii) is simply the reverse of clause (ii), yielding a biconditional.

On the other hand, (253) offers the possibility of altering the definition of c-command to include a wider (or narrower) range of configurations in which one might want government to hold. This is the approach Chomsky takes. Interestingly, the crucial arguments that lead Chomsky to adopt a different definition of c-command (and of government) all involve the ECP:

(254) ECP: [α e] must be properly governed

(255) Proper Government: α properly governs β if and only if α governs β
 (and $\alpha \neq$ AGR).

Consider first extraction from a position adjoined to VP in a language like Italian, a construction we have discussed a number of times in this study:

(256) chi_i pensi [$_S$ che [$_S e$ [$_{INFL}$ [$_{VP}$ [$_V$ mangia] e_i]]]];
 who you think that is eating

If (256) does not violate the ECP, then the V mangia must govern e_i . Under Aoun and Sportiche's definition of government in (249), or under Chomsky's definition in (253), if Aoun and Sportiche's definition of c-command is assumed, e_i is not governed by V. Chomsky (p. 166) suggests that the definition of c-command be revised, as in (257):

(257) α c-commands β if and only if

(i) α does not contain β

(ii) Suppose that $\gamma_1, \dots, \gamma_n$ is the maximal sequence such that

(a) $\gamma_n = \alpha$

(b) $\gamma_i = \alpha^j$

(c) γ_i immediately dominates γ_{i+1}

Then if δ dominates α , then either

(I) δ dominates β , or

(II) $\delta = \gamma_i$ and γ_1 dominates β .

Given Chomsky's definition of government in (253) and of c-command in (257), mangia in (256) does govern e_i . The configuration is that of (258):

(258) [_{VP} [_{VP₂} V . . .] NP]

Clause (ii) of (253) is satisfied, because VP_1 is a maximal projection, and dominates both NP (γ) and V (α). Clause (iii) of (253) is satisfied, since V c-commands NP: VP_2 dominates V (α in (257)), while VP_1 , a projection of VP_2 , dominates NP (α in (257)).

Chomsky considers another case in which (257) might be crucial, involving phrases like something broken. Since this analysis involves some further complications, we will ignore it here. It is adjunction to

VP in Italian that provides the main motivation for the definition in (257).

Now suppose we eliminate the ECP as an independent principle of grammar. In particular, suppose that there is no requirement that the trace in (256) be governed or properly governed. If we capture the effects of the ECP with Path Theory, then what is important is that the path between e_i and che contain or be contained by the path between INFL and COMP. In (256), this requirement is clearly satisfied. Therefore, there is no reason to suppose that e_i is governed by mangia in the adjunction construction in (256). We are free to retain Aoun and Sportiche's original definition of government for this case.

This conclusion is welcome for independent reason. Consider the following contrast between English and Italian:

- (259)a. [e_i^1 [_{VP}sembra [_S e_i^2 [_{VP}₁ [_{VP}₂ mangiare] Giovanni_i]]]]]
- b.*[it_i [_{VP}seems [_S e_i [_{VP}₁ [_{VP}₂ to eat] John_i]]]]]

This contrast is presumably due to Case theory.⁴⁶ In Italian, e_i^1 , or possibly a clitic binding e_i^1 , as in Safir (1982), transmits Case to Giovanni. In English, it cannot transmit Case to John. Suppose, however, that John is governed by the V eat. Why can't eat assign objective Case directly to John? Notice that eat and John are string-adjacent, as well as being in a relation of government, under the definitions in (253) and (257).

Obviously, if eat does govern John in (259)b, we shall not want for plausible explanations of the ungrammaticality of (259)b. The point is simply that no explanations need be sought if eat does not govern John. In a theory without the ECP, we are free to make exactly this assumption.

Finally, recall that consideration of familiar contrasts like (260) in the context of the ECP forces a further complication of the definition of government to (261) (Chomsky 1981a, 250):

(260)a. *the man who_i I know [_S, that [_S⁻ⁱ [_{INFL}, [_{VP} eats]]]]

b. the man who_i I know [_S, e_i [_S⁻ⁱ [_{INFL}, [_{VP} eats]]]]

(261) Consider the structure:

[_β . . . γ . . . α . . . γ . . .], where

- (i) α = X⁰ or is coindexed with γ
- (ii) φ is a maximal projection, if φ dominates γ then φ dominates α.
- (iii) α c-commands γ.

In this structure, α governs γ.

As we discussed at length in Chapter Three, Path Theory allows us to eliminate the odd disjunction of structural government with government by coindexation.

Path Theory thus allows a return to the simple definitions of government and c-command. The simplest definitions seem to us to be those of Sportiche and Aoun. Notice, however, that Path Theory does not tell us that Chomsky's revisions, or any others that might be made, are wrong. Path Theory simply eliminates the main motivations for Chomsky's revisions. In as much as Chomsky's revisions are complications, this result is desirable. Nonetheless, since Path Theory does not even refer to government or to c-command (but cf. Chapter Three, footnote 17) the gate is open to further exploration of these concepts, a task which Path Theory might simplify in unforeseen ways. This is always to be hoped: introducing a new subtheory of grammar, a new principle, should not only shed its own light on the

nature of grammar, but it should also help other subtheories and principles to shine more brightly. This was the case in Chapter Two: Chomsky's ECP, which the PCC later replaced, illuminated the nature of θ -marking and categorial selection. The degree to which Path Theory achieves more of this second type of success remains to be discovered in full.

5.2 Conclusions

The first two chapters were concerned with two subtheories of grammar: θ -theory and the theory of categorial selection. We argued that the Projection Principle belonged only to θ -theory, and not to categorial selection, with a range of consequences in Russian and, we suggested, other languages. Our main tool in this undertaking was Chomsky's ECP, with Case Theory also playing a major role. We finally concluded, based on a reexamination of some arguments of Grimshaw's (1979, 1981), that categorial selection should be subsumed by a broader semantic theory. This discussion, besides clarifying c-selection and θ -marking, and indicating broad consequences for the theory of empty categories, sets a task for the theory of lexical semantics: to discover the broader principles that are reflected in syntax as c-selection. We left the matter at this point.

In Chapters Three, Four and Five, we switched our attention to the ECP itself. We argued in detail that the main effects of the ECP fall under Path Theory. We looked first at constructions containing a single \bar{A} -bound empty category at S-structure, and showed how CTP effects (covered by the ECP), Subject Condition effects and Crossing effects all followed from a simple definition of paths and an extremely simple constraint: the PCC. We were able to compare our account of these effects with others, drawing particularly on data from a mixed Head First/Head Last language to

support a paths account of Crossing effects over a linear account.

In Chapter Four, we adapted ideas of Kayne (1982) to give an account in Path Theory of constructions with multiple gaps. We showed first that Kayne's ideas, applied to our framework, explain automatically a series of exceptions to the Subject Condition, and some surprisingly clear judgments on complicated Crossing sentences. We then extended our theory to coordinate structures. Postulating a single path linking conjuncts to the conjunction derived almost without problems much of Ross's Coordinate Structure Constraint, along with the pattern of ATB exceptions discussed in the work of Williams. Even the apparently odd subject/object asymmetry in matrix clauses of ATB constructions derived immediately from Path Theory.

Finally, in Chapter Five, we showed how the PCC and Path Theory extends to a number of other problems. We considered in greatest detail the properties of WH-in-situ, and showed how Path Theory accounts for a full paradigm of effects that are covered variously by the ECP and the Superiority Condition. The properties of WH-in-situ were also detected in Gapping structures, suggesting a new analysis of these constructions. Once again, Path Theory explains these properties. In brief, we suggested a paths account of some differences between indicative sentential complements and subjunctive complements, which in turn motivated some speculations about the properties of W-verbs. Various other problems were considered, with no firm conclusions drawn.

In our judgment, the explanatory success of Path Theory and its simplicity, taken together, suggest that something like Path Theory is indeed a subsystem of Universal Grammar. Many issues remain to be resolved; many avenues have not yet been explored; some of our auxiliary proposals and assumptions still demand a more precise or accurate presentation.

