

CHAINS AND ANAPHORIC DEPENDENCE
On Reconstruction and Its Implications

by

Andrew Barss
B.A. Hampshire College (1982)

Submitted in partial fulfillment
of the requirements for the
degree of

DOCTOR OF PHILOSOPHY

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August 1986

© Andrew Barss 1986

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

Signature of Author: _____

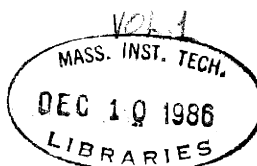
Andrew Barss
Department of Linguistics and
Philosophy. August, 1986

Certified by: _____

James T. Higginbotham
James T. Higginbotham
Thesis Supervisor

Accepted by: _____

Morris Halle
Morris Halle
Departmental Graduate Committee



HUM.

CHAINS AND ANAPHORIC DEPENDENCE

On Reconstruction and Its Implications

by

Andrew Barss

Submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

ABSTRACT

This thesis is concerned with developing an account within the Government and Binding (GB) theory of the grammaticality of such structures as (1), and exploring the implications of this account for the theory of empty categories, chains, and scope. The hallmark characteristic of such grammatical S-Structure representations as (1) is that the anaphor is outside the c-command domain of its understood antecedent. The basic anaphoric effect is termed connectivity.

1) [which of each other's friends][did the men see t]?

Chapter 1 is a brief overview of the necessary definitions presumed in the thesis, and an outline of the subsequent chapters. Chapter 2 introduces a large body of data which must be treated on a par with (1), and reviews and criticizes several existing proposals which have been made to account for (1). The chapter argues that the binding theory must apply to structures having the essential form of (1). We demonstrate that no treatment which involves lowering the anaphor into the c-command domain of the antecedent via "reconstruction" operations, or involves applying the Binding Theory at a level at which WH movement is not represented, can be maintained. Chapter 3 develops a revision of the binding theory, focusing on Condition A, which is capable of treating all the connectivity data in a unified way. The major formal construct proposed in the chapter is the chain accessibility sequence, essentially a path of nodes through which the potential antecedents for an expression are accessed. The revised binding theory is defined in terms of such sequences; as the name implies, the notion chain plays a prominent role. This approach to connectivity is developed in the spirit of the Path theory of Kayne (1983) and Pesetsky (1982). We also discuss properties of structures of the form of (1), but where the constituent containing the anaphor is predicative in nature. We shall see that the predicative nature of the constituent significantly constrains the possibilities of assigning the anaphor an antecedent. This chapter adopts, and argues in favor of, the Linking theory of binding introduced by Higginbotham (1983).

Chapter 4 focuses on the theory of empty categories, arguing that it is desirable to construct the theory so that no empty categories bear binding features (the features [+/- anaphoric] and [+/- pronominal] are thus restricted to overt categories). This proposal, which I term the No Features Hypothesis, departs from the characteristic treatment of ECs in GB theory. The chapter adopts Brody's (1985) proposals concerning the distribution of PRO and NP-trace. We adopt, and later extend, the Local Binding Condition (LBC) on A chains, argued by Rizzi (1982) to constrain the well-formedness of A chains. We reformulate it in terms of Linking theory, as the Chain Obviation Condition (COC), and argue that it holds of all chain types. This is shown to be a principle with considerable generality, subsuming the LBC, Condition C of the binding theory, and the anti-c-command condition on linking. Adopting the COC, along with the NFH, allows the elimination of the class R-expression from the inventory of binding types. It will be shown that the anti-c-command condition on parasitic gaps derives directly from the COC, with no stipulations. The chapter concludes with a defense of the proposal that the theory of anaphora must recognize anaphoric dependence and obviation as separate relations (as argued by Lasnik (1976), (1981), and Higginbotham (1985)).

Chapter 5 discusses constraints on the interpretation of sentences in which a quantificational NP is the antecedent of an NP-trace which it does not c-command. These considerations lead us to formulate a constraint on movement operations. The chapter also argues that the operations of WH-movement and QR are strictly ordered in the LF component.

Thesis Supervisor: James Higginbotham

Title: Professor of Linguistics and Philosophy

ACKNOWLEDGEMENTS

I have greatly enjoyed writing this dissertation, and I owe a debt of gratitude to everyone who has directly or indirectly aided in its development and completion. First and foremost, I thank the members of my committee, Jim Higginbotham, Howard Lasnik, and Noam Chomsky, for their generous assistance, friendly encouragement, and rigorous criticism. Together they have constituted an exceptional committee, generous with their time and (almost) endlessly patient in their willingness to scrutinize my sometimes half-formed ideas. Every section of this thesis has benefitted from their criticism. Nobody writing a thesis about binding theory could ask for a better group of people to work with.

Jim Gee was my first syntax teacher, and I count his lectures and tutorials at Hampshire College among the most valuable aspects of my student years. From Jim I learned to appreciate crisp and elegant argumentation, and clear teaching. The many discussions (and fine dinners) I have had with him over the past four years have been most enjoyable.

I worked on many of the topics considered here in my third year at MIT as well, and I would like to thank the many people who discussed earlier stages of the proposals with me: Luigi Rizzi, Anna Szabolcsi, Kyle Johnson, Mamoru Saito, Tim Stowell, Joseph Aoun, Adriana Belletti, Pino Longobardi, Pierre Pica,

Edwin Williams, and Ur Schlonsky. Discussions with Tarald Taraldsen, Richie Kayne, Richard Larson, and Robert May this year helped to clarify my thoughts on several key aspects of the thesis.

Sam Epstein has been my closest colleague during the writing of this thesis. He has had to listen to most of the proposals made in this thesis, sometimes over and over and over again, and his contribution to the overall thesis is enormous. He has assisted me in sharpening my argumentation, finding data, and everything else which goes into dissertating.

Writing a thesis is, thankfully, a very social process, and I have been enjoyably entertained by the friendship of Sam Epstein, Elaine McNulty, Doug Saddy (and his magnificent automobile), Viviane Deprez, Jim Gee, Kyle Johnson, Lonny Jarrett, Kelly Sloan, Naoki Fukui, Ur Schlonsky, Yosef Grodzinsky, Gabriel Segal, Diana and Dante Archangeli, Betsy Ritter, John Lumsden, Marc Ryser, Juan Uriagareka, Janet Nicol, Maggie Browning, and Betsy Sagey. I doubly thank everyone who has fed me dinners this year.

In previous years, Mario Montalbetti, Judy Kegl, Juliette Levin, Dinette Massam, Diana Archangeli, Mamoru Saito, and Richard Sproat helped make the department a lively and interesting place.

I would like to thank Maggie Carracino and Nancy Peters for

their friendly moral support during the last four years, and I thank Maggie and Wayne O'Neil for last-minute technical assistance.

Portions of the material in this thesis were presented in various stages of development to audiences at The University of Connecticut, M.I.T., and WCCFL, and I thank those audiences.

I have been supported, in my time at MIT, by a research assistantship from the Department of Linguistics and Philosophy, and a Graduate Fellowship from the National Science Foundation. I thank them for their support.

This thesis is dedicated to my family: Peter, Sara, Taiya, Chris, Myra, John, Parker, Ingram, Patchen, Joshua, Silas, Jacob, Anna, and especially to my mother Currie.

On with the show.

TABLE OF CONTENTS

Abstract	2
Acknowledgements	4
Table Of Contents	7
Chapter One: Introduction	13
Chapter Two: On The Nature of the Reconstruction Problem	17
2.0.1 Against NP-Structure	25
2.0.2 Against the Dual S-structure Hypothesis	33
2.1 Two Distinct Theories of Anaphor Connectivity	37
2.1.1 Lowering Theory	37
2.1.2 Non-Lowering Theory	38
2.2 On Lowering Treatments	40
2.2.1 Reconstruction and Strong Crossover	41
2.2.2 Problems with Reconstruction and Strong Crossover	44
2.2.3 A Possible Counterargument	47
2.3 Against The Quantifier decomposition Rule	50
2.3.1 Objection One: Unrestricted Quantification	50
2.3.2 Objection Two: The Theta Criterion and Saturation	57
2.3.3 Objection Three: The ECP and Rule (36)	62
2.4 Against a Lowering Treatment of Anaphor Connectivity	66
2.4.1 Against PP Lowering as a Solution	68
2.4.2 NP Lowering	81

2.4.3 At Which level does Condition A Apply?	85
2.5 Connectedness Theory and Anaphoric Paths	89
2.5.1 A Connectedness Treatment of Lexical Anaphora and Connectivity	95
2.5.2 Anaphoric Paths and Complete Functional Complexes	107
2.6 Problems with the Path Approach	110
Appendix	118
 CHAPTER THREE: Chains, Predicates, ,and Linking	 121
3.1 Chain Accessibility	122
3.1.1 Reference to Chains	124
3.1.2 Chain Accessibility and the Binding Theory	125
3.1.3 Some Differences	129
3.2 The Modified Binding Theory	132
3.2.1 The Multiple Binding-Domain Effect	133
3.3 NP-Movement and Reconstruction	137
3.3.1 Psych-Movement	140
3.3.2 Raising and Connectivity	147
3.3.3 Pseudo-Clefts and Raising	150
3.3.4 Raising and Pseudo-Clefts	158
3.3.5 Pesetsky's Analysis of Psych Movement	159
3.3.6 Ambiguities with Psych-Verbs	166
3.3.6.1 Passive	167
3.3.7 Against Successive Cyclic Binding	168
3.3.7.1 Binding at Intermediate Representations	168

3.3.7.2 Passive and Successive Cyclic Binding	171
3.3.7.3 Strict Cyclicity and feature Assignment	176
3.4 Binding Domains and a Predicate/Non-Predicate Asymmetry	186
3.4.1 Predicates and Ambiguous Antecedents	187
3.4.2 Hebrew and French	190
3.5 The Definition of Local Binding Domains: Chomsky (1981) and (1985)	195
3.5.1 CFCs and The Asymmetry	198
3.5.2 Theta Compatibility	200
3.6 Theta-Compatibility and Condition A	203
3.6.0.1 CFCs and Saturated Domains	206
3.6.1 An Alternative Proposal	213
3.6.2 Couquaux's Analysis of Predicative Structures	217
3.6.2.1 The OB Indexing System	220
3.6.2.2 Summary	227
3.6.3 Predicate Nominals and Small Clauses	228
3.6.4 The Necessity of the Theta-Compatibility Requirement	230
3.7 Reciprocals and Connectivity	233
3.7.1 Reciprocals and Their Interpretation	235
3.7.2 Antecedent Clash	241
3.7.3 Connectivity and a Redundancy in Indexing	245
3.7.4 Antecedent Clash in Indexing theory: Another Redundancy	252
3.7.5 Adverbial <u>Each</u>	259
3.7.6 Suspension of the <u>Each</u> Conflict	263
3.8 Layered Traces	271

3.8.1 Reciprocals and Layered Traces	276
3.8.2 The Blocking Effects	278
3.8.3 Moved Predicates	282
3.9 Appendix: On the Properties of Specificational Pseuso-Clefts	292
3.9.1 Anaphora	293
3.9.2 Other Facts	298
3.9.3 Three Accounts	301
3.9.3.1 Literal Reconstruction	301
3.9.3.2 Chain Formation	308
3.9.3.3 Predication	313
CHAPTER FOUR: On the Featural Content of Empty Categories	318
4.1 Review: EC Features in GB	321
4.2 What are the Binding Features of ECs?	326
4.3 Suppression of Strong Crossover	330
4.3.1 Binding at Different Levels of Representation: A Non-Solution	334
4.3.2 More Oddities	338
4.4 The Feature Matching Hypothesis	343
4.5 The FMH and A Chains	346
4.6 The Auxiliary Hypothesis	348
4.6.1 Chain-Internal Binding	348
4.6.2 Improper Movement	349
4.6.3 NP-Trace as an Anaphor	357
4.7 Chomsky (1985) and Condition (49)	357
4.7.1 The Redundancy of ECP and Condition A	363
4.7.2 The Ban Against Accidental Coindexing	373
4.7.3 Other Cases	379
4.7.4 Passivized PRO	387
4.7.5 A Redundancy Between (49) and Linking	390

4.8 Do ECs Have Features?	393
4.9 The No Features Hypothesis	407
4.9.1 Review of the Accessibility Condition	418
4.9.2 A' Dependent Pronouns	421
4.9.3 Generalizing the Local Binding Condition on A Chains	427
4.9.4 On Deriving the Anti C-Command Condition on Linking	430
4.9.5 Strong Crossover	436
4.9.6 On Indexing	438
4.10 Condition B	439
4.10.1 At What level Does Condition B Apply?	453
4.10.2 A Note On Clitic Trace	458
4.11 Reconstruction and Parasitic Gaps	460
4.12 Opacity in Parasitic Constructions	461
4.12.1 Complex Crossover and Parasitic Gaps	469
4.12.2 The Lack of Sloppy Identity	475
4.12.3 Obviation and Parasitic Gaps	477
4.12.4 The Relation between the Accessibility Condition and the Chain Obviation Condition	499
4.12.5 Crossover in <u>too</u> -Adjective Constructions	503
4.13 Deriving the Anti-C-Command Condition on Parasitic Gaps	507
CHAPTER 5: NP-Trace, Quantifier Scope, and A Constraint on Movement	517
5.1 Non-C-Commanded NP-Trace	519
5.2 Quantifier Scope and Lowering	524
5.3 Restricted Cases	531
5.3.1 Representational Constraints and (33)	534
5.3.2 The Interpretations of (26)	535

5.3.3 The Interpretations of (27)	538
5.3.4 On Answers and Variables	542
5.3.5 Crossed Binding	551
5.3.6 Idioms and Pleonastics	553
5.3.7 No LF Extraposition	555
 5.4 WH-In-Situ and Ordering of Rules at LF	 561
5.4.1 Review of Taraldsen (1981)	565
5.4.2 Proper Government at Intermediate Levels	569
 5.5 Conclusion	 574
 References	 576

Chapter One: Introduction

This is an extremely brief introduction to the rest of this thesis. I presume familiarity with the Government and Binding Theory (GB) of Chomsky (1981), (1985), (1986), and much other current work. Virtually all of the review of the technical aspects of GB needed to advance the argumentation of the thesis is present in the relevant chapters; therefore I keep my comments here brief.

GB is a formal theory of syntax, which is distinguished by its modular nature; it consists of a number of sub-theories, each concerned with what are considered to be separate domains of syntactic phenomena. This thesis is concerned essentially with the Binding Theory, that sub-theory which is concerned with defining well-formedness conditions on the referential interaction of nominal phrases. As the thesis develops, other issues in other sub-theories are discussed. In particular, one will find proposals concerning the Empty Category Principle, chain well-formedness, quantifier scope, constraints on syntactic derivations, and the featural content of empty categories in various parts of the essay.

There are two major, and quite different, frameworks of binding in GB research. the first is indexing theory, developed

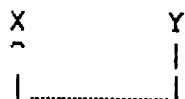
and explicated in Chomsky (1981). In this theory, NPs are assigned single integers, freely (a modification of the more complex indexing system of Chomsky (1980)). The Binding Conditions impose restrictions on the appearance of indexed NPs in a syntactic representation, relativized to two factors: the binding-type of the NP bearing the index, and the structural position of the NP.

GB recognizes two binding features: +/- anaphoric, and +/-pronominal. Together, they give rise to four possible binding types. Principle A of the Binding Theory (BT) constrains the distribution of +anaphoric elements; Principle B the distribution of +pronominal elements. Because these two conditions are complementary, no element which bears them both may occur in a position which forces it to be evaluated by the BT (e.g. a governed position). Thus, there are three binding types among overt NPs: anaphors, pronominals, and R-expressions, each governed by a different principle of binding. In Chapters 3 and 4, I develop a different formalization of the binding theory, which has two major consequences: the class R-expression is eliminated (along with Condition C), and the assignment of binding features is restricted to overt NPs. This has the result that there are only two binding types: pure anaphors (himself, each other,...) and pure pronominals (he, they).

Structural positions are of two primary types: A-positions, positions in a phrase marker to which a thematic role might be

assigned, and A'positions, positions to which a theta-role cannot be assigned. The theory of binding is essentially concerned with referential relations between two NPs both of which are in A positions.

The other framework of binding is linking theory, introduced by Higginbotham (1983). In this theory, links are assigned between two positions, where in the following representation X is the antecedent of Y, and Y is linked to X. Linking is fundamentally asymmetric, while coindexing is fundamentally symmetric.



The thesis will argue that linking is to be preferred over indexing. I will, however, occasionally revert to the use of indexing in examples where the question of the choice between theories is not under discussion, especially in chapter 5.

The dissertation as a whole is concerned with a certain class of problems in binding theory; although typically Conditions A, B, and C are sensitive only to relations between an NP and an NP which c-commands it, there are cases relevant to the binding theory in which this relation does not hold. These cases, and the interesting questions concerning them, are the topic of Chapters 2 and 3. Chapter 4 develops a new theory of empty categories, and chapter 5 discusses certain restrictions on quantifier scope and extraposition. All the chapters relate to

one another in a systematic way.

I presume the theory of phrase structure and successive cyclic movement advanced in Chomsky (1986). For the reader unfamiliar with this framework, the category S is taken to be the maximal projection of INFL, and is designated IP (I use S and IP interchangeably throughout the thesis). Constituents previously taken to be of category S-bar are regarded in this approach to be the maximal projection of COMP, and are designated CP. The two chief landing sites for WH-movement are SPEC (specifier) of CP, a distinct position from COMP, and adjunction to certain maximal projections, chiefly VP.

Chapter Two: On the Nature of the Reconstruction Problem

This chapter is concerned with reconstruction, the term given to both a general problem in the theory of anaphora exemplified in (1), and to one postulated solution to it.

(1)(a) [which pictures of himself][does John like e]?

(b) [[These photographs of himself],[John really likes e]]

In (1), an anaphor is contained within a phrase which is dislocated by WH-movement, Topicalization, or other movement to

a non-argument position (A' position).^{1 2}

1. More generally, as we shall see in later chapters, this phenomenon occurs even when the container of the anaphor is plausibly generated in its surface position, binding the empty category in argument position but never occupying that position at any stage of the derivation. For this reason, I shall temporarily abstract away from the analysis of Topicalization involving such base-generation which is proposed in detail in Chomsky (1977), since the analysis of the reconstruction effect will extend to both possibilities. See Chapter 4.

2. See Lasnik and Saito (forthcoming) for arguments that Topicalization is (for some cases) properly characterized as direct movement from a sentence-internal position to sentence adjunction, as in (i)-(ii).

(i) D-Structure: John said that[Mary likes Bob]).

S

(ii)S-Structure: John said that[Bob [Mary likes e]]

S

i

i

See also chapter (3) of this thesis for discussion.

What is of relevance to the Binding Theory is that such dislocation places the anaphor out of the c-command domain of its understood antecedent, with no loss of grammaticality. Indeed, as many authors have noted, the possibilities of grammatically assigning an antecedent to the anaphor include those which would exist had the dislocation not occurred in the first place, as we see in comparing (1a) to (3a) and (2) to (3b). In (2) and (3b), the noun phrase containing the anaphor has a Subject (whose and Mary respectively), which is obligatorily understood as the anaphor's antecedent.

(2). [whose pictures of himself][does John like e]?

(3). a) John likes some photos of himself

b) * John likes Mary's photos of himself

By contrast, there is no such 'Subject of NP' in either (1a) or (3a), which apparently allows John in both cases to antecede the anaphor. In some sense to be formally articulated, the

3

3. I use the term 'antecede' rather than 'bind' because, in the Binding Theory of LGB, binding requires c-command (A binds B iff A c-commands B, and A and B are coindexed), and one of the puzzles of the reconstruction effect is that (at least at S-Structure) the anaphor is not necessarily c-commanded by its antecedent (as in (1a,b)). Later, I shall use the linking notation for antecedence introduced by Higginbotham (1983), which encodes antecedence irrespective of c-command. This will make the relation rather clearly represented, without the definitional problem posed by indexing; I will, in later sections of this essay, argue that linking is a preferred theory of binding on formal grounds as well.

4

dislocated phrase is 'connected' to the trace, in that, as far as anaphor binding goes, the phrase behaves as if it were in the position of the gap. Because one prevalent solution to this problem is also termed "Reconstruction", I will refer to the general phenomenon as anaphor connectivity, following Pauline Jacobsen, reserving the term Reconstruction for one of the postulated treatments of the phenomenon. Reconstruction has been offered as a rule by several authors (see below) as a means of actually lowering the dislocated phrase into the position of the gap in the S-Structure-to-LF mapping, 'reconstructing' the D-Structure content of the trace.

In doing so, the apparent violation of Condition A is dismissed; the anaphor is neither c-commanded by its potential antecedent, nor in the same local domain as it, in the S-Structure representation (1a); but, the advocate of Reconstruction will say, the representation at LF is what is relevant to Condition A, and the anaphor can be put into a position at LF where it is locally c-commanded by its potential antecedent.

The intent of this Chapter, and the one that follows it, is to both fully characterize the connectivity problem, and to explore various ways of approaching it within the general framework of Government and Binding Theory. In this chapter, I shall examine

4. To use the terminology of Akmajian (1970) and Higgins (1973).

several existing proposals in the literature, and some of the data relevant to the problem, and argue that certain types of accounts must be abandoned for empirical reasons. In doing this, I will gradually converge upon one type of analysis, which suggests that Condition A of the Binding Theory -- that condition which regulates the distribution of anaphors -- must be reformulated to apply directly to S-Structure representations. The last section of this chapter is devoted to constructing one such analysis within the Connectedness framework of Kayne (1983, 1984), and showing certain problems with such an analysis. The subsequent chapter develops in full a formal revision of Condition A which avoids such problems, and which has a variety of other consequences as well.

The idea behind Reconstruction as a solution to this apparent problem is to deform the S-Structure representation which violates Condition A⁵ by lowering some portion of the WH phrase (crucially including the anaphor) into the c-command domain of the understood antecedent, as part of the derivation of LF, and then applying the Binding Theory at LF only. By 'putting back' the WH phrase into its D-structure position prior to the application of Condition A, the parallel in grammaticality between (1a) and (3a), and between (2) and (3b), is due to their having equivalent representations when Condition A applies.

5. See (22) below for Condition A.

(3a) and (3b) will have LFs identical to their S-Structures, and Reconstruction will transform (1a) and (2) into something like:

(4) a. (LF of (1a)):

[which (x): pictures (x)] [John likes [x of himself]]

(b) LF of (2):

[which (y): person(y)] [John likes [y's picture of himself]]

The exact form of the LF derived through Reconstruction crucially depends upon which part of the WH phrase is lowered by the rule, a question I shall consider in more detail shortly. This is not a trivial question; for example, in (4a), the variable inside the lowered NP is not a maximal projection (and indeed probably does not correspond to an S-Structure constituent), and so a rule which has (4a) as its output is not Structure preserving in the sense of Emonds (1970).

Williams and van Riemsdijk (1986) offer the following rule of Reconstruction:

(5) WH-Interpretation (Reconstruction)

[[...[...]] ...e ...
 COMP a b b ai COMP i

Where b is a WH-phrase, replace e_i with a_i , replace b

with

x_i , and place $?x_i N'$ in COMP, where N' is the head of b.

Or, if $b = \text{who}$, place $?x$ [$x : \text{person}$] in COMP.

This rule has as its output the LF representations in (6), which are rather similar to those representations in (4), with the exception that Williams and van Riemsdijk (for reason's that aren't clear to me) use subscripts on variables, which seems unnecessary, and they appear to advocate leaving the entire N' of the WH phrase in COMP as the restriction on the question operator. Doing so fails to completely lower the anaphor through reconstruction, as we see in (6a).

(6)

a. LF of (1a): [_ix : pictures of himself] John likes [_ix of himself

b. LE of (2):

[?y : person] John likes [y 's picture of himself]

For this reason, I will modify their '...place ?N' in COMP, where N' is the head of b...' with '...place ?N in COMP, where N is the head of b...'. This has the result that only the head noun remains in COMP, with the rest of the nominal, including crucially the complement 'himself', lowering into the position of

the trace. Clarifying these points gives us rule (7) :

6. Two other parts of this rule require some clarification. First, the final clause of their rule, 'or, if b = who, place ?x [x:person] in COMP', confers specialized status upon the quantifier who. This distinguishes who, which constitutes the entire WH-phrase, from other WH-phrases in which the lexical WH-element constitutes only the determiner, like which picture of Mary. It also distinguishes who from other WH-maximal projections like what, when, where, and why.

(7) WH-Interpretation (Reconstruction)

$$\begin{bmatrix} \text{COMP} & a & b & b & ai & \text{COMP} & i \end{bmatrix} \dots e \dots$$

Where b is a WH-phrase, replace e with a , replace b

with

x, and place $[?x \ N]$ in COMP, where N is the head of b.

Or, if b = who, place ?x [x:person] in COMP.

And finally the modified rule of reconstruction (7) will give us the LF representations (8a) and (8b) for the sentences of (1a) and (2):

1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808

This part of the Reconstruction rule seems unnecessary, since there is no reason why LF should express the lexical meaning of a single word. Abandoning this clause will still retain the basic operation of the rule, specifically its way of placing an anaphor in a WH-phrase like which picture of himself back into the c-command domain of its antecedent for an example like (1a).

Secondly, the rule substitutes a variable for a WH-phrase ('where b is a WH phrase,...replace b with x...') in the lowered constituent, and translates b into an informal logical notation ('place [?x N] in COMP'). This raises the question of what they mean by 'WH-phrase'. Suppose we interpret this term as 'the minimal maximal projection constituting or containing a WH-item'. For [whose pictures of Mary], b = who; for [which pictures of himself], b = [which pictures of himself]. But this yields the wrong result, since it is the purpose of this rule to permit himself to be lowered (along with other material), while allowing the interrogative to remain in COMP. For a constituent like [which pictures of himself], the rule would appear to be able to break up the NP into two parts: the complement PP, and the Det-N sequence, which does not form a constituent. For this latter reason, the rule is non-structure preserving.

(8)

- a. LF of (1a): [$?x$: pictures] John likes [x of himself]
- b. LF of (2): [$?y$: person] John likes [y 's picture of himself]

In (8a), the anaphor is c-commanded by its antecedent, within the governing category (GC) of the anaphor (the matrix clause), hence, on the assumption that Condition A applies only at LF, the sentence fully satisfies the relevant part of the Binding Theory. By contrast, the LF representation of (2), given as (8b), has a smaller GC for the anaphor, namely the NP [y 's picture of himself]. This NP has a Subject (the argument trace y), hence is the GC for the anaphor. Therefore, while in (8a) the anaphor may seek John as its antecedent, in (8b) Condition A requires that the anaphor be bound by the subject of the object NP, the variable y . Thus the hypothesized rule of reconstruction (7) will yield the right binding results for sentences (1a) and (2).

The general assumption in the literature which advocates such a process of reconstruction is that the variable⁷ left behind in the first instance of movement is the empty category into

7. I assume here that an empty category is a variable if it is casemarked, occurs in an A position, and is A' bound. In Chapter 4, I examine in detail the various existing proposals having to do with the definition of ecs, and the assignment of features to them, offering a modification of Michael Brody's recent proposals (the 'random characterization' hypothesis). The details of the analysis, which differ considerably from existing proposals, are irrelevant here.

which the anaphor and the accompanying material⁸ is lowered. Since, on this view, the lowered material moves back to its D-Structure position, there should be an absolute parallel between the possibilities for antecedence in the cases with movement and the cases without movement.

However, it is important to note that this parallel breaks down, as Weisler (1982), Langendoen and Battistella (1982), and Barss (1984) note, when cases of successive cyclic movement are considered. In the example (9), successive cyclic movement proliferates the possibilities for grammatically assigning an antecedent to the reflexive beyond what exists in the pre-movement structure. In sentence (9), either John or Fred may antecede himself.

(9) [[which pictures of himself]_i did John think [t'_i [Fred liked t_i]]]?

2.0.1 Against NP-Structure

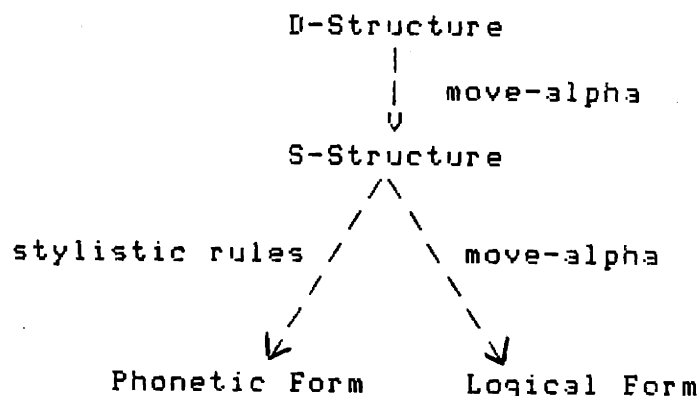
This fact (and a number of other related data of a similar type, to be considered in subsequent sections) eliminates in principle the approach to such examples as (1) given by van Riemsdijk and Williams (1982), in whose theory anaphor binding

8. The rule of reconstruction given as (7) lowers the entire complement to the head noun. Other rules of reconstruction, similar in intent to (7) will be discussed later; they each lower a different constituent, but each, quite crucially, is formulated to lower at least the anaphor.

occurs prior to WH-movement.

They offer a theory of grammar which is organized differently from the 'T-model' which was proposed in Chomsky and Lasnik (1977), and which has become standard in most studies within GB theory. The (relevant portions of) the T-model is given as (10).

(10)



D-structure is taken as a pure representation of predicates and their arguments, in the sense that the only NPs which occur at

this level are in theta-marked positions⁹. D-Structures are manipulated into S-Structures by the generalized process of

movement¹⁰, or move-alpha, where alpha may be any category, and

9. Ignoring, of course, predicate nominals, which occur in predicate positions. See Williams (1983), Higginbotham (1983) for discussion. See Larson (1985) for discussion of adverbial NPs, which also do not occur in theta-marked positions.

10. I defer discussion of the exact nature of 'movement' to forthcoming. Movement is of course a metaphor; no physical objects are actually changing positions. The nature of what the

there is no specification as to the landing site, the extraction site, or the category which moves, internal to the rule itself. Filters and constraints on representations will rule the derived representations as grammatical or ungrammatical, depending on the properties of the derived structure. Additionally, movement itself is constrained, subject to such conditions as the Strict Cycle Condition and Subjacency (in some views).

Two canonical cases of movement are NP-movement, illustrated in (11), and WH-movement, as in (12).

(11) John was arrested t
 i i

(12) [who [do you like t]]?
 i i

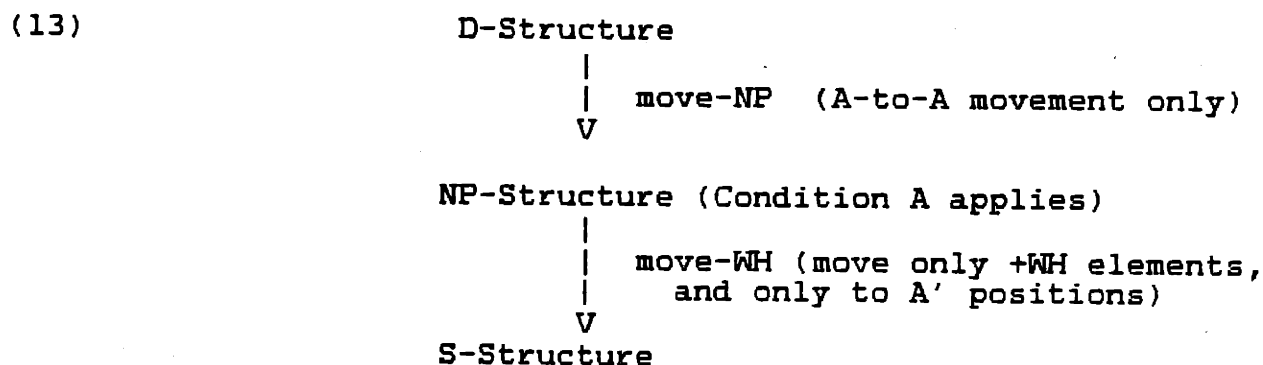
In the T-model, there is no real formal distinction between NP movement and WH-movement as movement processes; the two movements deriving (11) and (12) from their D-Structures are two instances of move-alpha. There are considerable differences between the two on other grounds: the EC in the two sentences is not the same, being an anaphor in (11) and an R-expression in (12); the EC is casemarked in (12) but not in (11); and the landing site of (11) is an A-position, meaning that is is a

 term is a metaphor for is of profound interest, and I shall devote part of chapter four to considering it in detail. A fully specified theory of the operations typically called movement has import for the theory of empty categories, the formulation of 'constraints on movement', among other things. See Lasnik and Kupin (1977) for a formal theory of movement operations, and Sportiche (1982), Brody (1984), and Lasnik and Saito (in progress) for discussion of some of the syntactic issues involved.

structural position which might be assigned a theta-role, while in (12) the landing site is an A' (non-argument) position.

In the alternative model of grammar that Williams and van Riemsdijk propose, the processes of NP-movement and WH-movement are considered to be formally different rules. NP movement moves an NP from an A position to an A position, while WH-movement moves a WH phrase to an A-bar position. Positing the two processes as two distinct rules allows W&vR to motivate the following model of grammar, which is termed the NP-structure

11
model . In this model, there is a level of representation NP-structure, derived from D-structure by the rule move-NP, which moves NPs from A positions to A positions. NP-structure representations feed the rule of WH-movement, which moves a WH-phrase into COMP. Such a separation of derivations is impossible in the T-model, of course.



11. The theory is designated the L-model (Linear-model) by W&vR, but the term NP-Structure theory, borrowed from the title of their article, is the one which has stuck, and it is the one I shall use here.

To give a sample derivation, contrasting the two models, consider the sentence 'who was John introduced to?'. On the

T-model, the derivation will proceed something like (14)¹²:

(14)

a) DS: L [[[[[[was]] introduced John L to who]]]]]
 CP C' C IP I' I VP PP VP

b) [[[[[John [[was]] introduced t [to who]]]]]
 CP C' C IP i I' I VP i PP VP

c) SS: L who [[[was]] John L [t]] introduced t
 CP k C' C j IP i I' I j VP i
 [to t]]]]]
 PP k VP

The final representation, (14c), is the S-Structure representation. The intermediate step, (14b), is just that --

an intermediate representation¹³ between two levels of representation, DS and SS. It is typically assumed that

12. To simplify things a bit, I ignore the question of the exact point at which was (the U-Structure INFL) moves to COMP, and I also ignore the step of the derivation in which the WH-phrase who adjoins to VP, as it would in the Chomsky (1986) theory.

13. It is not clear to me that GB as it is commonly understood is explicitly committed to the notion of such intermediate representations. However, I believe that the only coherent way in which 'movement' can be understood is that it is the association of pairs of such intermediate representations (subject to conditions, including the theta criterion and the projection principle, the Strict Cycle Condition, and the Superiority Condition), thus giving evidence for their existence. See Saito (1985), and Chapter four of this thesis for discussion of this matter.

well-formedness conditions hold of the syntactic structures at the designated levels of representation, but not at these intermediate representations (but see Lasnik and Saito (forthcoming), and Chapter Three and Four of this thesis for proposals that conditions hold of the intermediate representations, and their pairings, as well). Thus the binding theory might hold at LF and or at SS, but it will not apply directly to the representation (14b), since that is not a 'level of representation' structure .

In the Williams and van Riemsdijk model, the derivation of the sentence "Who was John introduced to?" proceeds in the same sequence, although the SS is derived by two different rules. the rule of NP-movement applies to the D-structure (15a) to produce the NP-structure (15b), which then is transformed via the rule of WH-movement into the S-Structure (15c).

(15)

- a) D-Str: [[[]C [[was]C introduced John [to who]]]
 CP C' C IP I' I VP PP VP
- b) NP-Structure: [[[]C John [[was]C introduced t [to who]
 CP C' C IP i I' I VP i PP
- c) S-Str: [who [[was]C John [[t]C introduced t [to t
 CP k C' C j IP i I' I j VP i PP

They propose that the Binding Theory, at least Condition A, applies at NP-structure. For the derivational sequence in (15),

 14. The ordering of (14b) and (14c) is given by the Strict Cycle Condition.

the representation evaluated for well-formedness by Condition A is (15b), since this is the NP-structure of the sentence. Condition A, in its common formulation (22), applies to lexical anaphors, such as reflexive and reciprocal NPs, and to anaphor created through movement, such as the post-verbal trace in (11).

All lexical anaphors are present at D-Structure, and thus will be present at all subsequent levels of representation. This follows from the constraint of Recoverability of deletion, which prevents the deletion of a lexical anaphor in the course of the derivation. All anaphors introduced into a sentence in the course of the derivation -- that is, all anaphors created through movement -- will be anaphoric traces, created through NP-movement. Thus, it will always be the case that all anaphors of a sentence will be present at NP-Structure; there cannot be any new anaphors introduced after NP-structure in this model, since all post-NP-structure movement leaves a WH-variable, not an anaphor. It therefore is appropriate that NP-structure be the level of representation at which Condition A applies.

Williams and van Riemsdijk assert that this model of the grammar provides a solution to the problem of anaphor connectivity. As a result of the NP-structure model, Condition A applies before any instance of WH-movement. For the sentence (1a), repeated below as (16), there will be the derivation given in (17).

(16) (a) [which pictures of himself][does John like e]?

(17) a. DS: [John likes which pictures of himself] ----> (by move-NP)

b. NP-S: [John_i likes which pictures of himself_i] ---> (by move
wt)

c. S-S: [which pictures of himself_i][does John_i like e]?

Crucially, Condition A applies to the representation (17b), which is the NP-structure of the sentence. In this structure, the reflexive anaphor himself is c-commanded, in its governing category, by John, and thus the sentence is grammatical as far as Condition A is concerned.

The alleged fact that WH-movement does not alter possibilities for grammatically assigning an antecedent to an anaphor, as seems to be suggested by the well-formedness of (16), is in a certain sense theorematic in the NP-Structure theory.

However, let us return now to example (9), which we will shortly see falsifies the NP-structure model of grammar, at least in its present form, and specifically indicates that this model of grammar provides no solution to the general problem of anaphor connectivity, which is one of the model's major

15
motivations .

Example (9) allows an interpretation in which John binds the

15. See later sections of this thesis (esp. 3.3) for further arguments against the NP-structure model.

reflexive, which is not possible in the minimally contrastive (18).

(9) [[which pictures of himself]_i did John think [t'_i [Fred liked t_i]]]?

(18) John_j thinks that Fred_i likes these pictures of himself_{i/*j}

This interpretation is not allowed in the NP-structure theory, since at the pre-WH movement representation at which Condition A applies (NP-structure) the two sentences are indistinguishable in relevant respects. The NP-structure for (9) is given below, as (19):

(19) [John think [[Fred liked [which pictures of himself]]]]

In (19), as in (18), the only possible antecedent for the anaphor is Fred, since the embedded clause is the governing category for the anaphor, and John is not within the governing category. To conclude, the NP-structure model of the grammar incorrectly predicts sentence (9) to be on a par with (18).

2.0.2 Against the Dual S-Structure Hypothesis

The same objection holds with respect to models of grammar, including Hale (1983) and Mohanan (1983), in which S-Structure is viewed as containing two simultaneous representations, Phrase Structure, in which WH-movement occurs (but NP-movement does not), and Lexical Structure, in which NP-movement occurs, but WH-movement does not occur. Let us call this the Dual

S-Structure Model. In this type of model, (21) is the pair of S-structure representations associated with the sentence having the D-structure (20):

(20) e was convinced John that Fred likes which picture of himself

(21) a. Lexical Structure at S-structure (derived from (20) via NP-movement):

John_i was convinced t_i Fred likes which pictures of himself

b. Phrase Structure at S-Structure (derived from (20) via WH-movement):

[which pictures of himself]_i was [convinced John that Fred likes e

Condition A applies to the Lexical Structure representations. In this model, which is indistinguishable from the NP-structure model in relevant respects, the Lexical Structure corresponds to the NP-structure of the Williams and van Riemsdijk theory; in both theories, this representation (whatever we call it) is unaffected by WH movement, and is the representation to which Condition A applies. The entailment is that for examples like (1a) and (9), the anaphor is predicted to be bound by the NP which might have bound it locally at D-structure.

As in the NP-structure proposal, this is offered essentially as a way to explain the grammaticality of (1) and such examples, and fails on exactly the same grounds: it cannot explain why successive cyclic movement makes it possible for the anaphor to be bound by an NP which is too far away from it to bind it at D-Structure.

For such considerations, I will consider the basic T-model of grammar to be correct, and focus on two issues: (1) is the rule of Reconstruction (as in (7)) an adequate and desirable way to handle the full range of data, and (2) what is the level of representation at which the binding principles apply.

To clearly lay out the issues: if a rule of reconstruction (something like the rule (7)) is to be invoked, then:

1. it must square with independently motivated constraints placed on the derivation and structure of LF representations; and
2. if Condition A, as formulated in standard GB theory
16
(Chomsky (1981); see (22) below), is adopted with no adjustments, then it must apply only at LF, not at S-structure.