Precisely because of this mixture of success and problems, we believe that the study of paths, a study begun by Kayne (1981a, 1982), will be a productive area for research.

At some point, it may be useful to pose the next question: why do paths and the PCC exist? We speculate that paths motivate a type of reanalysis inside a phrase marker. The existence of an empty category serves to "freeze" a portion of the phrase marker -- a notion due to the work of Culicover and Wexler (1977). If another empty category internal to this domain is bound outside of the domain, what will be violated are the normal "anaphoric island" conditions on binding into a word, discussed in Chapter Three. The PCC will thus reduce to the simplest principle of all -- the principle of lexical integrity. The development of this idea, which somewhat resembles the notion of "complete constituent" of Guéron (1981), awaits further work in Path Theory.

FOOTNOTES: CHAPTER FIVE

1. For some speakers, the distinction between who (nominative) and whom (objective) is particularly strong in situ. I will not indicate these distinctions, which are not relevant to our discussion.
2. This analysis derives ultimately from Bresnan (1970). See also discussion in Chomsky (1973).
3. It is also obligatory in relative clauses, which we do not discuss in this chapter.
4. We exclude from consideration WH-words in "echo questions". These have roughly the properties of names, and show none of the effects we will discuss below. On echo questions, see Hendrick and Rochemont (1982). We assume that echo WH-words do not undergo LF movement.
5. From (10), it only follows that WH may not be in an A-position. Nothing so far excludes adjunction to S, i.e. application of QR. This question is discussed by Aoun et al. The facts seem to suggest that WH must move to COMP, in order to receive an interpretation. There is no adequate explanation of this limitation, however, either in syntax or at LF.
6. In Higginbotham and May's formulation, absorption is a binary operation which can apply iteratively to pairs of quantifiers, and not to n quantifiers at once. This detail (which could belong to the interpretation of our cosuperscripting convention below) seems irrelevant to our discussion.

7. Perhaps we should replace "contains" by "immediately dominates" or some other expression of locality, to avoid ruling out:

(i) John, [_S' [_{COMP} whose desire that we ask where to go] we didn't understand e]

COMP of S' is [-WH], but contains [where].

8. Kayne actually argues that the trace of WH-in-situ obeys the Nominative Island Condition (NIC) of Chomsky (1980a). His NIC account translates easily into a framework in which the ECP subsumes some of the effects of the NIC. The same applies to a theory in which the PCC subsumes some of the effects of the ECP. In Kayne (1982), he argues, against Kayne (1981e), that the relevant effects are derivable at S-structure, but in a sense different from ours; cf. our fn. 20 below.

9. Kayne (1981e, fn. 8) following Kuno and Robinson (1972, 466), suggests that there is an improvement without that. (For Kuno and Robinson, the difference is between one star and two, although they generally dislike embedded WH-in-situ.) I do not find the effect very significant. More important, the same slight improvement when that is deleted can be found in the (a) sentences of (28)-(30), suggesting that the effect is independent of the PCC.

It should be noted throughout that judgments about WH-in-situ are often difficult, particularly because of interference from the echo interpretation. In all cases, it is helpful to make the judgments relative. Our alleged contrasts between "OK" and "*" may seem too bald; what is relevant, however, is that some contrast exists where we claim it does.

10. Superiority and the analysis we will be adopting both predict that (i) should be a grammatical outcome of (49) (cf. (44)b):

(i) about what topic did you tell who

Fiengo (1980) cites examples of this sort as grammatical. My own judgment is that they are not startlingly better than (50)b, with preposition stranding, but that there may be some improvement. The residual badness may suggest reconstruction of the preposition at LF, but I shall avoid this thorny topic here.

11. The observation that there is a residue of Superiority effects not subsumed by the ECP on any theory is independently noted by Hendrick and Rochemont (1982), who attribute the observation to P. Culicover.

Fiengo (1980) derives the Superiority Condition from a "priority Filter" applying at PF. This filter, in its final formulation (p. 126), is close to some of the formulations of the "Nested Dependency Constraint", discussed in Chapter Three.

12. This is not true in those languages in which all interrogative WH-phrases are in situ, e.g. the Chinese languages. In Chinese, the WH-word interpreted as the main interrogative word of a question is always in situ at S-structure, and has widest scope at LF, as discussed by Huang (1982; forthcoming). Huang (personal communication) indicates that there are no "Crossing effects" or other effects that we attribute to the PCC in Chinese questions. Perhaps the PCC does not apply in Chinese, as Huang suggests for the ECP. On the other hand, as N. Chomsky (personal communication) points out, Huang proposes a rationale for ECP failures in terms of the properties of INFL; no such explanation will explain the lack of Crossing effects.

13. It is important not to confuse this observation with Aoun et al's claim that WH-in-situ always takes widest scope in the clause to which it is assigned at LF. Our point is that it does not have to be assigned to the highest clause. For discussion of Aoun et al's observation, which motivates the claim that WH moves to COMP at LF, and does not undergo QR, see Huang 1982, Haik 1982.

14. I assume that locatives at least may occur under S. See our discussion of PPs under S in Chapter Three 4.2, where we suggest that they may also be included in VP. If where is taken to be in VP, then (56)b becomes a "Crossing effect" identical to (57)b. If it may be under S, then we must also rule (56)b out by referring to the path between INFL and COMP, as we do in the text.

If adverbials like where are always under S, it might be possible for the PCC to account for some contrasts observed by Huang (1982), although the attempt meets with various difficulties:

(i) ?what book_i do you know [_S, where_j John bought e_i e_j]

(ii) *where_j do you know [_S, what book_i John bought e_i e_j]

Sentence (ii) is grammatical only if where is construed with know and not with buy. As Huang notes, this contrast resembles familiar contrasts with extraction of a subject:

(iii) ?what book_j do you know [_S, who_j [e_j bought e_i]]

(iv) *who_j do you know [_S, what book_i [e_j bought e_i]]

If where is always under S, and not under VP, we explain the contrast between (i) and (ii) in the familiar way: the path from the trace of

where to its binder will overlap the path between INFL and COMP of the lower clause, but neither will contain the other.

We cannot as easily explain the same contrast in infinitivals, however:

(v) what book_i do you know [_S, where_j to buy e_i e_j]

(vi)*where_j do you know [_S, what book_i to buy e_i e_j]

We have alleged that there is no path from INFL to COMP in infinitives.

Also, like Huang, we cannot easily explain the absence of "that-trace" effects when where is extracted from a [-WH] clause. We expect a contrast between (vii) and (viii), but none seems to exist:

(vii) where_j did John say [_S, that Mary buried the hatchet e_j]

(viii) where_j did John say [_S, Mary buried the hatchet e_j]

Thus, it is not clear whether Huang's facts can be made to fall naturally under the PCC.

15. (58) probably does not interact with the binding theory, if we do not assume the definitions of variable and anaphor in Chomsky (1981a, 330):

(i) α is a variable if and only if it is locally \bar{A} -bound and in an A-position.

(ii) If α is an empty category and not a variable, then it is an anaphor.

Given these definitions, a WH-word in situ under (58) is an anaphor, and fails the Binding Theory requirement that it be locally A-bound in all cases. We must therefore assume the opposite convention, if anything

(as we did in Chapter Two):

(iii) α is an anaphor if and only if it is locally A-bound.

(iv) If α is an empty category and not A-bound, then it is a variable.

J.-R. Vergnaud points out that these definitions do not predict that arbitrary PRO is an anaphor -- necessary if the Binding Theory is to determine its distribution. Obvious modifications are possible (retaining (i)-(ii) and changing (ii) to read "if α is an empty category and not a variable or if α is free..."), which we will not consider in any detail here.

16. We have changed the tense of Huang's example slightly.

17. The (b) sentences of (63)-(64) are particularly acceptable with a light stress on the WH-in-situ.

18. The same effect shows up with the A/A cases of the Coordinate Structure Constraint applying to WH-in-situ in (60). These also get better when all conjuncts are WH-phrases:

(i) *I wonder who wrote which textbook and that novel

(ii) I wonder who wrote which textbook and which novel

Here the contrast may be explainable at LF. In (ii), but not in (i), the entire conjoined structure might undergo WH-movement together, as in:

(iii) *which textbook and that novel did he write

(iv) which textbook and which novel did he write

Extraction of just one conjunct in (i) will, of course, violate the A/A condition:

(v) *which textbook did he write and that novel

19. Kayne (1982) notes some extremely interesting cases in which multiple WHs do share a referential index, due to the predication rule in relative clauses. They show the subject/object asymmetries of ordinary WH-in-situ, but not the "saved" structures with a lower WH:

(i) ?John_i, [whose_i wife's desire that I visit whom_i]_j [_S e_j grew stronger] ... ✓

(ii) *John_i, [whose_i wife's desire that who_i visit me]_j [_S e_j grew stronger] ...

(iii) *John_i, [whose_i wife's desire that who_i consider whom_i foolish]_j [_S e_j grew stronger] ...

I have no judgments about Subject Condition or Crossing (pure Superiority) effects with this construction that have any clarity (they all seem fairly, but not absolutely, bad), but Coordinate Structure Constraint effects seem to show up; here "saving" effects are possible, due to the lack of c-command:

(iv) *John_i, [whose_i wife's [stories about Bill] and [pictures of whom_i]]_j [_S e_j had long ago started to bore us] ...

(v) ?John_i, [whose_i wife's [stories about whom_i] and [pictures of whom_i]]_j [_S e_j had long ago started to bore us] ...

20. This is what Kayne (1982) argues for also. Like us, he argues that S-structure is relevant for explaining the special properties of WH-in-situ. Unlike us, he proposes that the link between WH-in-situ and the WH in COMP with which it is absorbed is established at S-structure. This prevents his theory, as far as I can tell, from capturing the distinction between the properties of WH-in-situ at S-structure and at LF that we will argue to exist.

21. Actually, for many speakers, PCC effects are much weaker with "complex" WH-phrases like which man than with "simple" WH-phrases like who. This is noted in Chomsky (1973), who attributes the observation to R. Kayne, for Superiority effects, and by Chomsky (1981a) for pure ECP effects. We have no account for this difference (and cf. Fiengo 1980 for more discussion). We use "complex" WH-phrases in some of the examples in the following text to avoid excessive repetition of who and what, which tends to confuse judgments. Replacement of which man, etc, by who in situ may sometimes yield sharper contrasts. For some discussion, in an ECP framework, see May (forthcoming).

22. Hendrick and Rochemont (1982) note that whether has a number of other odd properties, and suggest that it is a [-WH] complementizer which can nonetheless introduce an interrogative. They observe, for example, that verbs that appear to take declaratives as well as interrogatives do not easily allow the interrogative to be introduced by whether:

- (i)a. *John wonders that Mary will plan a party
 - b. John wonders whether Mary will plan a party
- (ii)a. *John asked that Mary is planning a party
 - b. John asked whether Mary is planning a party
- (iii)a. John told Bill that Mary was planning a party
 - b.*John told Bill whether Mary was planning a party
 - c. John told Bill why Mary was planning a party
- (iv)a. John doesn't realize that Mary is planning a party
 - b. *John doesn't realize whether Mary is planning a party
 - c. John doesn't realize why Mary is planning a party

As they note, "affective" contexts make sentences like (iii)b better:

- (v)a. John didn't tell Bill whether Mary was planning a party
 - b. Did John tell Bill whether Mary was planning a party?

They suggest, following Grimshaw's (1979) theory of selection and sub-categorization (cf. Chapter Two), that the verbs in (i)-(iv) all select a *Q*, but that (iii) and (iv) subcategorize specifically for [+WH]. Given (v), I suspect that the distinctions are semantic in all cases, and do not refer to the distinctions between selection and subcategorization, which we have argued to be misleading. In any case, what is crucial to the argument presented below in the text is that COMPs containing whether not allow any WH-word to land there. The precise status of whether is not crucial.

23. This claim should be distinguished from Kayne's (1982) claim that WH-in-situ is, in effect, assigned scope at S-structure. See fn. 20.

24. A possible problem arises at LF with WH-in-situ inside a subject and inside a conjoined structure. Consider again (71)b and (73):

(71)b. who said that [for Bill to marry who] would be a surprise for whom

(73) which article proves which theorem and defends which theory

Clearly, these cannot have LFs in which only the smallest WH-phrase is moved:

(i) [who_i^m who_j^m who_k^m] [_S e_i said that [for Bill to marry e_j] would be a surprise [for e_k]]

(ii) [which article_i^m which theorem_j^m which theory_k^m] e_i [_{VP}₀ [_{VP}₁ proves e_j] and [_{VP}₂ defends e_k]]

Both representations violate the PCC. (i) should be a Subject Condition effect, and (ii) should violate the Coordinate Structure Constraint.

A reasonable solution for (71)b/(i) is to assume that the whole subject pied-pipes at LF:

(iii) [who_i^m [for Bill to marry whom]_j^m which theory_k^m] e_i said that e_j would be a surprise [for e_k]]

There are actual S-structures with such pied piping:

(iv) Mary, for Bill to marry whom would be a mistake

Such a solution is somewhat more dubious for (ii):

(v) [which article_i^m [prove which theorem and defend which theory]_j^m]
[_S e_i [_{INFL}' [_{VP} e]_j]]

This corresponds to no well-formed S-structure:

(vi) *this theory [prove which and defend which] this article does...

If (v) is impossible, I have no more constructive proposal to offer.

Other problems are presented by (vii), which is predicted to be worse than (viii) at S-structure:

(vii) John asked who saw what and wondered what had fallen on his head

(viii) John asked who saw Mary and wondered what had fallen on his head

If there is a contrast, it is not very clear. Notice that (vii) differs from other claimed violations of the Coordinate Structure Constraint, like (61)a-b in violating the PCC only at S-structure, where what in situ should create an infinite path. (61)a-b violate the PCC both at S-structure and at LF; since their ungrammaticality is weak to begin with, it is not surprising, perhaps, that (vii) should be undetectably ungrammatical.

25. Jaeggli (1980b), partially following work of Contreras, notes that certain focused NPs in the complement of factive verbs do show subject/object asymmetries. These verbs take subjunctive complements, a fact relevant to our discussion in the next section. We will suggest that the difference between the subjects of subjunctives and the subjects of indicatives under extraction amounts to the absence of an "escape hatch" through the nearest COMP. The facts cited by Jaeggli thus suggest an ECP (PCC) effect at LF: subjects can take wide scope if they can pass through COMP. It still appears to be the case that focused NPs in these sentences have no special properties at S-structure.

26. Also relevant here is Hankamer's (1973) "No Ambiguity Constraint" (NAC). This condition, perhaps perceptually based (see below), requires that the constituent interpreted as the first correspondent be the right-most eligible constituent. By this condition, the first correspondent in (120)b should be you and not this man. The reason we claim this constraint to be perceptual is that it can be fairly easily overridden. Conditions on strong parallelism, like our this man...that man, celui-ci...celui-là help to override the effects of the NAC quite effectively, in our judgment. Channon (1974) shows that in an inflected language like Russian, the NAC does not have any significant effects, presumably because surface Case marking disambiguates the relevant structures.

I suspect that what accounts for the question mark in (120)b is, as claimed in the text, a weak island effect of some sort. The apparent effects of the NAC are the results of trying to find interpretations for a gapping sentence that do not violate the island effect. Notice, as with many such island effects, that infinitives are weaker islands than tensed clauses. Compare (120)b with (i):

(i) [this man knows how to eat spaghetti], and [that man, macaroni]

(i) seems significantly better.

27. Better than (123)b is (i):

(i) [your sister prefers that John pick the committee], and [your brother, that Bill]

In (i), however, the correspondents are not your sister and John or that John. Rather, they are your sister and the S' that John pick the

committee, as in (ii):

- (ii) [your sister prefers that John pick the committee], and [your brother, that Bill pick the committee]

The second correspondent undergoes VP-deletion in the gapping conjunct, yielding (i). Cf. (iii):

- (iii) [your sister prefers that John pick the committee], and [your brother prefers that Bill]

VP-deletion does not entail do-support in a subjunctive clause, because subjunctives are [-tense]. Notice that (iv) is impossible, where the clause is indicative; its grammatical counterpart is (v):

- (iv) *[your sister said that John picked the committee], and [your brother, that Bill]
- (v) [your sister said that John picked the committee], and [your brother, that Bill did]

Also note that in French, which lacks VP-deletion, (i) is impossible:

- (vi) *[ta soeur préfère que Jean choisisse le comité], et [ton frère, que Pau.]