If, as I shall show below, no descriptively adequate rule of

16. I here assume a basic familiarity with the Binding Theory of Chomsky (1981). I believe that the modification of the binding principles proposed in Chomsky (1986) (in part following proposals of Huang (1983)) is a significant advance over the (1981) formulation. In the next chapter, I will be primarily concerned with a modification of the (1986) binding theory, modified to apply directly to S-Structure representations, without any lowering operation invoked to handle connectivity. Because much of that chapter is concerned with the precise formulation of Condition A (and B), a complete, in-depth review of the (1981) theory, its modification in Chomsky (1986), and its precedents in the EST, will be deferred until that chapter. Such a detailed review is unnecessary here, and would be redundant with the next chapter's.

Reconstruction which doesn't violate other properties of LF is possible, then:

1. another way of approaching the data like (1a,b) and (9) must be offered,
 2. the BT might apply exclusively at S-Structure (in that, if Reconstruction is rejected, the structural relation between an anaphor and its antecedent will be the same at S-Structure and LF), and
 3. the abandonment of Reconstruction may entail a revision of the Binding principles, since the principles will have to apply directly to structures isomorphic to (1a)¹⁷.
- (22) Principle A (Chomsky 1981): an anaphor A is bound in the minimal maximal projection containing A, a governor for A, and a Subject or AGR accessible to A.
- X is accessible to Y (for our purposes) if X c-commands Y.
X binds Y iff X and Y are coindexed, and X c-commands Y.
- X c-commands Y iff (i) neither dominates the other, and (ii) the first branching node dominating X dominates Y. (Reinhart 1976)
- X governs Y iff (i) X, a head, c-commands Y and either (ii) every maximal projection dominating Y dominates X [Aoun & Sportiche 1981] or (iii) Y is a head of a maximal projection Z which is governed by X in the sense of (ii) [Rizzi and Belletti 1983]

17. See chapter three.

2.1 Two Distinct Theories of Anaphor Connectivity

If we keep to the basic T-model of grammar, we are left with two approaches, what I will call a lowering theory and a non-lowering theory. Lowering theory is one which contains a rule of Reconstruction, something like rule (7).

2.1.1 Lowering Theory

In lowering theory, at least part of the dislocated phrase containing the anaphor in an example like (1a) or (9), is lowered at LF and substituted for one of the traces the WH-phrase binds at S-Structure. The result is that the antecedent of the reflexive actually c-commands it when Condition A applies. If we suppose (as Langendoen and Battistella (1981) suggest) that the landing site of Reconstruction can be an intermediate trace in COMP, then the lowering rule of Reconstruction can ambiguously apply to the sentence (9) (repeated here), giving either (23) or (24) as the output LF representations.

9) [[which pictures of himself] did John think [t' _i CP i
[IP Fred liked t _i]]]?

(23) [[which x: pictures] did_i John_S think_j [t' CP i S Fred_k

(25). Reconstruction Principle [Hornstein, p.101]

In a structure like

$$[\dots [\dots X \dots] \dots Y \dots t \dots]$$

Z_i
 i

X and Y can be related just in case such a relation is permitted with the phrase Z interpreted from the position marked by 't'.

18

Hornstein's condition is little more than a description of the data. We must both determine whether this general approach is correct, and then seek a treatment which does not stipulate as much, or which is at least formulated in terms of independently attested relations of grammar. Such a formulation will be the topic of Chapter Three.

18. For our purposes, and indeed Hornstein's, the general statement 'X and Y can be related...' is to be interpreted as 'X can be anaphorically dependent upon Y...'. There are many relations of grammar which aren't preserved in the structure described in (25), so that (25) taken literally is descriptively incorrect. An example is government; clearly, a verb does not govern PRO in subject position, even though it would if the NP had not moved:

(i) John tried LPRU (=X) to be arrested (=Y) t ?
i i

If the verb did govern PRO, then the sentence in (i) would be incorrectly ruled out by the PRO theorem, which blocks PRO from occurring in governed contexts (see Chomsky (1981)). Another case is the following:

(ii) [John_i seems [t'_i to have been arrested t_i]]

Take John = X, John = Z, and t'_i = Y. Clearly, John and t'_i can be related (through binding, and perhaps proper government), yet they certainly could not if John were in the position marked by t_i.

We must be rather careful here as to the role played by lowering operations in the two treatments. If a lowering treatment of connectivity is adopted, along with Condition A as formulated in (22), such lowering must be possible in all grammatical cases of anaphor connectivity. If a non-lowering theory is adopted, lowering at LF may or may not be possible. That is, the treatment of cases like (1a,b) and (9), on this latter view, will not depend upon any operation of lowering at LF. General lowering movements at LF will be consistent with this later treatment, but irrelevant to it. A non-lowering account of connectivity will be neutral with respect to lowering, while the former view requires that lowering be an available option at LF. (I shall show, in Chapter Three, that certain cases of lowering and movement must be blocked in the LF component.)

2.2 On Lowering Treatments of Connectivity: What Might The Reconstruction Rule Be?

I would like to suggest that lowering theory and non-lowering theory give rather different views of the structure of LF representations, that there are reasons to think that a theory involving general lowering is undesirable, and that a particular formulation of a non-lowering treatment (see Chapter Three) has other desirable consequences.

First, let us examine lowering theory more carefully. If lowering at LF is to account for the reconstruction data of (1) and (9), we must then ask what constituent it is which lowers.

2.2.1 Reconstruction and Strong Crossover

Reconstruction was originally suggested by Chomsky (1976), in order to account for the apparent Strong Crossover effect exhibited in the sentence (26) below. The interpretation in which he is understood as a variable bound by who is not

19
permitted .

(26) * [who_i se mother]_k [does he_i love e_k]?

If (26) is initially converted (between SS and LF) to (27) by

19. One might alternatively suppose that (26) is ill-formed because the variable will not be within the scope of the operator who. However, the availability of the bound-variable interpretation of him in (i) indicates that the possessor may have scope outside the NP containing it, plausibly due to further extraction, as in (ii).

(i) [every man]_i ['s mother]_j loves him_i

(for every man x, x's mother loves x)

(ii) [every man]_i [e_i 's mother]_j [t_j loves him]_i

See May (1977), (1985), and Higginbotham (1980), (1983). This indicates that (26) could have the LF (iii), in which who has scope over the pronoun.

(iii) [who]_i [e_i 's mother]_k [he_i loves t]_k

raising the interrogative possessor out of the WH phrase in COMP, and then lowering the remainder, namely [e's [mother]], down into the sentence-internal extraction site, as in (28), the illicit interpretation of (26) will be ruled out.

(27) [who] [e 's mother] [he loves e]
 i i k i k

(intermediate representation, derived from (26))

(28) LF Representation: [[who] [he loves le 's mother]]
 i i i

In the framework of Chomsky (1976), the derived structure (28) violates the Leftness Condition (a constraint formulated to rule out cases of Weak Crossover, like 'Who does the woman he loves admire?', in which the pronoun cannot be understood as a variable bound to who). This condition is given as (29) ²⁰.

(29) The Leftness Condition: A pronominal may not appear to the left of a variable with which it covaries.

In (28) the pronoun is to the left of the variable, and so the leftness condition rules out the representation in which they

 20. Chomsky (1977;201) gives it as the following:

(105) A variable cannot be the antecedent of a pronoun to its left.

Although (105) was formulated to rule out cases of weak crossover, it extends to these cases as well.

21
are coindexed.

In more recent terms, the structure (28) is ruled out by Condition C of the Binding theory, given in (30).

(30) Condition C: An R-expression must be A-free.

The class of R-expressions contains empty category variables, and the empty category in (28) is a variable. In the LF structure (28), the pronoun c-commands *e*. Thus, if they are coindexed a Condition C violation will result. Because the pronoun and the variable must be coindexed for the pronoun to be understood as a variable, Condition C will prohibit such an interpretation.

(31) is a subcase of (30), violated in the indexed LF structure (28).

(31): a pronominal *P* may not c-command a variable of a
quantifier *Q*, if

21. As a side note, Chomsky's article was one of the first works to argue for a level of representation distinct from S-Structure, namely LF, at which such relations as operator-pronoun binding would be encoded. The operation of reconstruction was both introduced to collapse strong crossover examples like (i) with apparent crossover cases like (26), and to justify this new level of representation.

(i) * Who does he love *e* ?
 i i i

It is only after the lowering operation, in the LF representation, that the pronoun is to the left of the coindexed variable.

P is to be understood as a variable bound to Q .

Condition C of course also rules out simpler cases of strong crossover, as in (32), in which there is no lowering involved.

(32). a. * who [does he love e]?

i i i

b. * he saw everyone

i i

(LF: [everyone][he saw e])

i i i

In each case ((26) and (32)), if obligatory lowering is utilized, the pronoun will c-command the object NP at LF, and so cannot be coindexed with any variable inside it, as would be needed for the bound variable interpretation of the pronoun.

2.2.2 Problems With Reconstruction and Crossover

Such an analysis of Strong Crossover (SCO) is shown to be insufficient by Higginbotham (1980), who observes a similar SCO effect in (33), where he cannot be understood as a variable dependent upon which man.

(33)a.) [which picture of which man][did he see e] (S-Structure)

b.) *[which man] [which picture of e] [he see e]

i i k i k

22. it should be noted that this corresponds closely to the second clause of Higginbotham's (1983) Accessibility Condition, which formulates the conditions under which a pronoun can be understood as a variable with respect to some quantificational NP Q. In Chapter Four, I examine Higginbotham's condition (which I believe to be the right condition on QNP-pronoun dependencies) closely; the subcase (31) will do here, for present purposes.

The problem for a lowering treatment of this example is that the phrase out of which which man_i moves in the derivation of LF paralleling (26-28), namely [which pictures of e]_i^{NP}, is itself an interrogative, and so must remain sentence-external, as in the final representation in (33b). The requirement that every WH-phrase have clausal scope forces (33b) to be the LF representation of the S-Structure (33a). In order to rule the sentence out as a simple Condition C violation, the representation (33b) would have to be transformed into (33c), in order that the empty category variable e_i will be c-commanded by the pronoun he_i.

(33c) [which man]_i [he]_i see [which picture of e]_{i k}]

But (33c) violates the requirement that all interrogative phrases be in COMP at LF (see Higginbotham (1980), (1983); Lasnik and Saito (1984) for discussion). Thus the SCO sense of the intended interpretation of (33a) must be ruled out by something other than Condition C. This raises the question of whether Lowering has anything to do with blocking cases of crossover like (26) and (32a).

As an alternative to lowering, Higginbotham introduces the filter in (34), which rules out the bound-variable interpretation of the pronoun in (33a), without the problem

b. LF : [every man]_i [every picture of e]_{i k} [he likes e]_{i k}

2.2.3 A Possible Counterargument

It might erroneously be supposed that a lowering account of SCO in examples like (26) and (33) and (35) could be retained by introducing a further rule into the syntax of LF, a rule which would (as part of the S-Structure-to-LF mapping) divide up any interrogative into two separate syntactic constituents. Such a rule is given in (36), and its operation illustrated in (37).

The S-Structure constituent [which man] is divided by the rule into two parts, the interrogative quantifier [which(x)] and the restrictive predicate [man(x)].

(36)
Quantifier Decomposition (c.f. Rule 1 of
Higginbotham and May (1981a)):

[Q N']	→	[Qx][x N']
[WH N']		[WHx][x N']

where [Qx] and [x N'] are separate syntactic constituents, as are [WHx] and [x N'].

- (37)a. (S-Structure) [which man][did he see e] ---> (by (39))
 b. [which x][x man][he saw e] ---> (by Lowering)
 c. (LF) [which x][he saw [x man]]

This hypothesized rule resembles the first part of van Riemsdijk and Williams's reconstruction rule (see (7)), except

that it, unlike their rule, makes the quantifier and the restriction separate syntactic phrases. For all cases of anaphor connectivity (in questions, at least), the anaphor is in the restriction on the interrogative. This means that, if rule (36) were permissible, and adopted, all cases of anaphor connectivity might reduce to lowering, followed by the LF application of Condition A. The important effect of the rule (36) is that the restriction can, by this hypothesis, be lowered independently of the interrogative, leaving the quantifier in scope position.

This result is shown in (37) and (38); crucially, (38e) is the LF representation which would be assigned to Higginbotham's (38a).

In both LF representations (37c) and (38e), the pronoun he c-commands [x man], and Condition C, as in (30), prohibits coindexing them.

38.a.) [which picture of which man]_i[did he see e]_k (S-Structure)

(38)a. [which picture of which man]_i[did he see e]_k ---->

b. [which man]_i [which picture of e]_k [he see e]_k ---->

c. [which x]_k[x man]_k[which y]_k[y picture of x]_k [he see e]_k ---->

d. [which x]_k[which y]_k[y picture of [x man]]_k [he saw e]_k ---->

e. [which x]_k[which y]_k[he saw [y picture of [x man]]]_k

((38e) is the LF representation of (38a), if we allow the rule

(39). (a) [which x]_i[x pictures of himself] [does John like e]? _i

(b) [which x]_i[does John like [x pictures of himself]]? _i

(Final LF representation, to which Condition C will apply)

I think, however, any such approach suffers from at least three serious problems, which render this extended Lowering account of anaphor connectivity unavailable in principle.

2.3 Against the Quantifier Decomposition Rule

2.3.1 Objection One: Unrestricted Quantification

First, if such decomposition were possible, we could then generate sentences with unrestricted quantifiers, which, it is generally assumed, do not occur in natural language (see Higginbotham and May (1982); May (1985)). For example, a sentence like (40) contains in direct object position the quantificational noun phrase 'most people', which consists of a quantifier 'most' and a head noun 'people'. The noun serves to restrict the portion of the domain of discourse over which the variable in the matrix will quantify (the variable in question is the trace left by assigning scope to the QNP, by May's rule QR). Assuming the LF structure (41) for the sentence in (40), following May (1977; 1985) with scope assigned to the quantificational NP as a whole, and assuming the analysis of

Determiner-Noun relations set forth in Higginbotham (1983), in which determiners, including such quantifiers as 'most', bind thematic grid positions in the nominal, we have a straightforward representation of the restricted

24
quantification .

24. A restricted quantification is expressed in the following formula:

(i) [Qx: P(x)] [Z(x)]

Here, P(x) is the restriction on the quantification. Given a domain D of individuals, the predicate P serves to restrict the domain over which the rightmost occurrence of the variable x ranges to the subset X of D consisting of the individuals in D which serve to satisfy P. (But see Higginbotham and May (1981b) for discussion of only, which is exceptional in this respect. This footnote is adapted in part from the appendix of their article.) To illustrate restricted quantification in natural language, consider a typical English sentence like (ii).

(ii) Many men run

(ii) expresses the statement paraphrased as 'many men x are such that x runs'. We assume May's rule of QR, which moves the quantified NP many men out of the matrix clause as in (iii), assigning it scope (defining scope as the c-command domain of the QNP). This movement leaves a syntactic trace t, which corresponds to the rightmost occurrence of x in (i). The VP (more precisely, INFL') has one open theta-grid position, and is so predicated of t. Thus, the syntactic constituent [t runs] corresponds part for part to [Z(x)] in (i).

(iii) [many men] [t run]

Many men contains two syntactic constituents, the quantifier many and the noun men. Following Higginbotham (1983), we suppose that men has one open theta-grid position, and is thus in essence a one-place predicate. The quantifier binds this theta-grid position, and so the QRed NP in (iii) corresponds part for part to the expression [Qx: P(x)] in (i). The quantifier and the N' which is a sister to it semantically restrict those individuals in D over which t ranges to those which are men. Which subset of D is chosen depends upon the (lexically specified) nature of Q. In (iii), then, the overall

(40). John likes most people

(41) [[most [people]] [John likes e]]
 S NP N' i S i
 <1*> <1>

In specific, the restriction is syntactically present as the N' of the QNP, and the semantic restriction takes place when the quantifier binds the single thematic grid position of 'people'. There are two open sentences, each bound by the quantifier: [John likes e], and [people]. They are distinguished by the fact that the variable of the first open sentence is an empty category, a syntactic trace created through the scope-assigning rule QR, while the variable of the second open sentence is a thematic grid position. A restricted quantification requires the presence of the two, separate, open sentences.

The basic operation of Reconstruction as offered by Chomsky and advocated by others has as one crucial step the substitution of part of the phrase in COMP for the syntactic trace left through movement. In the derivation (42) (paralleling (26-28)), for example, the NP [e's mother] is 'lowered into' the trace e, that is, substituted for it. This meets the requirements for substitution: the two phrases involved are of the same category,

 quantification expressed is true iff for the subset of individuals which are men, many of them run (equivalently, for many members of the subset, they run). It is rather important to my in-text discussion that there is a tight correspondence between the syntactic constituency of the LF representation of a sentence including a quantifier and the semantic constituents relevant for its interpretation. See the appendix to this chapter for discussion.

they are coindexed, and the one eliminated through such movement is empty, so that there is no violation of the principle of recoverability of deletion (as would happen if we did the substitution the other way around)²⁵.

(42)

- a. [CP [NP [which man] 's[mother]]
 does [he like e]]
 S i
- b. [CP [NP [which man] [t 's[mother]]
 does [he like e]]
 S i
- c. [CP [which man] does [he like
 [t 's[mother]]]]
 NP j N' i

This elimination of the trace e_i in deriving (42c) has no import for the quantification in the sentence (42a), since although that trace is a variable, it is not one which is relevant to the quantification. The actual quantificational NP is 'which man'. The first open sentence, as in (41), is

25. Lasnik and Saito (in progress) argue that the Proper Binding Condition (a modification of Fiengo's (1977) condition) must hold at every step of the syntactic derivation including those between DS and SS, and between SS and LF. Fiengo's PBC requires that a trace be to the right of a coindexed element; Lasnik and Saito's modification is put in terms of c-command, requiring (like Longobardi's (1985) Scope Condition) a variable to be bound by its operator. The PBC is met here, but, as I will show later, many examples of lowering fail to meet it, thus arguing against reconstruction as a lowering process.

represented by the head noun [man], whose theta-grid position (bound by which) makes the sentence open. The initial scope assignment of [which man] in (42b), moving it out of [which man's [mother]], creates another trace, t_i. When the constituent [t_j 's mother]_i is put back into the sentence from whence it came, the presence of t_j in the sentence [he loves t_j 's mother] makes that an open sentence. The elimination of e_i (the trace left in the original, pre-S-structure movement which derived (42a)) in this lowering operation is of no consequence to the quantification. Note that this will be so only in cases where the WH element which remains in COMP is itself an argument at S-Structure; being an NP with a lexical noun (which will have a theta-grid), it will contain one open sentence, and being an argument, it occupies a position which can be filled by a variable created by the first step in a derivation like (42).

The story is somewhat different in a sentence like (40), on a derivation entirely permitted by the hypothesised rule (36). (36) will permit (indeed, force, since Reconstruction must be obligatory) a derivation (43) in which [most people] is divided into two constituents, one consisting of the quantifier, the other of the restrictive predicate, in such a way that the restrictive predicate is capable of being lowered independently of the quantifier into the trace in the matrix sentence.

(40) John likes most people

(43)(a). [most [people]] [John likes e]
IP
<1* > <1 >

(b) (derived by the application of the rule (36) of
Quantifier decomposition):

[most x] [x people] [John likes e]
IP
<1* > <1 >

(c) (derived by lowering (substitution))

[most x] [John likes x people]
IP

(43c) is now an LF representation with an unrestricted quantification, and fails to meet the actual interpretation of the sentence (40). This is detailed below.

The two quantifiers of standard first-order logic, the existential and universal quantifiers, are such that restricted quantifications involving them can always be equivalently represented with unrestricted quantifications, with the two open sentences connected by one of the logical connectives of first-order logic. For example, 'some man is tall' may be represented as in (44a) or (44b):

- (44) a) [Ex: man(x)] [tall(x)] (restricted)
b) (Ex) [man(x) & tall(x)] (unrestricted)
c) [V(x): man(x)] [mortal (x)] (restricted)
d) (Vx) [man(x) ---> mortal (x)] (unrestricted)

The same is true of the quantifier 'every', as (44c,d) illustrate, although the connective is different in the two

cases.

A non-standard quantifier like 'most' has no equivalency between restricted and non-restricted uses. The restricted quantification (45) for the sentence (40) does in fact properly characterize the actual meaning of (40), paraphrasable as 'most things that are people are such that John likes them'.

(40) John likes most people

(45) [most (x): people(x)][John likes x]

Crucially, the LF representation of (40) given in (43c) fails to describe the meaning of the sentence in (40). As we have seen in (44), the logical connective used to connect two open sentences each with a variable bound by an unrestricted quantifier to achieve the equivalent to a restricted quantification varies depending on what quantifier is involved. (Conjunction is what is needed to relate the two open sentences in the unrestricted quantification representation of an English sentence containing the existential quantifier 'some' in order to get the meaning to come out right, as in (44b), while the implicational connective is needed for a sentence involving unrestricted universal quantification, as in (44d)).

However, there is no connective, or even a sequence of connectives, which can get things right with a non-standard

26
like most . The problem is actually even worse in
since there is only one open sentence (one variable bound
quantifier) to begin with. For this reason it is
difficult to even paraphrase what (44c) means.

Clearly, no such rule as (36) can be permitted, since it
permits unrestricted quantification .

27
Further, if we take LF to be a level of syntactic
representation, subject to certain constraints on form and
derivation, any rule or process deriving LF representations must
fall within the bounds of these constraints. This brings us to
the second problem with rule (36), which, I remind the reader,
was offered as a way to retain a purely reconstructionalist
approach to all cases of strong crossover, including
Higginbotham's cases.

2.3.2 Objection Two: The Theta Criterion and Saturation

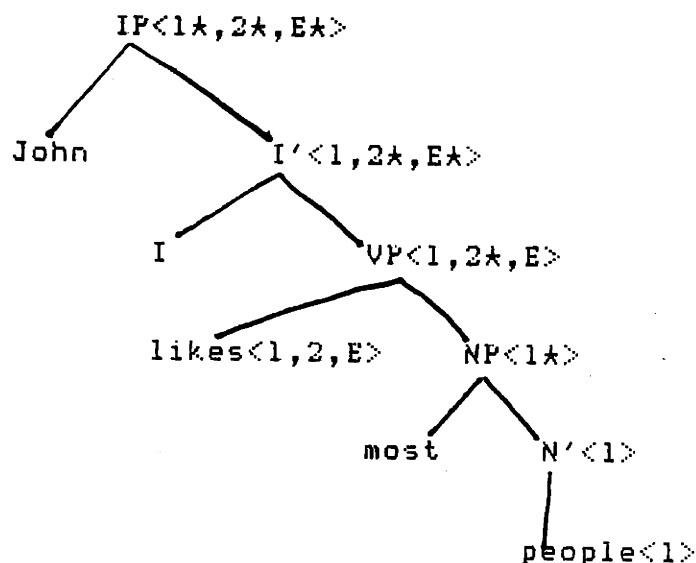
In the Surface Structure representation (46), the transitive
verb, in accordance with its lexically designated
subcategorization and/or theta-role assigning properties, takes

26. This is demonstrated by Cushing (1982). See the appendix
of this chapter for discussion.

27. See the appendix of this chapter for discussion. In the
appendix, I consider a way in which the restricted
quantification could be recovered from a 'reconstructed' LF like
(43c), and the implications this analysis has for the GB model.

an NP object. In the theory offered by Higginbotham (1985), the NP fills the theta-grid position (2) representing the verb's internal argument; this filling of theta-grid positions is notationally represented with a '*':

(46)



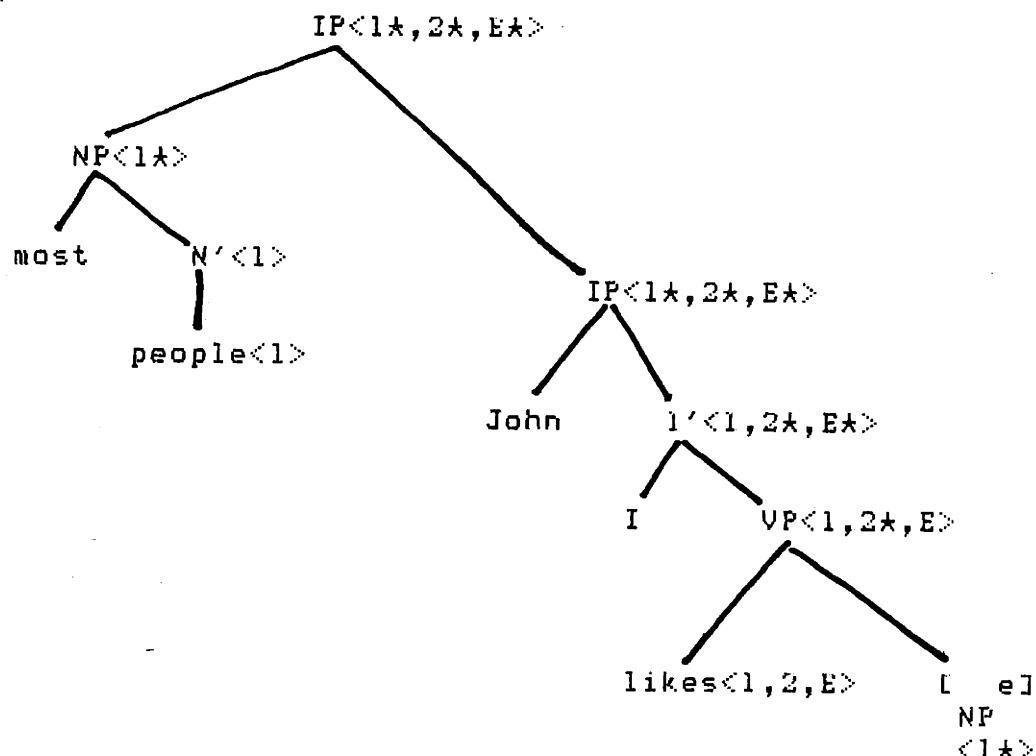
A constituent is saturated if every position in its theta-grid is filled.

Higginbotham proposes it as a fundamental condition that all

28. In this theory, developed in part following Stowell (1981), the verb is required to take an argument by the presence of a theta-position in its theta-grid. This argument will have to be a sister to the verb for thematic discharge to occur, and will have to be of a particular syntactic category due, presumably, to the specific nature of the theta-role. The relevance of the notion of strict subcategorization is reduced in such a theory, and it would be hoped that the category of the argument, in all cases, can be derived from a condition on the compatibility of a theta role with the NP with which it is associated. See Grimshaw (1979; 1981), Pesetsky (1985; 1986), Speas (1984; forthcoming), Higginbotham (1986), Fukui (forthcoming), and Rochette (forthcoming) for discussion.

arguments must be saturated (see also Rothstein (1983)). When the process of QR adjoins the QNP to its scopal domain in the derivation of LF, the NP carries along with it its internal structure, leaving behind a variable:

(47)



To satisfy the condition on argumenthood, which can be taken as

29. We note here, following a suggestion of Doug Saddy, that this conflicts with Chomsky's (1986) proposal that INFL theta-marks VP, since VP is not saturated and thus is not capable of being an argument. However, given Chomsky's revised definition of sisterhood, which permits the VP to directly theta-mark the subject NP, it will be the case that the VP is saturated at the same phrasal node --INFL' -- at which it serves as an argument of INFL, so it might be possible to reconcile the two proposals. For reasons of simplicity, I will follow Higginbotham's treatment of the INFL-VP relation, which is one essentially of binding.

a consequence of the Theta-Criterion, the phrase occupying the object position of the verb at LF must be thematically saturated. The trace in the LF representation above is thematically saturated, having no internal theta-roles

unsaturated (the theta grid on the variable, if there is one, can be taken as a copy of the one on its antecedent, the saturated NP [most people]), and so can saturate the (2) position of the verb's theta-grid.

However, consider the LF representation of sentence (40) if we were to permit into the grammar (obligatory) lowering after the application of the (obligatory) rule (36). That is, consider

30. The theta criterion requires that every thematic grid position be filled. A theta grid position can be filled (or discharged, in the terminology of Higginbotham (1983)), in a variety of ways, including theta-role assignment, theta-identification, and theta binding. As Higginbotham (1984) argues, since each of these takes place only under government, we can derive the statement that arguments must be saturated. Suppose a verb were to take an unsaturated argument, the unfilled position of the argument could never be discharged, since the potential filler would not be in a government relation with the argument. Thus a theta-criterion violation would occur. I am grateful to Margaret Speas for discussion.

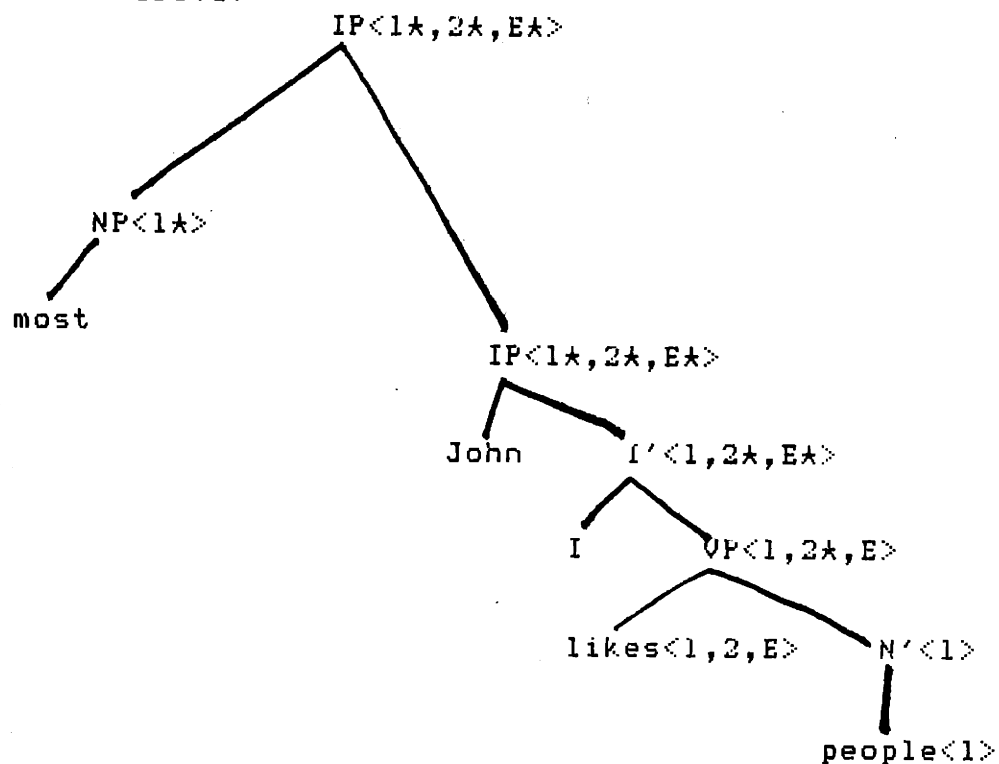
Jackendoff (1985) presents a number of interesting cases, including (i), which show that an argument may be multiply theta-marked. In (i), John be interpreted as both the actor (the one who initiates the event) and the theme (the rolled object). Crucially, the theta roles are all assigned to the same syntactic position. see Chomsky (1985) for discussion, and reformulation of the theta-criterion to reflect these facts.

(i) John rolled down the hill

31. This follows from the status of [e] as a variable.
NP

sentence (43c) in light of Higginbotham's condition. Since it is the quantifier 'most' which saturates the theta-grid position of the noun [people] and the N' [people] of (40), and since in (43b) (repeated below, with saturation indicated) the quantifier is structurally higher than the VP, the phrasal projection of 'people' is unsaturated at the point where it matters.

(48) (= (43c) above)



The representation above is ill-formed, since the N' [people]_{<1>} is not saturated, and thus cannot saturate the verb's <...2...> position. To put it slightly differently, it is an assumption of Reconstruction analyses that the lowered constituent has a variable in it, bound by the quantifier left in scope position. Within Higginbotham's theory, this variable

is the theta-grid position of the noun. By placing the N' into the position of the trace in deriving (43c/48), we are placing an unsaturated constituent into a position previously occupied by a fully saturated one, (namely the quantifier plus the N'). The theta criterion requires this position to be saturated at least at LF; the projection principle requires the argumenthood of the position to be uniform across levels of representation; and so this substitution is illicit.

This is why the only examples for which reconstruction appears to work -- like Chomsky's (26) -- are those in which the phrase moved to COMP contains a WH phrase which is itself an argument of some predicate included in the material in COMP.

Thus the rule of Quantifier decomposition (36) violates the theta-criterion, on one formulation, and the projection principle as well.

I now turn to a third problem which arises with a lowering process such as that made possible by (36). This problem is much more theory internal, yet is worthy of consideration.

2.3.3 Objection Three: The ECP And Rule (36)

The basic function of rule (36) is to separate a quantifier (in determiner position) from the rest of the NP, in order to lower the rest of the NP while still leaving the quantifier in a scope position. The objection raised in the last section had to

do with the view that this entailed lowering the thematically unsaturated N' into a position which requires thematic saturation. The theta grid of the noun is saturated by being bound by the determiner, on the projection from N' to N'', as we saw in (46)-(48). Since the lowered constituent in (48) has no determiner, I argued, the lowered element must be unsaturated.

Suppose, however, that the determiner is actually moved outside the NP, leaving a trace, and then the entire NP is lowered. This type of derivation, shown below in (49), differs from the one just considered in one crucial respect: the determiner's trace will remain within the lowered constituent, saturating it, and so the resultant representation will not

32

violate the theta-criterion .

This is essentially like the derivation offered by Chomsky (see (26)-(28)), with the exception of one crucial difference, which is that in Chomsky's examples the phrase moved in the first step is an argument of the NP in COMP out of which it moves, while in (49) the moved element is a determiner, not an argument.

(49)

Step 1: [most] [e [people]] [John likes e]]

? i NP i N' j IP i

32. This movement will, however, violate the Relativized A-over-A condition of Bresnan (1978).

Step 2: [most] [John likes [e [people]]]]]
 ? i IP NP i N' j
 <1*> <1>

The representation derived in Step 2 is the LF. The theta-criterion is satisfied, since [e [people]] is saturated (the theta-grid of its head noun is saturated). The problem that now arises is that in this representation, the empty category [e]_i bound by the determiner is not properly governed, in violation of the ECP. The question of whether lexical proper government can be eliminated, currently a topic of debate, is irrelevant, since this trace is not an argument of anything. The ECP requires this trace to be governed by its antecedent, which it is not. Therefore, a modification of rule (36), formulated to avoid the theta-criterion violation of the output of (36) (as we saw in the preceding section (3.2)), violates ECP. Both alternative rules violate the natural language constraint against unrestricted quantification.

Obviously, if reconstruction, e.g. LF-lowering, is the formal mechanism which accounts for the Strong Crossover effect in Chomsky's example, it must be obligatory. It must therefore be that the rule (36), if admitted into the grammar, must be obligatory as well. This means that there cannot be any natural language sentence with (i) WH determiners, or (ii) quantificational determiners. Any sentence with such determiners will end up at LF with a representation which violates the ECP, and which also violates a semantic condition

on natural language (that which blocks unrestricted quantification). Since obviously there are sentences with WH and quantificational determiners, it must actually be the case that no such rule as (36) exists, and that cases of Strong Crossover (including (38a)) are, exactly as Higginbotham proposes, to be ruled out by a representational constraint on LF which does not depend upon any type of lowering between S-Structure and LF.

This demonstrated, we now must consider whether cases of anaphor binding like (1) and (9) can be treated with lowering anyway, in spite of the fact that a representational theory of strong crossover, making reference to chains, is required. Having admitted a representational treatment of SCU, it appears to be of no conceptual cost to admit a representational theory of anaphor binding (something along the lines of (25)) which extends to the connectivity cases like (1) as well. Of course, this doesn't mean that we can now abandon a lowering account of connectivity a priori; rather, such a move can be made only on the grounds that it fails on its own merits. What follows in the next section is the same type of discussion of a lowering account of connectivity. In the discussion below I shall show that the cases of anaphor connectivity cannot be treated with a lowering theory, and that the theory of grammar must include a representational theory of this phenomenon.

2.4 Against a Lowering Treatment of Anaphor Connectivity

Consider again example (1a), a question which illustrates the basic phenomenon of anaphor connectivity:

1. (a) [which pictures of himself]_i does John like e_i?

Any lowering account of the grammaticality of that sentence with respect to anaphor binding will involve the following claims:

1. If Principle A of the Binding Theory is formulated as in Chomsky (1981), (see (22 above)), then it does not apply at S-Structure, since the S-structure (1a) violates Principle A.
2. As part of the derivation of the LF representation corresponding to the S-Structure (1a), some portion of the WH phrase which pictures of himself including the anaphor lowers to a position in which it will be c-commanded by John, the understood antecedent of the anaphor.
3. Condition A applies after this lowering operation, that is, at the level of LF, examining the representation derived through lowering, and ruling it grammatical since the anaphor is now c-commanded by its antecedent in a local domain, as required by Condition A.

There are, for the phrase which pictures of himself, four candidates for lowering: the entire WH-phrase, the N' pictures of himself, the PP of himself, or the anaphor alone. We will demonstrate below that none of the four can be lowered, thus arguing that a lowering treatment of anaphor connectivity is impossible, given other constraints of the grammar.

We observe, following Higginbotham (1980), that the WH phrase which pictures of himself cannot lower in toto between S-Structure and LF, since this will place the WH interrogative quantifier in a position where it ought not to be at LF, presuming, following Higginbotham (1980), Lasnik and Saito (1984), May (1985) and others, that all WH-interrogatives must

33

be in the matrix COMP (alternatively, matrix [SPEC, CP] at LF .

Further, following the discussion of the preceding section, we observe that the N' pictures of himself cannot lower, leaving which in COMP, since this will derive a structure which violates the requirement that natural language quantification be restricted, and violates either the theta criterion and the projection principle, or the ECP.

This leaves us with two alternatives: lower the PP, or lower

33. This means essentially that the SPEC of CP is the designated scope position (at least in English) for WH-phrases. Whether this is universal is still an open question; see Lasnik and Saito (1984), Chomsky (1986), Epstein (in progress) for discussion.

the anaphor in isolation.

2.4.1 Against PP lowering as a Solution

If we lower the PP by substituting it for the sentence-internal variable, the SS-to-LF derivation will proceed as follows:

(50)

a. S-Structure: [which pictures of himself] [does John like _i e]? _i

b. LF derived by PP-lowering:

[which pictures L e]] [does John like L of himself]]?
PP j i PP j

(50b) represents the outcome of such substitution. Note that the basic requirement of the unrevised Condition A (as in (22)) is met, since John locally c-commands the anaphor himself (where 'locally' is to be understood as 'within the governing category of the anaphor').

However, for other reasons, (50b) is an ill-formed LF representation. Since we have lowered an argument, namely the

object of the noun pictures³⁴, there must be a variable left

34. I am here ignoring the distinction between the PP and the NP inside it with respect to the thematic role assignment by the noun pictures. If proposals that of is simply a casemarker (see for example Chomsky (1985)) are correct, then this is a justifiable option.

(53)

[[which people] [[do]] you L t L wonder
 CP NP i C' C IP VP j VP
 [[which pictures of e] L [John L t L saw e]]]]]]]]]]
 CP NP i j C' IP VP j VP j

Here, the PBC is met. The sentence is of course not perfect. To my ear, it has the status of a subadjacency violation, in being unacceptable but quite interpretable (unlike an ECP violation). If it is indeed a Subadjacency violation, it suggests that wonder does not L-mark the NP [which pictures of which people],
 NP

forcing it and the CP dominating it to be barriers, causing this to be a strong Subadjacency violation.

(i) would be derived by first moving to whom to the intermediate SPEC position, then moving whom out of the PP, in effect stranding the preposition in the intermediate SPEC. The A-over-A condition fails to distinguish this illicit movement from the (relatively acceptable) extraction seen in (53).

The A-over-A Condition (A/A) is stated as follows (Chomsky (1973);p. 85) ['Conditions on Transformations'; reprinted in Chomsky (1977), Essays on Form and Interpretation; page number given is that of the (1977) publication]:

If a transformation applies to a structure of the form

[...[...]...]
 a A

where a is a cyclic node, then it must be so interpreted as to apply to the maximal phrase of the type A.

(A is maximal iff it is not dominated by another category of type A, within the cyclic domain a.)

In example (i), the two constituents of relevance, the two WH-phrases, are not only of the same category (NP), but they are both +WH, so even the most restricted version of the A/A is too strong, in incorrectly ruling out this example.

Since (52) can be replicated with PP movement, as in (54), with no improvement, it appears that the PBC must cover cases of PP movement as well. This means that, just as the constraint rules out (54), it must rule out (50b) as well.

(54)

* [[which pictures e] do you wonder[
 i j CP
 [of which people] l John saw e]]] ?
 PP i lP j

To sum up, the LF representation (50b) is ill-formed, and hence such derivations as (50a-b) cannot be offered as a solution to

37
 the connectivity problem .

There is a second problem with (50b), repeated here, which was derived through moving the PP of himself by substituting it for the variable left by overt WH movement of the WH NP which pictures of himself.