28. Additional evidence comes from Italian. Recall that free inversion in Italian can adjoin the subject to VP, placing it below INFL', and permitting long movement. An infinite path from such an adjoined subject will not violate the PCC, and the predicted contrasts with gapping occur:

- (i) ?so [quando [[_{VP} mangera Giovanni]]] e [quando, Maria]
 I know when will eat Giovanni and when Maria
- (ii) *so [quando [Giovanni [_{VP} mangera]]] e [quando, Maria]

(I am grateful to Luigi Burzio for these data.)

29. For some speakers, there is also a contrast with why, both in WH-in-situ constructions and in gapping constructions:

- (i) *he asked [why [I bought what]]
- (ii) *he asked [why I bought the macaroni], and [why, the spaghetti]

For these speakers, why provides an additional argument for the link between WH-in-situ and gapping constructions.

30. Interestingly, as noted by Ross, whether is ungrammatical in sluicing constructions:

- (i) perhaps somebody arrived, but I don't know who
- (ii) *perhaps somebody arrived, but I don't know whether

We have no way of relating this restriction to the restriction on whether in WH-in-situ and gapping constructions discussed earlier. No speaker appears to have a similar restriction with why, contra the gapping and WH-in-situ facts (cf. fn. 29):

- (iii) somebody arrived, but I don't know why

31. An obvious problem for the PCC arises, since the infinite paths emanating from the left, but not the right conjunct, should violate the PCC at S-structure. We discuss this problem at the end of the section, but no solution is offered. Ad hoc proposals are available, but nothing convincing.

32. Similar contrasts appear to arise in the case of VP-deletion, suggesting a similar analysis. Contra Williams and Sag, we consider

examples of the following type fully grammatical:

(i) I know [who Bill saw] and [who Bob did too]

(ii) I know [which books Bill gave to Mary] and [which books Bob did]

If we suppose that the VP from the first conjunct is copied into the second conjunct (much as in Williams' VP Rule), and assume our indexing conventions, these sentences are predicted to be grammatical. On the other hand, we do not accept:

(iii) *I know [which books Bill gave to whom] and [which books Bob did]

As in the gapping structures, whom in the first conjunct will move to COMP, leaving a variable in VP that will be copied into the second conjunct. In the second conjunct, this variable will be unbound.

VP-deletion, unlike gapping, probably does not crucially involve LF movement. We avoid the question, discussed by Williams, of whether VP-deletion and gapping belong to different grammatical components (discourse vs. sentence grammar).

33. Perhaps this is related to the fact that there appear to be no more than two slots in COMP at S-structure. There are languages that may allow doubly filled COMPs (e.g. some Scandinavian languages, popular French, etc.). but no languages with triply filled COMPs, as far as I know. (Sinicyn 1982a argues that some apparent examples in colloquial Russian do not in fact involve movement to COMP; cf. also Toman 1981.)

34. Our analysis is distinct from Sag's in assuming only a copy rule at LF, and not a deletion rule as well (in our terms); LF must "recover" the content of the deletion. The successful application of a recovery

procedure constitutes obedience to the Recoverability of Deletion Principle. We have argued for such a view ourselves (Pesetsky 1982, section 1.4), but it does not appear to be optimal. If, as seems likely, plausible conditions on LF duplicate all the proposed conditions on the deletion rule, there is no need both for a deletion rule and an LF "recovery" rule. Some of our conclusions from the discussion of ATB phenomena tend to suggest that deletion rules should be eliminated entirely from the grammar, or limited to very low level ellipsis or contraction rules, possibly local.

35. Rizzi and Kayne (1982) also note that wide-scope quantification of nessuno from preverbal subject position improves if a second negative word is added further down the tree, suggesting that nessuno may also have the properties we attribute to WH-in-situ at the beginning of this chapter. The possibility of sentences like (166)a and its Italian equivalents indicates that nessuno does not always have these properties. If nessuno or personne generated an infinite path, then it should never be acceptable in subject position unless licensed by a lower negation. Perhaps this is an instance of the link between absorption and infinite paths discussed earlier: when nessuno undergoes absorption it causes the formation of an infinite path; when it does not, it acts like a lexical category at S-structure and undergoes QR at LF.

36. The attentive reader will have noticed that the path between e_i and its binder in (180)a actually violates the PCC, since it overlaps the path between INFL and the higher verb desire, neither path containing the other. We deal with QR in the next section: we suggest that QR in

examples like this actually adjoins to VP or moves to COMP. This does not affect our discussion of examples like (182)b below, where the key violation of the PCC is unaffected by the choice of an adjunction site, beyond the question of clause-bound vs. long movement.

37. Interestingly, Kayne (1981a) also gives some examples of a subject/object asymmetry with wide scope in English, and chooses subjunctive sentences to exhibit them in:

(i) ?he's suggested that they write not a single term paper

(ii) *he's suggested that not a single term paper be written

I find the contrast weak, but am uncertain whether it is maintained in the indicative counterparts with wrote instead of write.

38. The subject/object asymmetry with for-infinitivals is also noted by Kayne (1981e), who gives examples with not a single along the lines of the examples cited in our footnote 37.

Postal (1974, 224) claims that wide scope is the only possibility in sentences like (202). In agreeing that wide scope is possible, his judgment does not conflict with our theory, but we find (202) ambiguous nonetheless. Perhaps Postal's choice of conditional mood for his example (I would prefer none of them to arrive late) creates a preference for wide scope. On the other hand, Postal claims that the corresponding examples with a for-infinitival is ambiguous, an intuition with which we strongly disagree.

39. The contrasts in (205)-(206), and an early version of the for-deletion analysis described below were discussed in Bresnan (1972, 153-177),

reporting joint work with H. Lasnik.

40. A problem arises if we push the analogy between TNS in COMP and W-verbs too far: we predict that PRO cannot appear in the complements to W-verbs. This counterfactual result follows from the "visibility" convention on empty categories that we introduced in Chapter Four, (155). If a higher W-verb contains a tense for its complement, then a PRO subject of its complement should generate an infinite path, beginning at S' and violating the PCC, contrary to our desires, with respect to the path between the W-verb and the INFL of its complement. We leave this as another problem for future research.

41. A number of our observations about subjunctives, suggesting that extraction is impossible from COMP of a subjunctive complement, could be accounted for under the ECP if we assume, with Stowell (1981), that empty categories in COMP are subject to the ECP, and if we assume that verbs that take subjunctive complements, unlike verbs that take indicative complements, do not properly govern into COMP (cf. Kayne 1981a, fn. 48). This explanation entails a rather specific asymmetry, however, since we must not rule out WH-movement from the subject position of small clause and (apparent) S'-deletion infinitival complements to W-verbs. If these complements have COMPs, then W-verbs must govern into these COMPs. If these complements do not have COMPs, then the asymmetry is less: W-verbs can govern across S, but cannot govern into COMP.

42. Another area in which the PCC yields insight with respect to subjunctives, at least in English, is in the area of complementizer deletion. In clauses with that complementizers, COMP appears to generate

an infinite path when that is missing. For this reason, that-deletion yields a subject Condition violation in tensed Ss, S'-deletion infinitivals and small clauses. The latter two cases have not been previously noticed, to my knowledge:

- (i) [*(that) John is a fool] shocks us
- (ii) ?I consider [[*(that) John is a fool] to prove my theory]
- (iii) ?I consider [[*(that) John is a fool] a national catastrophe]

The path from the empty COMP when that is deleted in (i)-(iii) violates the PCC with respect to the INFL-COMP path in (i) and the path from the θ -grid of consider in (ii)-(iii). In (iv), the path from an empty COMP would violate the PCC with respect to the path between the W-verb and the lower INFL; the ensuing violation is identical in form to that which results when WH-movement applies out of the COMP of a subjunctive clause:

- (iv) I prefer [*(that) [John come alone]]

The ECP can account for (i), as noted by Stowell (1981, 1982; cf. Guéron 1981), but not in any natural way for (ii) and (iii). (iv) could be accounted for under the theory outlined in footnote (41). The PCC, by contrast, seems to provide a unified account of that-deletion, with two puzzles: first, there is nothing wrong with that-deletion embedded in a subject, as in (v):

- (v) [that Mary claimed [(that) John is a fool]] shocked us

Second, the path from a deleted that is never united with any lower path from a deleted that -- i.e. there is no way to save that-deletion in (i)-

(iv). Perhaps this is because that lacks a referential index or superscript at S-structure.

For some discussion about why the absence of for does not yield similar effects in infinitives, cf. Stowell (1982). His analysis conflicts with some of our assumptions about the absence of tense of infinitives, but is worth trying to adapt to our framework, in view of the results obtained.

43. In this connection, one might ask about Dutch, where, as van Riemsdijk (1978a) shows, postpositions, but not prepositions, can be stranded under WH-movement. Since postpositions often differ in phonological shape from their prepositional counterparts, one might ask whether the postpositions are clitics in Dutch. Certainly it is the postpositions that cliticize to the WH-word waar 'what' and the demonstrative er 'that'.

44. Our explanation of preposition stranding, and of its absence in languages like French, is not "deep" -- that is, it does not follow from a deep principle of grammar like the ECP. We thus have no way of linking the presence or absence of preposition stranding to other phenomena, like the existence of S'-infinitivals, a connection proposed by Kayne (1981d). For some discussion of Kayne's analysis, see Chomsky (1981a, 294 ff.), and Stowell (1982). Many features of our discussion in this study depend on not accepting Kayne's analysis of S'-deletion infinitivals, under which they are actually full S's with a null complementizer.

45. This filter was proposed in yet another context by Vergnaud (1974).

46. Other accounts are available for the absence of inverted adjoined subjects in English. Note that (i) is also more impossible (as discussed in Chapter Two), while (ii) is possible:

(i) ??there seems to [_{VP} [_{VP} have eaten] a man]

(ii) there seems to [_{VP} have arrived a man]

If there transmits Case in (i), why can't it transmit Case in (ii)? Nonetheless, even if our argument against Chomsky's revision of the definition of government fails, we may still argue for Aoun and Sportiche's on the grounds of simplicity.

REFERENCES

- Anderson, S.R. (1977) "Comments on the Paper by Wasow," in P.W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Aoun, J. (1979) "On Government, Case Marking, and Clitic Placement," unpublished ms., MIT.
- Aoun, J. (1982) The Formal Nature of Anaphoric Relations, Doctoral dissertation, MIT.
- Aoun, J., N. Hornstein, and D. Sportiche (1982) "Some Aspects of Wide Scope Quantification," Journal of Linguistic Research 1.3.
- Aoun, J. and D. Sportiche (1982) "On the Formal Theory of Government," unpublished ms., MIT and UQAM [to appear, The Linguistic Review].
- Babby, L. (1975) "Impersonal Verbs and their Lexical Specification," Slavic and East European Journal 19, 182-187.
- Babby, L. (1980) Existential Sentences and Negation in Russian, Karoma, Ann Arbor.
- Babby, L. (1982) "Prepositional Quantifiers and the Direct Case Condition in Russian," unpublished ms., Cornell.
- Babby, L. and R. Brec̆it (1975) "The Syntax of Voice in Russian," Language 51, 342-367.
- Bach, E. (1977) "Comments on the Paper by Chomsky," in P.W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Baker, C.L. (1968) Indirect Questions in English, Doctoral dissertation, U. of Illinois at Champaign/Urbana
- Baker, C.L. (1970) "Notes on the Description of English Questions: the Role of an Abstract Question Morpheme," Foundations of Language 6, 197-219.
- Baltin, M. (1978) Toward a Theory of Movement Rules, Doctoral dissertation, MIT.
- Belletti, A. and L. Rizzi (1981) "The Syntax of "ne": Some Theoretical Implications," The Linguistic Review 1, 117-154.
- Bennis, H. (1980) "Coindexing and Complementizer-Trace Phenomena," paper presented at the 1980 GLOW Conference, Nijmegen, Netherlands; unpublished ms. U. of Amsterdam.
- Bergel'son, M.B. (1979) O Javlenii Sintaksičeskoj Redukcii v Russkom Jazyke, unpublished B.A. thesis (diplomnaja rabota), Moscow State U.
- Berwick, R.C. and A.S. Weinberg (1982) "Parsing Efficiency, Computational Complexity, and the Evaluation of Grammatical Theories," Linguistic Inquiry 13, 165-192.

Besten, H. den (1978) "Auxiliary Deletions and the Interplay between Local Deletive Rules and Filters," paper presented at the 1978 GLOW Conference, Amsterdam, Netherlands; unpublished ms., U. of Amsterdam.

Boguslavskij, I.M. (1979) O Sootnošenii Semantičeskix i Sintaksičeskix Svoystv Nekotoryx Ograničitel'nyx Častic v Russkom Jazyke, Doctoral dissertation (kandidatskaja dissertacija), Moscow State U.

Bordelois, I. (1974) The Grammar of Spanish Causative Complements Doctoral dissertation, MIT.

Borer, H. (1979) "Empty Subjects and Constraints on Thematic Relations," in J.T. Jensen, ed., Cahiers Linguistiques D'Ottowa 9 [Proceedings from the Tenth NELS Conference], University of Ottawa.

Borer, H. (1982) Parametric Variation in Clitic Constructions, Doctoral dissertation, MIT.

Borras, F.M. and R.F. Christian (1971) Russian Syntax, Second Edition, Oxford University Press, Oxford.

Bouchard, D. (1982), Doctoral dissertation, MIT.

Bresnan, J. (1970) "On Complementizers: Toward a Syntactic Theory of Complement Types," Foundations of Language 6, 297-321.

Bresnan, J. (1972) Theory of Complementation in English Syntax, Doctoral dissertation, MIT.

Bresnan, J. (1973) "Syntax of the Comparative Clause Construction in English," Linguistic Inquiry 4, 275-344.

Bresnan, J. (1975) "Comparative Deletion and Constraints on Transformations," Linguistic Analysis 1, 25-74.

Bresnan, J. and J. Grimshaw (1978) "The Syntax of Free Relatives in English," Linguistic Inquiry 9, 331-392.

Brody, M. (1982) "Anaphoric Dependencies and Conditions on Domains," paper presented at the 1982 GLOW Conference, Paris; unpublished ms., University College, London.

Burzio, L. (1981) Intransitive Verbs and Italian Auxiliaries, Doctoral dissertation, MIT.

Carlson, G. and T. Roeper (1980) "Morphology and Subcategorization: Case and the Unmarked Complex Verb," GLOT-Lexical Grammar, 123-164.

Channon, R. (1974) "The NAC and How Not to Get it," in E. Kaisse and J. Hankamer, eds., Papers from NELS V, Harvard.

Chomsky (1955-1977) The Logical Structure of Linguistic Theory, Plenum, New York.