 37. Kyle Johnson observes that this argument rests on the assumption that the PBC holds at LF, as well as at S-Structure. It has not been demonstrated that the PBC holds at LF. One construction raising a problem for this assumption is (i), in which the raised QNP everyone may take narrow scope with respect to the predicate likely. If May's (1977;1985) analysis of this phenomenon (which involves lowering the QNP into the complement of the predicate, as in (ii)) is correct, the empty category left behind in matrix subject position is free. This suggests either that the PBC does not hold at LF, or else that it does not hold of the type of empty category left by lowering. See (May 1985) and Chapter (5) of this dissertation for discussion.

(i) Someone is [likely [t to leave]]]
 i lP i

(ii) EC is [likely [someone [t to leave]]]]
 i lP lP i

(50b). LF derived by PP-lowering:

[[which pictures _{PP} _j _i]] [does John like [_{PP} _j of himself]]]?
 PP j i PP j

The problem is that this is a structure with vacuous quantification; by deriving (50b) through this substitution, we have in effect eliminated the variable which the WH phrase binds at S-Structure in (50a), with of himself now occupying the object position of like. Natural language, as Chomsky (1982) and Higginbotham (1986) emphasize, does not permit vacuous

38
 quantification . The only way to avoid vacuous quantification is to adopt a rule like (36), giving rise to a representation (56b):

(36). Quantifier Decomposition

(c.f. Rule 1 of Higginbotham and May (1982))

[Q N']	----->	[Qx][x N']
[WH N']		[WHx][x N']

 38. That is, an LF representation in which a quantificational NP does not bind a syntactic variable is ungrammatical; such syntactic variables may be traces, as in (ii), or resumptive pronouns, as in (iii). The ungrammaticality of (i) is due to the ban against vacuous quantification, since who is a quantifier, and binds no syntactic variable. Since WH phrases must be generable in COMP (at least in resumptive pronoun languages), (i) cannot be ruled out (as has sometimes been suggested) because there is no well-formed D-Structure for it.

(i) * Who does John love Mary?

(ii) Who do you see t ?
 i i

(iii) Who do you wonder whether his mother has arrived yet?
 i i

where [Qx] and [x N'] are separate syntactic constituents, as are [WHx] and [x N']

55

a. [many people] [t eat sandwiches] --->

b. [many x] [x people] [t eat sandwiches]

(56a) will be the intermediate representation derived by applying the rule (36) to (50a), prior to PP lowering. (56b) is the LF representation derived from (56a) through PP lowering.

(56)

a. [which pictures x]_j[x of himself]_i[John likes e] --->

b. [which pictures x]_j[John likes x of himself]

However, we saw in (section 3) that no rule of the form (36) can be permitted in the grammar.

And, finally, such substitution is also illicit since it substitutes a category (PP) for another (the trace e) with which it is contraindexed.

There is, however, another way in which the S-structure in (50a) can be converted by PP movement into an LF representation in which the anaphor himself is c-commanded by John. The derivation will involve moving the PP out of the WH phrase and adjoining it to some node c-commanded by John, that is, deriving

the representation through ³⁹extraposition of the PP, rather

39. I use the term extraposition here descriptively, in the approximate sense of Ross(1967) (who terms it PP-extraposition),

than substituting it for the clause-internal variable. (57a)

(57)

[which pictures of himself] ldoes John like e]?

[[which pictures L e]] [does
CP NP PP j i
John [L like e][of himself]]]]?
VP VP i PP j

(i). La man [CP who [IP I liked t']] arrived yesterday
CP k IP k i k

(iii). [NP pictures [PP of Mary]] were published
NP PP i k

The term extraposition is also used descriptively in much syntactic work for the pronoun-CP relation in sentences like (vi), which were presumed in earlier work (and are presumed in some current work) to be transformationally related to (v).

(vi) it is amusing [that Mary has purple hair]

Those advocating deriving (vi) from (v) postulate a (rightward) rule of Extraposition (see Rosenbaum (1967) for the original proposal); those advocating deriving (v) from (vi) postulate a rule of Intraposition (see Emonds (1970)). I presume that the relations between (i) and (ii) and between (iii) and (iv) are of a different sort than the relation between (v) and (vi). See Burzio (1981; 1985), Guéron and May (1984), Chomsky (1985), Pollock (1985), Johnson (1985), and Deprez (1986), and the references cited there for discussion.

Note that in this derived LF representation John c-commands himself, the goal of this proposal. The PP might adjoin to VP, as indicated in (57c), or it might adjoin to INFL', or (on some definitions of c-command) to IP. There are reasons why no such derivation as in (50a-c) is possible, for each choice of a specific landing site. For ease of exposition I will assume for the moment that the landing site is VP-adjunction, arguing against each type of adjunction in turn shortly.

This analysis is attractive, since it is the case that the LF in (57c) has a close counterpart, (58), which involves overt extraposition; (57a) and (58) would, on this view, simply differ in the component in which extraposition of the PP has taken place (SS-SS for (58); SS-LF for (57)).

(58) Which pictures does John like of himself?

We saw above that (50b) was ungrammatical for at least three reasons. Two of the reasons do not extend to (57c). (57c) is not a structure of vacuous quantification, since it leaves the variable e in the representation. Further, the derivation of i

(50b) violated a constraint on substitution, since it replaced an NP with index i with a PP with index j; this as well has not occurred in (57c), since no substitution is involved.

However, (57c) will be grammatical only if (1) the structure meets the Proper Binding Condition, which, as we saw earlier, is the third reason why (50b) is ungrammatical, and (2) such

extraposition is allowable in its own right in the S-structure-to-LF mapping. I shall treat these questions in turn.

The question of whether the PBC is met in (57c) rests largely on the selection of a landing site for PP, and on the definition of c-command. Taking S to be a maximal projection, it will not be the case, on any recently offered definitions of c-command, that the PP will c-command the trace inside [which pictures] if the landing site is VP-adjunction. This is so since there is both a branching node (several, actually: VP; I' IP; C'), and one maximal category (IP) dominating [of himself] which does not dominate e. Thus the landing site cannot be VP-adjunction.

i

40

For the same reasons, it cannot be [NEL' adjunction].

However, suppose that the PP in this example is extraposed from inside [which pictures ...] and adjoined to IP, giving us (59):

(59)

[[which pictures [e]] [does [[John
CP PP j i IP IP
[like e]] [of himself]]]]?
VP i PP j

In this representation, we will have to show two things to be

40. There are other reasons to suppose that adjunction to a non-maximal projection is prohibited; see Chomsky (1986), May (1984), and Fukui (forthcoming) for discussion).

true: that there is a defensible definition of c-command under which PP c-commands its trace inside the WH phrase, and that John c-commands the anaphor inside PP. If either of these fails to be true, then extraposition of PP is always going to violate the PBC, and cannot be adopted as a solution to the anaphor connectivity problem.

Take IP to be maximal. Then, the branching node definition of c-command (60) will have to be abandoned, since it will both block John from c-commanding himself, and block of himself from c-commanding its trace. The definition (61), by contrast, will permit the IP-adjoined PP to m-command its own trace. So, let us assume that it is the definition in (61), not (60), which is

41

relevant to the notion 'bind' ; i.e. c-command is equivalent to m-command. However, the definition (61) will still block John from c-commanding the anaphor, since IP is maximal, and includes John but does not include PP.

(60) A c-commands B iff every branching node dominating A⁴² dominates B, and A does not dominate B

(61) A m-commands B iff every maximal projection category which includes A includes B.

(62) X includes A iff every segment of X dominates A.

41. See Barss (1985), Chomsky (1985), Saito (1984) for arguments that the necessary relation is that in (60).

42. Equivalently, A c-commands B iff A is a sister to a node which dominates B.

(63) X is a projection category iff it is a set of nodes
 $(N_1, \dots, N_i, \dots, N_n)$, such that:

- (i) every N_i is a projection of the same head, and
- (ii) the set can be ordered such that for every N_i ,
 N_i dominates N_j , for every $j, i < j \leq n$.

Each N_i is a segment of the projection category X .

(64) X is a maximal projection category iff:
 X is a projection category, and each N_i is X'' .

Thus, on this 'extraposition of PP' approach to the connectivity seen in (50a) and (57a), we must take IP to be non-maximal. This will permit, on the m-command definition, both necessary c-command relations to obtain. CP will be the first maximal projection category including John, of himself, or e_i. John will c-command the anaphor, and thus bind it locally, as required by the unrevised Condition A, and the PP will c-command and hence bind its trace, as required by the PBC.

To sum up, the LF-extraposition analysis would seem to provide a solution to the connectivity of (50a), and give us a clue as to the proper definition of c-command (suggesting that the definition (61) is the correct one), and the status of IP as a non-maximal category.

However, if we examine an ambiguous case like (9), repeated here, we see that this analysis fails.

(9) which pictures of himself does John think Dan really likes e?

We saw earlier that this sentence is ambiguous, with either John or Dan capable of anteceding the anaphor. The extraposition analysis straightforwardly predicts the availability of the interpretation under which John can bind the reflexive; this will involve PP extraposition just as in (57), from the matrix SPEC phrase [which pictures of himself] to IP-adjunction in the main clause:

(65) which pictures of himself does John think Dan really likes e?
(S-Structure)

b) [which pictures [e]] [does [[John[think
PP j i IP IP VP

[t [Dan really likes e]]] [of himself]]]]?
CP i IP i PP j

(LF representation)

Here, John binds himself within the anaphor's governing category, and of himself binds its trace in the WH phrase.

Crucially, this analysis cannot provide a derivation of a well-formed LF representation in which Dan can antecede the anaphor. In order for this to occur, the PP of himself will have to extrapose and adjoin to the lower IP, as in (66):

(66) [which pictures [e]] [does [John[think
PP j i IP VP
[t [[Dan really likes e]] [of himself]]]]]]?
CP i IP IP i PP j

But, clearly, the PBC is violated in this representation. Even with the m-command definition of c-command (61), and the hypothesis that IP is non-maximal, the PP of himself is too far

down in the structure to c-command its trace (PP is dominated by the lower CP node, and the higher VP node, both of which are maximal). This PP fails to bind its trace, in violation of the

⁴³
PBC .

To conclude this section, we have considered two ways in which connectivity might be seen as the consequence of post-S-Structure PP movement from SPEC, in essence partially undoing pied piping. We have seen that independently motivated principles, especially the PBC, prevent this analysis from extending to the full range of cases. Finally, the definitions of c-command and the status of IP as a maximal or non-maximal projection are still open.

2.4.2 NP Lowering

This leaves us with one final candidate for lowering in an example like (1a), namely to lower only the NP himself, leaving all the other material in SPEC:

1. (a) [which pictures of himself] [does John like e]?
 i j i j

43. See Gueron and May (1984) for an empirically equivalent treatment of PP-extrapolation (not connectivity), put in terms of a failure of government between the moved constituent and the phrase out of which it moves; they argue that mutual government between these two categories is necessary at LF. See Chapter (5) of this thesis for review and discussion of Gueron and May's proposal.

This will derive the LF representation (67), if the 'lowering' movement is substitution, and (68) if the movement is adjunction, i.e. extraposition:

(67) [which pictures of e] [did John see himself]
 i j *i* *i*

(68) [which pictures of e] [did [[John see e] himself]]
 i j IP IP i j i

Such structures are ungrammatical for basically the same reasons as their counterparts involving PP extraposition are (see previous subsection). The LF representation (67) has a vacuous quantifier, and a free trace (e_i), and is furthermore

derived through an illicit substitution operation. The LF representation (68) has a free trace, and, at least on one definition of c-command (the branching node definition (60)) a free anaphor.

I therefore conclude that the anaphor connectivity effect cannot be explained in terms of any lowering operation at LF. Reconstruction, as originally proposed by Chomsky to handle certain cases of Strong Crossover, is an insufficient mechanism for dealing with this type of phenomenon.

The task now is to seek a proper formulation of a theory which conforms to the general requirement that it apply directly to such representations as (1a), without any type of lowering involved. As suggested earlier, this will entail either modifying the formulation of Condition A, permitting an anaphor

to in fact not be c-commanded by its understood antecedent, or modifying the definition of binding (so as to not require c-command) or modify one of the definitions of c-command, so that in (1a) the phrase John will actually c-command the anaphor, and thus bind it. This last suggestion is actually proposed by Chomsky (1981;p??).

Suppose the definition of 'c-command' is altered from (69) to (69'), leaving (70) unchanged, or that (70) is changed to (70'), leaving (69) unchanged.

- (69) A c-commands B iff every branching node dominating A dominates B.
- (70) A binds B iff A c-commands B, and A and B are coindexed.
- 69') A c-commands B iff for some node X dominating B, there is a chain C containing X and Y (X and Y not necessarily distinct), such that every branching node dominating A dominates Y.
- 70') A binds B iff (i) A and B are coindexed; and either (ii) or (iii).
- (ii) A c-commands B (by definition (69))
- (iii) X dominates B, A and Y are members of the same well-formed chain, and A c-commands Y.

To illustrate (69'), consider the following structure.

(69'') [which pictures of himself] [does John like e]
 j i j i

By definition (69'), John will c-command himself in this

structure. Take John to be A, and himself to be B; the chain in question is the pair ([which pictures of himself], e). Take the head of this chain (the WH phrase which pictures of himself) to be X, and finally take the empty category e_i to be Y. In (69''), every branching node dominating A (John) dominates Y (the empty category), and so John c-commands the anaphor himself.

With the adoption of either (69') or (70'), it will be the case that John binds himself in (1a). Condition A of the Binding Theory can remain unchanged.

However, this proposal falsely predicts that in a sentence like (71), there must not be coreference between he and John, since by the proposed adoption of (69') or (70') the pronoun will bind John in violation of Condition C.

(71) which of Mary's stories about John and Sue does
 he deny?
 e_i

(compare: * He denies all of Mary's stories
 e_i
 about John and Sue)
 i

The proposal also falsely predicts that a sentence like (72) will not violate the Proper Binding Condition, since the variable e_i will now be bound by which people, since which people c-commands e_i by (69'), thus binding it by (70), or fails to c-command it but binds it by (70').

(72) * [[which pictures of e] [[do] [you
 CP NP i j C' C IP
 [t [wonder [[which people] [
 VP j VP CP NP i C'
 [John [t [saw e]]]]]]]]
 IP VP j VP j

Thus it would appear that the definition of c-command and binding (in the indexing theory of anaphoric dependence) must remain unaltered. This means that it is Principle A of the Binding Theory which must be modified to permit such examples as (1a), and it is this task to which I turn.

2.4.3 At Which Level Does Condition A Apply?

It was observed earlier that a lowering theory of connectivity is committed to the view that anaphor binding is essentially an LF phenomenon, since it is only after the post-S-Structure lowering operation that the anaphor will be c-commanded by the understood antecedent. If it is the case that the LF representation of (1a) is essentially isomorphic to the S-Structure, as the rejection of the rule of Reconstruction would imply, then this raises anew the possibility that Condition A applies solely to S-Structure.

Before turning to the formal modification of Condition A, I wish to point out that there is some empirical evidence that anaphor binding is an SS phenomenon. Consider the sentences

(73) and (74) :

(73) John wonders where Mary bought what

(74) John wonders what Mary bought where

It has been observed that, although (74) is perhaps slightly less acceptable than (73), it is nonetheless a relatively grammatical sentence, as we can see by comparing (74) to (75), for example.

(75) is presumably an ECP violation (see Lasnik and Saito (1984) for discussion).

(75) * John wonder what Mary bought why

(c.f. 'John wonders why Mary bought what').

We have observed earlier that there is a 'multiple binding domain' effect in such examples as (9), in which the anaphor, having been moved through WH movement to a position outside the c-command domain of the lower subject John, can have Dan as its antecedent (a more formal discussion of this effect is developed in section (3.5), and in Chapter 3). This effect is exhibited as well in (76).

(76) Dan wonders which pictures of himself John is unaware of

Now consider (77) vs. (78).

(77) Dan wonders which pictures of himself John bought where

44. See Fiengo (1980; chapter 1), Huang (1982; chapter 7), Lasnik and Saito (1984, fn.12) for discussion.

In the section below, I will sketch a view of the connectivity effect which is an extension of recent work by Richard Kayne, principally a suggestion advanced at the end of his article 'Connectedness' (Kayne 1983), which proposes that instances of lexical anaphora are governed by the same condition of grammar, his Connectedness Condition, as governs WH-extraction. The Connectedness Condition was originally proposed to deal with island constraints on WH movement, serving as an alternative to the ECP and the Subjacency Condition of Chomsky (1981) and other work. I shall show that although such a Connectedness approach to lexical anaphora is desirable, in that it would (partially) subsume lexical anaphor binding under a theory which has far-reaching consequences for the theory of WH movement, it nonetheless appears not to be viable. The section below thus can best be viewed as a demonstration that the (sub)theories governing the distribution of trace and the distribution of lexical anaphors must be distinct.

2.5 Connectedness Theory and Anaphoric Paths

A central notion of Kayne's theory is the 'g-projection' (government projection), a path or more formally a set of nodes in a phrase marker which project upwards from some point in the tree, where the structural relation of government plays a crucial role in this projection. I will adopt the definition of

g-projection suggested by Longobardi (forthcoming) .

(82) G-Projection (adapted from Longobardi (forthcoming), extended with some differences from Kayne 1983):

definition: a node X is a g(overnment)-projection of a node Y iff:

- i) X is an X-bar projection of Y or of a g-projection of Y
- or ii) X is a head which governs Y or a g-projection of Y to the right

Longobardi's formulation states basically that g-projections may extend upward through a maximal projection X to a dominating node only if X is governed, from the left (Kayne's original definition allowed projection through a maximal projection on a right branch regardless of whether it is governed).

G-projections, and particularly sets of them as in (83), play a fundamental role in the Connectedness Condition of (84), which is essentially Kayne's alternative to the ECP (and, on the

45. Longobardi reformulates the definition from Kayne's original formulation in order to account for the differences between extraction from an S' complement and extraction from a sentential adjunct, thus subsuming much of Huang's CED (Condition on Extraction Domains) under the CC. Kayne's original formulation rested on left vs. right branching of phrase structure, incorrectly predicting as grammatical extraction of an object from an adjunct, as in example (87) in the text. The remarks made here in the text concerning the incorrect predictions made by adopting the Connectedness Condition for anaphors are dependent upon Longobardi's definition of g-projections. Similar problems arise, however, with Kayne's definition, although I will not demonstrate them explicitly here.

definition of (82), to the Subjacency Condition as well). The basic sense of the CC is that, given an empty category EC α , bound by an antecedent X, there must be a set of g-projections of EC α such that they extend from the node immediately dominating EC α up to the node dominating the antecedent. In this sense, unifying all the g-projections will build an unbroken path which extends from EC α to X, thus 'connecting' the trace and its binder. Failure of such connection results in ungrammaticality (see Kayne (1983) for details).

- (83) a g-projection set for Y is Y and some set of g-projections for Y.

46

- (84) Connectedness Condition

an empty category Y, A' bound by X, must be such that a member of its g-projection set dominates X.

(When this relation obtains, we say that Y and X are connected.)

To quickly illustrate, consider (85), (86), and (87).

In (85), the long-distance extraction of an object is licit, since the matrix S node is a g-projection of the trace (and further dominates an unbroken sequence of them, down to the trace), and since COMP governs IP from the left, CP is as well; who is therefore connected to the trace, as required. In (86), a typical case of illicit overt left-branch extraction, the g-projection set stops at the subject NP node (i.e. this

46. This is simplified to some extent, from Kayne (1983).

(85)

```

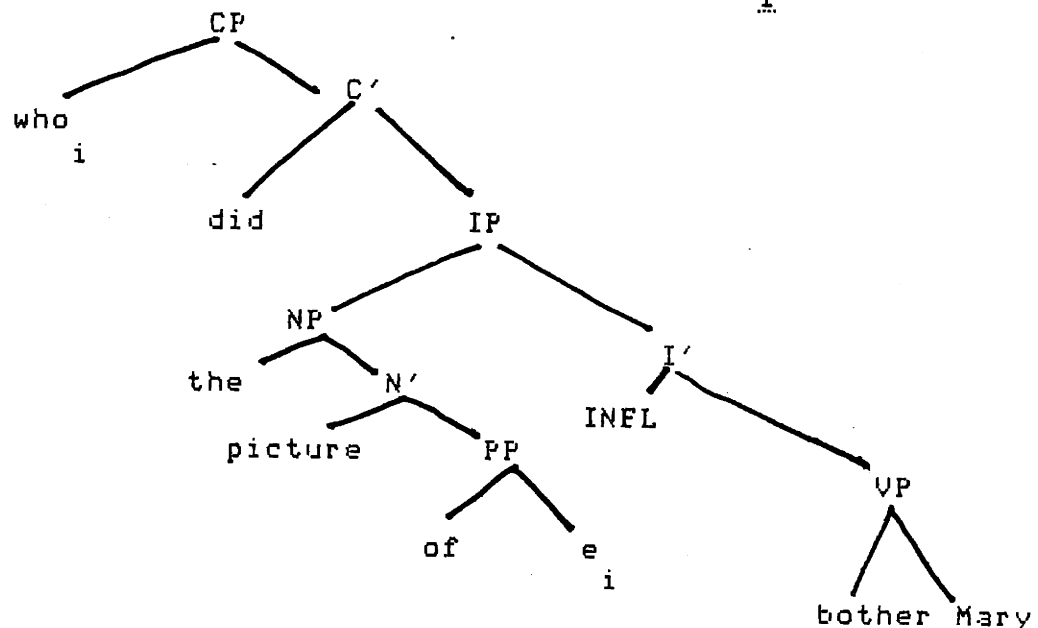
graph TD
    CP --> who[who]
    CP --> IP1[IP]
    IP1 --> you[you]
    IP1 --> I_prime1[I']
    I_prime1 --> INEL1[INEL]
    I_prime1 --> VP1[VP]
    VP1 --> think[think]
    VP1 --> CP_star[CP*]
    CP_star --> that[that]
    CP_star --> IP_star[IP*]
    IP_star --> Mary[Mary]
    IP_star --> I_prime_star[I']*
    I_prime_star --> INEL2[INEL]
    I_prime_star --> VP2[VP]
    VP2 --> INEL3[INEL]
    VP2 --> VP3[VP]
    VP3 --> likes[likes]
    VP3 --> e[e]
  
```

IP = a g
who is

(86)

- 92 -

g-projection of e stops at
NP; who not connected to e



Similarly in (87), the g-projection set of the trace is
upwardly bounded by the adjunct PP node⁴⁷, causing a failure of

47. The internal structure of such adjuncts, and the structural position in which they occur, are unclear matters. There is evidence from such phrases as (i) that before is a preposition, and thus that the phrase it heads is of category PP, even as in (ii).

- (i) [PP before [NP the meeting]]
- (ii) [PP before [Mary left town]]

A 'clausal' adjunct (of course, if the whole adjunct is a PP, it is not clausal; but such adjuncts as (iii), or that in (87), contain clauses (IPs), and so the name 'clausal adjunct' is useful in distinguishing these types of adjuncts from adjuncts like why, when, how, and where, or those of the form [PP for [NP what reason]] may contain a PRO subject, as in (iii):

connectedness.

(iii) John left [after [PRO eating breakfast]]

Since prepositions are governors, and PRO is excluded from governed positions, this would indicate that the adjunct in (iii) cannot have the structure (iii'), since in this structure after governs PRO:

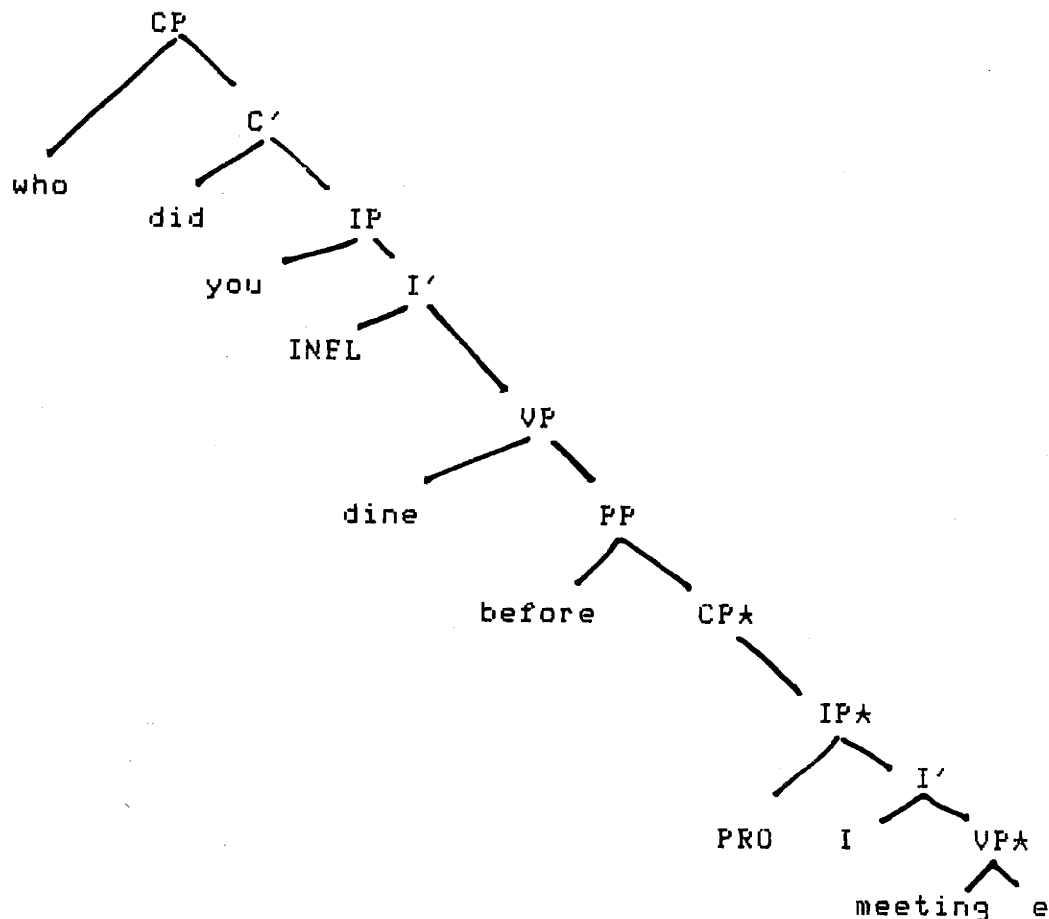
(iii') [after [PRO eating breakfast]]
PP IP

Rather it would appear that there must be some node dominating the IP, and not dominating the preposition, which blocks this government. One possible candidate is CP (S'); in (iv) is a commonly assumed structure for clausal adjuncts:

(iv) [after[[PRO eating breakfast]]]
PP CP IP

If this is the correct structure, it is a puzzle that there should never be an overt complementizer in the COMP position of this CP. See Stowell (1981) for a proposal that such structures lack COMP.

(87) *who_i did you dine [PP before meeting t_i]?



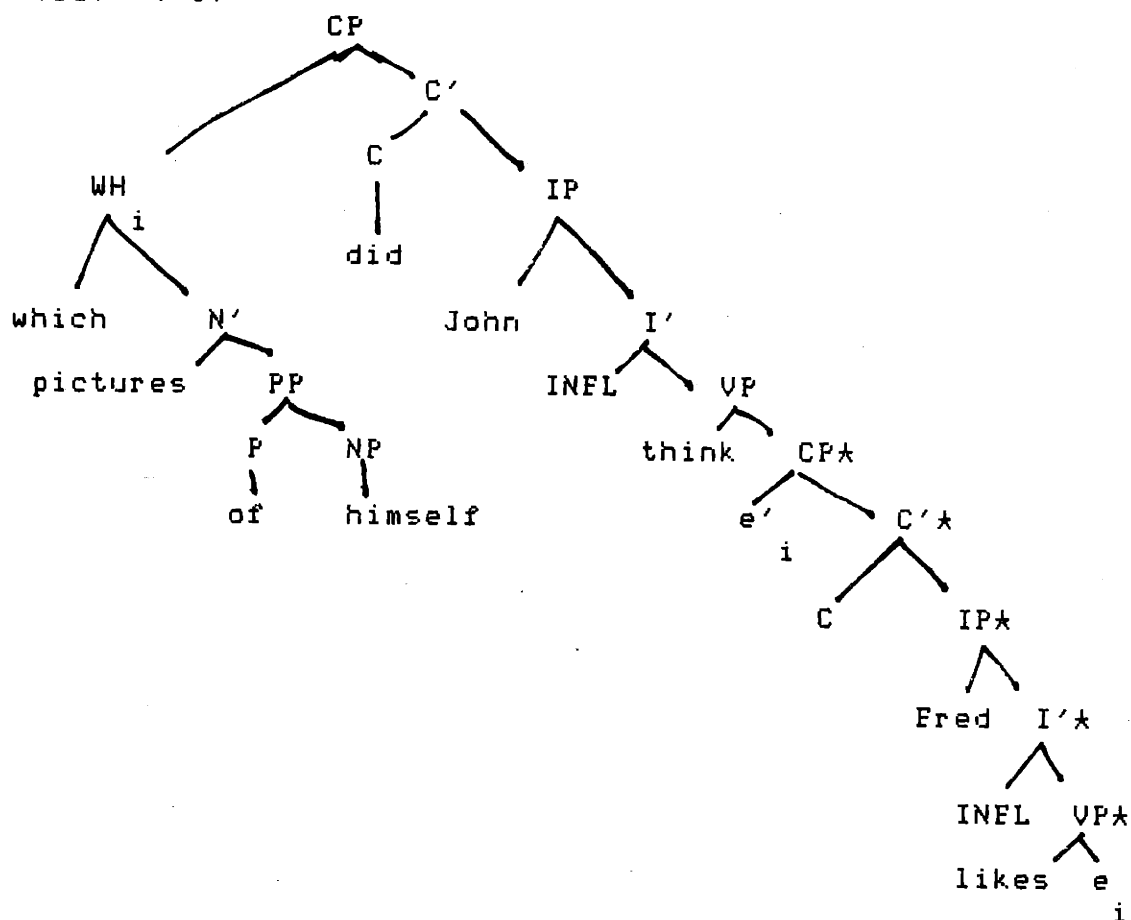
2.5.1 A Connectedness Treatment of Lexical Anaphora and Anaphor Connectivity

At the end of his original article on 'Connectedness', Kayne suggests that it might be possible to formulate a condition similar to the CC for lexical anaphors. I explore this possibility here, paying particular attention to the feasibility of such an approach for anaphor connectivity ('reconstruction').

Consider again the example of anaphor connectivity given in (88), which exhibits the 'multiple binding domain effect', in that either John or Fred may grammatically serve as an antecedent for the anaphor:

(88) [which pictures of himself][does John think [e'[Fred likes e]]?

(88) (=9)



WH = a g-projection of himself
WH is connected to e', e
John connected to e'
Fred connected to e

Notice here that we have the following Connectedness-type relations: (i) the WH phrase is connected to both of the empty

categories ; (ii) the root node of the WH phrase is a g-projection of the anaphor, since the anaphor is structurally governed by the preposition of (or the noun pictures), and we may reach WH_i by going upward from the maximal projection of this governor to WH_i without going through a left branch or an ungoverned maximal projection; and finally (iii) John connects to the intermediate trace e', Fred to the lower one, e.

Suppose we take (89) as an initial formulation of a condition which will permit himself to access either John or Fred as its antecedent, through the g-projection and connectedness relations which hold between them.

(89) Initial Proposal

The antecedent B of an anaphor A must connect to a node E, to which a g-projection of A connects.

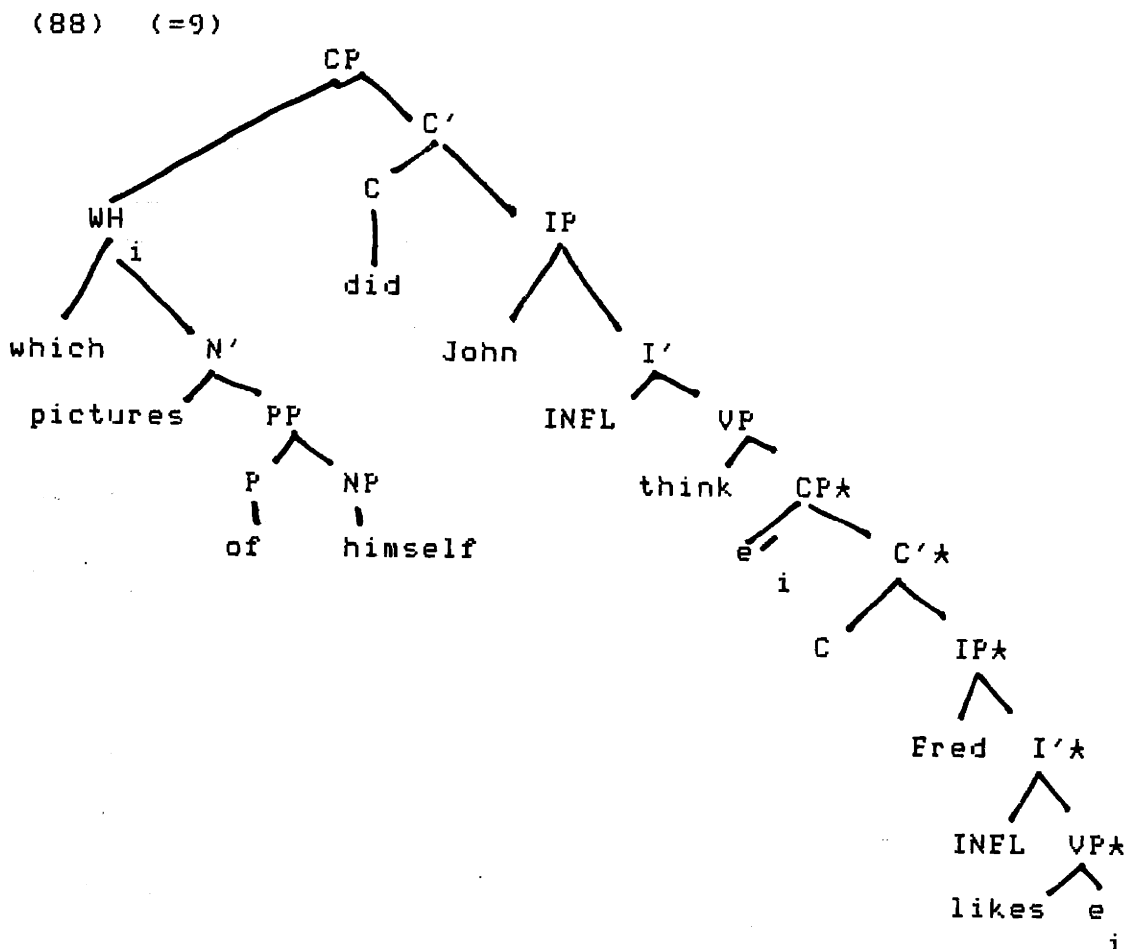
The basic idea of (89) is that there must be one node in the tree to which the antecedent and some g-projection set of the anaphor both connect. In a sentence like (88), these nodes where two connectedness paths converge are the two empty categories.

48. In order to satisfy the Connectedness Condition, the lowest trace must be connected to the higher trace, and the higher trace must be connected to the WH phrase. Nonetheless, it is true for this example that both empty categories are (separately) connected to the WH phrase, even though the lowest trace e_i need not be in order for the CC to be satisfied.

Since paths through the tree are sets of individual nodes, we can refine (89) as (90).

- (90) An anaphoric path P for an anaphor A is a set of nodes such that:
 P is the union of some set of g -projections for A (G_1, \dots, G_n) and some set of g -projections of nodes to which G_i are connected (in the sense of the Connectedness Condition; cf. (82)-(84).)

To illustrate, consider (88) again.



G_1 = a g -projection set for himself = {of, PP,

pictures, N', WH }
i

G2 = a g-projection set for $e' = \{ \text{think}, \text{VP}, \text{INFL}, \text{INFL}', \text{IP} \}$
i

G3 = a g-projection set for $e' = \{ \text{likes}, \text{VP}^*, \text{INFL}^*, \text{INFL}'^*, \text{IP}^* \}$
i

anaphoric path 1 (through which John is accessed) = G1 U G2

anaphoric path 2 (through which Fred is accessed) = G1 U G3

Here, G1 is a set of G-projections for himself; the ECs⁴⁹ are nodes to which the topmost G-projection of the anaphor, namely WH, connect; G2 and G3 are sets of g-projections for the ECs.
i

G1 meets the definition of anaphoric path, as does the union of

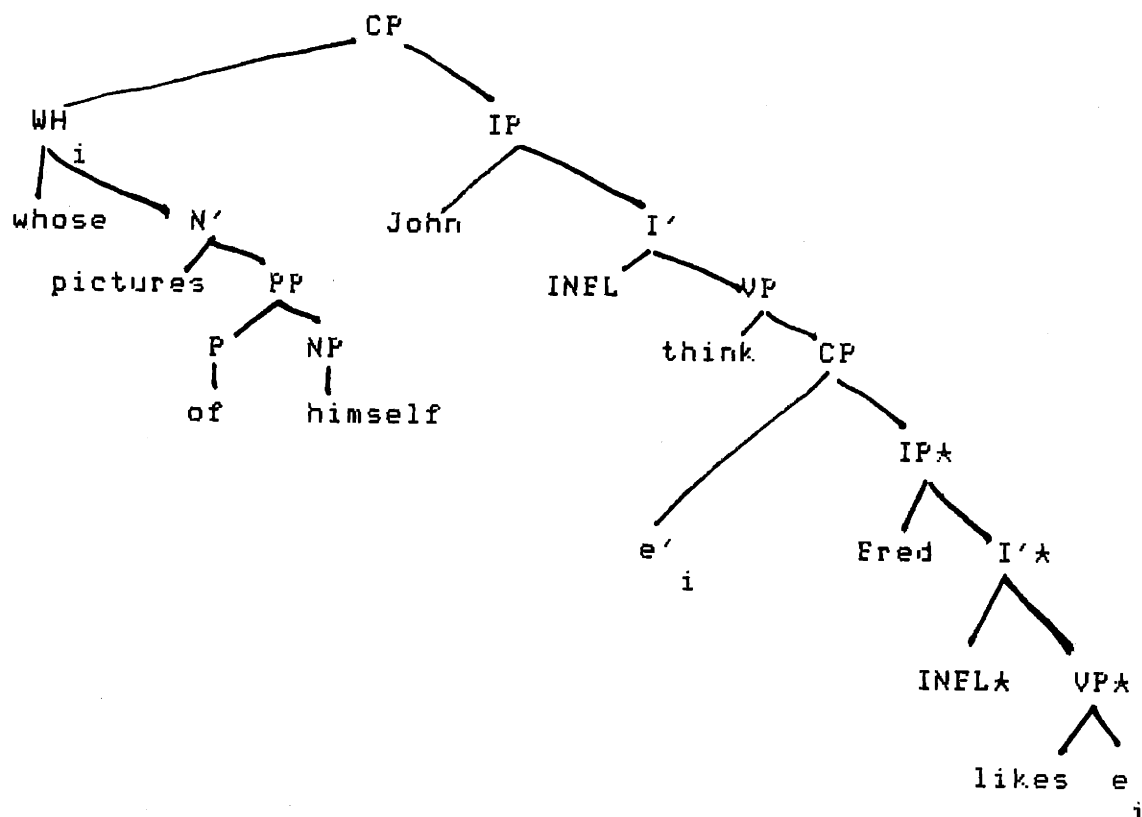
G1 and G2, and the union of G1 and G3⁵⁰. G2 by itself does not meet the definition, since it does not contain a g-projection

49. The actual representation of the S-structure (88) might contain two other traces, one adjoined to the matrix VP, the other to the lower VP. At this point in the discussion these need not be mentioned, although for the purposes of completeness it should be noted that were these traces present, there would in effect be two ways in which the anaphor could access either of the two NPs John or Bill as a local antecedent. Thus for an example like this, the role of the intermediate trace in SPEC of CP is the same as the role of a VP-adjoined trace. This is not always so. In a later section I will demonstrate that the model of anaphor connectivity developed here gives strong evidence for the existence of VP-adjoined traces at S-Structure. For ease of exposition, I will suppress discussion of these traces here.

50. Of course, these three do not exhaust the set of anaphoric paths for the anaphor. $\{P, PP, N'\}$ is an anaphoric path; $\{P, PP, N', e', V, VP\}$ is another, and so forth.

for the anaphor. The same holds for G3. Now, consider the structurally identical sentence (91). We observe that there is a striking difference between (91) and (88) in terms of the possibility of associating himself with either John or Bill. (91) demonstrates that the presence of an overt potential binder inside the WH phrase blocks the potential association of himself with John or Bill as grammatical antecedent; the multiple binding domain effect disappears here.

(91) whose pictures of himself does John think Fred likes ?



Here, we have the same anaphoric paths for the anaphor himself as in (88); including Path 1 (consisting of G1), Path 2 (the union of G1 and G2), and Path 3 (the union of G1 and G3). The empirical difference between (88) and (91) is that in (91) the subject of the WH phrase, whose, must be understood as the antecedent of the anaphor. Its presence forces the anaphor to be bound within the WH phrase, as we would expect. The absence in (88) of any such subject in the WH phrase permits the anaphor to seek along Path 2 or Path 3 to access John or Bill as an antecedent. What the contrast in (88) and (91) shows, in essence, is a type of SSC effect familiar from such non-connectivity cases as (92a) and (92b):

- (92) a. John likes the pictures of himself
- b. *John likes Mary's pictures of himself

For (92), the notion of 'closest' potential binder can be stated purely in terms of c-command⁵¹, while this is not possible for connectivity examples like (88). However, the development of the analysis of (88) in terms of anaphoric paths will extend easily, without modification, to such examples as (92a,b).