- Chomsky, N. (1964) Current Issues in Linguistic Theory, Mouton the Hague
- Chomsky, N. (1965) Aspects of the Theory of Syntax, MIT.
- Chomsky, N. (1968) Language and Mind, Harcourt Brace Jovanovich, New York.
- Chomsky, N. (1970) "Remarks on Nominalization," in R.A. Jacobs and P.S. Rosenbaum, eds., Readings in English Transformational Grammar, Ginn, Waltham, Massachusetts [also in N. Chomsky (1972) Studies on Semantics in Generative Grammar, Mouton, the Hague].
- Chomsky, N. (1973) "Conditions on Transformations," in S. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle, Holt, Rinehart and Winston, New York [also in Chomsky (1977b)].
- Chomsky, N. (1976) "Conditions on Rules of Grammar," Linguistic Analysis 2, 303-351 [also in Chomsky (1977b)].
- Chomsky, N. (1977a) "On WH-Movement," in P.W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Chomsky, N. (1977b) Essays on Form and Interpretation, Elsevier North-Holland, New York.
- Chomsky, N. (1980a) "On Binding," Linguistic Inquiry 11, 1-46.
- Chomsky, N. (1980b) Rules and Representations, Columbia, New York.
- Chomsky, N. (1981a) Lectures on Government and Binding, Foris, Dordrecht.
- Chomsky, N. (1981b) Some Concepts and Consequences in the Theory of Government and Binding [published 1982, Linguistic Inquiry Monograph 6].
- Chomsky, N. and H. Lasnik (1977) "Filters and Control," Linguistic Inquiry 8, 425-504.
- Chomsky, N. and H. Lasnik (1978) "A Remark on Contraction," Linguistic Inquiry 8, 268-274.
- Chvany, C. (1975) On the Syntax of BE-Sentences in Russian, Slavica, Cambridge.
- Comrie, B. (1973) "Clause Structure and Movement Constraints in Russian," in C. Corum, T.C. Smith-Stark, and A. Weiser, eds., You Take the High Node and I'll Take the Low Node: Papers from the Comparative Syntax Festival, Chicago Linguistic Society, Chicago.
- Corbett, G.G. (1979) Predicate Agreement in Russian, U. of Birmingham, England.
- Crockett, D. (1976a) "The Distributive po in Contemporary Russian," in T.F. Magner, ed., Slavic Linguistics and Language Teaching, Slavica, Cambridge.

- Crockett, D. (1976b) Agreement in Contemporary Standard Russian, Slavica, Cambridge, Massachusetts.
- Culicover, P.W. and K. Wexler (1977) "Some Syntactic Implications of a Theory of Language Learnability," in P.W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Emonds, J. (1977) "Comments on the Paper by Lightfoot," in P.W. Culicover, T. Wasow and A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Engdahl, E. (1981) "Parasitic Gaps," paper presented at Sloan Workshop on the Processing of Unbounded Dependencies, U. of Massachusetts, Amherst; unpublished ms., U. of Massachusetts, Amherst.
- Farmer, A. (1980) On the Interaction of Morphology and Syntax, Doctoral dissertation, MIT.
- Fiengo, R. (1974) Semantic Conditions on Surface Structure, Doctoral dissertation, MIT.
- Fiengo, R. (1977) "On Trace Theory," Linguistic Inquiry 8, 35-61.
- Fiengo, R. (1980) Surface Structure: the Interface of Autonomous Components, Harvard.
- Fiengo, R. and J. Higginbotham (1981) "Opacity in NP," Linguistic Analysis 7.4.
- Fodor, J.D. (1978) "Parsing Strategies and Constraints on Transformations," Linguistic Inquiry 9, 427-473.
- Franks, S. (1982) "Case and Control in Polish," paper presented at the Cornell Conference on Government-Binding Theory, July 1982; unpublished ms., Cornell.
- Freidin, R. (1978) "Cyclicity and the Theory of Grammar," Linguistic Inquiry 9, 519-549.
- Freidin, R. and L. Babby (forthcoming) "On the Interaction of Lexical and Structural Properties: Syntactic Case Assignment in Russian," unpublished ms., Cornell and McGill.
- Freidin, R. and H. Lasnik (1981) "Disjoint Reference and WH-trace," Linguistic Inquiry 12, 39-53.
- Gazdar, G. (1981) "Coordination," Linguistic Inquiry 12, 155-184.
- George, L. (1980) Analogical Generalizations of Natural Language Syntax, Doctoral dissertation, MIT.
- Goodwin, W.W. (1894) A Greek Grammar, Macmillan, London.

- Greenberg, J. (1963) "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements," in J. Greenberg, eds., Universals of Language, MIT.
- Grimshaw, J. (1979) "Complement Selection and the Lexicon," Linguistic Inquiry, 10, 279-426.
- Grimshaw, J. (1981) "Form, Function, and the Language Acquisition Device," in C.L. Baker and J. McCarthy, eds., The Logical Problem of Language Acquisition, MIT.
- Groos, A. and H. van Riemsdijk (1981) "Matching Effects in Free Relatives: A Parameter of Core Grammar," in A. Belletti, L. Brandi, and L. Rizzi, eds., Theory of Markedness in Generative Grammar, Scuola Normale Superiore di Pisa, Pisa.
- Grosu, A. (1973) "On the Nonunitary Nature of the Coordinate Structure Constraint," Linguistic Inquiry 4, 88-91.
- Guéron, J. (1981) "Logical Operators, Complete Constituents, and Extraction Transformations," in R. May and J. Koster, eds., Levels of Syntactic Representation, Foris, Dordrecht.
- Hačk, I. (1981) "On Clitic en in French," Journal of Linguistic Research, 2.1.
- Hačk, I. (1982) "Indirect Binding," paper presented at the 1981 GLOW Conference, Paris; unpublished ms., MIT.
- Haiman, J. (1974) Targets and Syntactic Change, Mouton, the Hague.
- Hale, K. (1979) "On the Position of Walbiri in a Typology of the Base," unpublished ms., MIT [distributed by Indiana University Linguistics Club].
- Halle, M. (1980) "Formal vs. Functional Considerations in Phonology," distributed by Indiana University Linguistics Club [also appeared in A Festschrift for Oswald Szemerényi, John Benjamins, Amsterdam].
- Hankamer, J. (1973) "Unacceptable Ambiguity," Linguistic Inquiry 4, 17-68.
- Hankamer, J. (1974) "On WH-indexing," in E. Kaisse and J. Hankamer, eds., Papers from NELS V, Harvard.
- Harlow, S. (1981) "Government and Relativization in Celtic," in F. Heny, ed., Binding and Filtering, MIT.
- Hendrick, R. (1979) On Nesting and Indexical Conditions in Linguistic Theory, Doctoral dissertation, UCLA.
- Hendrick, R. and M. Rochemont (1982) "Complementation, Multiple WH and Echo Questions," unpublished ms., UC Irvine and U. of North Carolina at Chapel Hill.

- Higginbotham, J. (1980) "Anaphora and GB," in J.T. Jensen, ed., Cahiers Linguistiques d'Ottawa 9 [Proceedings from the Tenth NELS Conference], University of Ottawa.
- Higginbotham, J. (1981) "The Logic of Perceptual Reports," unpublished ms., MIT.
- Higginbotham, J. (forthcoming) Structures and Meanings.
- Higginbotham, J. and R. May (1981) "Questions, Quantifiers and Crossing," The Linguistic Review 1, 41-80.
- Hornstein, N. and A.S. Weinberg (1981) "Case Theory and Preposition Stranding," Linguistic Inquiry 12, 55-91.
- Huang, C.-T. J. (1982) Logical Relations in Chinese and the Theory of Grammar, Doctoral dissertation, MIT.
- Huang, C.-T. J. (forthcoming) "Move WH in a Language without WH-Movement," The Linguistic Review, to appear.
- Ingria, R. (1981) Sentential Complementation in Modern Greek, Doctoral dissertation, MIT.
- Iomdin, L.L. (1980) "K Postroeniju Formal'noj Modeli Sovremennogo Russkogo Jazyka: Operator Sintaksičeskogo Soglasovanija dlja Opredeľitel'nyx Konstrukcij," in H. Āim (X. Yjm), ed., Logiko-Semantičeskie Voprosy Iskusstvennogo Intellekta, Učenyje Zapiski Tartuskogo Gosudarstvennogo Universiteta, Tartu State U., USSR.
- Jackendoff, R. (1977) X̄-Syntax: A Study of Phrase Structure, MIT.
- Jaeggli, O. (1980a) "Remarks on To Contraction," Linguistic Inquiry 11, 239-245.
- Jaeggli, O. (1980b) On Some Phonologically Null Elements in Syntax, Doctoral dissertation, MIT [to appear as Topics in Romance Syntax, Foris, Dordrecht].
- Jaeggli, O. (1981) "A Modular Approach in Romance Agentive Causatives," unpublished ms., USC.
- Jakobson, R. (1936-1971) "Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus," in Selected Writings, vol. 2, 23-71, Mouton, the Hague.
- Kalakuckaja, L.P. (1971) Ad"ektivacija Pričastij v Sovremennom Russkom Literaturnom Jazyke, Nauka, Moscow.
- Kayne, R. (1975) French Syntax, MIT.