The definition of anaphoric path given above allows us now to

51. A is a potential binder for an NP B iff A is an NP, and A c-commands B. A is the closest potential binder for B iff A is a potential binder for B and there is no Y, a potential binder for B, such that A c-commands Y and Y does not c-command X.

state the locality restrictions on anaphor-antecedent relations in a quite simple way. (93) is the definition of a local anaphoric path, which replaces the 'governing category' of Chomsky(1981).

(93) a set of nodes P is a local anaphoric path for anaphor A iff:

- i) P is an anaphoric path for A
- ii) some node N in P immediately dominates a subject accessible to A
- iii) no proper subset of P meets (i) and (ii).

Like (22), this definition makes reference to 'accessible subject'; the presence of the accessible subject is crucial in delimiting the top of a binding domain. In this respect, the definition above is similar to the definition of Governing

52

Category of Chomsky(1981) The intent of making reference to 'accessible subject' in the definition of governing category is to capture the fact that a subject NP does not induce a Specified Subject Condition effect when it contains the anaphor. The crucial contrast is (93'a) and (b):

(93')

(a) [John [thinks [that [[pictures
 IP VP CP IP* NP*
 [of himself]]] are on sale]]]
 PP

Here, NP* contains the anaphor, and is not accessible to it. In

 52. See the definition of accessible in (22).

(b), this is not the case, and so Mary apparently forces the lower IP to be the maximal binding domain for the anaphor.

(b) * [John [thinks [that [Mary
 IP VP CP IP*
 [likes [pictures of himself]]
 VP* NP*]]]]

We can adopt the definition of 'accessible' directly as formulated in (Chomsky (1981)). Alternatively, we may define it as (c):

- (c) A subject X is accessible to Y with respect to an anaphoric path P iff
- i) P is an anaphoric path for Y; and
 - ii) X is not a member of P; and
 - iii) X is dominated by a member of P.

To quickly illustrate, in (93'a) the anaphoric path extending from himself to John is {himself, of, PP, N', NP*, IP*, CP, VP, I', IP}. The subject NP* is a member of this path, so it is not accessible to the anaphor. Thus, there is no accessible subject for himself with respect to the shorter path {himself, of, PP, N', NP*, IP*}; and hence this path is not a local anaphoric path. In fact, {himself, of, PP, N', NP*, IP*, CP, VP, I', IP} is the only anaphoric path for himself which meets the definition (93), hence the grammaticality of the example.

In (93'b), the local anaphoric path is {himself, of, PP, NP*, VP*, IP*}, since Mary is an accessible subject by definition (c). Therefore, John fails to be close enough to the reflexive

to satisfy (89).

Since we will soon abandon reference to accessible subjects from Condition A, following the basic line of reasoning in Chomsky's (1985) theory of binding, we do not illustrate this in detail.

The specialized status accorded subjects, and the entire notion of 'accessible subject' in the (1981) theory of binding has been criticized (see Huang(1983), Chomsky (1985)), and the definition of anaphoric path just given is open to the same criticism.

Because of the familiarity of the Binding Theory of Chomsky (1981), we shall illustrate the operation of the revised binding principle A as defined in (94), then modify it to be in accordance with the recently modification of the Binding Theory presented by Chomsky (1985), which eliminates reference to Subjects or accessibility of subjects.

A local anaphoric path is the domain in which an anaphor must find an antecedent, leading us to formulate (94), which is intended to replace Condition A of the Binding Theory.

(94) Locality Condition (replaces Condition A of Chomsky 1981)

An anaphor A must be coindexed with an antecedent B which is a sister to (equivalently, mutually c-commands) a node in a local anaphoric path for A.

In our example (88), there are two local anaphoric paths: Path 1, consisting of the unions of G1 and G2, allowing John to serve

as the antecedent, satisfying (94); and Path 2, the union of G1 and G3, which permits the coindexation of Fred and himself to satisfy (94). More precisely, Path 1 consists of {of, PP, pictures, N', WH_i, think, VP, INFL, INFL', IP}. Path 2 consists of {of, PP, pictures, N', WH_i, likes, VP*, INFL*, INFL'*_i, IP*}.

Crucially, neither path is a subset of the other, hence both qualify as local anaphoric paths. Note that in this example G1 (the set of nodes {of, PP, pictures, N', WH_i}) is not by itself a local path, since no node in it immediately dominates a Subject (that is, a potential Binder of the anaphor).

The members of the g-projection sets and the anaphoric paths for (88) are listed below:

G1 = a g-projection set for himself = {of, PP, pictures, N', WH_i}

G2 = a g-projection set for e'_i = {think, VP, INFL, INFL', IP}

G3 = a g-projection set for e_i = {likes, VP*, INFL*, INFL'*_i, IP*}

Path 1 = G1 U G2 = {of, PP, pictures, N', WH_i, think, VP, INFL, INFL', IP}

Path 2 = G1 U G3 = {of, PP, pictures, N', WH_i, likes, VP*, INFL*, INFL'*_i, IP*}

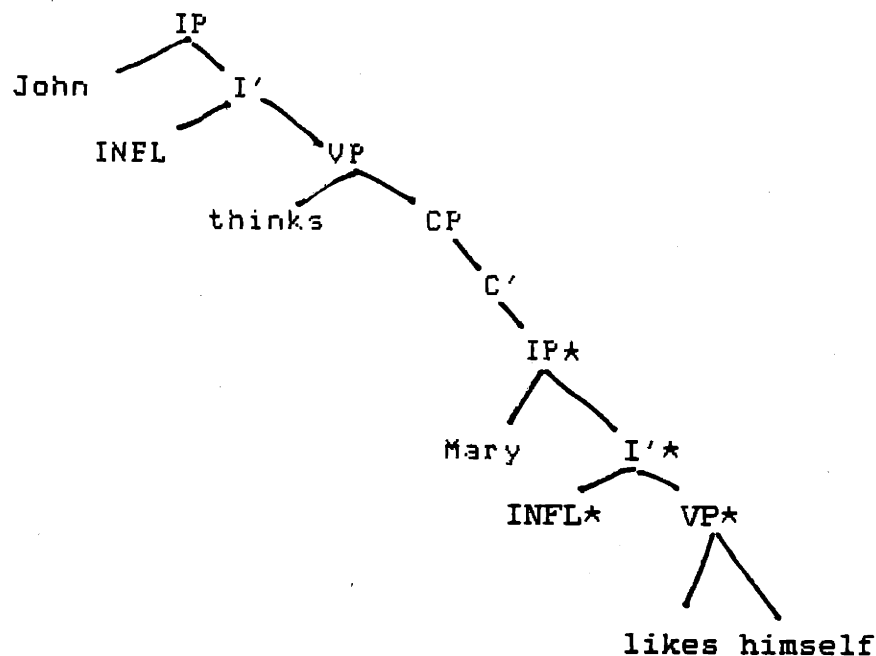
The situation is different in (91). Here, G1 is a local anaphoric path, since WH_i immediately dominates a potential binder (the Subject whose). Now, neither Path 1 nor Path 2 will

be local anaphoric paths for himself, since they are the unions of G1 --itself a local anaphoric path -- and some other non-empty set of nodes in the tree. That is, G1 is a third anaphoric path in the example (91) (Path 3), and Path 3 is a proper subset of Path 1 and of Path 2, thus by (93iii) neither Path 1 nor Path 2 is a local anaphoric path. In (91), the locality condition (94) will then force the anaphor to be coindexed with who, the correct result.

This locality restriction (94) of course applies to examples with no dislocation as well. In an example like (95), although both John and Mary are dominated by members of anaphoric paths for himself, Path 1 and Path 2 respectively, only Path 1 is a local anaphoric path (since Path 1 is a proper subset of Path 2), and so himself cannot be bound to John. Binding the

reflexive to Mary is prohibited⁵³, and so the sentence is ungrammatical.

53. Presumably by a feature mismatch in the gender of the two NPs.



Path 1 = {likes, VP*, INFL*, INFL'*, IP*}

Path 2 = {likes, VP*, INFL*, INFL'*, IP*, thinks, VP, INFL, INFL', IP}

Only P1 is a local anaphoric path.

2.5.2 Anaphoric Paths and Complete Functional Complexes

Chomsky (1985) introduces a modification of the Binding Theory which dispenses with the notions of Accessibility, SUBJECT, and other (perhaps stipulatory) parts of the (1981) binding theory. The emphasis on domains of thematic role assignment (Chomsky's Complete Functional Complexes) in this new theory has striking support from anaphor connectivity, as we shall see in the next chapter. In that chapter, I attempt a reformulation of condition A which relies heavily on the (1985) formulation; my discussion in that chapter will involve a review of the

technical and conceptual nature of this revised binding theory, and so I will not review it here, assuming familiarity on the part of the reader.

The relevant portions of the theory are as follows:

- (96) *A syntactic domain is a Complete Functional Complex (CFC) (Chomsky (1985; p. 169) if all grammatical functions compatible with its head are realized in it; the complements necessary, by the Projection Principle, and the subject, which is optional unless required to license a predicate, by definition.

- (97) Definition of a Binding Theory Compatible Indexing
(ex (231), p. 171)

Given expression E with the indexing I, I and the pair (A,B) [A a phrase in E, and B a subdomain of E] are compatible with respect to the binding theory if A satisfies the binding theory in the local domain B under the indexing I.

That is,

I is BT-Compatible with (A,B) if:

- (A) A is an anaphor and is bound in B under I
- (B) A is a pronominal and is free in B under I
- (C) A is an r-expression and is free in B under I

- (98) Licensing condition for A governed by C in expression E with an assigned indexing I (ex. 232, p. 175):

for some B such that (i) or (ii), I is BT-compatible with (A,B):

- (i) A is an R-expression and (a) if A heads its chain or (b) otherwise

a) B = E

b) B is the domain of the head of the chain of A

- ii) A is an anaphor or pronominal and B is the least CFC containing C for which there is an indexing J BT-compatible with (A,B).

We earlier defined local anaphoric paths in terms of the (1981) Binding Theory. These definitions can be easily recast in terms of the (1985) revision.

(99) Definition of Anaphoric Path (revised)

An anaphoric path P for an anaphor A is a set of nodes $(N_1, \dots, N_i, \dots, N_n)$ such that:

- (i) P is the union of some set of g -projections for A ($G_1 \dots G_n$) and some set of g -projections of nodes to which G_i connect.
- (ii) N_n is the root node of a Complete Functional Complex.
- (iii) some N_i in P is a sister to (mutually c -commands) an NP B .

(100) Definition of Local Anaphoric Path (revised)

a set of nodes P is a local anaphoric path for anaphor A iff:

- (i) P is an anaphoric path for A
- (ii) there is no proper subset of P which is an anaphoric path for A .

(101) Locality (Licensing) Condition on Anaphors
(replaces Condition A of Chomsky 1985)

An anaphor A must be coindexed with an antecedent B which is a sister to (equivalently, mutually c -commands) a node in a local anaphoric path for A .

As in the earlier characterization, for a sentence like (102),

there will be but one local anaphoric path for the anaphor; the condition (101) will require that the anaphor be coindexed with John in this example.

(102) Bob thinks that [John likes himself]

For a sentence like (88), there will be two local anaphoric paths, and the licensing condition (101) will be satisfied if either John or Bill is coindexed with the anaphor. Hence the ambiguity of the example (88).

2.6 Problems With The Path Approach

The conceptual appeal of the approach outlined above, itself a development of an initial proposal by Richard Kayne, is that it is couched within the set of formal relations (X' projection, government, and connectedness) which are of independent use in the grammar, and specifically in the theory of WH movement as developed by Kayne. The anaphoric path treatment of anaphor connectivity developed above thus is appealing in the broad sense, since it is of the general form that we argued any treatment of connectivity must be (e.g. applying directly to representations in which the anaphor is not necessarily within the c-command domain of the antecedent), and in the technical sense, since it introduces no new primitive relations, and builds the few new relations introduced out of technical devices

already made available in Kayne's theory. Further, since both WH-trace relations and antecedent-anaphor relations have to do with referential dependence, it is perhaps not surprising that the theory of locality in WH movement should be so fundamentally tied to the theory of locality in lexical anaphora.

However, there are reasons to suppose that the theory of anaphoric paths developed above is incorrect. I will show below that there is a consistent problem with the theory as developed, in that, since any anaphoric path (local or not) contains g-projections of the anaphor, and a set of g-projections is upwardly bounded by an ungoverned maximal projection (see discussion of (85-7) above), the theory predicts that the binding paths for anaphors are similarly bounded. I will show that this prediction is incorrect, and thus that the account of anaphor binding and locality in WH movement must be kept more distinct than in any approach which depends upon connectedness. This of course does not in any sense falsify Connectedness theory; it merely gives an indication that lexical anaphora is to be handled differently. This demonstration concludes this chapter. The next chapter develops a theory which, like the anaphor path/connectedness theory, is capable of applying to the S-structure representations of connectivity examples, but which does not have the problems inherent in the Connectedness treatment of these phenomena. The theory to be developed in the chapter to come has some other empirical advantages as well.

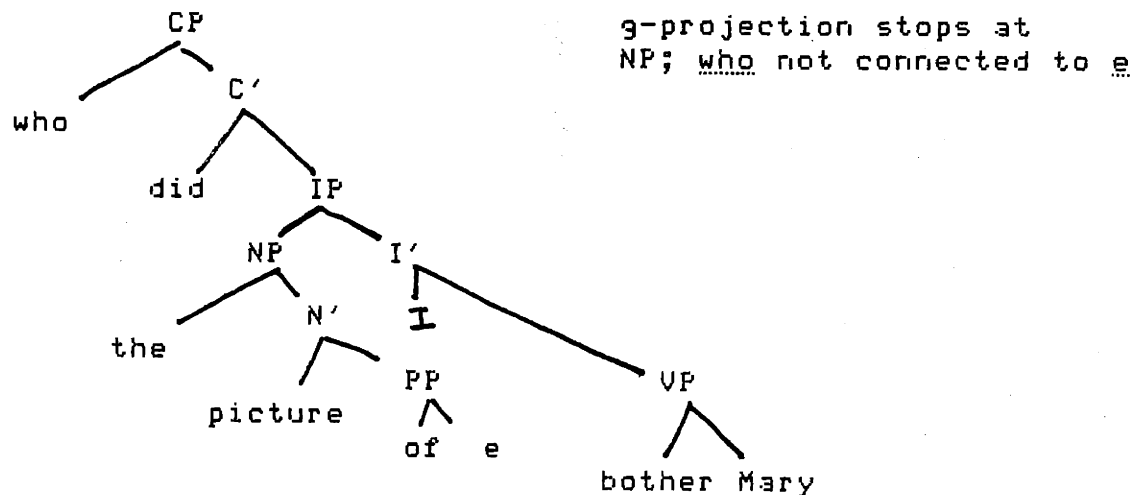
The Anaphoric Path approach quite accurately captures the connectivity facts -- the multiple binding domain effect under successive cyclic WH-movement, as seen in (88), and the blocking nature of a subject within the WH-phrase, as in (91), without recourse to LF lowering of any type. The theory thus conforms to the general form that it was earlier argued it must.

Ideally, we would want such conditions and definitions as (99), (100), (101) to apply without exception to any representation. This means that any syntactic representation which blocks WH movement due to a failure of Connectedness will block anaphor binding as well. This is not the case, although some facts do support it, making the topic worth addressing in some detail.

Let us first examine the contrast in grammaticality between (103) and (104).

(103)

*who_i did [[the picture of t_i] bother Mary]?



(104) The men think that photographs of themselves are ugly

In the Connectedness theory, (103) is ungrammatical because the structure fails to satisfy the CC. The topmost g-projection for the EC is the subject NP, since this subject is not governed from the left. Since the WH phrase is the local binder of the EC (there is no position for an intermediate trace here) the CC requires that the WH phrase be a sister to a g-projection of the EC, which does not occur here. The WH phrase is not a sister to the topmost g-projection, nor to any lower one.

Most speakers of English find (104) perfectly acceptable. The theory of anaphoric paths developed earlier exactly predicts this. ((104) is repeated below, with structure added.) First, the NP the men connects to the lower INFL'. This is so because INFL'' and INFL' are g-projections of INFL'*₁, and the men is a sister to INFL'. The subject NP of the lower clause pictures of themselves is a g-projection of the anaphor themselves (indeed

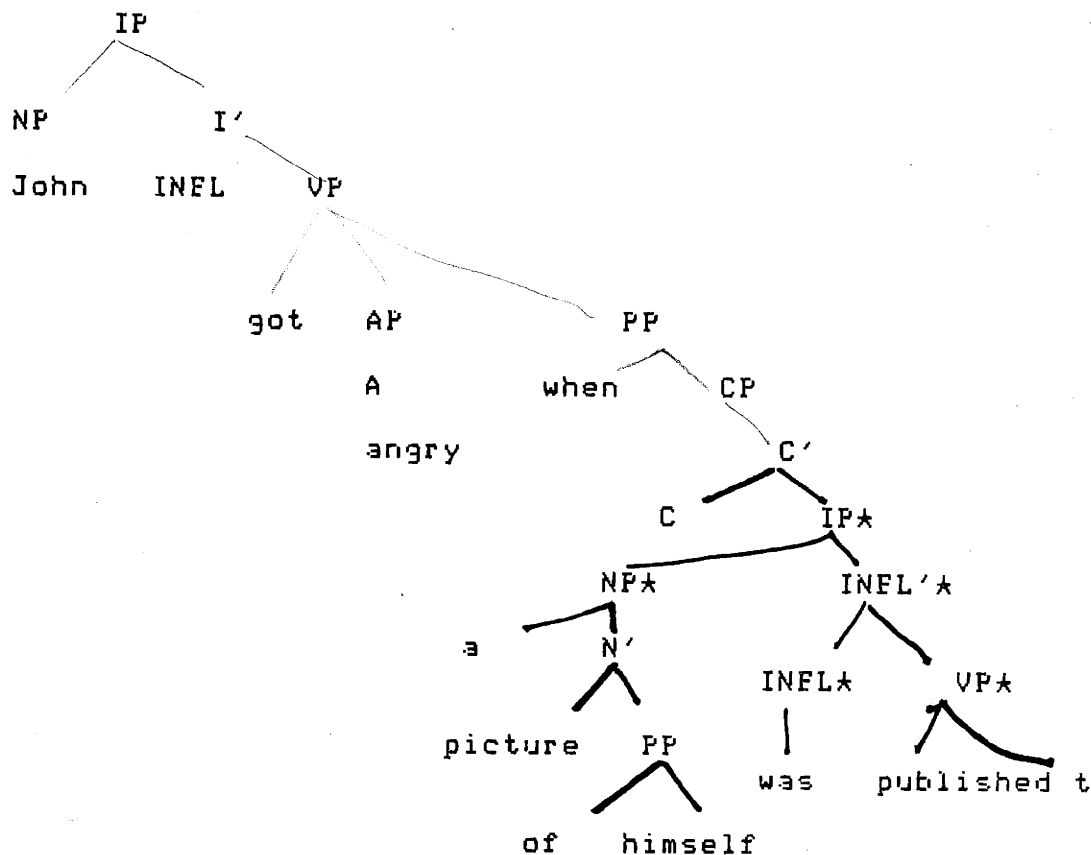
the maximal g-projection of it). Since both the men and this NP connect to the same node, namely INFL', there is an anaphoric path between the anaphor and its antecedent. Thus the locality requirement (101) is met, and the sentence is correctly predicted grammatical. Interestingly, the formulation of (101), and the definition of anaphoric path, independently motivated for cases of connectivity, correctly give the contrast between the left-branch extraction (103) and the licit anaphor binding into a left-branch seen in (104).

(104) The men think that photographs of themselves are ugly

[[The men] [INFL [think [[[[photographs
INFL' VP S' S NP* N'
[of [themselves]]]] [INFL* are ugly]]]]]
PP NP INFL'*

Now consider the sentence below, modified from an example in Jacobsen and Neubauer (1976).

(105) John got angry when a picture of himself appeared in the paper



G1 = maximal g-projection set of himself = {of, N', NP*}

INEL'* = a node to which NP* connects

54

G2 = maximal g-projection set for INEL'* = {IP*}

I concur with Jacobsen and Neubauer that this example is quite grammatical. Now consider the prediction made by the anaphoric path theory. As in (104), the topmost g-projection of the

54. It is unlikely that the empty COMP C governs IP strongly enough for the g-projection set to continue beyond IP* to CP (and thus on to PP). Even, if it did, the maximal g-projection set for the node INEL'* would only be {IP*, C, C', CP, PP}, and this still fails to reach the antecedent John.

anaphor himself is the subject NP*. The locality condition (101), along with the definition of anaphoric path requires that this node NP* connect to a node E that has a g-projection which is a sister to John. This is not so. NP* connects to INFL'*. The maximal g-projection set for INFL'* only goes as far up as IP*; and IP* is not a sister to John. Therefore, there is no anaphoric path for the anaphor (IP does not dominate an NP capable of anteceding the anaphor).

Since IP* is not a sister to John, there is a failure of connectedness, and the theory developed earlier predicts this sentence to be as ungrammatical as (103), which it clearly is not.

Indeed, (104) should be as ungrammatical as:

(106) ⁱWhat did John get angry when pictures of t_i were published?

This asymmetry between anaphor binding and overt WH movement indicates that these two phenomena cannot both be collapsed under connectedness theory. For this reason, I will suppose that the revision of Condition A needed for the anaphor connectivity cases must not be put in these terms. A distinct reformulation is offered in the next chapter.

APPENDIX: On Restricted Quantification and The Structure of LF

In this chapter, I discussed such cases as (1), and argued that literal lowering at LF of [most pictures of himself] or a sub-part of it clashed with the fact that natural language has restricted (and only restricted) quantification. Assuming that quantifiers are assigned scope through QR, it is readily apparent that lowering the entire QNP to the position of t and leaving it there (as in (2)) will clash with the scope requirements of the quantifier.

(1) [most pictures of himself] [John likes t]

(2) [John likes [most pictures of himself]]

I argued in the text of this chapter that lowering the N' pictures of himself to t, leaving most in scope position, will result in an LF representation representing unrestricted quantification. The LF so derived will be (3).

(3) most [John likes [pictures of himself]]

In this appendix I embellish upon this point; as made in the chapter the argument must be qualified to some extent.

As I pointed out, the rule of QR, taken together with Higginbotham's theory of the saturation of thematic-grid positions, gives a quite elegant result: each constituent at LF,

for a sentence like (4), corresponds to one of the semantic expressions relevant to the restricted quantification. The LF of (4) is (5) (irrelevant details omitted).

(4) Most people run

(5) [most [people]] [t run]
 NP<1a> N' <1> IP

A restricted quantification can be represented as a pair of ordered formulae as in (6).

(6) [Qx: P(x)] [G(x)]

Qx is the quantifier, P(x) its restriction; together, they form a restricted quantifier which restricts the domain over which x ranges in G(x) to the subset X of the domain D of individuals which satisfies P. The relation between D and X is determined by Q. We shall refer to [G(x)] as the matrix. The restricted quantifier, semantically, determines the values of x in the matrix; syntactically, the QNP binds the trace in the matrix. This is a most transparent relation between syntactic form and semantic interpretation.

If we are willing to give up this isomorphism between LF constituents and semantic expressions, it is perfectly possible to define an algorithm which can recover restricted quantification from an LF representation like (3). Below I sketch one way of doing this, adapted from Cushing (1982). (Cushing presents the argument, summarized in the text of this chapter, that there is no formula of first-order predicate logic

using only unrestricted quantification which can express the restricted quantification of a sentence like (4).)

Cushing presents a formalism for restricted quantification which allows us to define an algorithm which can recover the necessary pair of ordered formulae from an LF in which the restriction is part of the matrix, as in (3). Briefly, Cushing's representation of a restricted quantification takes the following form:

7) $Qx (P(x); G(x))$

The left formula is the restriction, the right one the matrix. For simplicity, we shall use the abbreviation F1 for the restriction, and F2 for the matrix.

On this view, restricted quantification is not directly represented at LF (in the sense that the two formulae are not separate at LF, as they are in (5)), but it is recoverable from LF. The two formulae can be derived via the following algorithm:

(8) Given a syntactic representation E having the form

$$\begin{array}{c} [\dots [Q N'] \dots] \\ E \end{array}$$

where Q binds a position T in the thematic grid of N, then construct an interpretation of E as follows:

- i) copy Q as a sister to E;
- ii) construct an ordered pair of formulae (F1; F2) for Q (giving the quantificational structure $Q(F1; F2)$) in the following way:
 - a) Take x as the variable associated with Q;

- b) F1 consists of the minimal maximal projection dominating Q in the expression E, with x substituted for Q;
- c) F2 consists of the expression E, with x substituted for the minimal maximal projection dominating Q in E.

To illustrate, consider the S-Structure representation (9), derived from (4) by moving most to IP-adjunction.

(9) [[most[people]] run]
 IP NP<1x> N' <1>

The algorithm gives us (10) by (8i), (11) by (8iia), (12) by (8iib), and finally (13) by (8iic).

(10) [most[[most[people]] run]]
 IP IP NP<1x> N' <1>

(11) [most x[[most[people]] run]]
 IP IP NP<1x> N' <1>

(12) most x (people (x);

(13) most x (people (x); x run)

The quantification operates left to right, with people(x) functioning as the restriction on most x, and x run functioning as the matrix.

Thus, the case made against lowering the N' of a QNP in the text of this chapter was somewhat overstated. However, the May/Higginbotham theory is to be preferred at least on conceptual grounds, as it provides an account giving a maximally transparent relation between the syntactic form of LF and its interpretation. Cushing's algorithm is more complicated, adding to the theory a mechanism unneeded in the alternative theory.

Chapter Three: Chains, Predicates, and Linking

In the preceding chapter, I argued that the correct way to approach the reconstruction problem is to revise condition A of the binding theory (and perhaps other conditions on antecedence as well) so that it applies directly to S-Structure representations. The cases under consideration in that chapter illustrate the anaphor connectivity effect. This proposal involves abandoning the view that every licitly bound anaphor is c-commanded by its antecedent, in favor of a perhaps more technically complicated condition, but one which has greater empirical coverage while retaining the formal precision of the LGB theory and its modification in Chomsky (1985).

The first two sections of this chapter develop an approach to the connectivity effect which has these properties; it is similar in spirit to the anaphoric path approach developed in the preceding chapter, but avoids the problems for the formalism noted there. I introduce chain accessibility sequences, similar to the anaphoric paths of the preceding chapter, but less directly related to Connectedness theory.

Section 3 discusses NP-movement in the context of the connectivity effect, observing (in part following Belletti and Rizzi (in progress)) that connectivity occurs in cases of

NP-movement, as well as WH-movement. This section discusses the proper treatment of 'psych-movement', and presents an argument against the successive cyclic application of Condition A, which, if maintainable, would provide an alternative to the framework developed in this thesis.

In the fourth section, I will refine this approach to anaphor connectivity in order to capture some interesting asymmetries between predicates and non-predicates. In the fifth section, I will consider in some detail reciprocal binding and reciprocal interpretation, modifying to some extent the treatment of reciprocals in Higginbotham (1981), (1985). This section observes that the 'multiple binding-domain effect' discussed in chapter 2 has consequences for reciprocal interpretation as well. These facts provide a strong argument that the Linking theory of Binding (Higginbotham 1983, 1985) is to be formally preferred over indexing theory, and gives a further argument against successive cyclic binding. The final part of this chapter considers the 'layered trace' approach to binding.

3.1 Chain Accessibility

In this section, I will offer an approach to anaphor connectivity which is rather close to the Connectedness analysis which ended chapter 2. The approach offered here differs from the preceding one principally in the role played by government.

The framework suggested in (2.5) proposes that anaphors are bound to antecedents which are related to them through anaphoric paths; anaphoric paths are parasitic upon well-formed connectedness paths relating a WH-trace to its antecedent. This latter type of path is crucially sensitive to either directionality of branching, or to government. As we saw at the end of ch. 2, anaphor binding is insensitive to these factors. This suggests that the correct account of anaphor binding will depart to a certain extent from Connectedness theory. The framework I will develop in this chapter remains, however, conceptually similar to Connectedness theory, in making use of paths throughout the tree to delimit the local relation between an anaphor and its potential antecedents.

In the preceding chapter, I introduced a modification of Condition A of the Binding theory, based within the Connectedness theory of Kayne (1983). This modification had two central attributes: it states Condition A as a condition capable of applying directly to S-Structure representations, like (1), and it allows 'binding domains' to be sets of nodes defined over the phrase marker. This latter aspect permits the ambiguity of sentences like (2) to be considered as deriving from a subset condition, independently motivated for simple cases of anaphor locality like (3). Essentially, an anaphor will have to access its antecedent through a path no subsequence of which will access another antecedent.

- 1) which picture of himself does Bob like ?
- 2) which pictures of himself does Bob think Joshua likes ?
- 3) Bob thinks Joshua likes this picture of himself.

As we saw in (2.6), the anaphoric path analysis has an undesirable attribute as well, namely that it overly restricts the distribution of anaphors in structures where they are separated from their antecedents by a structurally complex left branch, or an ungoverned maximal projection. My intent in this section is to modify the anaphoric path account of anaphoric connectivity in such a way as to eliminate reference to government and directionality of branching. This will avoid the problems arising in the anaphoric path account, while still retaining the core approach to connectivity present in that analysis.

3.1.1 Reference to Chains

As observed in (2.5), not every instance of coindexation is sufficient to allow the continuation of an anaphoric path. Rather, this coindexation is limited to coindexation between members of the same chain. This is why the definition of anaphoric paths was given in terms of Connected elements, assuming that connectedness holds of members of a chain. This particular formalization of the binding theory was seen to be technically incorrect, for two reasons: first, it predicts that anaphors cannot be separated from their antecedents by an ungoverned maximal projection, and that they cannot be deeply

embedded within a left branch which does not contain their antecedents. Therefore, I will here abandon reference to Connected elements, and instead make direct reference to chain co-membership in the modified Condition A.

The effect of Condition A, whatever its correct formulation, is to require that an anaphor have a antecedent bearing some specifiable structural relation to the anaphor. The later sections of this chapter will consider in some detail the interpretation of reciprocal expressions, and argue that certain facts make it desirable to adopt the linking theory of binding. Chapter 4 also provides evidence for linking as the proper system of notationally encoding binding relations. In anticipation of this, I will construct the definitions below in terms of linking. The appendix to this chapter gives the formulation of the licensing condition for anaphors and pronominals in indexing theory.

3.1.2 Chain Accessibility and the Binding Theory

I propose the following as essentially the proper formulation of Conditions A and B of the binding theory (certain changes will be discussed and made in later sections, and in Chapter 4).

4) Chain Accessibility Sequence (definition)

S is a well-formed chain accessibility sequence for an NP A only if:

- i) A is a
1

ii) some a_i is a projection of the governor of A

iii) for every pair (a_i, a_{i+1}) , either (1) or

(2):

1) a_{i+1} immediately dominates a_i

2) (a_i, a_{i+1}) is a link of a well-formed

(A' or A) chain

5) Chain Accessibility (definition)

B is chain accessible to A through an accessibility sequence $S = (a_1, \dots, a_i, \dots, a_n)$ such that:

B is a sister to some a_j in S

(6) Binding Theory Compatible Linking (Definition)

L, a set of assigned links, is binding-theory compatible wrt. an expression A and an accessibility sequence S for A iff:

i) for A = an anaphor, A is linked under L to a antecedent B which is chain-accessible to A through S.

ii) for A = pronominal, A is not linked under L to an antecedent B which is chain-accessible to A through S.

1. Equivalently, B is immediately dominated by some a_j in S, and B is not any a_k in S.

2. See chapter 4 for modification in terms of obviation, following Higginbotham (1985). Higginbotham argues that Condition B and C are not properly put in terms of strict linking. The remainder of this chapter is concerned almost exclusively with cases of anaphor binding and justification of

(7) Licensing Condition

for A = anaphor or pronominal, A is licensed only if:

- i) there is a chain accessibility sequence S for A, such that there is assigned a BT-compatible linking L for (A,S); and
- ii) there is no proper subsequence S_j of S, such that S_j is a chain-accessibility sequence for A, and there is a possible BT-compatible linking L' for (A, S_j).

This formulation takes the approximate form of the Chomsky (1985) theory, with a few exceptions. Chain accessibility sequences are similar to the anaphoric paths of chapter 2. They are sets of nodes (paths), beginning with the anaphor (or pronominal) in question, and extending up to a point very close to the antecedent; for a particular NP, there are potentially many well-formed accessibility sequences, hence there are many potential local antecedents for the NP. The definition of BT-compatible linking, and the licensing condition (modifications of the definition of BT-compatible indexing, and the Licensing Condition, of Chomsky (1985)), serve to define, for a structure E, which of the NPs in E might act as antecedent to an anaphor (and which cannot act as an antecedent to a pronominal). The licensing condition requires that one of these potential antecedents actually be the antecedent of the

the notion of chain accessibility sequence. Discussion of the proper formulation of Condition 8 would take us too far afield at this point, and is deferred until the next chapter.

anaphor. (10ii) parallels Chomsky's least CFC, with a crucial difference: it allows two (or more) potential antecedents to be equally close to the anaphor.

For purposes of expositional clarity, I would like to define one more term, minimally chain accessible.

(8) B is minimally chain accessible to A iff:

(i) B is chain accessible to A (through some chain accessibility sequence S)

(ii) there is no proper subsequence S* of S through which some B' is chain-accessible to A.

That is, B is minimally chain accessible to A iff it is an NP which might serve as an antecedent to A in satisfaction of the licensing condition, when A is an anaphor; or if it is an NP barred by the Licensing Condition from being an antecedent to A, when A is a pronominal. For an anaphor, the actual antecedent is chosen from the set of minimally chain accessible antecedents. A pronoun cannot be linked to any NP included in the set of minimally chain accessible NPs.

The next subsection briefly describes certain differences between (4-8) and Chomsky's (1985) theory. I then continue the discussion of (4-8) with some examples of how it works .

3

3. The reader will note that (7) contains no reference to R-expressions; there is no (modified) Condition C in this system. I defer discussion of Condition C until the next chapter, where it will be argued that Condition C effects are to be explained in terms of two separate conditions: a modification of Higginbotham's (1983) anti-c-command condition on antecedence

3.1.3 Some Differences

Chomsky's theory is constructed to license those anaphors which are c-commanded by their antecedents. For this reason, the minimal binding domain can be defined in terms of the notion 'subtree'. This is not possible in our framework, in that the relation between the anaphor and antecedent is put in terms of sequences of nodes which do not form a subtree.

This permits, in a way that Chomsky's theory does not, the ambiguity of (2) to be accounted for as an ambiguity in the locality restriction. Because John and Rob are accessible to the anaphor through different chain accessibility sequences (CASs), they do not interfere with one another in terms of locality.

Secondly, we depart from (Chomsky (1985)) (and from the 1981 theory) in changing the role that an anaphor's (or pronominal's) governor plays in the Binding Theory. In Chomsky (1985), the licensing condition applies to governed anaphors and pronominals. The system developed above applies to all anaphors and pronominals, regardless of whether they are governed. The

(linking), and a modification of the Local Binding Condition on chains (Chomsky 1981, Rizzi 1982, Lasnik 1985, Epstein 1985). Neither of these conditions will refer to R-expressions, and in this sense there is no counterpart of Condition C. While the BT (4-7) specifically makes reference to anaphors and pronominals, the framework developed in this chapter and the next makes no formal reference to R-expressions.

chain accessibility framework requires every anaphor and pronominal to have a BT-compatible link to an antecedent⁴; BT-compatible linking is defined in terms of CASSs. Thus all anaphors and pronominals must have CASSs, and such sequences are well-formed only if they contain a projection of the anaphor's governor. In turn, this forces all anaphors and pronominals to have governors. This is, clearly, a stronger requirement than (1985). To put the contrast concisely, in Chomsky's theory, an anaphor or pronominal is subject to the licensing condition only if governed; in our system, the anaphor or pronominal is licensed only if governed.

A significant consequence of this formulation is that there cannot be any anaphoric pronominals, i.e. there is no PRO. The BT of Chomsky (1981; 1985) gets the restrictions on the distribution of PRO (the empty category subject of certain infinitival clauses and gerunds) to follow theoremtically from the Binding Theory, and from the status of PRO as being both an anaphor and a pronominal. This explanation is unavailable in our account. In chapter Four, we adopt (with certain

4. the sentence in (i) is, of course, grammatical.

(i) He left

Obviously, there is no linking here at all. This means that we must understand 'possible linking' to include the null case, in which there is no linking assigned at all. For (i), since the pronominal is not linked to any chain-accessible antecedent, the linking is BT-compatible.

modifications) Brody's (1985) account of the distribution and nature of those ECs which occur in such ungoverned contexts, an account which is cast, not in terms of the binding theory, but in terms of Case and chain well-formedness.

To conclude this discussion of government and its role in defining binding domains, consider the sentences below:

11) * John believes [_i him _i to be nice]

12) * John _i likes him _i

In Chomsky's theory, the GC for the pronoun must contain its governor. This is a desirable result, since otherwise coreference would be allowed in these examples. If the Binding Theory did not require that the GC for the pronominal contain its governor, then in (11) the clausal complement _{IP} [him to be nice] would be the GC; and the GC for the pronoun in (12) could

be taken simply as [_{NP} him]. Both are structural domains in

which there is a BT-compatible indexing under which the pronoun is free. In the (1981) theory, these domains are barred from being GCs for the pronouns because they lack accessible subjects. The (1985) theory dispenses with reference to accessible subjects.

In light of these cases, the role of government in the Chomsky

(1985) Binding Theory is strengthened; this is a desirable result, since (as Chomsky (1981) acknowledges) reference in the (1981) binding theory to a governor in defining governing category serves only to gain the distribution of PRO.

On our formulation of the binding theory, an anaphor or pronominal is licensed only if there is a well-formed CAS for it; and by (4ii), a CAS is well-formed only if it contains a projection of the governor of X. In (11), then, the pronominal is required to have a CAS which includes a projection of the governor of the pronoun, namely the exceptional-casemarking verb believe. The CAS will extend up to the matrix IP node. Under the assigned coreference relations the licensing condition is violated, since John is minimally accessible to the pronominal which it binds.

3.2 The Modified Binding Theory

The licensing condition on anaphors as modified in (7) requires that the anaphor in (1) be linked to a chain-accessible antecedent. Given the linking in (1'), this condition is satisfied, since John is chain-accessible to the anaphor through the sequence $S_1 = \{\text{himself}, PP, N', NP, e, VP, I', IP\}$.

[which [pictures [of himself]]] does [John[[like e]

NP N' PP IP I' VP

Further, John is minimally chain accessible --local enough to the anaphor to be a licit antecedent -- since there is no subsequence of S through which another antecedent is accessible.

Now consider the ambiguous sentence (2), in which there are two possible grammatical antecedents for the anaphor. This sentence illustrates the effect we earlier described as the multiple binding-domain effect. In our framework, this effect has formal status: due to the configuration of the sentence,

there are two accessibility sequences for the anaphor, and the two NPs are 'local binders' through these sequences, Bob through one, John through the other.

2) [[which [pictures [of himself]]] does [John
 CP NP N' PP IP
 [[think [e [Joshua [likes e']]]]]] ?
 I' VP CP IP* I'* VP*

chain = ([which pictures of himself], e, e')

Here, there are two accessibility sequences, neither of which is a subsequence of the other, through which John and Joshua are accessible. Consider first John. The relevant accessibility sequence S_1 is {himself, PP, N', NP, e, CP, VP, I', IP}. John is a sister to I', and the sequence meets the definition of accessibility sequence.

The accessibility sequence S_2 through which Joshua is accessible is {himself, PP, N', NP, e', VP*, I'*, IP*}. Joshua is a sister to I'*, and so is accessible to the anaphor.

Crucially, neither of these sequences is a proper subsequence of the other. Thus the two NPs are each minimally chain accessible to the anaphor, and either of the linkings in (9) and (10) will satisfy Condition A:

17) John told Joshua about himself

18) [John [[told Joshua [about himself]]]]
 IP I' VP PP



19) [John [[told Joshua [about himself]]]]
IP I' VP PP



I', IP}. By the definitions given in (4,5), Joshua is accessible through a sequence S (himself, PP, VP), which is a proper

is minimally chain accessible to the anaphor, and hence ought to be its only possible binder. Incorrectly, the revised licensing condition (7) excludes John as a potential antecedent to the anaphor.

- 136 -

requirement must be carried over to the accessibility framework, and motivates the addition of one final well-formedness constraint on chain accessibility sequences.

4. (iv) a is the root node of a Complete Functional
 n
 Complex.

With this addition, S in (18-19) is not a well-formed
 2
 accessibility sequence. Indeed, S is the only well-formed
 1
 sequence, and so John and Joshua are both accessible through the same accessibility sequence, and so each are minimally accessible to the anaphor. Thus the addition of (4iv) captures the ambiguity of (3).