- Kayne, R. (1980) "Extensions of Binding and Case Marking," Linguistic Inquiry 11, 75-96.
- Kayne, R. (1981a) "ECP Extensions," Linguistic Inquiry 12, 93-134.
- Kayne, R. (1981b) "Unambiguous Paths," in R. May and J. Koster, eds., Levels of Syntactic Representation, Foris, Dordrecht.
- Kayne, R. (1981c) "Binding, Quantifiers, Clitics and Control," in F. Heny, ed., Binding and Filtering, MIT.
- Kayne, R. (1981d) "On Certain Differences between French and English," Linguistic Inquiry 12, 349-371.
- Kayne, R. (1981e) "Two Notes on the NIC," in A. Belletti, L. Brandi and L. Rizzi, eds., Theory of Markedness in Generative Grammar, Scuola Normale Superiore di Pisa, Pisa [written 1979].
- Kayne, R. (1982) "Connectedness," unpublished ms., Université de Paris VIII.
- King, H.V. (1970) "On Blocking the Rules for Contraction in English," Linguistic Inquiry 1, 134-136.
- Klima, E. (1964) "Negation in English," in J.A. Fodor and J.J. Katz, eds., The Structure of Language: Readings in the Philosophy of Language, Prentice-Hall, Englewood Cliffs, New Jersey.
- Koopman, H. and D. Sportiche (1982) "Variables and the Bijection Principle," unpublished ms., UQAM and MIT.
- Koster, J. (1978a) "Why Subject Sentences Don't Exist," in S.J. Keyser, ed., Recent Transformational Studies in European Languages, Linguistic Inquiry Monograph 3.
- Koster, J. (1978b) Locality Principles in Syntax, Foris, Dordrecht.
- Kuno, S. (1976) "Gapping: A Functional Analysis," Linguistic Inquiry 7, 300-318.
- Kuno, S. and J.J. Robinson (1972) "Multiple WH-Questions," Linguistic Inquiry 3, 463-487.
- Kuroda, S.-Y. (1972) "The Categorical and Thetic Judgement (Evidence from Japanese Syntax)," Foundations of Language 2, 153-185.
- Lakoff, G. (1970) "Linguistics and Natural Logic," Synthese 22.1-2.
- Lakoff, G. and S. Peters (1966) "Phrasal Conjunction and Symmetric Predicates," in Mathematical Linguistics and Automatic Translation, Harvard Computation Laboratory, Report No. NSF-17 [also in D.A. Reibel and S.A. Schane, eds., Modern Studies in English: Readings in Transformational Grammar, Prentice-Hall, Englewood Cliffs, New Jersey.

- Lasnik, H. (1982) "Binding or Pragmatics? The Nature of Condition C," paper presented at the Cornell Conference on Government-Binding Theory, July 1982; unpublished ms., U. of Connecticut at Storrs.
- Lasnik, H. and R. Freidin (1981) "Core Grammar, Case Theory, and Markedness," in A. Belletti, L. Brandi, and L. Rizzi, eds., Theory of Markedness in Generative Grammar, Scuola Normale Superiore di Pisa, Pisa.
- Lasnik, H. and J. Kupin (1977) "A Restrictive Theory of Transformational Grammar," Theoretical Linguistics 4, 173-196.
- Levin, J. (1982) "Free Relatives and the Non-Matching Parameter," unpublished ms., MIT.
- Levin, L. (1981) "Lexical Representations of Quirky Case in Icelandic," unpublished ms., MIT.
- Levin, L. (forthcoming) "Sluicing: A Lexical Interpretation Procedure," in J. Bresnan, ed., The Mental Representation of Grammatical Relations, MIT.
- Levin, L. and J. Simpson (1981) "Quirky Case and Lexical Representations of Icelandic Verbs," in Proceedings of the Seventeenth Regional Meeting of the Chicago Linguistic Society, Chicago.
- Lightfoot, D. (1977) "On Traces and Conditions on Rules," in P.W. Culicover, T. Wasow, and A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Manzini, M.-R. (1981) "On Control and Control Theory," unpublished ms., MIT.
- Marantz, A. (1981) On the Nature of Grammatical Relations, Doctoral dissertation, MIT.
- Marcus, M. (1980) A Theory of Syntactic Recognition for Natural Language, MIT.
- May, R. (1977) The Grammar of Quantification, Doctoral dissertation, MIT.
- May, R. (forthcoming) Logical Form: Its Structure and Derivation.
- Miller, G. and N. Chomsky (1963) "Finitary Models of Language Users," in D. Luce, R. Bush, and E. Galanter, eds., Handbook of Mathematical Psychology, vol. 2, Wiley, New York.
- Neidle, C. (1982) The Role of Case in Russian Syntax, Doctoral dissertation, MIT.
- Neidle, C. (forthcoming) "Case Agreement in Russian," in J. Bresnan, ed., The Mental Representation of Grammatical Relations, MIT.
- Nichols, J. (1981) "Prominence, Cohesion, and Control: Object-Controlled Predicate Nominals in Russian," unpublished ms., UC Berkeley.

- Obenauer, H.-G. (1976) Etudes de Syntaxe Interrogative du Français, Niemeyer, Tübingen.
- Obenauer, H.-G. (1981) "Le Principe des Catégories Vides et la Syntaxe des Interrogatives Complexes," Langue Française 52, 100-118.
- Padučeva, E. V. (1974) O Semantike Sintaksisa, Nauka, Moscow.
- Perlmutter, D. (1971) Deep and Surface Structure Constraints in Syntax, Holt Rinehart and Winston, New York.
- Perlmutter, D. (to appear) "The Unaccusative Hypothesis and Multiattachment: Italian Evidence," unpublished handout, UC San Diego.
- Pesetsky, D. (1982) "Complementizer-Trace Phenomena and the Nominative Island Condition," The Linguistic Review 1, 297-343 [earlier version with same title in Proceedings of the Ninth NELS Conference: CUNYForum 6, Queens College, Flushing, New York, 1979].
- Peškovskij, A.M. (1956) Russkij Sintaksis v Naučnom Osveščenii, 7th edition, Gos. Učebno-Pedagogičeskoe Izdatel'stvo Ministerstva Prosveščenija RSFSR, Moscow.
- Picallo, C. (1982) "The INFL Node and the Pro-Drop Parameter," unpublished ms., MIT and CUNY Graduate Center.
- Plann, S. (1980) Relative Clauses in Spanish, U. of California Publications in Linguistics 93, UCLA.
- Pollock, J.-Y. (1981) "On Case and Impersonal Constructions," in R. May and J. Koster, eds., Levels of Syntactic Representation, Foris, Dordrecht.
- Postal, P. (1969) "Anaphoric Islands," in R.I. Binnick, A. Davison, G.L. Green, and J.L. Morgan, eds., Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, Chicago.
- Postal, P. (1974) On Raising, MIT.
- Postal, P. and G. Pullum (1978) "Traces and the Description of English Complementizer Contraction," Linguistic Inquiry 9, 1-29.
- Postal, P. and G. Pullum (1982) "The Contraction Debate," Linguistic Inquiry 13, 122-138.
- Ravič, R.D. (1971) "O Vyboře Padeža Prjamogo Dopolnenija pri Perexodnyx Glagolax s Otricaniem v Russkom Jazyke," in R.I. Avanesov, ed., Fonetika, Fonologija, Grammatika: k 70-letiju A.A. Reformatskogo, Nauka, Moscow.
- Reinhart, T. (1981) "A Second COMP Position," in A. Belletti, L. Brandi, and L. Rizzi, eds., Theory of Markedness in Generative Grammar, Scuola Normale Superiore di Pisa, Pisa.