3.3 NP-Movement and Reconstruction

Up to this point, we have been considering connectivity in cases of WH-movement, following the long-standing observation that in a representation like (20), the anaphor inside the phrase X can be referentially dependent upon Y if (i) there is no possible binder for the anaphor inside X , (ii) Y c-commands t , and (iii) $[\dots]$ is related to t through WH-movement, or,
 X
 more abstractly, a well-formed A' chain. The last aspect of the structure -- that the anaphor's container and the EC c-commanded by the anaphor's antecedent are both members of the same chain

SOME TEXT ON THE FOLLOWING PAGE(S) IS
ILLEGIBLE ON THE ORIGINAL MATERIAL.

3.3.1 Psych-Movement

Belletti and Rizzi (forthcoming) develop an extensive analysis of 'psych-movement' constructions, crucially asserting that NP-movement occurs in them. In the following discussion, I will adopt their analysis, which allows us to consider the apparently exceptional anaphora of psych-movement to simply be one example of connectivity, falling directly under the Binding Theory, as formulated in (4-8).

Postal (1970) observes that a certain class of verbs, the psychological predicates, appear to be syntactically transitive, yet have an odd behavior which sets them apart from other transitive verbs. With respect to certain syntactic processes, especially anaphora, the object of the sentence behaves as if it is the subject, and vice versa. Thus a sentence of the form (24) will typically (that is, with a transitive, non-psych verb) be ungrammatical due to a failure of the necessary relation between himself and NP* (a failure of c-command, in contemporary terms). However, a psych verb used in the sentence will permit it to be grammatical, as in (25).

24) [...anaphor] [V NP*]
NP VP

*Each other's parents visited the men

25) a. Each other's parents annoy the men

picture of himself bothers Dan

Postal proposes a process of "Psych-Movement", which has essentially the effect of making the post-verbal NP function as subject of the clause at some level of representation to which the binding condition governing anaphors applies. This rule is restricted in its operation to apply only to sentences with psych verbs. Postal does not formulate the rule, but it is clear what would be necessary. At some level of representation, say, LF, the object of the psych verb (the men in (25a), Dan in (25b)) will have to move to an A-position, where it c-commands

7
the anaphor. Such a rule is impermissible within the restricted theory of GB, since it would give rise to a violation of the projection principle.

Belletti and Rizzi offer an alternative solution to this problem, drawing upon extensive evidence from Italian, as well as the binding facts in English. The core of their proposal is that sentences like (25a,b) have derived subjects; the NP which occurs in subject position at S-Structure will be internal to the VP at D-Structure, and will raise (to get Case) from object position in the DS-SS mapping. The apparent syntactic object of (25b), John, is in essence in indirect object position. Belletti and Rizzi's proposed derivation proceeds as in (26):

7. The landing site must be an A position, since the antecedent of an anaphor must be located in an A position.

(26)

DS:[e [[bothers [this picture of himself] John]
IP V'' V'

SS:[[this picture of himself] [[bothers t] John]
IP i V'' V' i

On this view, psych verbs are very much like unaccusative or ergative verbs (see Burzio (1981), Levin (1982)), in that they assign a theta role but no Case to direct object position, forcing NP-movement as in (26). (Psych verbs differ, on Belletti and Rizzi's analysis, essentially in that the psych verbs take

8

an indirect object as well) .

9

The movement creates a chain ([pictures of himself], t) , associating the casemarked argument [pictures of himself] with the trace occupying the theta-marked object position. Belletti and Rizzi observe that the account of anaphor connectivity making reference to A' chains (originally proposed in Cinque

8. Barsa and Lasnik (1986) discuss the problem of anaphor in double object constructions like (i) and (ii), observing that the linearly first object appears to asymmetrically c-command the second NP:

(i) I showed the students [each other's books] (ii) *I showed [each other's friends] the students

This indicates that the underlying complement structure of psych verbs proposed by Belletti and Rizzi is apparently confined to this class of verbs. I know of no verb in English which takes two case-marked objects, where the second object may bind into the first. That is, English has no grammatical sentences like (iii). This is an interesting gap in the English lexicon.

9. Or, alternatively, the theta criterion is satisfied iff a chain ([pictures of himself], t) exists.

(1982), and extended in Barss (1984)) can be generalized to chains of either type ¹⁰. The grammaticality of such examples as (25a) and (25b) is now predicted, precisely for the class of verbs that exhibit this type of movement. The binder John will not c-command the anaphor, but it will c-command the trace, satisfying the licensing condition. The restriction of such movement to a lexical class is not a complete mystery, since the movement is forced for case-theoretic reasons, and it is the case that certain lexical classes of verbs assign a theta role but no case to their objects.

It has sometimes been suggested that such examples as (25a,b) need not conform to the core principle governing anaphors. Rather, on this alternative view, the anaphor is permitted to freely seek an antecedent in its linguistic context when certain properties hold of the anaphor's position. That is, anaphors are intrinsically non-referential, and so must have an antecedent somewhere, but the antecedent, under certain circumstances, need not c-command the anaphor. One such proposal along these lines is Manzini's (1983) analysis of control and binding, which has the consequence (as noted by Mohanan (1985)) that when an

10. In our terms, Belletti and Rizzi's argument holds that (4iii2) should simply refer to 'chain' rather than a particular type of chain. In earlier work on the connectivity problem, I proposed a formalism similar to that developed in this thesis, but which referred only to A' chains. Belletti and Rizzi's work led me to abandon such specific reference to one type of chain, as is reflected in (4ii2). I wish to thank them for helpful discussion.

anaphor occurs in a lexically ungoverned position, e.g. specifier of NP, and that NP occurs in a lexically ungoverned position (e.g. subject position), the anaphor has no binding domain; this exempts the anaphor from Condition A, allowing it to pick up a non-c-commanding antecedent. In a sentence like (27) (Manzini's (82)), the anaphor may 'corefer freely', subject only to the necessity of having a linguistic antecedent.

27) each other's pictures would please the boys

As noted by Mohanan (1985), this proposal fails on several grounds. To give one of Mohanan's objections, (28) (Mohanan's 11a) is ungrammatical, yet satisfies the requirement that there be a plural NP present in the sentence with which the reciprocal

11,12

might be construed, namely their.

11. This has also been noted elsewhere, recently by Georgi (1985), and indeed was noted in Postal (1971).

12. A side note on Mohanan's proposal, with which I do not entirely agree. He claims that the phenomenon in (27), which might be characterized as the relaxation of c-command, is restricted to sentences in which the NP containing the anaphor is headed by a picture noun. He also claims that the phenomenon exhibited by such sentences as (i), where the anaphor and its antecedent are contained in different clauses, is similarly restricted to sentences where the containing NP is headed by a noun of the picture class.

(i) The boys thought that [each other's pictures] were on sale (Mohanan's (9d); Manzini's (80))

He cites the following pair (his (13a,b) as 'ungrammatical for most speakers':

ii) The boys thought that each other's girlfriends were pretty

iii) [each other's girlfriends] upset the boys

I find (ii) and (iii) as fully grammatical as (i) and (27), as do the speakers I have consulted. An example identical to (ii) is given as grammatical by Chomsky (1976), and many other authors have cited similar examples as grammatical.

Johnson (1985) observes that psych verbs optionally theta-mark their surface subjects, giving rise to the ambiguity of (iv):

iv) The girls annoyed the boys

On one interpretation, the girls deliberately did something which we might call annoying the boys; on the other interpretation, the girls' mere existence annoyed the boys, and they were not consciously involved in any event of annoyance. Johnson observes that in the first sense, the girls is the agent of the clause, while on the second, the theme. He further suggests that the NP-movement derivation proposed by Belletti and Rizzi only occurs for the latter type of reading.

Assuming that, the binding effect can occur only when the subject (containing the anaphor) is interpreted non-agentively. This is the natural interpretation of (i), since pictures cannot act agentively or initiate events. It is much more easy to construe (iii) volitionally, and so this may lead to a certain indeterminacy in the grammaticality judgement.

Let us suppose, however, that a large number of speakers do entirely reject (ii) and (iii) as Condition A violations. Then we must ask what type of explanation can or cannot be admitted into the theory of binding. Mohanan's characterization of this split would seem to be purely descriptive:

If we treat (13a-b) as the core facts, all we need to say is that in some dialects of languages like English, a small class of picture noun phrases allows the anaphor contained in the subject (with the relevant added restrictions) to take as its antecedent the object of the same clause or an element of the immediate matrix clause.

First, the former cases (in which the antecedent may be 'the object of the same clause') are restricted to sentences with psych-verbs; this is Postal's original observation. Secondly, in the cases like (i), the antecedent must c-command the anaphor; consider for example (v).

v) * Mary was persuaded by John's sister that pictures

28) *Each other's pictures would please [their professor]

Belletti and Rizzi's proposal explicitly captures this fact. Since only the experiencer argument of the verb, never a sub-part of it, c-commands the object trace, only it can serve as the antecedent of the anaphor inside the subject NP.

We may now tentatively conclude, following Belletti and Rizzi, that the apparent exceptions to Condition A illustrated in psych-movement sentences are due to the occurrence of NP-movement in the DS-SS mapping, and that the Binding Conditions (4-7) are properly put in terms of chains (of either A or A' type), not restricted only to A' chains. This is, of course, not so much an extension as a simplification; there is no need to distinguish different chain-types as far as anaphor connectivity goes.

We shall now examine further cases of NP-movement, and show that they give further support for the conclusion that A-chains are involved in connectivity effects. This section will conclude with a brief discussion of Pesetsky's (1985) analysis of connectivity in psych-movement sentences, which is quite different from that suggested by Belletti and Rizzi.

of himself were on sale

It would seem, then, that the phenomenon in (27) is not the peripheral phenomenon Mohanan suggests, falling almost outside the purview of formal research, but rather a quite core phenomenon, one which obeys quite strict restrictions which may be revealed by the methodological approach used elsewhere.

3.3.2 Raising and Connectivity

The operation descriptively known as raising is illustrated in (29); I shall adopt the standard GB view of this construction, in which the verb seem takes an infinitival sentential complement, and governs but does not casemark the subject of this complement. Since seem is not an exceptional casemarker, the overt NP John (the underlying subject of the complement) must move in the DS-SS mapping to a casemarked position, in order to meet the requirements of the Theta-Criterion or the

13
Chain Condition .

13. Chomsky (1985) states the Theta Criterion as in (i), where CHAIN refers to chains, and to expletive argument pairs (see chapter four of this thesis for review and discussion):

- (i) A CHAIN has at most one O-position; a O-position is [must be;AB] visible in its maximal CHAIN. (Chomsky's (168; p. 135)

Visibility is constrained by Chomsky's condition (167; p. 135):

- (ii) A CHAIN is Case-marked if it contains exactly one Case-marked position; a position in a Casemarked chain is visible for O marking.

Finally, the Chain Condition (ex.(171); p. 137) restricts Casemarking to the head position in a chain:

- (iii) If $C = (a_1, \dots, a_n)$ is a maximal CHAIN then a_1 occupies its unique O-position and a_n its unique Case-marked position.

Movement of John is forced by these three conditions. The subject position of the complement clause is O-marked, but not Casemarked; in order for this O-position to meet the

29) John seems to be tall

30) DS: seems [John to be tall]
IP

SS: John seems [t to be tall]
IP

~
| |

The chain (John, t) meets the conditions on chains, in having one theta position and one case position and one argument (see footnote ()). Since seem governs its complement's subject, the necessarily ungoverned element PRO¹⁴ is prohibited from this position, accounting for the ungrammaticality of (31).

theta-criterion, it must be visible; it is visible only in a Casemarked chain, thus John moves to the casemarked (but 0') subject position of seem, giving rise to the well-formed chain (John, t).

14. In the view of empty categories set forth by Chomsky (1981), (1982), (1985), and adopted in much of the GB literature, PRO is an anaphoric pronominal, subject to Conditions A and B of the binding theory if it is governed. Since no element can simultaneously satisfy both conditions, PRO is necessarily ungoverned.

As observed at the beginning of this chapter, the formulation of the binding theory I give requires all anaphors and pronominals to be governed in order to be licensed; a consequence of this is that there can be no anaphoric pronominals, governed or ungoverned. In the next chapter, I develop an alternative view of empty categories which builds upon Brody's (1985) approach. In Brody's theory, there is no feature [+/- pronominal] for empty categories, and so there is no PRO. Rather, on Brody's view, an EC which heads its chain must be ungoverned. Hence, empty categories occupying the subject position of infinitival clauses, gerunds, and so forth, since they head their chains, must be ungoverned. On this view, example (31) is ill-formed since the EC is governed, hence must not head its chain; and the theta criterion can only be satisfied by the existence of a chain headed by the EC. See the next chapter for a review of Brody's theory and my extensions of it.

31) it seems [PRO to be tall] (meaning 'it seems that someone or other is tall')

We observe first that if the subject of a psych-verb sentence is raised, there is no decrease in grammaticality, even if the

subject contains an anaphor. 15

321

It seems that[_{IP} [these pictures of each other] [_{UP} bother t them]]

33) [these pictures of each other] seem [IP t' to [VP bother t them]]

On the hypothesis that connectivity holds of A-chains, advanced by Belletti and Rizzi for psych-movement, this is to be expected. There is a triplet chain (these pictures of each

15. In an indexing theory, the indexed representations are as follows:

32') It seems that[[these pictures of each other]
 IP k i
[bother t them]]
 VP i k

33') [these pictures of each other] seem [t']
 k i IP i
to [bother t them]]
 VP i k

Please see the appendix for a version of (4-7) within indexing theory; this modified Binding Theory will allow (32'-33'), as well as all the other grammatical examples of connectivity in NP-movement environments.

other], t', t), the experiencer argument of bother c-commands one member of the chain (the lowest trace, and thus the linking of each other to them satisfies the modified binding theory. Assuming the Belletti and Rizzi proposal for psych-movement, which gives (34) as the structure for (33), the accessibility sequence through which them is accessible to each other in (33) is: {each other, PP, N', NP, t, V', VP*, I', IP*}.

34) [[these[pictures[of [each other]]]] [[seem
IP NP N' PP I' VP
[t'[to [[bother t] them]]]]]
IP* I' VP* V'*

Note that if we were to adopt some other account of connectivity in simple psych-movement sentences like (25) -- Manzini's or Mohanan's for example -- this additional preservation of grammaticality under raising is totally unexpected.

3.3.3 Pseudo-Clefts and Raising

Higgins' (1976) rich study of connectivity examines cases of this phenomenon in the English pseudo-cleft (PC) construction. This construction is illustrated in (35) and (36). Higgins identifies (in part following Akmajian (1971)), two distinct types of PCs, each associated with a rather different

16
interpretation .

35) [what John bought] was that table over there

36) [what John bought] was ugly

(35) is an instance of a specificational pseudo-cleft, while
(36) is a predicational pseudo-cleft. Taking the free relative
subject of (35), what John bought, to denote an individual

16. The term 'pseudo-cleft', originally due to Jespersen, arose
in the literature notes the construction's similarity to clefts:

i) (cleft) It was that book that I bought

ii) (PC) what I bought was that book

17. The subject of (35) has the internal structure of a CP. It
is however, interpreted as a definite description. This has
sometimes been taken to indicate that the free relative is
dominated by an NP node, as in (i), since NPs are the typical
instance of definite descriptions. This seems undesirable from
a syntactic viewpoint, since the structure so assumed violates
typical X' structure -- the NP node is not the projection of any
noun.

(i) [[what [John [bought t]]]]
NP CP IP VP

An alternative way to reconcile the expression's internal
syntactic structure with its semantic interpretation would be to
suppose that the WH phrase is interpretable as a definite
determiner. The IP inside the CP will act to restrict this
quantifier. Thus the expression (ii) maps directly onto an
interpretation like (iii):

ii) [what [John bought e]]
CP i IP i

iii) [the x [John bought x]]

 It has been observed that WH extraction from inside a free relative is impossible.

iv) I know that this car was what Bob gave to Sue

v) * [who do [you [t' ' ' [know [t' ' [that
 CP IP VP VP CP C'
 [this car [was [what [Bob [t'
 IP VP CP IP VP
 [gave e to t]]]]]]]]]] ?
 VP

This cannot be taken as evidence that the free relative is syntactically an NP, since the Subjacency principle will bar this extraction anyway (assuming the formulation of Subjacency in Barriers, and assuming the structure in (v), movement from t' to t' ' crosses a barrier (CP, which inherits barrierhood from IP, crossed on the same instance of movement)). Furthermore, extraction out of a definite NP is in general blocked, so that the definite interpretation of a free relative, whatever its topmost syntactic node, will block extraction from it.

Further evidence that a free relative is interpreted as a definite expression, not an indefinite, can be adduced from the following paradigm:

- vi) a. I consider Dan a fool
 b. * I consider Dan the fool
 c. * I consider that table what John bought
 d. I consider Dan to be the fool
 e. I consider that table to be what John bought

- vii) a. There was a man in the room
 b. * There was the man in the room
 c. * There was what John bought in the room

A free relative is excluded from predicate position inside a small clause (vi, bc) and from there constructions, which exhibit the definiteness effect (prohibiting definite expressions from occurring there; see Jenkins (1976, Safir (1981), Higginbotham (1985)).

Note that the exclusion of a free relative from predicative position, in (vi), indicates that these expressions are not interpreted as predicates. This contradicts the proposal of Williams (1983) that they are. See Appendix 2 of this chapter

(the thing that was bought by John), (35) specifies (or identifies) the identity of that thing, informing us in particular that the thing in question was that table over there. (36) on the other hand given us no information as to the identity of the bought object, but does assert that a certain

18

property holds of it, namely ugliness .

It thus seems appropriate to designate the post-copular material in a PPC as a predicate, differing neither in form nor function from the predicate of a sentence like (65):

37) That table is ugly

The existence of PPCs is therefore properly seen as nothing more than a result of the permissability of interpreting phrases with the internal structure of CPs with WH-phrases in SPEC as definite descriptions.

The exact function of the post-copular material in a specificational pseudo-cleft, and the nature of the relation between it and the free relative subject, is still an open question. As we shall shortly see, certain connectivity effects suggest that this relation is of a very different sort than

for further discussion.

18. It should be noted that the suffix 'ever' may be appended to the WH element in the free relative of only the predicational pseudo-cleft:

63') *whatever John bought was that car over there

64') whatever John bought was ugly

predication.

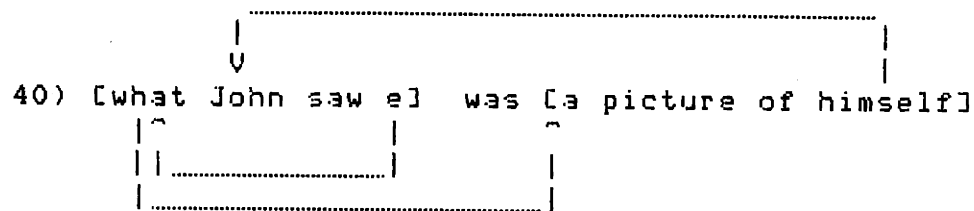
As a cautionary note, the distinction between PPC and SPCs is often blurred by the fact that certain constituents, including indefinite NPs, may function in pseudo-clefts in two quite distinct ways. They may function as predicates, predicating a property of the free relative subject (see (36)), or as foci of SPCs, in which case they are understood as bearing the same relation to the subject inside the free relative as they would if they were in the position of the empty category (39).

39) what George became was important to his city

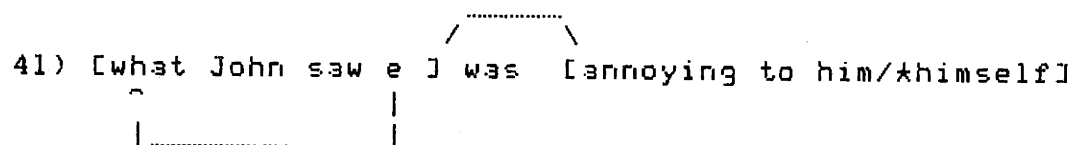
On one interpretation, (39) is synonymous with 'George became important to his city'; the AP [important to his city] is understood as predicated of George directly, as if it were in the position of the EC inside the free relative. On another interpretation, George became something -- say, police commissioner -- and it is that which he became which is important to the city. On this latter interpretation, [what George became] denotes something --his office -- and the property [important to his city] is predicated of that, not of George directly. Rather, George is important to his city in virtue of what he became.

The most striking asymmetry between the two types of PC, whatever the formal nature of this difference may be, is the presence of connectivity in specificational sentences. Consider

the contrast between (40) and (41):



Predication



What we see in (40) is the by now familiar connectivity environment. The sentence is grammatical; the anaphor is not c-commanded by its understood antecedent at SS, but the antecedent does c-command a trace which is anaphorically related to a constituent containing the anaphor. We suppose, following essentially Cinque (1982) and Barss (1984), that the phrase [a picture of himself] heads an A' chain containing the EC, the chain being (a picture of himself, what, e). John commands e, and thus is chain-accessible to the anaphor himself, and may therefore antecede it, satisfying the Binding Theory (4-7).

By contrast, there is no such chain relation between annoying to himself and e in (41), a predication pseudo-cleft, and so the connectivity effect does not occur.

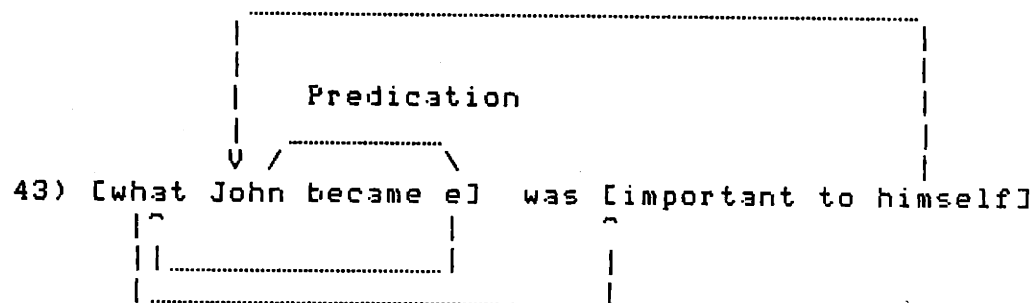
Let us now consider a case which ought to be ambiguous between a predication reading and a specificational one, like (39).

When the AP contains an anaphor, the pseudo-cleft receives only a specificational interpretation (the following example is adapted from Higgins (1970)):

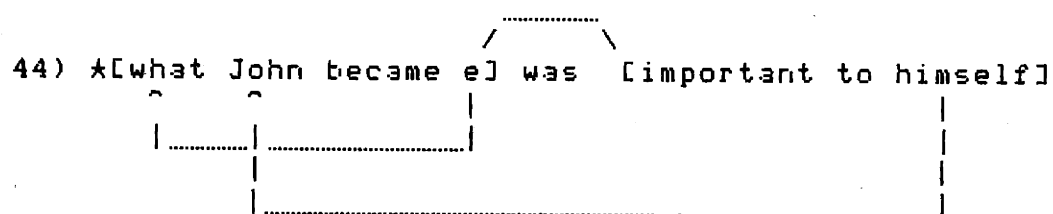
42) what John became was [important to himself]
AP

(42) can only (grammatically) mean 'John became important to himself', not 'John became something (say, the best teacher in town), and that role was important to John'. This is expected, on our analysis, since only on the specificational reading is there a chain associating important to himself and e. Connectivity occurs only when there is a chain, one member of which is c-commanded by the anaphor's antecedent; to grammatically link himself to John requires this chain, which forces the specificational reading.

To be more explicit, (42) admits of two distinct S-structure representations (paralleling (40) and (41)), differing in the antecedence and predication relations involving important to himself.



Predication

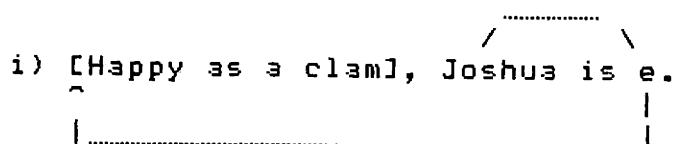


(43) gives rise to the specificational interpretation; there is a chain (important to himself, what, e), the head of which determines the value of e, (eg. 'specifies' the value of the trace), allowing predication to occur between e and John inside the free relative ¹⁹. Because this chain exists, John is (minimally) chain-accessible to the anaphor himself, and so the link between John and himself satisfies the binding theory.

By contrast, the link between himself and John in (44) violates the Binding theory, since in this structure [important to himself] is predicated of the entire free relative [what John saw e]. No chain associates [important to himself] with the trace c-commanded by John, hence John is not accessible to the anaphor. (What fixes the value of the trace in (44) is what,

19. This is much as in the Topicalization sentence (i), where the left-shifted AP binds the trace e in a chain (happy as a clam, e), allowing predication to occur IP-internally.

Predication



interpreted here (as suggested above) as a definite quantifier;
(what, e) is the the ECs maximal chain.)

The above analysis of specificational pseudo-clefts is sketchy, and the proposal that there is a chain relating the focus expression and the trace is radical enough to require further justification. In Appendix 2 of this chapter, I discuss pseudo-clefts in more detail. For the purposes of this discussion, let us suppose that some syntactic relation holds between the trace and the focus phrase which permits a chain accessibility sequence for an anaphor inside the Focus to be constructed through the trace, even if the chain analysis is incorrect. In the next section, we show that connectivity is preserved in pseudo-clefts when the focus is raised; this is expected, given our incorporation of A-chains into the definition of CASs.

3.3.4 Raising and Pseudo-Clefts

To return to raising, observe that a SPC may be inverted, as in (45):

45) This picture of himself is what John likes e best

The NP this picture of himself may raise, as in (46-7), with no
20
decrease in grammaticality.

20. The mere permissibility of raising this NP demonstrated that it is the syntactic subject of the sentence, and thus that the

- 46) This picture of himself seems [t' to be what John likes t best]
IP
- 47) these pictures of each other seem[t' to be [what[John and Mary
like best t]]

The grammaticality of these examples, within the framework developed here, is due to the existence (at S-Structure) of two well-formed chains, which together relate the NP containing the anaphor, namely this picture of himself, to a trace c-commanded by John. The first chain is the A' chain (t', what, t), and the second is the A-chain (this picture of himself, t'). The casemarked trace t is c-commanded by the NP John, and the anaphor is contained within the ultimate antecedent of this trace.

Assuming that (47) has the constituent structure (48), the chain accessibility sequence through which John and Mary is minimally chain accessible to the reciprocal is: {each other, PP, N', NP, t', t, VP**, I'**, IP**}.

- 48) [[these[pictures [of each other]]] [[seem
IP NP N' PP I' VP
[t [to [be [what [[John and
IP* I'* VP* CP C' IP**
Mary [[like t best]]]]]]]]]]]
I'** VP**

3.3.5 Pesetsky's Analysis of Psych-Movement

Pesetsky 1985 develops another analysis of the binding problem

'inversion' of (45) is not a stylistic operation.

exhibited in sentences with psych verbs, one which crucially differs from that of Belletti and Rizzi in not assuming movement of the surface subject from inside the VP. Thus while Belletti and Rizzi propose (49) as the appropriate structure for the sentence, with a derived subject, Pesetsky holds that something more or less like (50) is correct.

49) [IP [NP pictures [PP of himself]] [I' [VP [please t] John]]]

|

| |

50) [IP [NP pictures [PP of himself]] [I' [VP please John]]]

Pesetsky proposes a different analysis of anaphor connectivity in psych-movement sentences than the one just outlined. Like the preceding derived-subject analysis, his treatment attempts to subsume the anaphor binding in such sentences as (49)-(50) under a general treatment of connectivity or 'reconstruction'. This section reviews his analysis and suggests that the preceding treatment (based within Belletti and Rizzi's characterization of psychological predicates) is to be preferred on theoretical grounds.

As Pesetsky observes, psychological predicates sometimes appear with CP complements, as in (51):

51) This book pleased John to read

Invariably, the sentential constituent is a control structure, with PRO controlled by the experiencer of the psych verb (John

in this example).

Pesetsky assimilates such structures as (51) to tough movement sentences, like (52):

52) This book is easy [for Mary] [PRO to read]
PP

Like sentences of the form (51), TM sentences have infinitival complements.

When the post-adjectival NP (the benefactive argument) is
overt, it obligatorily controls PRO²¹.

21. Epstein (1984) observes that there is strong reason to consider PRO to be controlled in sentences of the form (52) even when the benefactive argument is not overt, as is the case in (i). This claim runs counter to the standard assumption that sentences like (i) exhibit 'arbitrary control', which is assumed to involve assigning a generic or arbitrary interpretation to PRO itself (see Chomsky (1980), 1981; Willimas 1980; etc.)

i) It is hard [PRO to play baseball]

To cite one of Epsteins central arguments, the sentence (i) has an understood benefactive argument, understood essentially as a universal quantifier over human individuals. As Epstein observes, there are in essence two 'arbitrary' arguments in (i), not one. (i) may be paraphrased as (ii).

ii) For any x, it is hard for x [x to play baseball].

This fact -- that there are two non-overt arguments -- is missed in a theory which treats the (i) as having no benefactive argument, and an arbitrarily interpreted PRO.

Crucially, the quantificational benefactive has matrix scope, just as anyone does in (iii), indicating that the benefactive occurs in the main clause in both cases.

(iii) It is hard for anyone [to play baseball]

The analysis of (i) which proposes that the interpretation of

Pesetsky remains neutral on the question of whether the TM subject is moved from inside the CP (the earliest suggestion on this matter), or is generated in place, following the analysis of TM suggested by (Chomsky (1981); (1986))²² which proposes that that there is movement of an empty operator internal to the CP:

53) [[This book] is [fun [for Mary]
 NP i AP PP j
 [OP [PRO to read t]]]
 S' i IP j i

Pesetsky further notes that in TM examples there is anaphor connectivity between the benefactive NP and an anaphor inside the derived subject, as in:

54) [[Stories about themselves] are [unpleasant
 S NP j i AP
 [for the candidates] [OP [PRO to hear t]]]
 PP j S' i IP j i

 PRO in (i) is to be given by assigning it the index arb can capture neither the scope fact, nor the existence of two arguments in a sentence like (i). Epstein concludes that the non-overt benefactive in (i) is an empty category, specifically pro, which is interpreted as a universal quantifier. This pro controls PRO, which is then understood as a bound variable. This captures the scope fact, and the presence of two arguments, as well as the important fact that the two variables covary. The conclusion of Epstein's argument is that PRO is controlled in many cases in which it is assumed that there is no control, including (i).

22. Chomsky (1977) makes this same basic suggestion, except that the fronted operator is overt (what, who, etc) and then deleted. The empty operator analysis eliminates the need for deletion, proposing instead that there is a phonetically null operator, represented OP, among the class of syntactic phrases permitted to move to SPEC of CP.

Theory applies. Secondly, the apparent absence of the CP at S-Structure is due to the operation of a deletion rule.

Pesetsky's analysis has several undesirable attributes. First, it is not the case that every psych-verb sentence is paraphrasable with a sentence in which an overt CP complement appears. Second, the analysis is committed to the view that a sentence like (56) is as fully grammatical as (57), which is not true.

56) ??? This book pleases me to write/buy/read

57) This book pleases me

Thirdly, his analysis involves wholesale deletion of a clause. This would appear to violate the Projection principle, and the principle of recoverability of deletion.

The status of deletion rules in GB theory is reduced from that in EST; the development of GB theory has seen a shift toward viewing many processes previously considered to be deletion (under identity) as interpretive processes. However, the type of deletion proposed by Pesetsky is intrinsically more powerful than any of the other deletion phenomena of English, like gapping, VP deletion, ellipsis, and so forth. This is so because the deletion of the CP on Pesetsky's analysis is not under identity; the material deleted is not recoverable from the syntactic structure. For this reason alone, the analysis of binding in such sentences like (49-50) offered by Pesetsky is to be barred from a restricted theory like GB in its current form.

The semantic point Pesetsky makes is that sentences like (49-50) often have an understood, but unmentioned, proposition.

As he puts it:

Consider also the meaning of a sentence like (4a)[our (50); AB]. The pictures in question are not annoying in a situational vacuum. They are annoying in the context of some other relation that obtains between them and the politicians (c.f. Lee (1971)): they are annoying to look at or to hang on the wall or to have to discuss. ...I therefore suggest that examples like...[our (50); AB] actually involve a deleted infinitival clause which has undergone Tough movement. The exact clause deleted depends on discourse factors.

He then addresses the problem that such sentences may be uttered in isolation, without a previous discourse to support the deletion:

"In these circumstances, some prototypical or contextually likely embedded infinitive is simply assumed and deleted."

This seems to me rather similar to the fact that the sentence (58) has an understood instrument.

(58) I broke the window

But this fact is no reason to suppose that the sentence in (58) is derived through deletion from (59).

(59) I broke the window with a rock

Rather, it is a fact about breakings that they typically involve an instrument of some sort. Similarly, it is a fact about experiences which bring about a psychological change in one that they often arise through some activity involving the

experiencer and the entity (theme) which is involved in the change. This is no reason to suppose that this relation is syntactically specified in a sentence like (49-50) as part of the sentence. Just as breakings can sometimes occur without an instrument, as expressed in (60), pleassings can occur without accompanying activities, as in (61).

60) I broke the window by accidentally letting it fall.
I broke the window -- I forgot to fix the sashcord.

61) Books please me, in all sorts of ways.
Books please George, although I don't know why.

Therefore, it seems desirable to consider the anaphoric dependency between the anaphor and its antecedent in (49) to be licensed through a means like that made available in the Belletti and Rizzi analysis.

3.3.6 Ambiguities with Psych Verbs

Consider the following grammatical sentences:

62) George thinks that these pictures of himself annoy Susan

|
|.....|

63) George thinks that these pictures of herself annoy Susan

|
|.....|

Both of these sentences are grammatical. This fact is predicted in our analysis, as there are accessibility sequences through which George and Susan are each minimally accessible to the anaphor.

On the Belletti and Rizzi analysis, these sentences have the structure (64) (where X is himself or herself):

(64) [George [[thinks[[[[these [pictures
 IP I' VP CP C' IP* NP* N'*
 [of X]]] [[[please t] Susan]]]]]]]
 PP* I'* VP* V'*

S₁ = {X, PP*, N'*, NP*, I', IP*, C', CP, VP, I', IP}

S₂ = {X, PP*, N'*, NP*, t, V', VP, I'*, IP*}

George is accessible to the anaphor through S₁, Susan through

S₂. Since neither sequence is a proper subsequence of the other,

both NPs are minimally accessible to the anaphor.

3.3.6.1 Passive

Observe that in the passive sentences corresponding to (62) and (63) only the lower subject, Susan, is accessible to the anaphor.

65) George thinks that Susan would be pleased by these pictures of herself

66) * George thinks that Susan would be pleased by these pictures of himself

(adapted from Jacobsen and Neubauer(1976; ex. ??)

This is expected, since in the structure (67) only Susan is minimally chain accessible to the anaphor. This is so because the sequence through which Susan is accessible is a proper subsequence of the sequence through which George is accessible

(indices used for clarity).

(67) [George [[thinks[[[Susan [would be
 IP I' VP CP C' IP* j I'*
 [[pleased] t] [by [these
 VP* V'* j PP NP*
 [pictures [of X]]]]]]]]
 N'* PP* i

$S_1 = \{X, PP^*, N'^*, NP^*, PP, I'^*, IP^*, C', CP, VP, I', IP\}$

$S_2 = \{X, PP^*, N'^*, NP^*, PP, I'^*, IP^*\}$

Susan is accessible through both sequences, crucially through S_2 . George is accessible through S_1 . Since S_2 is a proper subsequence of S_1 , George is not minimally accessible to the anaphor, hence the ungrammaticality of (67).

3.3.7 Against Successive Cyclic Binding

3.3.7.1 Binding at Intermediate Representations

It has sometimes been suggested that cases of anaphor connectivity like (68), where the anaphor is not c-commanded at S-Structure by its antecedent, can be handled by allowing Condition A to apply successive cyclically, that is, to intermediate representations in the DS-SS mapping.

- (68) a. John wonders [which pictures of herself] [Mary likes t]
 b. Which pictures of himself does John think t that Mary criticized t?

(68a) has the D-structure (69); (40b) has a representation (70)

intermediate between DS and SS:

69) John wonders Mary likes which pictures of herself

70) John think [[which pictures of himself] [Mary criticized]]

Suppose we were to adopt Chomsky's licensing condition for anaphors as formulated (see (Chapter 2, exx. 96-98)), and allow it to apply to any representation between D-structure and S-structure (including these representations themselves). The licensing condition is a positive filter -- it allows an anaphor to be licensed, or accepted as grammatical, when it meets certain requirements. The requirement imposed by the licensing condition is that it be c-commanded by a coindexed NP within the smallest domain in which it could be so bound. On this view, the licensing condition will accept (license) the anaphor in (68a) in virtue of its D-structure (69), and it will license the anaphor in (68b) in virtue of its intermediate representation (70). The fact that the anaphor in (68a,b) is not c-commanded by its antecedent is of no matter, since at one representation this necessary relation occurred.

On this 'successive-cyclic binding' view, the licensing condition will apply at any representation in the derivation; as formulated, it would be a global condition, since it is satisfied if one (arbitrary) representation meets it. In order to avoid globality, let us suppose that a feature, call it +B, is assigned to an anaphor whenever it meets the requirements of the licensing condition. This rule is given as (71). The

S-structure filter in (72) will serve to rule out sentences with anaphors that are never bound, like (73).

71) Rule of +B Assignment

Assign +B to an NP Y (optionally), if Y is bound in the least CFC B for which there is a possible indexing under which Y is bound in B.

72) * [+ anaphor]
 NP
 -B

73) * [John's mother] saw himself

In such a framework, anaphor binding is quite technically similar to the formulation of the ECP proposed by Lasnik and Saito (1984) and adapted by Chomsky (1986)²⁴.

I will now show that this analysis cannot be maintained, suggesting that the restriction of the KI to S-structure representations is correct.

24. The successive-cyclic approach to binding has one major drawback -- it requires that in all cases of anaphor connectivity, the anaphor's container actually be moved from the position occupied by the trace which is c-commanded by the antecedent. This is quite difficult to maintain for such examples as (40, 43). See Chomsky (1981; 144, fn. 79) and Burzio (1981). In the remainder of this section, I give an argument against successive cyclic binding based on examples of movement, so this issue is independent of the validity of the successive cyclic analysis, which will be shown to be untenable for other reasons. See also section (3. 8) of this chapter.

3.3.7.2 Passive and Successive Cyclic Binding

In a typical passive sentence, the operation of attaching passive morphology effects two changes: it blocks assignment of case to the verb's argument (forcing NP-movement to subject position), and absorbing the theta-role assigned by the active verb to subject position. This latter change allows the NP-movement to be grammatical, moving the verb's object into a theta-less position.

74) Bob kissed Mary

75) Mary was kissed t [by Bob]
 +Case -Case
-0 +0

The status of the by-phrase is a topic of debate. Verbal passive sentences have the same number of understood thematic arguments as the corresponding active sentences. This is true even when there is no overt by-phrase:

76) Mary was kissed

In both (75) and (76), it is understood that there is a kisser (the actor, or perhaps agent, of the kissing). In (75), the by-phrase in some sense expresses the actor role, but there is still an understood actor in (76), which lacks a by-phrase. This suggests that the agent theta role is assigned to some syntactic element present even in such sentences as (76); let us follow Baker, Johnson and Roberts (in progress) in supposing

that the passive morpheme is assigned the verb's external O-role, and syntactically encliticizes onto the verb, and is assigned the verb's accusative case. The by-phrase, then, acts to fill in the semantic content of this argument, but is not O-assigned directly.

VP-deletion, a standard test of constituency, can target a passivized VP, deleting the by-phrase as well:

77) Mary was kissed by Bob, and John was _____ too.

Let us take this as evidence that the by-phrase is not external to the core VP, plausibly adjoined to it (78a). VP-anaphora indicates that the by-phrase is not outside the VP, as in (78b):

78) a. $\begin{matrix} \text{IP} & & \text{I'} & & \text{VP VP} \end{matrix}$ Mary was kissed t] by John]]]]

b. [IP Mary L was [VP kissed t] by John]]]

We will now examine the possibility of adopting successive cyclic binding, considering such cases as (79).

79). The men were kissed t by each other's wives.

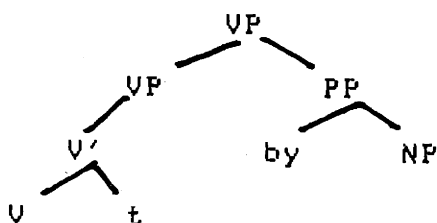
For concreteness, let us assume VP-adjunction for the by-phrase, as in (80).