- Revzin, I.I. (1978) Struktura Jazyka kak Modelirujuščej Sistemy, Nauka, Moscow.
- Riemsdijk, H. van (1978a) A Case Study in Syntactic Markedness, Foris, Dordrecht.
- Riemsdijk, H. van (1978b) "On the Diagnosis of WH-Movement," in S.J. Keyser, ed., Recent Transformational Studies in European Languages, Linguistic Inquiry Monograph 3.
- Riemsdijk, H. van (1980) "On Theories of Case: the Case of German Adjectives," unpublished ms., University of Amsterdam and MIT.
- Rizzi, L. (1978a) "Violations of the WH-Island Constraint in Italian and the Subjacency Condition," in C. Dubuisson, D. Lightfoot, and Y.C. Morin, eds., Montreal Working Papers in Linguistics, vol. 11.
- Rizzi, L. (1978b) "A Restructuring Rule in Italian Syntax," in S.J. Keyser, ed., Recent Transformational Studies in European Languages, Linguistic Inquiry Monograph 3.
- Rizzi, L. (1982) "Negation, WH-Movement, and the Null Subject Parameter," in Issues in Italian Syntax, Foris, Dordrecht.
- Rosen, C. (1981) The Relational Structure of Reflexive Clauses: Evidence from Italian, Doctoral dissertation, Harvard.
- Ross, J.R. (1967) Constraints on Variables in Syntax, Doctoral dissertation, MIT.
- Ross, J.R. (1969) "Guess Who?" in R.I. Binnick, A. Davison, G.L. Green, and J.L. Morgan, eds., Papers from the Fifth Regional Meeting of the Chicago Linguistic Society, Chicago.
- Ross, J.R. (1970) "Gapping and the Order of Constituents," in M. Bierwisch and K. Heidolph, eds., Progress in Linguistics, Mouton, the Hague.
- Rotenberg, J. (1978) The Syntax of Phonology, Doctoral dissertation, MIT.
- Rothstein, S. (1981) "On Preposition Stranding," unpublished ms., MIT.
- Rouveret, A. and J.-R. Vergnaud (1980) "Specifying Reference to the Subject: French Causatives and Conditions on Representations," Linguistic Inquiry 11, 97-202.
- Ružička, R. (1963) "O Transformacionnom Opisanii Tak Nazываемyx Bezličnyx Predloženij v Sovremennom Russkom Literaturnom Jazyke," Voprosy Jazykoznanija 12, 22-31.
- Ružička, R. (1981) "Markirovannost' Primenitel'no k Parnosti Otricatel'nyx Predloženij v Sovremennom Russkom Jazyke," Russian Linguistics 6, 1-14.

- Safir, K. (1981) "On the Status of Null Expletive Elements," unpublished ms., MIT.
- Safir, K. (1982) Syntactic Chains and the Definiteness Effect, Doctoral dissertation, MIT.
- Sag, I. (1976) Deletion and Logical Form, Doctoral dissertation, MIT.
- Salamanca, D. (1982) "Remarks on Spanish Subjunctives," unpublished ms., MIT.
- Schein, B. (1982) "Non-Finite Complements in Russian," in A. Marantz and T. Stowell, eds., MIT Working Papers in Linguistics, vol. 4.
- Schein, B. (forthcoming) "Small Clauses and Predication," unpublished ms., MIT.
- Selkirk, E. (1972) The Phrase Phonology of English and French, Doctoral dissertation, MIT.
- Selkirk, E. (1977) "Some Remarks on Noun Phrase Structure," in P.W. Culicover, T. Wasow, A. Akmajian, eds., Formal Syntax, Academic Press, New York.
- Sidorov, V.N. and I.S. Il'inskaja (1949) "K Voprosu o Vyraženii Sub^Vekta i Ob^Vekta Dejstvija v Sovremennom Russkom Literaturnom Jazyke," Izvestija AN SSSR: Otdelenie Literatury i Jazyka 8.4, 343-354.
- Simpson, J. (1982), Doctoral dissertation, MIT.
- Sinicyn, M.G. (1982a) "Strange Questions in Slavic Languages," abstract in GLOW Newsletter 8, Foris, Dordrecht.
- Sinicyn, M.G. (1982b) "The Feature Coincidence Filter," unpublished ms., ul. Oficerskaja 1-3, Kaliningrad-100, USSR 236000.
- Sjoblom, T. (1980) Coordination, Doctoral dissertation, MIT.
- Sportiche, D. (1981) "Bounding Nodes in French," The Linguistic Review 1, 219-246.
- Steele, S. cum al. (1981) An Encyclopedia of AUX: A Study of Cross-Linguistic Equivalence, Linguistic Inquiry Monograph 5.
- Steriade, D. (1981) "If-Complements," unpublished ms., MIT.
- Stowell, T. (1980) "Complementizers and the Empty Category Principle," in V.A. Burke and J. Pustejovsky, eds., Proceedings of the Eleventh Annual Meeting of the North-Eastern Linguistic Society, U. of Massachusetts, Amherst.
- Stowell, T. (1981) Origins of Phrase Structure, Doctoral dissertation, MIT.
- Stowell, T. (1982) "The Tense of Infinitives," Linguistic Inquiry 13, 561-570.

- Stowell, T. (forthcoming) "Subjects Across Categories," unpublished ms., MIT [to appear in The Linguistic Review].
- Suprun, A. (1959) O Russkix Čislitel'nyx, Kirghiz State U., Frunze, USSR.
- Svedova, N.Ju., ed. (1970) Grammatika Sovremennogo Russkogo Literaturnogo Jazyka, Nauka, Moscow.
- Taraldsen, K.T. (1980) "On the NIC, Vacuous Application and the That-Trace Filter," distributed by Indiana University Linguistics Club.
- Taraldsen, K.T. (1981) "The Theoretical Interpretation of a Class of Marked Extractions," in A. Belletti, L. Brandi, L. Rizzi, eds., Theory of Markedness in Generative Grammar, Scuola Normale Superiore di Pisa, Pisa.
- Vergnaud, J.-R. (1974) French Relative Clauses, Doctoral dissertation, MIT.
- Vergnaud, J.-R. (1982) Dépendance Formelle et Niveau de Représentation en Syntaxe, thèse de doctorat d'état, Paris [to appear, Benjamin, Amsterdam].
- Vergnaud, J.-R. and M.-L. Zubizarreta (1982) "Mixed Representations and Virtual Categories," in A. Marantz and T. Stowell, eds., MIT Working Papers in Syntax, vol. 4.
- Williams, E. (1975) "Small Clauses in English," in J. Kimball, ed., Syntax and Semantics, vol. 4, Academic Press, New York.
- Williams, E. (1977a) "Across-the-Board Application of Rules," Linguistic Inquiry 8, 419-423.
- Williams, E. (1977b) "Discourse and Logical Form," Linguistic Inquiry 8, 101-139.
- Williams, E. (1978) "Across-the-board Rule Application," Linguistic Inquiry 9, 31-43.
- Williams, E. (1980) "Predication," Linguistic Inquiry 11, 203-238.
- Williams, E. (1981a) "On the Notions 'Lexically Related' and 'Head of a Word'," Linguistic Inquiry 12, 245-274.
- Williams, E. (1981b) "Argument Structure and Morphology," The Linguistic Review 1, 81-114.
- Williams, E. (1982) "Transformationless Grammar," Linguistic Inquiry 12, 645-653.
- Woolford, E. (1981) "Comments on Grimshaw (1981)," in C.L. Baker and J. McCarthy, eds., The Logical Problem of Language Acquisition, MIT.

Zaliznjak, A.A. and E.V. Padučeva (1979) "Sintaksičeskie Svojstva Mestoimenija kotoryj," in T.M. Nikolaeva, ed., Kategorija Opredeľennosti-Neopredeľennosti v Slavjanskix i Balkanskix Jazykax, Nauka, Moscow.

Zubizarreta, M.-L. (1982) On the Relationship of the Lexicon to Syntax, Doctoral dissertation, MIT.

Zubizarreta, M.-L. (forthcoming) "Theoretical Implications of Subject Extraction in Portuguese," unpublished ms., MIT [to appear, The Linguistic Review].