80) The men were [_{VP} [_{VP} kissed] t [by each other's wives]]



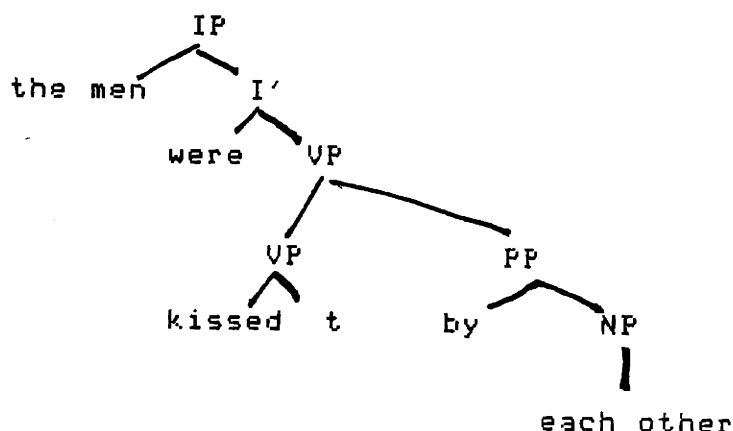
The VP has this hierarchical structure.

81)



What is crucial is that the by-phrase is outside the c-command domain of the D-structure position of the trace of the passivized NP. This will be so on the branching node definition of c-command, assumed here, or the m-command definition, since t is included in a maximal projection (VP) not including the PP.

82)



Now let us again consider examples (65) and (66), in which an anaphor inside the by-phrase must have the passivized subject as its antecedent.

The difference in grammaticality between (65) and (66) is

quite important, since it is only in the derived representation of (66) that Mary prevents George from being minimally accessible to the anaphor. Therefore, the binding theory (at least that part which is concerned with anaphors) must be restricted to apply only to the sentence (66) as it is, with Mary in subject position. If the licensing condition for anaphors were allowed to apply prior to the movement of Susan to subject position, (66) would be grammatical. That is, condition A does not apply successive cyclically. (I shall return to this issue later).

The question of passive in psych verbs is complex, since the by-phrase expresses the argument (Theme) corresponding to an internal argument of the active verb. This is in contrast to all other passive structures, where the by-phrase corresponds to the originally external argument. For this reason, I illustrate the point with standard passive (see Fiengo (1980) for a similar point).

Consider (83a) and (b). As we see, the anaphor inside the by-phrase must be anteceded by the derived subject, not the next highest subject.

83)

- a. John thinks that [the men were [kissed t] by each other's wives]
- b. * The men think that [John was [kissed t] by each other's wives]

Now consider the D-structure of (83b).

(83c)

[The men [[think [[
 IP I' VP CP C'
 [[was [[[kissed John]] [by
 IP* I'* VP* VP* V'* PP
 [each other's wives]]]]]]]]]
 NP*

The following accessibility sequences exist in this structure:

(84)

S₁ = {each other, NP*, PP, VP*, I'*, IP*, C', CP, VP, I', IP}

S₂ = {each other, PP*, N'*, NP*, PP, VP*, I'*, IP*}

The men is accessible through S₁. No NP is accessible through S₂, and so we can disregard this accessibility sequence, since (although it is well-formed) it does not identify any potential binder for the anaphor. In this D-Structure, John is accessible through no accessibility sequences, since it is not a sister to any of the nodes which occur in S₁ (or S₂) in this pre-movement structure. Therefore, were the licensing condition (7) to be able to apply at D-structure alone, the sentence should be grammatical -- in this structure, the men is minimally accessible to the anaphor, hence is its 'closest potential binder'. In terms of the [+B]-assignment mechanism outlined above, the feature +B could be assigned at D-structure to the anaphor each other, since it is bound by the men in the minimal CFC (the matrix clause) in which it could be bound. Since John does not c-command the anaphor at D-Structure, it is not a

potential binder.

Recall that in the S-structure representation (66) George was blocked from acting as the antecedent of the anaphor, because Susan is minimally accessible to the anaphor through all the sequences. In the D-structure (83c), crucially, the men is accessible to the anaphor through a sequence (S₁) through which John is not accessible. Therefore, in (83c), the men is minimally accessible to the anaphor, and (83b) is falsely predicted grammatical.

This can only lead to one conclusion -- that the licensing condition for anaphors cannot apply prior to NP movement, and so cannot apply to the DS representation (83c). (This point is made by Jacobsen and Neubauer 1976, in a different framework, and the same result carries over to our framework.)

3.3.7.3 Strict Cyclicity and Feature Assignment

What we now must do is determine how to derive the descriptive statement:

- 84) The licensing condition for anaphors applies after all instances of NP-movement.

If we were to suppose that the Binding Theory applies at S-structure, and not to any intermediate representations in the D-structure -- S-Structure mapping, (84) is derived. There is, however, a way to retain the proposal that anaphor licensing

does occur successive cyclically, to intermediate representations, and still derive (84).

This is to consider the rule of [+B] assignment to be subject, along with movement rules, to the Strict Cycle Condition (Chomsky (1973); see chapter 5 of this thesis for review, and discussion). Observe that in (83), the feature +B is assigned to himself (thus allowing the anaphoric relation between George and the anaphor to be licensed) on the matrix IP cycle, crucially before NP-movement raises Mary from object to subject position. This instance of NP-movement occurs strictly on the lower IP cycle. This movement would violate the Strict Cycle Condition.

Chomsky's (1973) discussion of the SCC argues for it as a constraint on movement rules. The move hypothesized here would extend the phenomena governed by this constraint to processes which do not have the form of movement rules, specifically the rule of +K-feature assignment.

Let us consider in a bit more detail what would be involved in extending the SCC to anaphoric processes.

If the SCC is a constraint on movement, it can be formulated as in (85):

(85)

The Strict Cycle Condition on Movement

No instance of movement can associate positions X and Y, dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which

is also a cyclic node.

If the rule (71) of +B assignment (part of the licensing process for anaphors) is also subject to the SCC, then the SCC can be taken in its original form (Chomsky 1976; 97):

(86) (Chomsky's (51))

No rule can apply to a domain dominated by a cyclic node A in such a way as to affect solely a proper subdomain of A dominated by a node B which is also a cyclic node.

The result of (86) is that a rule which could be applied on cycle B must be applied on that cycle, if at all. As Chomsky puts it, 'rules cannot in effect return to earlier stages of the cycle after the derivation has moved to larger, more inclusive domains.' (p. 97).

This would be violated in (83c). When +B assignment applies to assign the feature to each other, it is applying on the matrix IP cycle, since this is the only cyclic node which dominates the men. In order for (83c) to be transformed into a well-formed S-Structure, in which John has moved to the embedded subject position, the derivation would involve a rule (move-alpha) operating on the lower cycle, after a rule (+B assignment) has applied on the higher cycle. This is precisely what the SCC blocks.

So, let us hypothesize that the Binding Theory involves the rule of +B assignment, which applies optionally throughout the

derivation of S-Structure, and is subject to the SCC.

Thus we rule out (83b) as a violation of either the binding theory filter (72), or a violation of the SCC. If +B is assigned to the reciprocal by the men, on the matrix IP cycle, then the SCC precludes moving John to the subject position of the lower IP. If John is not moved, then the chain containing this NP will have no case, in violation of the case Filter.

If the feature +B is not assigned to the anaphor, the filter (72) is violated. Thus we can derive (84) as a result of considering the rule of feature assignment to be governed by the SCC. We can maintain the successive-cyclic approach to anaphor binding, and explain the ungrammaticality of (83b).

Crucially, the rule of feature assignment can interact with WH-movement. Consider (87).

(87)

D-Structure: John wonders [[Mary is [how proud of herself]]
S-Structure: John wonders [[how proud of herself] [[Mary is e]]

Herself can be assigned +B under local binding by Mary. Then, the WH-phrase how proud of herself can move to the embedded SPEC of CP position. No strict cycle violation ensues.

The mechanisms outlined above (the rule (71) and the filter (72)) seem to be necessary in any account of binding which

25. This use of the SCC in a sense returns to an earlier conception of anaphora, as a structure-dependent rule.

asserts that the binding principles apply successive cyclically. The rule of feature assignment is necessary, to record the proximity of an anaphor to a locally c-commanding antecedent. The filter is necessary to impose restrictions on the derived representations. If the rule of feature assignment was not adopted, successive cyclic anaphor binding would have to be global. Condition A, applying at S-structure, would have to be able to look back at every intermediate representation between DS and SS, ruling out the sentence only if no intermediate representation exists in which the anaphor is bound in its governing category.

However, there is one problem with this analysis, giving empirical evidence against this approach to anaphor connectivity.

Consider the following example:

88) * Mary thinks that John wonders which pictures of herself are
on the table

This has the following structure:

89) * [Mary [[thinks[[that[John [wonders
IP 1' VP CP C' IP* VP*
[[which[pictures [of herself]]] [CP* NP N' PP i C'*
[t are on the table]]]]]]]]]
IP' i

On the accessibility sequence framework, this sentence is ruled out straightforwardly (assuming that S-Structure is the only representation relevant to the licensing condition (7)).

There are two accessibility sequences through which Mary is accessible to the anaphor, S1, and S3.

S1 = {herself, PP, N', NP, CP*, VP*, I'*, IP*, C', CP, VP, I', IP}

S2 = {herself, PP, N', NP, CP*, VP*, I'*, IP*}

S3 = {herself, PP, N', NP, t, IP', C'*, CP*, VP*, I'*, IP*, C', CP, VP, I', IP}

S4 = {herself, PP, N', NP, t, IP', C'*, CP*, VP*, I'*, IP*}

John is accessible to the anaphor through all four sequences²⁶, crucially through S₂ and S₄. These are proper subsequences of the sequences through which Mary is accessible. Thus, John is minimally accessible to the anaphor, and Mary is not.

There is a derivation of the S-Structure (89) which is not ruled out by any independently motivated constraints, and in which the anaphor himself could acquire the feature +B within the framework of successive cyclic binding. That is, this system of feature assignment will incorrectly permit the sentence (88) to be grammatical. The derivation goes as follows.

First, from the D-Structure (90), derive (89) as an intermediate representation through applications of WH-movement. Nothing is violated.

26. There are actually two others, if we add the embedded C'* and CP* to S2 and S4: S5 = {herself, PP, N', NP, CP*, VP*, I'*, IP*, C', CP}, and S6 = {herself, PP, N', NP, t, IP', C'*, CP*, VP*, I'*, IP*, C', CP}. These need not concern us here.

90) [Mary [[thinks[[that[John [wonders
 IP I' VP CP C' IP* VP*
 [[[which[pictures [of herself]]] are
 CP* C'* IP' NP N' PP
 on the table]]]]]]]]]

89) [Mary [[thinks[[that[John [wonders
 IP I' VP CP C' IP* VP*
 [[which[pictures [of herself]]] [[t are on
 CP* NP N' PP i C'* IP' i
 the table]]]]]]]]]

Then, move the WH phrase to the next highest SPEC of CP position, that in the complement of think. This yields (91).

91) [Mary [[thinks[[which[pictures [of herself]]]
 IP I' VP CP NP N' PP i
 [that[John [wonders [t [[t are on
 C' IP* VP* CP* i C'* IP' i
 the table]]]]]]]]]

At this point, the anaphor herself is c-commanded within its GC by Mary, and so the feature +B may be assigned to the anaphor. This gives (92).

92) [Mary [[thinks[[which[pictures [of herself]]]
 IP I' VP CP NP N' PP +B i
 [that[John [wonders [t [[t are
 C' IP* VP* CP* i C'* IP' i
 on the table]]]]]]]]]

Let us assume that both CP and IP are cyclic. The application of the feature assignment rule (71) takes place on the matrix IP cycle. To now lower the WH phrase from the SPEC position it occupies in (92) directly to the lower one cannot occur, since this movement would take place on the embedded CP cycle, and thus is on a lower cycle than the previous rule. Such a movement would violate the SCC.

However, nothing blocks movement of the WH phrase further upward, to the matrix SPEC position. This gives (93):

93) [[whichL pictures [of herself]]] [Mary
CP NP N' PP +B i IP
[[thinks [[that [John [wonders [t [
I' VP CP C' IP* VP* CP* i C'*
[t are on the table]]]]]]]]]
IP' i

The derivation is now on the matrix CP cycle, the highest cycle. The final step in the derivation lowers the WH phrase to the lowest SPEC of CP, giving us the S-Structure (94), which is identical to (89) except that the anaphor has now been assigned the feature +B.

(94) [Mary I L thinksI [thatI John [wonders
 IP I' VP CP C' IP* VP*
 [[whichI pictures [of herself]]] [
 CP* NP N' PP [+B] i C'*
 [t are on the table]]]]]]]]]
 IP' i

Therefore, the sentence will pass the filter (72). This derivation, if it is not ruled out by anything else, falsifies the successive binding hypothesis (within the set of assumptions made here, of course).

Does any other constraint on derivations or representations rule out (89)-(94)? Apparently not. In (89,90), the WH phrase has been moved from a +WH specifier position to a -WH specifier position. This is precisely what Chomsky (1976) suggests occurs

on the derivation of (95) :

(95) ? what did you say John wonders who bought

The ECP is satisfied in (89-94), in that every trace in the derived S-Structure representation (94) is in a configuration of proper government. Finally, the selection properties of the verbs are met; in (94) wonder has a WH phrase in the specifier of its complement, and the specifier of the complement of think is not a WH phrase. Therefore, it appears, the derivation violates nothing. We conclude that what is amiss in this derivation is the assignment of the feature +B, which permits the anaphor of the derived SS to satisfy the binding theory. Without this mechanism (88) is excluded as a binding theory violation. Therefore, we conclude that the Binding Theory does not work in this fashion, but instead applies only to SS representations, as we have proposed earlier. This requires something like the chain accessibility sequence framework.

However, it might be supposed that the derivation of the ungrammatical S-Structure (89) proceeds exactly as just suggested, except that there is no rule of +B assignment. Such

27. The derivation involves, basically, leaving who in situ at S-Structure, moving it to the SPEC position in the SS-LF mapping. What is extracted through the lower SPEC position, successive cyclically, up to the matrix SPEC position. Thus, the derivation involves moving what from a +WH SPEC position (in the complement of wonder), to a -WH SPEC (in the complement of say).

3.4 Binding Domains and a Predicate/Non-Predicate Asymmetry

In this section I will examine two phenomena which bear on the questions of how to define binding domains, and what the relation is between binding domains and the theta-criterion. A binding domain is that part of a phrase marker within which an anaphor's antecedent must be located, and within which a pronominal's c-commanding antecedent cannot be located. In the theory of binding advanced in Chomsky (1985), the binding domain (or governing category (GC)) is the minimal CFC within which there is a BT-compatible indexing for the NP. Within our framework, the binding domain is what I have termed a chain accessibility sequence; unlike the GC of Chomsky (1981; 1985), such a binding domain is not a subtree, but rather a set of nodes in the tree.

The phenomena of relevance here have to do with the multiple binding effect, shown in such examples as (104), and an apparent suspension of this effect when the container of the anaphor (X, in the schema (50)) is a predicate, eg. an AP or VP. An illustrative case is (96) vs. (97):

96) John wonders how proud of herself Mary became

97) * John wonders how proud of himself Mary became

After presenting the data in more detail, I shall advance two

100) John wonders [[which pictures of himself][Mary saw e]]

| |

101) John thinks that Mary likes that photograph of herself

| |

102)*John thinks that Mary likes that photograph of himself

| |

103) John thinks that this photograph of himself, Mary likes e

| |

As has been demonstrated in earlier sections, this effect (seen in (99-100) and (103)) occurs as a result of there being some non-singleton chain C, one member of which structurally dominates the anaphor. The chain in (99)-(100) is (which pictures of X-self,e). In the framework proposed here, the existence (at SS) of a n-member chain gives rise to (at least) n chain accessibility sequences, and thus the possibility of more than n possible antecedents for the anaphor. Both John and Mary are minimally accessible to the anaphor, through different sequences. We might expect that this effect will hold true for all cases, but this is in fact not true.

Sentences (101) and (102) exhibit the typical Specified Subject Condition effect; the anaphor in the embedded clause is required (by the licensing condition) to find its antecedent within the domain of the lowest subject, which does not occur in (102), hence the ungrammaticality of the example. Sentences

(100) and (103) are examples of connectivity where the anaphor's antecedent is in a position in which it would not qualify as an antecedent for the anaphor had WH movement or topicalization not taken place. Thus we see, in this contrast, the fact that WH movement (or, more generally, movement to an A' position) of a phrase containing an anaphor proliferates the possibilities for binding beyond what would exist without movement. Descriptively, the movement adds another binding domain to the structure, adding John to the set of possible grammatical antecedents.

Examples like (99) and (100) were noted originally by Jacobsen and Neubauer (1976). The same basic effect was noted by Langendoen and Batistella (1981) and Weisler (1982), for examples in which the WH phrase containing the anaphor was moved out of the c-command domain of both possible antecedents, as in (104):

(104) which pictures of himself/herself did John say Mary likes?

In Barss (1984) I noted that this effect occurs in all constructions exhibiting the basic type of connectivity seen in (1), including Topicalization, restrictive relatives, clefts, and pseudo-clefts.

The second phenomenon I shall address here is the fact that the multiple binding domain effect does not occur when the dislocated constituent is a predicate. When a predicate containing an anaphor is dislocated, as in (105)-(108), the only

possible antecedent for the anaphor is an NP which c-commands
 28
 the anaphor prior to movement .

- 105) John wonders [how proud of herself][Mary is e]
 106) * John wonders [how proud of himself] [Mary is e]
 107) John thinks that [proud of herself],[Mary certainly is e]
 108) * John thinks that [proud of himself],[Mary certainly is e]

This latter fact will require some modification of the licensing
 condition for anaphors (7).

3.4.2 Hebrew and French

29
 This same asymmetry occurs in Hebrew and French .

30
Hebrew

- 109) John lo-yode's eize tmunot sel acmo Mary maxra
 John not-know which picture of himself Mary sold

28. This basic fact was noted independently by the author, in December 1984, and by Cinque (1984). Aoun and Clark (1984) refer to Cinque's article, citing his example (69), which they repeat as (14b). They claim, incorrectly, that Cinque's observation that Bill (the lower subject) is the only possible antecedent for the reflexive is false. (14b) [how proud of himself] does [John think [Bill was e].
 i S1 S2 i

29. I am grateful to Ur Schlonsky for the Hebrew data, Viviane Deprez for the French data, and to both for discussion.

30. One Hebrew speaker I consulted found both (109) and (110) ungrammatical, and indicated that he found the internal structure of the WH-phrase unacceptable. He did, however, accept (113) as grammatical while rejecting (114). This is the contrast crucial for this discussion.

- 110) John lo-yode'a eize tmunot^v sel acma Mary maxra
John not-know which picture of herself Mary sold
- 111) John xosev se-et-ha-tmunot^v sel acmo Mary ohevet
John thinks that-ACC-DEF-pictures of himself Mary likes
- 112) John xosev se-et-ha-tmunot^v sel acma Mary ohevet
John thinks that-ACC-DEF-pictures of herself Mary likes
- 113) John taha kama ge'a be-acma hi Mary
John wonder-past how proud-fem in herself is Mary
'John wondered how proud of herself Mary is'
- 114) *John taha kama ge'e be-acmo hi Mary
John wonder-past how proud-masc in himself is Mary
'John wondered how proud of himself Mary is'
- 115) John xosev se-ge'a be-acma hi Mary
John thinks that-proud of herself Mary is
- 116) *John xosev se-ge'e be-acmo hi Mary
John thinks that-proud of himself Mary is

French

- 117) Pierre a rendu [Marie fiere d'elle meme]
SC
Pierre made Marie proud of herself
- 118) Pierre a rendu [Marie fiere de lui]
SC
Pierre made Marie proud of him
- 119) *Pierre a rendu [Marie fiere de lui-meme]
SC

* Pierre made Marie proud of himself

120) ??Fiere de son fils Pierre l'a rendu Marie
proud of her sons Pierre made Marie

121) ??Fiere d'elle meme Pierre l'a rendu Marie
Proud of herself Pierre made Marie

122) * Fier de lui-meme Pierre l'a rendu Marie
* Proud of himself, Pierre made Marie

The construction illustrated in (120-122) is marginal, as indicated by the ?? judgement on (120). But, crucially, the sentence in (121) is no worse than (120), while (122) is completely unacceptable. This indicates that the distinction between (121) and (122) is essentially that seen in the corresponding English and Hebrew examples.

Further, this seems to be a distinction between predicates and non-predicates, rather than a distinction between arguments and non-arguments, given the contrast in (123) and (124) (examples (65a,b) of Jacobsen and Neubauer (1976)):

123) John thought that Mary would faint when a picture of
himself was published

124) John thought that when a picture of himself was published,
Mary would faint.

(123) is grammatical only when the adjunct clause is interpreted as modifying the matrix verb. Presuming that adjuncts are dominated by the IP which they modify, this requires the following representation for the grammatical interpretation (123):

123') [John [thought [that Mary would faint]] [when a picture of himself was published]]

Here, John c-commands the reflexive, with no closer possible binder, hence the sentence is grammatical, as expected. However, (123) is not grammatical when the adjunct clause is interpreted as modifying the complement clause, as in (123''):

123'') * [John [thought [that [Mary would faint [when a picture of himself was published]]]]]

Here, Mary c-commands the anaphor (within the domain of John), and hence must act as its antecedent, blocking John from being the antecedent.

What is interesting is that when the adjunct clause is preposed, as in (124), the sentence is grammatical when John binds the anaphor and the adjunct clause modifies the complement. Given that the complementizer that precedes the adjunct in (124), it must be the case that the adjunct is within the complement CP, presumably adjoined to the IP, as in (125):

125) [John [thought [that [^{IP} [when a picture of himself ^{IP} was published], [Mary would faint]]].

Since Mary no longer c-commands the adjunct, and thus no longer c-commands the anaphor within it, John is now available as a local antecedent for the reflexive.

In our terms, there is now a chain accessibility sequence through which John is minimally accessible to the anaphor.

The contrast in grammaticality between (124) and (123'), and the configurational difference between them, is like the contrast between (102) and (103), another example of the multiple binding-domain effect.

It appears, then, the proper way to descriptively characterize the asymmetry is to say that the multiple binding domain effect occurs when a constituent that is not a predicate is moved. Note that this will require a formal analysis of clausal adjuncts, as in (123-124), which classes them as non-predicates.

If we were to adopt the position that syntactic connectivity is to be handled by literal Reconstruction (a position which was shown in Chapter 2 to be untenable), then it would have to be the case that the operation of Reconstruction --lowering a dislocated constituent in the SS-LF mapping, into its D-Structure position-- is optional for arguments and adjuncts, and entirely obligatory for predicates. If we hold to the position which I assume is true, namely that the binding condition governing anaphors applies to S-structure representations, then it will still have to be the case that a distinction between predicates and non-predicates is made in calculating the domain within which an anaphor's antecedent must be located. One approach to this is given in the next section, where I attempt to reduce this predicate/non-predicate asymmetry to the independently motivated notion of 'Complete Functional

The contrast in grammaticality between (124) and (123'), and the configurational difference between them, is like the contrast between (102) and (103), another example of the multiple binding-domain effect.

It appears, then, that the proper way to descriptively characterize the asymmetry is to say that the multiple binding domain effect occurs when a constituent that is not a predicate is moved. Note that this will require a formal analysis of clausal adjuncts, as in (123-124), which classes them as non-predicates.

If we were to adopt the position that syntactic connectivity is to be handled by literal Reconstruction (a position which was shown in Chapter 2 to be untenable), then it would have to be the case that the operation of Reconstruction --lowering a dislocated constituent in the SS-LF mapping, into its D-Structure position-- is optional for arguments and adjuncts, and entirely obligatory for predicates. If we hold to the position which I assume is true, namely that the binding condition governing anaphors applies to S-structure representations, then it will still have to be the case that a distinction between predicates and non-predicates is made in calculating the domain within which an anaphor's antecedent must be located. One approach to this is given in the next section, where I attempt to reduce this predicate/non-predicate asymmetry to the independently motivated notion of 'Complete functional

Complex' introduced by Chomsky (1985).

3.5 The Definition of Local Binding Domains: Chomsky (1981)
and (1985)

In this section we briefly review Chomsky's (1985) formulation of the Binding Theory, in order to introduce the conceptual and empirical support for this formulation.

Condition A of the binding theory, as formulated by Chomsky in (1981), requires that an anaphor be bound in a local domain. The domain in question is the governing category, defined as in (127).

126) Condition A: an anaphor is bound in its governing category

127) B is the governing category (GC) for X iff B is the minimal maximal projection containing X, a governor of X, and a SUBJECT accessible to X.

Y is a subject accessible to X iff:

- i) Y is a Subject or AGR which c-commands X
- ii) assignment of the index of Y to X would not violate the i/i filter:

* [...Z ...]
i i

128) The Root Clause Condition:

The matrix clause is the governing category for a governed element which has no smaller governing category.

Chomsky's (1985) modification of the Binding Theory is repeated in (129):

(129) Principles A, B, C (Chomsky 1985):

(a.) for α an NP, and B a local domain, an indexing I is binding theory compatible with (α, B) iff:

- (A) α is an anaphor and is bound in B under I
- (B) α is a pronominal and is free in B under I
- (C) α is an r-expression and is free in B under I

(b.) Licensing Condition

for a category α governed by G in the expression E with indexing I , α is licensed only if:

for some B such that (i) or (ii), I is BT-compatible with (α, B) :

- (i) α is an r-expression and (1) if α heads its chain,
(2) otherwise:

(1) $B = E$

(2) B is the domain of the head of the chain of α

- (ii) α is an anaphor or pronominal and B is the least Complete Functional Complex (CFC) containing G for which there is an indexing J that is BT-compatible with (α, B) .

This reformulation permits the elimination of several stipulatory parts of the (1981) theory. First, the root clause condition (128), which was added to the theory to rule out sentences like (130), can be abandoned:

(130) *All pictures of PRO are on sale

Without (128), the binding conditions (126) cannot rule the sentence out, since, although the anaphoric pronominal PRO is governed in (130), it has no governing category, since there is no SUBJECT accessible to it. (128) stipulates that the matrix clause becomes the GC for PRO in (130) by default, and so the sentence violates either Condition A or B of the Binding

Theory . The (1985) theory, on the other hand, rules out (130) with no additional stipulation. PRO is governed here, and is thus subject to the licensing condition (129b). There is no binding theory compatible indexing for PRO as a [anaphoric element] in (130), and so PRO is not licensed as an anaphor. Although Condition B is satisfied in (130), Condition A is violated.

Secondly, the notion of SUBJECT is eliminated, a desirable result, since it is not clear (as Chomsky and others have noted) that NPs occurring as subjects form a natural class in any other respects with the agreement morphology AGR.

And, finally, the (1985) theory permits a straightforward treatment of some cases problematic for the 1981 theory, like (131)-(132). These examples are both grammatical, showing that there are positions in which a bound anaphor and a bound pronoun can each grammatically occur.

131) The men_i like [their_{NP} parents_i]

132) The men_i like [each other_{NP} 's parents_i]

The basic problem for the 1981 theory is that the object NP must be the GC for the pronoun, but not for the anaphor. Since

31. PRO will violate Condition A if it is locally free, as it is required to be by Condition B; and it will violate Condition B if it is locally bound, as Condition A requires.

GC is defined in the (1981) theory (see (127)) without regard to the type of NP for which it is a GC, this requires abandoning a uniform definition of GC for pronouns and anaphors, thus complicating the theory (see Huang 1983). In the 1985 theory, the object NP is a Binding Domain for the pronoun in (131), since (i) the NP is a CFC, and (ii) there is an indexing in which the pronoun is free in that domain. Since there is no indexing under which the anaphor in Specifier position in (132) is bound within the NP, the NP is not a GC for the anaphor, and so the anaphor may seek its antecedent in the next highest CFC, namely the matrix clause. The NP acts as a GC for a pronominal but not for an anaphor, derived from the different requirements anaphors and pronominals have in needing or not needing an antecedent.

The distinction between the 1981 Binding Theory and the 1985 Binding Theory which is of importance to the predicate/non-predicate asymmetry described earlier in this section is the introduction of the notion of Complete Functional Complex.

3.5.1 CFCs and The Asymmetry

To summarize the data introduced in this section so far:

(133)

- a. When an anaphor is embedded within a moved argument, its antecedent can be determined with respect to the S-Structure position of the anaphor, its D-Structure position (as recorded by trace theory, or by the constructed chain), or any intermediate position (c.f. (104)).

- b. When the moved constituent containing the anaphor is a predicate, there is no such multiple choice of antecedents; the antecedent must c-command the anaphor within the anaphor's GC at D-Structure, i.e. be such that it would grammatically bind the anaphor had movement not occurred.
- c. Moved Clausal Adjuncts: the effect remains as for moved arguments, as in (a). The asymmetry is thus between predicates and non-predicates.

We have hypothesized that the distinction between these examples has to do with the status of the anaphor's containing phrase as a predicate. Two questions must now be confronted. First, is this a correct generalization, or is the difference due to the syntactic category of the dominating phrase? And second, how is the distinction to be formally captured?

The first question can be answered by examining the multiple binding domain effect in sentences where the dislocated element

32

is a predicate nominal .

Consider the following contrast:

- 134) a. John thinks that a critic of herself, Mary became
- b. Mary thinks that critics of each other, John and Bill became
- 135) a. *John thinks that a critic of himself, Mary became
- b. * John and Bill think that critics of each other, Mary and Sue became (with coreference between John and Bill and each other)

32. I am grateful to Adriana Belletti and Jim Higginbotham for posing this question, and providing the test case.

This seems to resolve the question in the direction assumed earlier. The dislocated phrase in (134)-(135) is categorically indistinct from that in (100)-(101), being NP in both cases. What differs in the two sets of examples is the function of the NP. In the case where the NP is a D-structure argument, the multiple binding effect obtains; when the NP functions as a predicate, as in (134)-(135), the anaphor must have as its antecedent an NP which is minimally accessible to it at D-Structure.

3.5.2 Theta-Compatibility

We earlier added the condition (4iv) to the definition of accessibility sequence, for the reasons given in section (3.2) (essentially, to permit the ambiguity of 'John told Bill about himself').

- (4iv) a [the final member of a chain accessibility
 ⁿ
 sequence] must be the root node of a Complete
 Functional Complex.

This condition as stated cannot block John from grammatically anteceding himself in (97) or (106) (repeated here with bracketing and O-grids added) since there is an accessibility sequence through which John is minimally accessible to the anaphor, and whose final member is the root node of a CFC³³.

33. I here use an asterisk to distinguish different occurrences of the same labelled node; eg. IP* is the embedded IP node.

This sequence (S_1) is well-formed by the definition (4).

(=106)

136) *L JohnL L wondersL L L how L proud
 IP 1' VP CP C' AP<1,2*> A'<1,2*>

[of himself]]]] [Mary [is [ec]]]]]]]]]
 PP IP* \langle 1*,2* \rangle 1'* \langle 1,2* \rangle

sequence $S_1 = \{\text{himself}, \text{PP}, \text{A}', \text{AP}, \text{C}', \text{CP}, \text{VP}, \text{I}', \text{IP}\}$

The sequence through which Mary is accessible to the anaphor is $S_2 = \{\text{himself}, \text{PP}, \text{A}', \text{AP}, \text{e}, \text{I}'^*, \text{IP}^*\}$. This is not a proper subsequence of S_1 , and so John is minimally accessible. Thus

the incorporation of the notion of CFC in (4iv) is insufficient to capture the predicate/non-predicate difference.

It is however possible to alter the definition of accessibility sequence in a different way, one which will subsume the effects of (4iv), and account for this predicate/non-predicate asymmetry at the same time.

Observe that the accessibility sequence S_2 through which Mary is accessible to the anaphor in (136) contains the node AP \langle 1,2* \rangle , a node with a thematic grid which is not fully

 These have no formal status; they should not be confused with the asterisks to the right of theta-grid positions, which do have formal status, indicating that this position is discharged. The latter notation is adopted from Higginbotham (1985).

saturated. In a declarative sentence like (137), the AP $\langle 1, 2^* \rangle$ is not a complete functional complex, because not '...all the grammatical functions compatible with the head are realized in the domain...' (Chomsky (1985)).

137) [John is [proud of him]
 IP $\langle 1^*, 2^* \rangle$ AP $\langle 1, 2^* \rangle$

Thus the binding theory of Chomsky (1985) takes IP, not AP, to be the minimal binding domain for the pronoun in (137), forcing disjoint reference between John and him. (In our terms, $\{ \text{him}, \text{PP}, \text{A}', \text{AP} \}$ is not a well-formed CAS since AP is not the root node of a CFC).

Suppose we adjust the definition of CFC somewhat, and put it in terms of the saturation of thematic grids.

(138)

Definition: X is a Complete Functional Complex if the thematic grid of X contains no undischarged positions (equivalently, X is saturated).

By this definition, AP in (137) is not a CFC, since its external theta position is not discharged. IP is the only CFC containing the pronoun, and thus the pronoun must not have John as an antecedent. The correct results carry over with the modified definition.

3.6 Theta-Compatibility and Condition A

Suppose we formulate the definition of chain accessibility sequence (part of the well-formedness condition on the distribution of governed anaphors and pronominals), in the following way.

4) Chain Accessibility Sequence (definition)

S is a well-formed chain accessibility sequence for an NP A only if:

- i) A is a₁
- ii) a₂ is a projection of the governor of A
- iii) for every pair (a_i, a_{i+1}), either (1) or (2):
 - 1) a_{i+1} immediately dominates a_i
 - 2) (a_i, a_{i+1}) is a link of a chain
- iv') for every a_i in S, if a_i bears a thematic grid G, either G is saturated or there is a member a_j of S which bears a saturated grid G' identical to G in all other respects.

(Compare with (4) above). Clauses (4i-iii) carry over identically to the new definition. (4iv') is added, replacing (4iv). (4iv) in essence forces a certain compatibility between chain accessibility sequences and the theta-criterion; let us

then informally refer to (4iv') as the theta-compatibility requirement.

In order to see the effects of (4iv') (which replaces (4iv)),
 34
 consider again the contrast between the following examples .

139) [[John [[wonders [[which
 IP I' VP CP NP<1*>
 [pictures [of himself]] [Mary[
 N'<1*> PP IP* I'*
 [saw [ec]]]]]]]]
 VP*

140) [[*John[[wonders [[how
 IP I' VP CP AP<1,2*>
 [[proud [of himself]]] [Mary
 A'<1,2*> PP IP*<1*,2*>
 [is [ec]]]]]]]]
 I'*

Consider first (139). The analysis of anaphor connectivity developed earlier has the result that there are two distinct accessibility sequences, through which the anaphor can seek an antecedent. The sequence {himself, PP, N', NP, CP, VP, I', IP} is that which permits the matrix subject John to be accessed in (i); another sequence, {himself, PP, N', NP, ec, VP*, I'*, IP*} permits the lower subject to be accessed. Crucially, both of these sequences meet the revised definition above. This captures the ambiguity of sentences of the form (139).

 34. In these, and other examples, I suppress mention of the internal 0-grid position in nouns, which is saturated by the determiner (see Higginbotham 1983), (1985)). I also suppress annotation of theta grids on certain projections, where they aren't germane to the matter under discussion.

As we have just seen, sentences in which the moved phrase containing the anaphor are predicative are not similarly ambiguous. In a sentence like (140), the anaphor must seek as its antecedent the lower subject.

A principal difference between predicates and arguments is that the former are not fully saturated. The effect of (4iv') is that when the moved phrase is a predicate, the accessibility sequence for the contained anaphor must have as a member the node at which the predicate's theta-grid becomes fully saturated. The sequence {himself, PP, A', AP, CP, VP, I', IP} does not meet this requirement, since the sequence has A' $\langle 1, 2^* \rangle$ and AP $\langle 1, 2^* \rangle$ as members, but not IP* $\langle 1^*, 2^* \rangle$. Thus the only well-formed accessibility sequence for the anaphor in (140) is {himself, PP, A', AP, ec, I', IP}. Mary is minimally accessible through this sequence, but John is not; and the only accessibility sequence through which John is minimally accessible is ill-formed, by (4iv'). Thus the suppression of the multiple-domain effect is captured.

This requirement (4iv') subsumes Chomsky's requirement that only Complete Functional Complexes (CFCs) may be binding domains.. In a sentence like (137), the VP cannot be taken as the binding domain (in the framework of Chomsky (1985)) of the pronoun, because it is not a CFC. In the terms of the framework developed here, the accessibility sequence must be {him, VP, I',

IP}, since the sequence {him, VP} contains a node with an unsaturated theta-grid.

3.6.0.1 CFCs and Saturated Domains

The question now arises as to whether there are any differences between Chomsky's distinction between a domain in which the GFs compatible with the head of a domain are all represented (a CFC), and domains in which they aren't (a non-CFC), and the distinction made in (4iv') between thematically saturated and thematically unsaturated domains.

There are two potential distinctions, which will be treated here.

First, a chain accessibility sequence is a set of nodes -- a path -- rather than a subtree or sub-phrase marker, as a CFC is. However, in the sense that the proper distinctions can be made in an explicit way within both frameworks, no particular issues arise here.

Secondly, Chomsky's definition places emphasis on the
35
'grammatical functions compatible with the head of a domain', while the theta-compatibility requirement (4iv') places emphasis on the presence or absence of undischarged thematic grid

35. Taking Grammatical Functions to be Subject, Object, Indirect object.

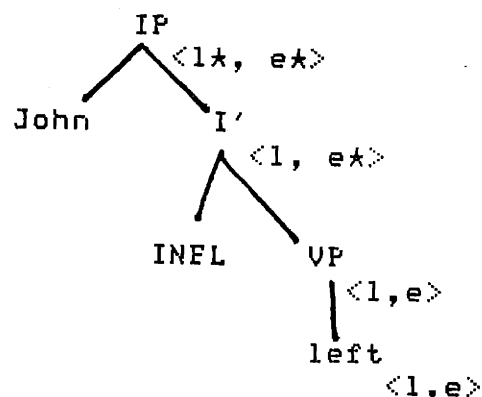
positions. To a large extent, these have the same coverage. In a simple sentence like (141) for example, the GFs compatible with the verb include its object, and its subject.

141) [John [IP<1*, 2*, e*> I'<1, 2*, e*>
 [VP<1, 2*, e> nominated <1, 2, e> himself]

The verb's thematic grid has three positions, that which is discharged by the object, that which is discharged by the subject, and the event position, which is discharged by INFL. For such sentences, only IP is a CEC, and only IP is a node with a fully saturated theta grid.

There is at least one structure in which the two notions of CEC and saturated domain diverge. Recall that in Higginbotham's framework there are thematic grid positions which are discharged by some means other than being assigned to structurally overt constituents. A central example of this is the event position, which in the typical case is discharged by INFL (specifically, I assume, by being bound by INFL). This is illustrated below:

(142)



For Chomsky, it is the presence of John in IP which makes IP the CFC. John is in subject position, and is compatible with the two heads present in IP, INFL and the verb. The subject functions as the specifier of IP, and as the logical subject of the verb. On the other hand, both NP and INFL contribute to the saturation of IP's theta grid, in that both discharge (in different ways) one of the positions in the grid.

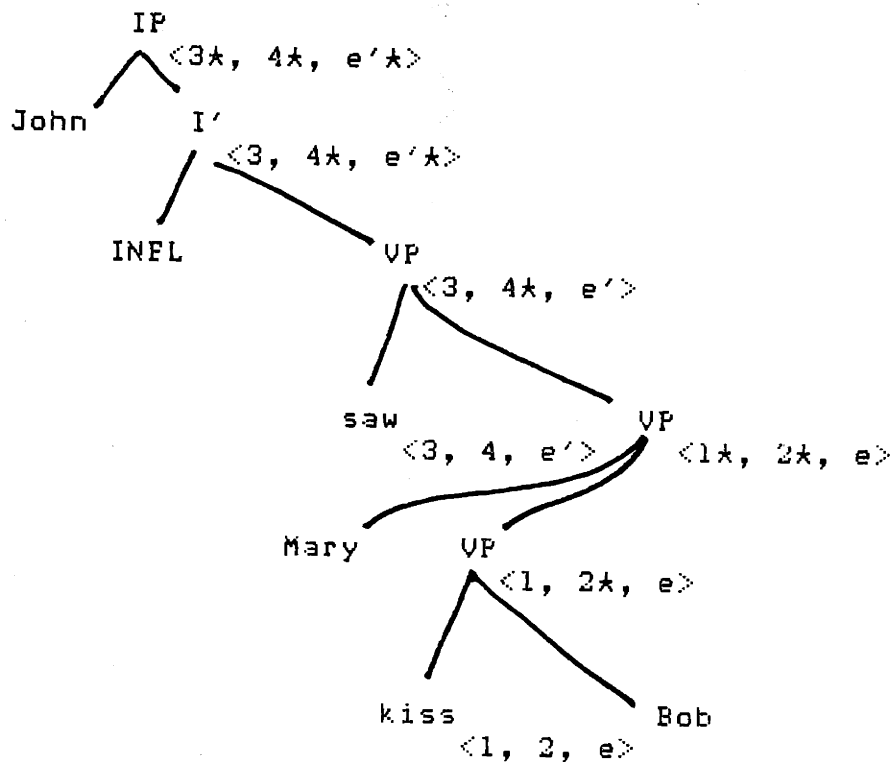
Now suppose that we were to leave the event position of the verb open. This would make no difference for IP's status as a CFC, since the event position has nothing to do with Grammatical Functions. On the other hand, were this position to be open, IP would no longer have a saturated theta grid, shifting its status within the accessibility sequence framework.

In a typical declarative clause, this state of affairs will not arise. However, this is precisely what obtains in non-inflected perception verb complements, in the analysis of Higginbotham (1983). This analysis proposes that the perception verb semantically selects for an event, taking the open expression [John leave] as its argument. The lack of inflection in such complements now follows, since it is only by omitting the INFL node that the event position may remain open (assuming that INFL always discharges the event position of its sister, when INFL is present).

The structure Higginbotham assumes is something like the

following:

(143)



I assume that the complement is not of category IP, since there is no INFL to project such a node. What is of relevance here is the thematic grids of each part of the structure.

The saturation of positions 2 (the logical object of kiss), 1 (the subject of kiss), e' (the event position of see), and 3 (subject of see), is straightforward. It is the saturation of the positions e (event position of kiss), and 4 (object of saw) which deserves some attention.

Higginbotham's analysis of non-inflected PVCs has two fundamental assertions. First, that the complement of the perception verb does not act as an embedded clause at LF, but

rather a definite description of an event, and that this semantic unit has scope over the verb see at LF, obligatorily. These two results are achieved by existential closure of the event position of the verb in the complement. This causes the sequence [John leave] to be treated as a quantified phrase, which obviously does not have propositional interpretation. The quantifier will be assigned scope via QR, and so the complement will have matrix level scope. The SS in (144) will have the interpretation (145):

(144) I saw John leave

(145) [$\exists e$: e is an event & leave(John,e)] I saw e

A variety of semantic facts about such sentences are derived from this basic analysis (see Higginbotham (1983b)).

A syntactic question which is left open in this analysis is basically whether the constituent [John leave] is thematically saturated at S-Structure.

The answer to this will determine the answer to the question I posed earlier: is a CFC to be identified with a syntactic node with a fully saturated theta grid?

I shall assume, following Deprez (1985), that the sequence
36
 [John leave] is a projection of the verb leave, a 'small

 36. Deprez further suggests that an inflected PVC, as in 'I saw John leaving', is sentential, being a projection of INFL. A number of distinctions between null PVCs and inflected PVCs in

clause' in the sense of Stowell (1981) and others. It would seem that if the existential closure of the event position is achieved only at LF, such structures will violate the theta-criterion and the projection principle; a position is open at S-Structure but closed at LF. If we suppose existential closure of the event position at all levels of representation (thus fulfilling the projection principle and the theta-criterion), then a NI perceptual report will be a thematically saturated domain, and the notion CFC can be equated with that of a domain whose root node is thematically saturated. I therefore conclude that it is likely that the formulation of the binding conditions as presented in (4) can be

 English and Romance follow from this analysis; see Deprez (1985) for details.

One of Higginbotham's conclusions is that the impossibility of passive movement from a non-inflected PVC follows from its status as a quantifier; at LF, (i) transformed to (ii) by QR:

i) * John was seen [t leave]

ii) [t leave] John was seen

Higginbotham analyses this as a binding theory violation, since the anaphoric trace is not c-commanded by its antecedent. Deprez notes a problem in this analysis, since (ii) conforms to the basic structure of connectivity constructions, and ought to be grammatical as far as the revised binding theory goes. Deprez develops an analysis in which (i) is barred as an ECP violation.

Connectivity holds for NP-trace, as the grammaticality of (iii) shows:

iii) [how likely [t to win]] is John ?

|
|.....|

See Chapter (4) for discussion.

retained, adopting (4iv').

If it is the case that the correct account of unsupported clauses involves considering them to have an undischarged event position at SS, then we must slightly alter the theta-compatibility requirement (4iv') (which I repeat below) to (4iv'') in the following way:

- 4iv') for every a_i in S, if a_i bears a thematic grid G, either G is saturated or there is a member of a_j S which bears a saturated grid G' identical to G in all other respects.
- (4iv'') for every a_i in S, if a_i bears a thematic grid G, either G is saturated in every position which is to be saturated by thematic assignment or there is a member a_j of S which bears a saturated grid G' identical to G in all other respects.

This distinguishes domains which are saturated except for having an open event position, and domains which are open with respect to theta roles which will ultimately be assigned to syntactic constituents, as the AP in (140) is. Only the latter will fail to be thematically complete in the sense important for binding domains.

To conclude this subsection, I have proposed that Chomsky's CFC constraint on binding domains can be translated in a natural way into a requirement that a node in an accessibility sequence be fully saturated, or have a co-member which is fully saturated. This permits the necessary distinctions to be made

in such contrasts as (99)-(100) vs. (96)-(97), by looking at the properties of the nodes in the sequence themselves, with no reference to the internal structure of the syntactic material that they dominate in the full phrase marker.

The particular formulation I have offered subsumes the reference to Complete Functional Complexes made by (4iv) of the binding theory, as well as capturing the predicate/non-predicate asymmetry identified earlier.

3.6.1 An Alternative Proposal

There is another possibility, one which makes the theta-compatibility requirement (4iv') unnecessary (although it would require re-adopting (4iv), which was subsumed under (4iv')).

I have assumed that the structure of (146) is as in (147).

146) Dan is proud of himself

147) [Dan is [[proud
 IP<1*, 2*> <1, 2*>AP A'<1,2*> <1 , 2>
 [of himself]]]]]
 PP

However, suppose the structure is actually as in (148).

148)
 [Dan is [t [[proud
 IP<1*, 2*> AP<1*, 2*> AP<1, 2*> A' <1 , 2>

[of himself]]]]]
PP

This would give such sentences as (146) the basic structure of raising sentences, where Dan appears inside the projection of proud at D-Structure, raising to the subject position in order to get Case. Such a proposal is made for predicative sentences (those with the linear form [NP be AP]), by Rizzi (1982). (Another possibility is that t and the AP are sisters, dominated by VP; we shall consider this (suggested by Couquaux (1981)), as well as the structure in (148), below.)

For all such cases, on the hypothesis that something like (148) is the correct S-Structure, the overt anaphor occurring as the complement of the adjective will be bound by the trace. The highest instance of AP will be thematically saturated, and will have a possible binder for the complement anaphor.

Now consider one of the examples used to hypothesize a predicate/non-predicate distinction, the observation which ultimately led to the theta-compatibility requirement (4iv'). Such an example as (140) will have this structure now:

149) [Parker[wonders [[how [t
IP VP CP AP<1*,2*> AP<1*, 2*>
[[proud [of himself]]]] [AP<1, 2*> A' <1, 2> PP IP<1*, 2*>
Dan is [e]]]

I earlier suggested that the accessibility sequence which would permit Parker to be minimally accessible to the anaphor was ill-formed, in violation of the proposed (4iv'). On the

hypothesized structure (149), the sequence will be {himself, PP, A', AP, AP, CP, VP, IP}, and will actually
 $\langle 1, 2 \rangle \quad \langle 1, 2^* \rangle \quad \langle 1^*, 2^* \rangle$
conform to (4iv').

But now observe that there will be a chain associating the pair (Dan, t). This is necessary, since only through such a chain can the theta-criterion be satisfied with respect to the 'subject' theta-role of proud, which is discharged by t. Dan is an argument in a non-theta position, with Case, and t is a non-argument, in a theta-position lacking Case.

If this is so, then the linking in (150) will violate the theta-criterion.

150) Parker wonders [[how t proud of himself]] [Dan is e]
 $\begin{array}{c} \wedge \qquad \qquad \qquad \wedge \qquad \qquad \qquad \wedge \\ | \qquad \qquad \qquad | \qquad \qquad \qquad | \\ | \text{-----} | \text{-----} | \end{array}$

We presume the following condition to be true of any well-formed

37
chain :

(151)

In any chain $C = (a_1, \dots, a_i, \dots, a_n)$, for every

37. See Chapter Four for discussion. We essentially remain neutral on the important question of whether chains are strictly formed under movement, or are formed at S-structure when certain relations hold between NPs (like (151)), or are formed entirely freely, improper choices being filtered out by well-formedness conditions. The latter alternative seems to be more in the spirit of the GB approach to syntactic operations, but there may be empirical reasons to prefer one approach to chain formation over another. See Rizzi (1982), Chomsky (1986) for discussion.

pair a_i, a_{i+1}, a_{i+1} is linked to a_i .

The only chain structure for (150) consistent with (151) is (Parker, t) and (Dan). The first chain contains two theta-positions, the latter none, in violation of the theta-criterion.

The only way in which himself can be understood as coreferent with Parker is if Parker is an antecedent of the reflexive. The reflexive cannot link directly to Parker, since it is forced already (by the binding theory) to link to the trace, and reflexives cannot have split antecedents. If the phrase Parker is the antecedent to the trace, then either we have the linking in (150), which violates the θ criterion or (151)³⁸, or we have the linking in (152), which is again ill-formed, since traces cannot have split antecedents.

152) * Parker wonders [how t proud of himself]] [Dan is e]

38. The theta criterion requires that the chains (Dan, t) and (Parker) be formed. The chain (Dan, t) violates (151), since the trace is not linked to Dan. The chains which are well-formed by (151) are (Parker, t) and (Dan); they are ill-formed by the theta-criterion, since the first has two theta-positions and the latter none.

3.6.2 Couquaux's Analysis of Predicative Clauses

Indeed, an analysis of the phrase structure of predicative clauses along these lines has been offered by Couquaux(1981), in a discussion of French predicative constructions like (153):

153) Pierre est gentil
 'Pierre is nice'

(Rizzi (1982) makes a somewhat similar suggestion, to be discussed below).

The surface order of these clauses is as in (154), where V* ranges over être 'to be', sembler 'to seem', paraître 'to look (AP)', demeurer, rester 'to remain', and others.

154) NP V* AP

Couquaux argues that the sentences have the structure (155) (derived via NP-movement from (156)), rather than (157).

155) [NP [V t AP]]
 IP VP

 | |
 | |

156) [e [V NP AP]]
 IP VP

157) [NP [V AP]]
 IP VP

Couquaux adduces two pieces of evidence for this position. First, in such predicative constructions as (158), the clitic en

can be extracted from the surface subject, as shown in (159)³⁹ ;
 this extraction is limited to predicative sentences, as the
 ungrammaticality of (161) attests.

158) La preface de ce livre est trop flatteuse
 The preface of this book is too flattering

159) La preface en est trop flatteuse
 the preface of it is too flattering

160) L'eclatement de ce pneu provoquerait un grave accident
 the bursting of this tire would provoke a severe accident

161) (=Couquaux's (42b))

* L'eclatement en provoquerait un grave accident
 the bursting of it would provoke a severe accident

This type of cliticization differs from other case in French,
 which typically occur from inside the VP. Couquaux argues that
 no special rule of en-cliticization need be posited if we assume
 that the D-Structure of (159) is as in (156); the clitic will
 move from inside the object NP prior to being moved to subject
 position. The ungrammaticality of (161) is thus explained,
 since this sentence is generated with [l'eclatement en] in
 NP

subject position, blocking clitic extraction .

 39. Couquaux presents evidence that this is actually clitic
 movement.

40. That is, clitic movement in this case will have to proceed
 from inside the subject, in contrast to the clitic movement in
 (158)-(160). Movement from inside subjects is generally
 prohibited, perhaps due to the ECP or a constraint against

Couquaux's second piece of evidence for (155) as the S-Structure representation of predicative clauses comes from the fact (noted originally by Kayne (1975)) that reflexive clitics cannot occur in such sentences.

(162) (=82a) Pierre est cher à lui même
Pierre is dear to himself

(163) (=79a) * Pierre s'est cher
Pierre to-himself is dear

Crucially, this is a restriction only on reflexive or reciprocal clitics, as a pronominal clitic can occur in such a sentence:

(164) (=83) Odile lui est chère
"Odile is dear to him (or her)"

Couquaux's analysis of the impossibility of a reflexive clitic here involves treating such predicative constructions as involving raising, thus subsuming the contrast in (162) and (163) under the general fact that NP-movement and reflexive clitics are incompatible (Kayne 1975, Rizzi 1982). Below, I summarize his analysis, then give an alternative account within the GB framework, proposed by Rizzi (1982). (Rizzi's analysis will be incorporated into the discussion of empty categories and chain well-formedness in Chapter 4).

extraction from a left branch. See Rizzi and Belletti (1981), Kayne (1983) from discussion.

3.6.2.1 Couquaux's Analysis: the OB Indexing Analysis

Couquaux's account of (163) is based within his modification of the framework of indexing proposed by Chomsky (1980), henceforth the OB framework. In what follows, I will give a short summary of the OB indexing system, and the basic sense of Couquaux's proposal; the reader is referred to Couquaux (1981) for details.

In OB, a non-anaphoric NP may be assigned two indices: a referential index *R* (an integer), freely assigned or assigned under movement; and an anaphoric index *A*, consisting of the set of referential indices borne by all c-commanding NPs. Anaphors receive only referential indices. In the example below, the referential index is to the left of the semicolon, the anaphoric index to the right, in {} (this departs minimally from the OB notation).

In an example like (165), for example, John is freely assigned the referential index *i*; since John c-commands the anaphor himself and him, *i* is assigned as the anaphoric index of the anaphor and the pronoun. The interpretation of anaphoric indices is that an NP bearing an anaphoric index {...*i*...} must be disjoint in reference from any NP bearing the referential index *i*. (See Chomsky (1980) for details.)

165) a. John told himself about him
 i; ;{*i*} j{*i*}

The binding conditions alter the content of the R- and A-indices of pronouns and anaphors. Chomsky terms the anaphoric index of a pronoun and the referential index of an anaphor their designated indices. When an anaphor or pronoun X is free in an opaque domain Z from the index i (meaning that X is not c-commanded in the opaque domain by another NP Y which bears the referential index i), then i is deleted from its designated index. (Opaque domains are the domains of Tense and Subject). The effect of the binding conditions is to eliminate the referential index of an anaphor which is free in its opaque domain; this results in ungrammaticality. The effect for pronouns is that they are permitted to overlap in reference with an NP which c-commands it, as long as that NP is not within the pronoun's opaque domain. If the pronoun is not free in that minimal domain, then it must be disjoint in reference from any NP bearing i as its referential index. (See Chomsky (1980), Lasnik (1981) for details.) Thus (165a) is converted to (165b).

165) a. John_i told himself_i about him_{j< i>}

The anaphor is not free in the clause from the index i, since John_i c-commands it; thus it is not subject to the binding conditions, and is allowed to retain its referential index. The pronoun is also not free in its opaque domain from the index i, hence this index does not delete from its anaphoric index. This

in (165b) the pronoun must be interpreted as disjoint in reference from John (and himself).

Couquaux's analysis introduces another indexing convention which has the effect given in (166) (not his exact formalization):

(166)

For an anaphor X c-commanding an NP Y bearing the referential index i, delete i from the referential index of X.

Consider again (163), assuming a version of Couquaux's proposed

41
structure :

163) * [Pierre se [est t [gentil]]]
IP VP AP

The index on the trace is copied from that of Pierre. se is coindexed with Pierre by rules of construal, giving (167). (166) deletes i from the referential index of se (since se c-commands the trace, which bears the index i) giving (168).

167) [Pierre se [est t [gentil]]]
IP i; i; VP i; AP

168) [Pierre se [est t [gentil]]]
IP i; VP i; {i} AP

41. Couquaux additionally makes other modifications to the OB indexing system, which need not concern us here. In the discussion below, I modify his analysis to reflect the OB system, with the added (166), which reflects the intent of his proposal.

Couquaux argues that this violates the binding theory, as the anaphoric clitic is left without an index.

To abstract away from the indexing system, the sense of Couquaux's analysis is that an anaphor cannot bind a coindexed element. His analysis is problematic. This treatment of (163) would seem to also rule out (169), with himself coreferent with he (and with John):

169) John told himself that he should have a sandwich

Rules of construal and free indexing will produce (170a), and then Couquaux's rule will produce the ungrammatical (170b).

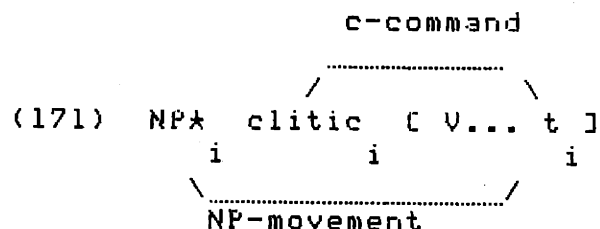
170a) John told himself that he should have a sandwich
 i; i; i;

b) *John told himself that he should have a sandwich
 i; i;

Couquaux's analysis is also based within his extension of the OB indexing system, and makes use of mechanisms (specifically index erasure) not available within GB.

Rizzi (1982) proposes that sentences of the abstract form (171) are ill-formed due to a locality condition on chains. He suggests that (172) is a well-formedness condition on chains⁴²; the intervening reflexive clitic prevents (NP_A ,t) from constituting a well-formed chain.

42. This condition is given as one of four properties of chains by Chomsky (1981), but Chomsky does not present arguments for the necessity of adopting (172) as a well-formedness constraint.



(172) Local Binding Condition on Chains (Rizzi (1982))

$C = (a_1 \dots a_n)$ is a chain iff, for $1 \leq i \leq n$, a_i is the local binder of a_{i+1} .

A is a binder for B iff A and B are coindexed, and A c-commands B.

A is the local binder of B iff A is a binder of B, and there is no Z, such that Z is a binder of B and Z is not a binder of A.

The theta criterion demands that in the sentence (171), (NP_i, t) form a chain. This is not a well-formed chain by (172), since NP_i is not the local binder of t; the clitic is a binder for the trace, and does not bind NP_i. This blocks the NP_i from being the local binder. (See Rizzi (1982) for many other

examples, across a variety of NP-movement constructions). This problem does not arise in a sentence with a non-anaphoric clitic, like (164), since here the pronominal clitic does not

43. Also see Lasnik (1985) for arguments in favor of such a condition based on English. Epstein (1986) discusses the definition of local binder. Finally, the next chapter of this thesis generalizes (173) to certain cases of Strong Crossover, as well as to standard cases like (i) analyzed as Condition C violations in Chomsky's (1981), (1985) theory.

(i) * He_i likes John_i

bind the trace, allowing Odile to be the clitic's local binder, and thus allowing the necessary chain (Odile, t) to be well-formed. Similarly, no ungrammaticality arises in (170) (a problematic case for Couquaux's analysis of (163)), since each NP forms a singleton chain.

Rizzi's condition explains the incompatibility of anaphoric clitic movement and NP-movement of any sort. He notes that his analysis extends to cases in Italian similar to Couquaux's examples, and suggests, like Couquaux, that NP-movement occurs in predicative clauses. The phrase structure he suggests is different from that of Couquaux; Rizzi characterizes the post-copular material in such sentences as a small clause.

(173) (=47b) * Loro si sono simili
 They to-each other are similar

NP si essere [e [e]]
 i i SC i A' i

If we are to extend the Couquaux/Rizzi proposal that such Romance predicative sentences have derived subjects to English, then the predicate/non-predicate asymmetry in the multiple binding domain-effect is explained, as observed earlier (cf. the discussion of (149)-(151) above). This is so only if we accept the notion that the post-copular material forms a constituent (small clause), and is not a sequence of two independent constituents, as Couquaux proposes.

To see this, consider (149) again.

149) [Parker [wonders [[how [t
 IP VP CP AP<1*,2*> AP<1*, 2*>
 [[proud [of himself]]]]
 AP<1, 2*> A' <1, 2> PP
 [Dan is [e]]]
 IP<1*, 2*>

In this representation, the entire AP, crucially including the trace, has been moved; this follows Rizzi's characterization of copular verbs as taking SC complements. If this is the correct structural representation, the lack of ambiguity follows: the reflexive is linked to the trace, and the trace is linked to Dan, following essentially the local binding condition on chains, recast in terms of linking.

But suppose Couquaux's analysis is correct, and the copular verb does not take a small clause complement, but rather takes two complements, an NP and the AP:

174) Parker [is t [proud of himself]
 VP AP

If this is so, then (149) is not the correct structure for (14), but (175) is:

175) [Parker [wonders [[how [proud
 IP VP CP AP<1,2*> A' <1, 2>
 [of himself]]]] [Dan
 PP IP<1*, 2*>
 [is t [e]]]
 AP<1, 2*>

In this structure, the trace of Dan remains in the lowest clause, while the AP containing the reflexive is moved to the higher SPEC position, out of the c-command domain of Dan and the trace t.

Now, the anaphoric dependence of himself upon Dan will have to be accounted for by the theta-compatibility requirement (4iv'). Without this, the linking of the reflexive to Parker is not blocked.

3.6.2.2 Summary

To conclude this discussion, we have observed an asymmetry in connectivity between cases where the dislocated phrase containing the anaphor is a predicate, and cases where it is not. When it is a predicate, the anaphor is bound by the subject of the clause in which it occurred at D-structure.

We have considered two proposals each of which capture this asymmetry. In the first, the definition of chain accessibility sequence is formulated to force the anaphor's antecedent to be the D-structure subject of the moved predicate. This part of the definition capitalizes on an inherent distinction between predicates and other expressions in terms of their thematic saturation. This requirement subsumes the clause (4iv), which is the parallel in our system to Chomsky's (1985) requirement that only Complete Functional Complexes constitute potential binding domains.

On the second proposal, predicative sentences are viewed as having derived subjects, following Couquaux (1981) and Rizzi (1982). We believe this raising proposal to be correct, although its implications for the proper treatment of the

asymmetry have not yet been determined.

If a WH-moved AP contains the trace of the derived subject, then the predicate/non-predicate asymmetry in connectivity reduces to independent properties of the structure. The theta-compatibility requirement (4iv'') can be abandoned, although we will have to re-introduce (4iv), which was subsumed under (4iv'').

If Couquaux is correct, and the trace of the derived subject in a predicative sentence is not part of the AP, then it does not move along with the WH-moved AP in such cases as (140). If this is correct, then we will have to adopt the theta-compatibility requirement on chain accessibility sequences.

We now turn to a discussion of predicate nominals and small clause structure.

3.6.3 Predicate Nominals and Small Clauses

The suppression of the multiple binding domain effect (MBDE) is not confined to moved APs; it occurs with predicate nominals as well, although the judgements are somewhat subtle.

Consider this example (see also (134-135)):

176) * John thinks that an admirer of himself, Mary became

If we are to characterize this suppression of the MBDE as a

consequence of raising in predicative constructions, then we shall have to say that in (176) the phrase Mary would have to have raised out of the VP. If we adopt Rizzi's proposal that copular verbs take small clause complements, we must first determine what small clauses are. We assume, following Stowell (1983) and Chomsky (1986), that small clauses are of the form (177):

177) [NP [...]]
 XP XP

I consider [Dan [silly]]
 AP AP
 I consider [Dan [a fool]]
 NP NP

That is, small clauses are projections of lexical heads, with base-generated adjoined NPs, which receive a theta-role from the head. The NP subjects of the SCs are governed by the higher verb, which then can assign Case to them.

In a case like (178), Dan has raised to subject position; let us assume that the copula does not assign Case, forcing this movement.

178) Dan is [e [a fool]]
 NP NP
 ┌ ┐
 └──────────┘

This analysis gives (176) this structure:

179) John thinks that [[t [an [admirer of himself]] [Mary became]]]
 IP NP NP N' IP

The restriction in anaphor binding in such cases with predicate nominals will now fall under the treatment(s) formulated for AP-movement.

3.6.4 The Necessity of the Theta-Compatibility Treatment

I shall now present evidence which suggests that even if we adopt Couquaux's and Rizzi's proposal that predicative sentences involve raising (which I believe to be correct, given the clitic facts), and Rizzi's SC analysis of such structures, the theta-compatibility requirement on accessibility sequences -- the first of the two analyses of the predicate/non-predicate asymmetry offered above -- must be adopted.

We saw above that in order to explain the connectivity facts in cases of moved predicates, we must either adopt (4iv'), or suppose that the trace of the derived subject must move along with the rest of the AP containing the predicate (see (149)).

Now consider the simple case (180):

180) [[how silly] do [you [consider [Dan [e]]]]]?
 CP AP IP VP AP AP

In this sentence, we have moved, not the whole AP (the entire small clause), but only part of it, the part containing the head, its A' projection, and its specifier (how). The WH-movement obligatorily leaves the 'subject' behind.

181) * [[Dan [how silly] do [you [consider [e]]]]]
 CP AP AP IP VP AP

We suggest that the ungrammaticality of (181) is due to Case theory; the NP Dan (which constitutes a chain) needs case, which is assigned to it in (177) by consider). The WH-movement puts Dan in a position in which it cannot be assigned Case, and so the chain is not visible for theta-role assignment, and thus a theta-criterion violation arises (cf. footnote (11)).

Now consider WH-movement of an AP in a sentence where the SC subject has raised to matrix subject position, along the lines of the Couquaux/Rizzi proposal.

182) a. D-Structure: [[[is [Dan [how proud of himself]]]]]
 CP C' IP AP AP

b. Intermediate representation, derived via NP-movement:

[[[Dan is [t [how proud of himself]]]]]
 CP C' IP AP AP
 |
 |_____|

c. S-Structure 1:

[[how proud of himself] [is [Dan [t [e]]]]]
 CP AP C' IP AP AP

d. S-Structure 2:

[[t [how proud of himself] [is [Dan [e]]]]]
 CP AP AP C' IP AP

If we are correct in supposing that (181) is out for reasons of Case, then (182d) is grammatical, since the trace t does not need Case.

(182c) parallels (180); the 'core' AP is WH-moved, leaving the

subject t behind. The theory, in essence, allows two well-formed S-Structures to be derived from the underlying representation (182a). The crucial one for our purposes is (182c).

By the same line of reasoning, (184) is a well-formed S-Structure of (183):

183) * John wonders how proud of himself Mary became

184) [IP John wonders [[CP AP<1, 2★> how [A'<1, 2★> proud
[of himself]]] [[C' IP★<1★, 2★> Mary is [t
PP AP<1, 2★> e]]]]]
[AP<1, 2★> e]]]]]

The trace of Mary is left downstairs, the reflexive has no binder internal to the WH-phrase. Thus, the sentence can only be excluded if the accessibility sequence {himself, A', AP, CP, VP, I', IP} is ill-formed; John is minimally <1,2★>

accessible to the anaphor through this sequence, and so the link between himself and John satisfies the Binding Theory if this accessibility sequence is well-formed.

The theta-compatibility requirement on accessibility sequences, (4iv'), rules out this sequence, since AP is a member of the sequence, and IP★ is not. I therefore conclude that (4iv'') is to be adopted as part of the definition

of accessibility sequences. The predicate/non-predicate asymmetry in connectivity is due to the thematically unsaturated nature of predicates, and is independent of the NP-movement of the subject in predication sentences.

In the next section, I consider cases of anaphor connectivity which are manifested in sentences which have A chains. The basic point of this section is that such cases fall directly under the formulation of Condition A offered above, with no changes.

3.7 Reciprocals and Connectivity

This section examines reciprocal interpretation in connectivity constructions. We are concerned with several issues, among them the relation between syntactically assigned antecedence relations and interpretation; and the choice between linking theory and indexing theory. This section presumes the linking framework; advantages of this notation will be discussed.

In particular, evidence will be adduced in favor of linking, and in favor of the approach to anaphor binding developed earlier in this chapter. The evidence will emerge from considering two semantic constraints on the distribution of the reciprocal each other. I shall initially focus on a blocking


effect Higginbotham (1981) terms 'antecedent clash'. In the second subsection, I shall review a second blocking effect, and then observe that these two blocking effects appear in declarative sentences but disappear in connectivity sentences of the appropriate form; this provides an argument that the structure upon which reciprocal interpretation is computed is isomorphic to SS in relevant respects. The disappearance of the blocking effects is unexpected in a theory which utilizes LF lowering (reconstruction) as a means to account for reciprocal binding and interpretation.

The English reciprocal NP each other is an anaphor, subject to the binding theory; within the framework of this chapter, it is required to link at S-Structure to an NP which is minimally chain accessible to it. Thus we see a typical SSC effect in (185), and a suppression of the effect in (186):

185) *The men believe that Mary painted pictures of each other
 | |

*The men believe that the women love [each other's friends]
 | |

 | |
 | |
 186) Which pictures of each other do the men think that Mary
 painted?



Which of each other's friends do the men believe that the women love?

Independent of the reciprocal's syntactic status as an anaphor, it has a specific semantic interpretation as well.

3.7.1 Reciprocals and their Interpretation

Higginbotham (1981) presents a discussion of the interpretation of reciprocals, and ambiguities which sometimes arise with reciprocals; this analysis is partially amended in Higginbotham (1985). The focus of the subsequent discussion is on a peculiar constraint on reciprocal interpretation that Higginbotham terms antecedent clash; we shall argue that the suspension of this constraint in connectivity contexts is a strong argument in favor of determining these reciprocal

44

interpretations on structures essentially isomorphic to S-Str , and thus a strong argument against lowering as a preliminary step in determining the structure interpreted by semantic rules. We will argue that no reformulation of Higginbotham's (1985) algorithm for assigning values to reciprocal expressions

44. By this, I mean that the rules of reciprocal interpretation apply to syntactic representations in which no literal 'reconstruction' has occurred. In Chapter 2, I suggested that the binding theory applies to S-Structure; this proposal will be discussed and modified in Chapter 5. the basic conclusions of Chapter 2 concerned the undesirability of reconstruction will be maintained in Chapter 5.

is necessary, even for connectivity structures, given the approach to connectivity developed earlier.

We take as a starting point the discussion of reciprocal and pronominal interpretation of Higginbotham (1985). Higginbotham proposes (187) and (188) as principles constraining the assignment of values to linked pronominals and anaphors:

(187) Suppose a pronominal X is linked to Y in an expression E; then,

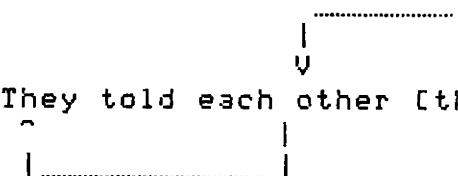
(=Higginbotham's (78)) Include some values of Y among those of X.

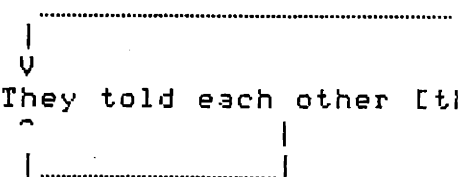
(188) Suppose a reciprocal R is linked to a plural term Y, denoting objects a, b, ...; then,

(=81)) For each a such that $v(a, Y)$ [eg. a is among the values of Y; AB], the values of R are those b such that $v(b, Y)$ and $b \neq a$.

To illustrate, consider (189) (Higginbotham's (68), adapted from Finer (1984)), which can have the two linkings (190) and (191):

189) They told each other [they had better leave]

190) They told each other [they had better leave]


191) They told each other [they had better leave]


For convenience, let us take they to denote a pair of

individuals, x and y. The interpretive principles (187) and (188) give (192) as the interpretation of (190):

(192)
x told y that y had better leave, and
y told x that x had better leave.

(191) has two distinct interpretations, again given by (187) and (188); the ambiguity is due to the possibility of understanding the lower pronominal on a group reading (as in (193)), or a

45

distributed reading (as in (194)) :

193) x told y that x&y had better leave, and
y told x that x&y had better leave

194) x told y that x had better leave, and
x told y that x had better leave

This ambiguity is due to the looseness of the notion 'include

45. As Higginbotham observes, this same ambiguity arises in cases like (i) (his (72)):

(i) They said they would leave

(i) can mean, 'x said that x and y would leave, and y said the same thing' (the group reading of the lower pronominal), or 'x said that x would leave, and y said that y would leave' (the distributed reading). Cases like this actually have additional readings. Consider (ii):

ii) They announced they would run for office

It could be the case that they made a joint announcement, or that they each separately made an announcement; and they might be running together (on a slate) for office, or each running separately. Thus (ii) is four ways ambiguous. See Higginbotham (1981), and the review of this in section () of this thesis, for details.

The expression which fixes the value of a reciprocal must

48

c-command it, within a local domain ; the expression which

at hand.

48. Howard Lasnik observes a problem raised by the contrast of interpretations between (i) (which is discussed in Higginbotham (1981), and (ii):

(i) They said they criticized each other

(ii) [After they took the microphone], they criticized each other

(i) is (at least) two ways ambiguous. Take they to denote John and Mary, then (i) can mean either (iii) or (iv):

(iii) John said [Mary criticized John and John criticized Mary], and

Mary said [Mary criticized John and John criticized Mary]

(iv) John said John criticized Mary, and
Mary said Mary criticized John

The simple clause (v) can mean only (vi), essentially the interpretation of the embedded clause in (iii).

(v) They criticized each other

(vi) [Mary criticized John and John criticized Mary]

Thus it would appear that the interpretation of the embedded clause in (iv) is allowed by the existence of the matrix subject they in (i). Lasnik observes that this pronominal must c-command the reciprocal, since (ii) lacks the interpretation (ix), having only (vi) and (viii):

(vi) After John and Mary took the microphone, John criticized Mary and Mary criticized John

(ix) After John took the microphone, John criticized Mary and after Mary took the microphone, Mary criticized John

Lasnik observes that this indicates that in (i), on the interpretation (iv), it is the higher pronominal which determines the value of the reciprocal; each other is 'reciprocal' with respect to this higher pronoun, not the

fixes the value of a pronominal need not c-command it, and can be indefinitely far away from it. These facts, in Higginbotham's analysis, need not be stipulated in the semantic interpretive rules (187) and (188); they follow from the fact that the rules feed off antecedence relations, which are licensed (in the syntactic component) by the binding theory.

To emphasize this point, the interpretation of a reciprocal R is determined by the NP Y to which R is linked. As is well-known, anaphors cannot have split antecedents -- in the linking theory, an anaphor can link to only one NP. The binding theory forces Y to be a local c-commander of the reciprocal R

Thus the fact that the reciprocal's semantic antecedent must

lower one. Further, this antecedent must c-command the reciprocal, as the unavailability of the interpretation (ix) (similar to (iv)) for (ii) shows. (In the framework of Higginbotham (1981)), this follows, although it does not in the framework of Higginbotham (1985), which I assume here.) This suggests that the expression which fixes the value of a reciprocal must c-command it but not necessarily in a local domain. Along these lines, the introductory clause of (188) might be amended to:

- (x) Suppose a reciprocal expression is transitively linked to a c-commanding plural term Y...

Unfortunately, this makes reference to c-command. In forthcoming work, Lasnik and collaborators attempt to derive the fact that the plural term which fixes the reciprocal's value must c-command it from Chomsky's (1985) proposal that anaphors must move to A' positions close to their antecedents at LF; the c-command condition then follows from the fact that the trace left behind must be bound by the anaphor. For purposes of discussion, I will assume (188) without modification; the main points made in this chapter will carry over to a theory which modifies (188) along the lines explored by Lasnik.

49. For non-connectivity cases, that is.

c-command it follows neatly from Higginbotham's proposal that the semantic antecedent simply is the syntactic antecedent (but see footnote ()). We will take this point up again in a later section.

This aspect of the system is quite important when we consider cases of anaphor connectivity which involve reciprocals. Such cases also reveal an advantage of the linking system over the indexing theory.

3.7.2 Antecedent Clash

Higginbotham (1981) observes and discusses the odd status of examples like (13) ⁵⁰:

198) The men watched each other hitting each other'

The sentence satisfies the binding theory, with the assigned linking in (199):

199) The men watched each other hitting each other'

^		^	

Nevertheless, (198) is not an acceptable sentence of English. Higginbotham's judgement, with which I concur, is that sentences like (198) are semantically odd, and in particular the value of the lowest reciprocal is indeterminate. The sentence's -----

50. I use the prime notation simply to be able to distinguish the different occurrences of 'each other' and 'they'; it has no formal significance.

interpretation oscillates between (200) and (201):

200) x watched y hitting y, and y watched x hitting x.

201) x watched y hitting x, and y watched x hitting y.

Higginbotham observes that the defining characteristic of such anomolous sentences is that the lower reciprocal each other' has two antecedents, the men and each other, and these two antecedents have different values. He terms this property antecedent clash. He further notes that (202) does not have this anomolous interpretation:

202) They promised each other that they' would love each other'

What sets (202) apart from (198)? In (202), the lower reciprocal has clashing antecedents, the local antecedent they' and the higher each other. (202), however, has another antecedent, the matrix subject they. In (202), then, the lowest and the highest antecedents of each other' do not clash. Higginbotham terms the closest antecedent the proximate antecedent, and the structurally highest one -- which ultimately fixes the values of all the other NPs in (202) -- the evaluative antecedent.

51. It is important to emphasize that these sentences are not simply ambiguous, but that the interpretation of the reciprocal is genuinely indeterminate. See Higginbotham (1981) for discussion.

Within the linking theory , we may define these as follows:

- 203) When X is linked directly to Y, then Y is the proximate antecedent of X.
- 204) When a series of links determines X to be dependent upon Z, and Z is not linked to any W, then Z is the evaluative antecedent of X.

We may then formulate the constraint against antecedent clash quite simply:

- 205) For an anaphor X, the values of the proximate and ⁵³ evaluative antecedents of X must be identical .

52. Higginbotham (1981) uses indexing theory. (205) is identical to his formulation of this constraint, although the definitions of proximate and evaluative antecedent are different; proximate antecedent is the local binder, evaluative antecedent is the binder which is not itself bound. My (203) and (204) adapt these notions to the linking framework, using essentially the definitions of antecedence given by Higginbotham (1985).

53. Higginbotham (1981) proposes this as a constraint on the interpretation of all anaphors, citing such examples as (i) as deviant:

- (i) They watched each other hitting themselves

Antecedent clash arises because themselves has a proximate and evaluative antecedent which do not agree in value.

Howard Lasnik observes that this effect is weakened for certain cases, when the relevant anaphor is reflexive. Thus (ii) seems better than (iii) (although still deviant, to my ear):

- (ii) They believe each other to have hit themselves

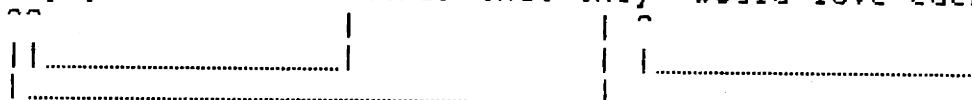
- (iii) They believe each other to have hit each other

This suggests that the antecedent clash constraint might properly be restricted to reciprocals.

We further note that the effect disappears when the anaphor in question is an empty category, either PRO or NP-trace, both of which are classed as anaphors in GB.

(205) is violated in (198-199). However, (206) is a possible linking for (202); this gives rise to the interpretations in (207), which do not violate the constraint (205).

206) They promised each other that they' would love each other'



(207) a. x promised y that x would love y, and
y promised x that y would love x

b. x promised y that x would love y and y would love x, and
y promised x that x would love y and y would love x

Note that the linking in (208) is grammatical (cf. (191)), but it yields the interpretation (209), which violates the constraint against antecedent clash.

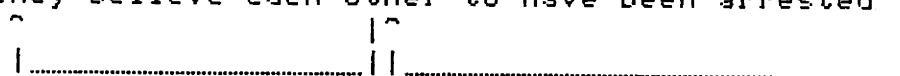
(208)

They promised each other that they' would love each other'

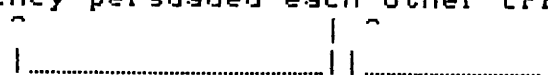


(209) a. x promised y that y would love x, and
y promised x that x would love y.

(iv) They believe each other to have been arrested t



(v) They persuaded each other [PRO to leave]

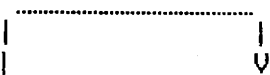


In the next chapter, I develop an alternative theory of empty categories, which holds that neither NP-trace nor PRO are anaphors.

3.7.3 Connectivity and a Redundancy in Indexing

Consider the following case, repeated from above:

210) Which pictures of each other do the men think that Mary painted?



Within the binding theory formulated in (4-8), this sentence is syntactically well-formed; the anaphor is linked to an NP (the men) which is minimally chain-accessible to it, and so is licensed by the Licensing Condition (7).

The reciprocal will be assigned a value by the rule given above; we do not need to complicate rule (188) in any way. This result does not hold true in indexing theory.

Within indexing theory, the rule (187) for interpreting pronominals could be restated thus:

(187') Suppose a pronominal X is coindexed with Y in an expression E; then, include some values of Y among those of X.

No particular problems arise.

However in restating (188), we essentially have to repeat the binding theory internal to the rule, or at least make explicit reference to the binding theory.

For non-connectivity cases like (210), we might re-phrase

(188) as follows in indexing terms:

(188') Suppose a plural term Y , denoting objects a, b, \dots , is the local binder of a reciprocal phrase K ; then,

For each a such that $v(a, Y)$ [eg. a is among the values of Y ; AB], the values of K are those b such that $v(b, Y)$ and $b \neq a$.

(210) They love each other

This is somewhat suspicious. In Higginbotham's linking treatment, the fact that the expression Y must c-command the reciprocal, but need not c-command the pronominal, is simply a consequence of the fact that rules (187) and (188) feed off syntactic representations. In (188'), on the other hand, we need to explicitly state that Y must c-command the reciprocal. Since the syntactic nature of each other as an anaphor, and its semantic properties, are independent, this addition to the rule for interpreting reciprocals is rather stipulatory.

The problem becomes worse when we consider a connectivity case, in which the reciprocal is not in fact c-commanded by the expression Y which determines its value. (186) is such a case. (186) will have the representation (211) within indexing theory:

211) [Which pictures of [each other]] do [the men]
 i j i
think[_it that [Mary painted e]?]
 i i

Here again, the reciprocal's value is determined in terms of the NP the men. But, this violates (188'), since the men does not

c-command the reciprocal. Hence, we need to embed within (188') the fact that Y must be the NP with which R is coindexed in satisfaction of the Binding Theory. That is, we have seen that in connectivity sentences, an anaphor is not in fact c-commanded by its antecedent. Nonetheless, there is a quite formal difference between such grammatical cases as (211) (or the many other examples of connectivity discussed in this dissertation), and ungrammatical cases like (211'), a difference we have characterized in terms of chain-accessibility sequences.

(211') [pictures of [the men]_i] fell on [each other]_i

Similarly, we have seen that in non-connectivity cases the semantic antecedent of an anaphor (the NP which determines its value) must c-command it (cf. discussion above, and footnote ()), but that in connectivity cases like (211) this requirement is relaxed. Since, by hypothesis, the rules for assigning interpretations to anaphors are distinct from conditions filtering assignments of syntactic antecedents to them, this means that the special type of relaxation of c-command must be encoded both in the binding theory and in the rules for interpretation. This is a redundancy between the syntactic and semantic conditions on anaphors. As was pointed out above, the redundancy exists in indexing theory, but not in linking theory. In sections to come, I further examine the redundancy. For the time being, let us consider how to reformulate the rule for assigning an interpretation to a reciprocal expression.

(212) and (213) are two possible formalizations.

(212) Suppose a plural term Y , denoting objects a, b, \dots , is coindexed with a reciprocal phrase R , in satisfaction of the Binding theory; then,

For each a such that $v(a, Y)$ [eg. a is among the values of Y ; AB], the values of R are those b such that $v(b, Y)$ and $b \neq a$.

213) Suppose a plural term Y , denoting objects a, b, \dots , is coindexed with a reciprocal phrase R , and is minimally chain-accessible to R ; then,

For each a such that $v(a, Y)$, the values of R are those b such that $v(b, Y)$ and $b \neq a$.

This is very problematic, giving specialized status to coindexation which satisfies the binding theory; and in order to do this, we need to re-state the binding theory internal to a semantic rule of interpretation. Within linking theory, the specialized antecedence relation between an anaphor and its (binding-theory satisfying) antecedent is already given, so we need only make reference to chain accessibility once, in the binding theory. the relaxation of c-command for interpretation in (211) follows, since the interpretation of reciprocals, in Higginbotham's formulation, is sensitive to the value of the NP to which the anaphor is linked..

The antecedent clash effect is suspended in connectivity cases

54
of a certain type. Consider the following contrasts :

54. Higginbotham (1981) asserts that the antecedent clash effect obtains in clefts, roughly of the form (214b). I disagree, as

- 214) a. † Joyce and Shaw believed each other to revile all of each other's plays
- b. It was each other's plays that Joyce and Shaw believed each other to be proud of
- (i.e. 'Joyce believed Shaw to be proud of Shaw's plays, and Shaw believed Joyce to be proud of Joyce's plays'.)
- 215) a. † John and Mary forced each other [PRO to talk to each other's parents]
- b. which of each other's parents did John and Mary force each other to talk to? (answer: John forced Mary to talk to her mother, and Mary forced John to talk to his father).
- 216) a. † They watched each other hitting each others' friends
- b. which of each other's friends did they watch each other hit?
- 217) a. † They saw each other painting rectangular pictures of each other
- b. what kind of pictures of each other did they [t [see[each other paint t'?

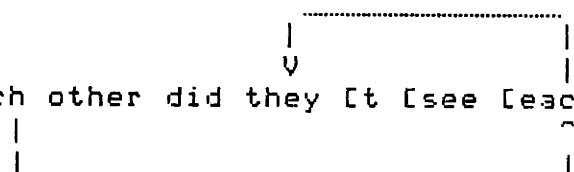
The antecedent clash effect disappears (although the sentences remain difficult due to their complexity) when the lower reciprocal is carried outside the c-command domain of the other reciprocal. This fact receives a simple treatment within the linking approach to reciprocal interpretation and antecedent clash. Let us consider, for example, (217b). The question (217b) can be answered by (218):

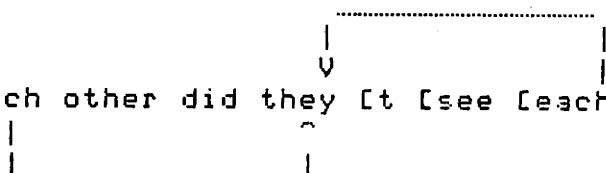
218) Well, Mary saw John paint big pictures of him, and John saw Mary paint miniature pictures of her.

(217) has two grammatical linking structures. I assume, -----
far as example (214b) goes.

following Chomsky (1986), that the movement of the WH-phrase adjoins to VP, and that this movement may leave a trace, *t* in

(217). (I return to this question later. ⁵⁵).

(219) 

(220) 

Each other' links to they in both cases. By the revised Binding theory, each other -- the anaphor inside the WH-phrase -- may link either to each other', or to they. Both of these NPs are minimally accessible to each other, through different accessibility sequences. Note that this is a strong argument in favor of VP-adjunction.

The linking in (219) will give rise to antecedent clash, since each other' is the proximate antecedent of each other, they is the evaluative antecedent, and the two have different values.

55. The suppression of antecedent clash in such an example provides evidence for Chomsky's VP-adjunction proposal, if my treatment of the general suppression of the clash effect is correct. If there were no VP-adjoined trace, then there would be no way, in the framework advanced here, for they to be assigned as the grammatical antecedent of each other.

(221) [what kind of pictures of x] did y see x paint], and
[what kind of pictures of y] did x see y paint]?

The (b) examples are rather complex, since they are questions containing two reciprocals. The antecedent clash effect occurs with all anaphors, as Higginbotham (1981) observes (but cf. footnote ()):

The suppression of antecedent clash is seen quite sharply in the contrast between (222) and (223), and (224)-(225):

224) # They heard each other telling stories about themselves

225) which stories about themselves did they hear each other telling?

|²² |
| || |

3.7.4 Antecedent Clash in Indexing Theory: Another Redundancy

As we have just seen, the formulation (205) of the constraint against antecedent clash requires no supplementation to handle the suppression of this effect in cases of connectivity. Let us now consider how we would have to formulate the counterpart of (205) in indexing theory.

We might formulate proximate and evaluative antecedence in the following way:

- 226) Y is a proximate antecedent of X iff Y and X are coindexed, and Y is minimally chain-accessible to X.
- 227) Y is the evaluative antecedent of X iff Y and X are coindexed, Y is chain accessible to X, and there is no Z such that Z and Y are coindexed and Z is chain accessible to Y.

(205) can be then reformulated as (228):

- 228) for an anaphor X, X must have a proximate antecedent W whose value is identical to that of X's evaluative antecedent.

Finally, the rule for assigning values to anaphors can be as follows:

- (229) for X an anaphor, and Y, a proximate antecedent for X, assign values to X as follows:
- i) if X = reciprocal, for each a such that $v(a, Y)$, the values of X are those b such that $v(b, Y)$

and $b \neq a$.

- ii) if $X = \text{reflexive, PRO, or trace}$, for each a such that $v(a, Y)$, the values of X are those b such that $v(b, Y)$ and $b = a$.

In (217b) for example, both they and each other are proximate antecedents to the anaphor each other; hence (228) is satisfied, since one of these they has a value identical to that of the evaluative antecedent, which is also they in this case.

Now consider examples of the type discussed in section (3.3), in which the phrase containing the anaphor is a predicate. As we saw, the anaphor inside the moved predicate must have as its antecedent the subject of the clause containing the variable:

230) John wonders how proud of each other the men became

231) *The men wonder how proud of each other John became

The antecedent clash effect is not suspended in such cases.

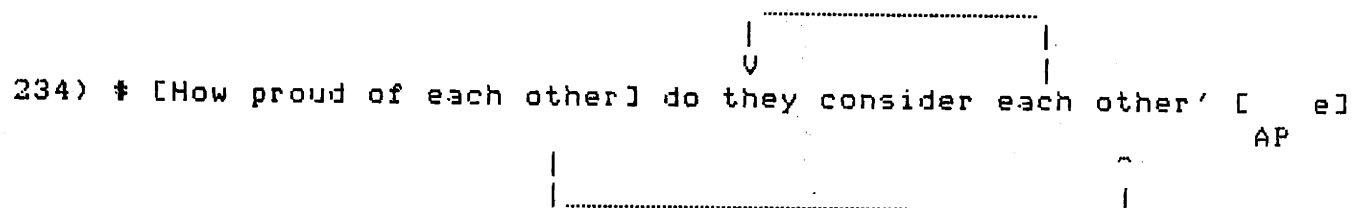
232) *the men know how proud of each other they consider each other' to be t

233) * [How proud of each other] do they consider each other' [e]
AP

(234) is the only grammatical linking available for (233), given

the discussion and analysis of section (3.3) ⁵⁶.

56. Given the constituent structure (i), each other' is a minimally chain accessible antecedent for each other through the well-formed sequence (ii); but they is minimally accessible through only the sequence (iii), which is ill-formed, by the theta-compatibility requirement (4iv').



The interpretation given by this linking will violate (205).

Similarly, in the indexing-theoretic formulation of antecedent clash, given as (44), (235) will give rise to an interpretation violating this constraint.

235)
* [How proud of each other] do they consider each other' [e]

Only each other'_j is minimally accessible to each other through a well-formed chain accessibility sequence; they is not (see fn. ()). Since the indexing theory prohibition against antecedent clash has the licensing condition for anaphors

(i) [[CP AP<1, 2★> How [A'<1, 2★> proud [of each other] PP
[do [they [t [consider[each other' [e]
C' IP VP VP AP★<1★, 2★> AP

(ii) (each other, PP, A' <1, 2★>, AP <1, 2★>, e, AP★ <1★, 2★>)

(iii) (each other, PP, A' <1, 2★>, AP <1, 2★>, t, IP <1★, 2★>)

Thus the only antecedent which is accessible to each other through a well-formed accessibility sequence is each other'; and this will give rise to antecedent clash.

embedded within it, this fact rules out the interpretation of (235) as anomolous.

However, an interesting problem arises for the indexing-theoretic formulation of the rule for assigning values to anaphors, and the constraint on such assigned values which rules out antecedent clash. In Higginbotham's theory, these are two distinct semantic principles; one assigns values to anaphors, the other rules out certain assignments of values as anomolous. We have earlier formulated them in indexing theory --following the distinction between them in Higginbotham's framework -- as two distinct principles. For convenience, I repeat them.

Prohibition Against Antecedent Clash

- 228) for an anaphor X, X must have a proximate antecedent W whose value is identical to that of X's evaluative antecedent.

Anaphor Value Assignment

- 229) for X an anaphor, and Y, a proximate antecedent for X, assign values to X as follows:

- i) if X = reciprocal, for each a such that $v(a, Y)$, the values of X are those b such that $v(b, Y)$ and $b = a$.
- ii) if X = reflexive, PRO, or trace, for each a such that $v(a, Y)$, the values of X are those b such that $v(b, Y)$ and $b = a$.

Recall that 'proximate' and 'evaluative' antecedent must be defined in indexing theory in terms of accessibility sequences; this then repeats a crucial part of the (syntactic) binding

theory within the semantic principles.

Consider now the example (236), which is similar to (235); it has a moved predicate, forcing each other' to be the only accessible antecedent for each other. Unlike (235), it has an additional pronominal they, which here is the evaluative antecedent.

236) * They wonder how proud of each other they'
 i i i
 consider each other'
 i

By the definition of chain accessibility sequence, only each other' can be the proximate antecedent of each other, and they' is the proximate antecedent of each other' (see the definitions in (226-227)). By the interpretive rule (229), we arrive at this interpretation (take they to denote John and Mary):

(237)
John wonders [how proud of John] John considers Mary, and
Mary wonders [how proud of Mary] Mary considers John

This interpretation violates (228), since the proximate and evaluative antecedents of each other have different values. Crucially, (228) must make reference to 'proximate' antecedents, since otherwise this interpretation cannot be ruled out. Thus we see a double redundancy; the notion of proximate antecedent is stated first in the binding theory, and also in the two separate principles (228) and (229) of the interpretive component.

Once again, it should be emphasized that the formulation of

reciprocal interpretation and the prohibition against antecedent clash within indexing theory must have a central aspect of the licensing condition for anaphors -- the core of the binding theory -- embedded within it. Given this massive redundancy between syntactic and semantic principles, I suggest that linking has fundamental formal advantages over indexing theory.

57

No such redundancy exists in the linking theory formulation .

There is another possibility. As I have shown, there is a redundancy in the indexing theoretic treatment of the distribution and interpretation of anaphors. I have resolved this in one direction, namely by abandoning the indexing framework in favor of linking. This treats the structural relation between an anaphor and its antecedent as a purely syntactic matter, with the rules of interpretation feeding directly off syntactically licensed antecedence relations (links). This makes no reference to configurational syntactic

57. In other words, linking is to be preferred on grounds of simplicity; for an anaphor X linked to Y in accordance with the binding theory, Y determines the value (via semantic rules) for X, in all cases. In indexing theory, for an anaphor X whose syntactic relation with Y satisfies the binding theory, Y determines the value of X. The additional complexity of indexing theory arises in trying to keep track (in the interpretive component) of which NP Y was the one which satisfied X's needs as an anaphor in the syntactic component. Linking encodes this directly (we need only assume that linkings carry over (along with the rest of the syntactic structure) into the semantic component). Indexing does not encode the relation, and so the structural relation between X and Y must be re-calculated in the semantic component. This seems both redundant and rather complicated, in comparison with linking.

structure in the rule of interpretation.

One might suppose, however, that the redundancy is to be resolved in the other direction: the structural relation between the anaphor and its antecedent is to be encoded (as in (229)) in the interpretive rule itself, and the binding theory as a syntactic system independent of interpretation is to be abandoned. When an anaphor fails to have a minimally chain-accessible antecedent, the sentence will be barred by the interpretive rule: the anaphor will receive no value.

Such an approach is falsified, in essence, by Higginbotham (1985), who points out that the sentence (238) (his (12)) is perfectly understandable, albeit not acceptable as a natural (eg. grammatical) English sentence.

(238) John wishes that Mary would visit himself

In (55), the anaphor is assigned a value, namely John, although it violates the principle governing the syntactic relation between an anaphor and its antecedent. This shows that the solution to the redundancy just hypothesised cannot be maintained. Suppose himself is linked to John. Then, by the rule (229ii), John is assigned as the value of the reflexive, correctly, since the sentence simply means 'John wishes that Mary would visit John'. This linking will violate the syntactic binding theory, since John is too far away from the anaphor for the linking to satisfy the binding theory (the presence of Mary

causes the embedded clause to be the binding domain for the anaphor). This is a satisfactory result: the sentence is understandable yet ungrammatical, and by hypothesis violates a syntactic well-formedness condition but can be interpreted by semantic rules.

On the alternative approach, where an anaphor can be assigned a value only if linked to a local NP which c-commands it, the reflexive would not be assigned a value, and (238) would not be interpretable; this is the wrong result.

3.7.5 Adverbial each

Dougherty (1974) observes the incompatibility of the adverb each with reciprocal NPs⁵⁸ :

239) * They each like each other

I here follow, in its essentials, the treatment of the ill-formedness of (239) offered by Gillon (1984), incorporating aspects of Higginbotham's (1981) discussion of the interpretation of pluralities and reciprocals, which is an extension of that in (Fiengo and Lasnik (1973)). To summarize Higginbotham's proposals, a plural NP denotes a plurality; the

58. See Gillon (1984) for arguments that this occurrence of each (and, in similar sentences, both and all) is adverbial. Gillon's analysis departs from earlier treatments (Fiengo and Lasnik (1973), Maling (1976)), which consider each to be a quantifier which has 'floated' out of the subject NP through a minor transformation.

plurality (a 'set-like object') has members. Higginbotham defines partition as follows:

- 240) A subplurality Q of a plurality P is a plurality each member of which is a member of P.
- 241) A covering of a plurality P is a family C of subpluralities of P, such that every element of P belongs to some member of C.
- 242) A partition of P is a covering of P such that the members of P are pairwise disjoint.

He then offers (243) as an interpretive principle governing the assignment of truth values to sentences whose subjects are plural NPs:

243) (=his (17))

[NP VP] is true whenever the
S plural
plurality P denoted by NP admits a partition
C such that [A is VP] is true for every A in C.
S

To give one of his examples, consider (244) and (245) (his (13) and (14)):

244) Gilbert and Sullivan wrote operas

245) Handel and Wagner wrote operas

Both of these sentences are true, but in different ways. The pair of men Gilbert and Sullivan together wrote operas, while Handel and Wagner each, separately, wrote operas. This ambiguity is captured by (243). the subject NPs of (244) and (245) are plural, and denote two-member pluralities. These each can be divided into two different partitions:

- 246) a. { {Gilbert, Sullivan} }
 b. { {Gilbert}, {Sullivan} }
 247) a. { {Handel, Wagner} }
 b. { {Handel}, {Wagner} }

The sentence (244) is true, since the property represented by wrote operas holds of each member (the sole member) of the partitioning (246a). (245) is true since the property holds (separately) of each member of the partitioning (247b). Higginbotham terms partitions of the type (246a, 247a) holistic, those in (246b, 247b) atomic. A holistic partition has the entire plurality as its sole member; an atomic partition is a set of singletons, with each member of the plurality a separate member of the partition.

In these cases, the division of the plurality into partitions is free. However, other properties of the structure may dictate a particular choice for the plurality. One such case, as Higginbotham (following Fiengo and Lasnik) observes, is when the VP contains a reciprocal expression. The essential idea is that in a sentence conforming to the schema (249), the presence of the reciprocal forces the members of the partition of the plurality denoted by the subject to have at least two members.

- 248) John and Mary love each other
 249) [NP [V reciprocal]]
 S plural VP

For the sentence (248), this will force the partition to be { {John, Mary} }, disallowing {John}, {Mary}.

To return to adverbial each, Gillon demonstrates that this adverb also imposes a restriction on the partitioning of plurality associated with the the subject, but a different one than each other does. Consider (250):

250) Mary and Sue each wrote operas

(250) is true just in case Mary wrote operas, and Sue also wrote operas. Thus each forces the choice of partition to be such that each member of the partition is a singleton (Higginbotham's atomic partition).

59

Thus, Gillon observes , (239) is semantically ill-formed. The adverbial each requires the partition of the plurality denoted by the subject to be an atomic partition, while the reciprocal each other requires each member of the partition to

60

have at least two members. .Let us descriptively term this

59. Essentially the same point is made by Fiengo and Lasnik (1973).

60. More formally, Gillon takes verbs to be generalized propositional functions, taking values in some plurality cover of the set denoted by the subject. A cover of a set Y is a family X of subsets of Y, such that each member of Y is in at least one of the subsets. A partition of Y is a cover of Y, where each member of Y is in exactly one set in X (that is, no member of Y is in two sets in X). Finally, a plurality cover of Y is a cover X such that no set in X is a subset of another. See Gillon (1984; chapters 4, 5) for details.

The adverbs all, both, and each restrict the interpretation of the relevant subject, each by requiring the function to take singleton sets as values. The interpretation of the reciprocal argument each other, however, requires that the function assigned to the verb take as values sets which can be

semantic conflict each-conflict.

We observed in the preceding subsection that the antecedent clash effect is suppressed in certain connectivity cases. We shall see here that the same is true of the type of semantic ill-formedness illustrated by (239). When each other is moved out of the domain of each, the sentence becomes acceptable; I will argue that this is a direct result of the enriched possibilities for assigning an antecedent to the moved anaphor.

3.7.6 Suspension of the Each-Conflict

First observe that the conflict just reviewed does not hold simply in sentences of the exact form (251); (252) is as anomalous as (239).

(251) [NP each [V each other]]
 IP plural VP

(252) * They each painted pictures of each other

Let us suppose that the adverbial each links to the subject of its clause; this gives (252) the linking (253) ⁶¹.

partitioned into at least two distinct sets, i.e. non-singleton sets. Thus (239) is semantically ill-formed, because it cannot meet the conflicting requirements of the two each expressions.

61. This assumption is not crucial to the argument, but it makes formulation of the interpretive rules relatively straightforward. Suppose we abandon the idea that adverbial each actually links to the subject of its clause. Then, we could recast (254) as (254'):

253) * They each painted pictures of each other



It is then natural to view the constraints imposed upon the subject by the two each elements as to be conditions on their antecedents. More explicitly,

254) for a plurality P denoted by an NP X, to which each is linked, the values of X must be taken from the atomic partition of P.

255) for a plurality P denoted by an NP X, to which each other is linked, the values of X must be members of a partition Q

254') for a plurality P denoted by an NP X, if each is a constituent of the clause of which P is the subject, the values of X must be taken from the atomic partition of P.

(254) would seem to be superior to (254'), given the appearance of each in such cases as (i)-(iii) (adapted from Burzio (1981)):

- (i) I assigned the diplomats [one interpreter each]
- (ii) One interpreter each was assigned to the diplomats
- (iii) * One interpreter each took the tests

The plural NP with which each is associated in (i) is not in subject position, thus (254') would fail to capture the fact that each here has the same effect as in (250), requiring that we understand [the diplomats] on the individuated reading.

The treatment of each as an anaphor, implicitly assumed in the text, seems well-supported by the facts of (ii) and (iii). In (iii), the plural NP does not c-command each, thus the linking of the latter to the former would presumably violate the binding theory. In (ii), although each is not c-commanded by the diplomats, the NP one interpreter each moves from VP-internal position. This is somewhat reminiscent of the facts concerning psych-movement discussed above, and may be reducible to a modification of the framework developed here. I leave such considerations for further work.

of P such that each member of Q has at least two
62
members .

Thus we expect that the conflict illustrated in (239) and (253) will arise only when each and each other have the same proximate antecedent. Thus (256) is well-formed.

256) [John and Mary] each said that they love each other



i.e. "John said that [he loves Mary and Mary loves him], and Mary said that [she loves John and he loves her]."

Now, consider the contrast between (257) and (258) (recall also (214-217)).

257) * John and Mary said that they each admire
each other's parents

258) which of each other's parents did John and Mary
say that they each admire?

62. (255) follows from the truth condition formulated by Fiengo and Lasnik (1973) for sentences containing reciprocals bound by non-quantificational antecedents (elementary reciprocal sentences). (The formulation is quoted from Higginbotham (1981)):

- (i) for a plurality P having at least two members, and R a relation, P is R-reciprocal if any two distinct members of P stand in the relation R.
- (ii) An elementary reciprocal sentence correlated with the relation R is true if there is a partition Q of the plurality P denoted by the antecedent of the reciprocal such that every element of Q is R-reciprocal.

From (ii), a truth value can be assigned to an elementary reciprocal sentence if the partitioning of the plurality denoted by the antecedent (the subject, in my cases) is non-singleton.

(255) is phrased the way it is to minimally contrast with (254).

(A: John said that they each admired Mary's mother, and
Mary said that they each admired John's father.)

(257) is unacceptable for the same reasons as (253); both each
and each other are linked to they, resulting in semantic
conflict.

On the other hand, in (258) WH-movement of which of each
other's parents allows each other to link directly to John and
Mary (such a link in (257) would be syntactically ill-formed),
in virtue of one of the intermediate traces. The potential
answer given in (258) fits with our analysis; each other takes
John and Mary as its immediate antecedent, hence forcing the
interpretation under which John and Mary said independent
things.

258) which of each other's parents did John and Mary say

[t_i that [they each admire e_i]]?
CP C' IP

(answer: "John said that they each admired Mary's mother,
and Mary said that they each admired John's
father")

Further, we can now see the effects of the theta-compatibility
requirement in sentences with each. Consider the sentence
(259):

(259)

John and Mary wonder [how proud of each other] they ought to be

The linking shown is the only one consistent with the theta-compatibility requirement. (Linking each other to John and Mary violates the binding theory, since the accessibility sequence through which this would be permitted is ill-formed.)

The linking in (259) allows the interpretations in (260).

260)a. John wondered [how proud of Mary he ought to be and how proud of him Mary ought to be], and Mary wondered [how proud of John she ought to be and how proud of her John ought to be].

b. John wondered [how proud of Mary he ought to be], and Mary wondered [how proud of John he ought to be].

This is the same ambiguity seen in (191), due to the two ways of assigning a value to they. Crucially, in both interpretations of (259), they is the semantic antecedent of each other; this is so because the WH-phrase is a predicate, forcing the linking in (259), and the anaphor's value is determined in relation to the NP to which it is linked.

We note a sharp contrast between (261) and (262):

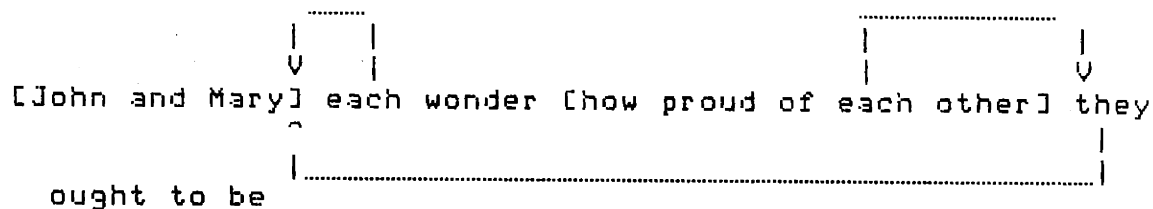
261) John and Mary each wondered [how proud of each other] they ought to be

262) *John and Mary wondered [how proud of each other] they each ought to be

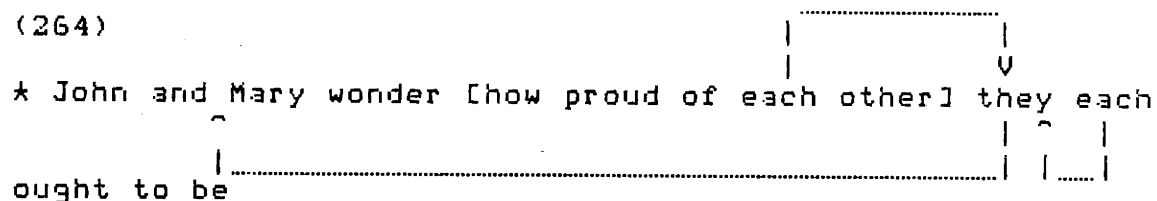
This is expected, given the formulation (254,255) of the

constraints that each and each other place on the interpretation of their antecedents. These two elements place semantic constraints on their proximate antecedents, that is, the NPs to which they are linked. The binding theory, crucially involving the theta-compatibility requirement (4iv'), forces the linking to be as in (263) and (264):

(263)



(264)



In the ungrammatical (264) each and each other have the same proximate antecedent, hence the conflict between (254) and (255). In (261, 263), they do not share proximate antecedents, hence the acceptability of the sentence.

This suppression of each conflict in such examples as (258) provides a powerful argument against any treatment of anaphor connectivity involving reconstructive lowering of the anaphor at LF which did not also involve paths (our accessibility sequences) connecting an anaphor and its antecedent; such a treatment could not in any way handle the distinctions between

(258) and (262).

Thus we see that the use of the linking formalism in specifying the binding theory, the semantic principles assigning values to anaphors, and constraints on assignments of values to anaphors (the antecedent clash prohibition (205)) and to plural NPs ((254) and (255)), taken together with the chain-accessibility framework to binding, allows a quite simple account of a wide range of data. Formulation of the same principles in indexing theory -- even with the chain-accessibility formalism -- leads to redundancy and great complexity.

As a final comment, recall the discussion of successive cyclic binding developed earlier; recall that we saw that it makes incorrect empirical predictions. It should be noted that the facts of reciprocal interpretation discussed above have non-trivial implications for this view of binding, even if it could be maintained. The basic motivation for successive cyclic application of the licensing condition for anaphors and pronominals is that we could retain the strict c-command requirement between an anaphor and its antecedent. That is, we could adopt Chomsky's (1985) formulation of this condition, without the technical amendments necessary within the framework (4-8) developed here.

Our conclusion, based on such examples as (257 vs. 258), and (214-217), is that when the rules of interpretation apply to a

sentence containing a reciprocal, certain information has to be recoverable from the representation - chiefly, which NP satisfies the binding theory for the reciprocal. We have seen that such recovery follows directly from the linking theory, and can also be recovered -- albeit with a great deal of redundancy -- in indexing theory as well. The information is recoverable in indexing theory only if we adopt the essential chain-accessibility framework, allowing the semantically and syntactically relevant antecedent of the reciprocal to not c-command it.

In the successive-cyclic view of binding, the NP which c-commands the reciprocal anaphor must c-command it when the feature of +B is assigned. This means that we will have also

63

assign values to the reciprocal successive cyclically as well , deriving a semantic representation of a sentence in parallel with the syntactic derivation. While this is not incoherent, it

63. This is so because the successive cyclic view is based within indexing theory. We saw above that, within indexing theory, we can recover which NP was the binding-theory-satisfying binder of the anaphor if we build chain accessibility into the semantic rules of reciprocal interpretation. This option is inconsistent with the successive cyclic view, since this position on anaphor binding is based on the hypothesis that the anaphor must actually be c-commanded by its antecedent when the binding theory licenses it. Since the +B-assignment framework does not encode which NP it was that the anaphor was bound by, and since this framework also does not acknowledge anything like chain-accessibility, then the remaining option must be adopted: for an anaphor A, A is syntactically licensed (by the binding theory) and semantically interpreted at the same point in the syntactic derivation, when the anaphor is locally c-commanded by a coindexed NP.

falls outside the basic GB model, in which rules of semantic interpretation feed off LF representations, which in turn are linearly derived from S-Structure.

3.8 Layered Traces

We shall now consider one final approach to the anaphor connectivity effect, one which involves the notion 'layered trace'. On this view, movement leaves a phonetically null category, but one which has all the internal structure of the moved expression. This is the approach advocated by Rizio (1981) and Longobardi (1979), among others.

Movement, on the standard trace theory view, involves two operations: copying the moved phrase into the landing site (such copying taking place into a base-generated position, thus constituting substitution, or else creating a new position, through the restricted process of adjunction), and replacing the target item with a trace. Trace is typically considered to be a phonetically null category of the same syntactic category (NP, VP, etc.) as the moved item, which bears certain features (+/-anaphor, +/- pronominal, [-WH], agreement features, etc.). A typical case is illustrated in (265-7).

265) [e [[John [INFL [likes
 CP C' IP I' VP
 [which pictures of himself]]]]] ----->
 NP

On the layered trace view, (270) has the S-structure (271):

271) * Who do you wonder [who[Mary loves who]]?
 +D +D

Thus selection features will have to be restricted to the heads of chains. Similarly, the layered trace view raises the question of which of the three instances of who in (271) is the operator; the answer is that only the highest one, again, the head of the chain, is the operator. As long as the theory recognizes chains as formal objects, no insurmountable problems arise.

Consider again the connectivity example (268).

268) [[which pictures of himself] [[John [INFL
 CP NP C' IP I'
 [[which pictures of himself] [likes [which
 VP NP +D VP NP
 pictures of himself]]]]]
 +D

On the layered trace account, the fact that one of the three anaphors is bound is sufficient to satisfy the Binding Theory for both. One difference between the two systems is that the layered trace approach allows the Binding Theory to require that an anaphor be c-commanded by its antecedent. This in turn allows the binding theory to be put in terms of indexing, instead of linking, since the special relation of antecedence is defined in terms of c-command, in indexing theory.

This account requires some further formalization, offered below. As we shall see, this approach is empirically equivalent

to the chain accessibility sequence framework, insofar as both theories are capable of covering all the cases of anaphor binding in which the overt anaphor is not c-commanded by its antecedent.

The facts about reciprocal interpretation considered in the preceding section will show that the notion chain still must be formally referred to in the statement of the 'layered trace' binding theory. I shall discuss the exact form that the various parts of reciprocal interpretation will have to take.

Consider the contrast between (272) and the ungrammatical (273) (I use indexing for clarity):

272) [CP [NP which pictures of himself] [C' [IP John [INFL I']]]]
 [VP likes [NP which pictures of himself']]]]]
 +D

273) * [each other _i 's mothers] said that Mary believes that the men _i
 like each other' _i

In (272), we have a pair of anaphors, which together satisfy the Binding theory, as one of them is bound in its governing category by John. We might conjecture the following:

(274) For a set T of coindexed anaphors (A_1, \dots, A_n) , T is licensed iff some A_i in T is bound in its binding domain (the least CFC in which there is a possible BT-compatible indexing for A_i).

X binds Y iff X c-commands Y , and X and Y are coindexed (Chomsky (1981)).

This condition is overly permissable; it allows (273) to be grammatical, since one of the pair (each other, each other'), namely each other', is bound in its binding domain. The clear difference between the two is that in (272), but not (273), himself is a copy of himself'. (Copying takes place as part of movement.) This difference must then be formally reflected in the definition of T in (274). The way to do this is to explicitly acknowledge the copying:

(275) Definition

$T = (X_1, \dots, X_n)$ is a copy-set iff every X_i
⁶⁴
 and X_j in T occupy the same position in members of a chain.

Insofar as this definition makes no reference to coindexation, it will carry over to the linking framework, as well. In the linking system, one member of a T-set will have to link to an NP which c-commands it is its governing category.

To illustrate, the relevant chain for (272) is ([which pictures of himself], [which pictures of himself']_{+I}). The two anaphors in (272) constitute a copy-set, since they each occupy the same position (object of of) inside the WH phrases in which they occur.

This suffices to distinguish (272) from (273). In (273), the

64. Or address, in the sense of Vergnaud (1982).

two anaphors do not constitute a copy-set, since ([each other's mothers], [each other]) is not a well-formed chain, nor is (each other, each other'). Hence, the two anaphors must independently satisfy the Binding Theory. each other in (273) is free in violation of the relevant part of the Binding Theory (274).

3.8.1 Reciprocals and Layered Traces

Now let us consider the import of reciprocal interpretation, in particular the two constraints of each-conflict and antecedent clash, for the layered trace view.

As argued above, such contrasts as (276) and (277) indicate that the rules of semantic interpretation must be able to read off the syntactic representation which of the NPs in the sentence is the proximate antecedent of each other and each. That is, it must be discernable from the LF-structure which NP served as the binding-theory satisfying antecedent of the anaphoric elements. Given the simplicity of doing so in the linking theory, and the redundancy and complexity of indexing theory, we argued that linking is to be adopted.

(276)

$$\begin{array}{c} \text{---} \\ | \quad | \\ \text{V} \quad | \\ | \quad | \\ \text{---} \end{array}$$

*Mary wondered [which of each other's friends] they each like

$$\begin{array}{c} | \\ \text{---} \end{array}$$

(277)

John and Mary wondered [which of each other's friends] they each like

The same fact is of course true in the layered trace framework.

(276)-(277) will have this structure.

278) Mary wondered [which of each other's friends] they each
liked [which of each other's friends']

+D

279) John and Mary wondered [which of each other's friends] they each
liked [which of each other's friends']

+D

Here, the problem takes on a different character. Since we have two reciprocal NPs, the members of the T-set (each other, each other'), we need to keep track of which one of the anaphors is semantically active. What we will need to say is roughly the following: given the T-set (each other, each other'), one of the members has its value fixed by an anaphoric relation to another NP; then this value is shared by the other members of the set.

The basic motivation for the layered trace approach to binding is to be able to retain the requirement of strict c-command between an anaphor and its binder. This is seen as an advantage of the system. This is really the entire motivation for layered trace, and we may suppose that the layered trace framework stands or falls depending upon whether this strict c-command hypothesis can be maintained.

In this spirit, the interpretive rules for anaphors are to be

stated in terms of strict c-command; an anaphor will be assigned a value determined by the value of an actual binder of the anaphor (as opposed to a non-coccommanding, chain accessible NP, as in the framework I develop above). With this assumption, we may formulate the interpretive rules as follows:

Anaphor Value Assignment

(280) for X an anaphor, and a member of a T-set I, and Y, a local binder for X, assign values to the members of I as follows:

- i) if X = reciprocal, for each a such that $v(a, Y)$ [eg. a is among the values of Y; AB], the values of the members of I are those b such that $v(b, Y)$ and $b \neq a$.
- ii) if X = reflexive, PRO, or trace, for each a such that $v(a, Y)$, the values of the members of I are those b such that $v(b, Y)$ and $b = a$.

To quickly illustrate, (281) is converted to (282); each other is taken as X in (280), and the value of each other' is given for free, since the two reciprocals form a T-set.

281) [which pictures of each other] did John and Mary like
[which pictures of each other']
+D

282) [which pictures of John] did Mary like [which pictures of John] and [which pictures of Mary] did John like [which pictures of Mary]

3.8.2 The Blocking Effects

We have seen that there are two 'blocking effects' for reciprocals: the first is antecedent clash, which prevents a reciprocal from having a proximate and evaluative antecedent which are not identical in value; and each-conflict, which

prohibits adverbial each from having the same proximate antecedent as each other, due to conflicting requirements they place on their antecedents. Semantically, it is the proximate antecedent which fixes the value of the reciprocal items.

In the layered trace view of binding, formulated in terms of indexing, 'proximate antecedent' is equivalent to local binder, a c-commanding coindexed NP in the GC for the anaphoric element. If we are forced away from this equivalence, there is no reason to adopt the layered trace framework.

Thus in the layered trace view, the semantic constraints must be stated in terms of the values of local binders, for the layered trace framework to have any advantages over the chain-accessibility approach.

For non-connectivity cases, in which the T-set for any anaphor has one member, we may state the blocking effects as follows.

Prohibition Against Antecedent Clash

- 283) for an anaphor X, X must have a local binder W whose value is identical to that of X's evaluative antecedent (highest binder).

Partition Constraints

- 284) for a plurality P denoted by an NP X, by which each is locally bound, the values of X must be taken from the atomic partition of P.
- 285) for a plurality P denoted by an NP X, by which each other is locally bound, the values of X must be taken from a partition Q of P such that each member of Q has at least two members (i.e. a non-atomic partition).

However, as soon as we consider cases of connectivity, in which the T-set will have multiple members, we must revise these constraints on a par with (280).

We have seen a suspension of each conflict in cases like (287).

286) * John and Mary think that they each resemble friends of each other

287) John and Mary wonder which of each other's friends they each resemble

(i.e. John wonders which of Mary's friends they each resemble, and Mary wonders which of John's friends they each resemble)

Within the framework developed in the preceding section, this is due to the fact that in (287) each other is linked to (and hence has its value fixed by) John and Mary, while each is locally linked to they. Hence there is no each conflict, and the indicated interpretation of (287) is permitted.

In the layered trace framework, (287) has this representation:

288) John and Mary wonder which of [[each other] 's friends] they each resemble which of [[each other'] 's friends]

The T-set contains (⁶⁵each other, each other') .

65. The option of leaving traces at each landing site could enrich the structure, and thus extend the T-set. For example, a layered trace could be left adjoined to the lowest VP, following the proposals of Chomsky (1986) that WH-movement may involve VP-adjunction. This doesn't matter here, and I will keep to the simplest possible structure.

Notice that in the lower clause we have the same configuration as in the lower clause of (286). The presence of each and each other in the same clause gives rise to each-conflict in (285), but not in (288). It therefore appears that [each other] doesn't count, in some formal sense, for the partitioning constraint (285), for the interpretation of the sentence (287).

The simplest way to handle this is to suppose that the value of the T-set (each other, each other') in (288) is determined in virtue of the local relation between each other and John and Mary. This will amount to basically the same characterization as given in section 3, in the linking approach. Thus the reciprocal in the lower clause of (288) is not necessarily relevant to the assignment of values to the T-set, while in (286) it is, since it will be the only member of its T-set.

More formally, we may revise the rules given above to express this.

Anaphor Value Assignment

(289)

(A) for X an anaphor, and a member of a T-set I, and Y, a local binder for X, assign values to the members of I as follows:

- i) if X = reciprocal, for each a such that $v(a, Y)$, the values of the members of I are those b such that $v(b, Y)$ and $b \neq a$.
- ii) if X = reflexive, PRO, or trace, for each a such that $v(a, Y)$, the values of the members of I are those b such that $v(b, Y)$ and $b = a$.

(B) Term X the designated member of its T-set.

Partition Constraints

290) for a plurality P denoted by an NP X, by which each is locally bound, the values of X must be taken from the atomic partition of P.

291) for an instance of each which functions as the designated member of its T-set:

for a plurality P denoted by an NP X, by which each other is locally bound, the values of X must be taken from a partition Q of P such that each member of Q has at least two members (i.e. a non-atomic partition).

(289A) is the same rule as (280); (289B) adds a new definition, giving specialized status to the member of a T-set whose value is externally fixed, and whose value sets the value of the entire T-set. (291) is the constraint (285), amended to make use of this notion. (290) is unchanged.

Returning now to the grammatical sentence (288), with the T-set (each other, each other'), we see that no each-conflict arises, because each other is the designated member of its T-set, and so places its restrictions on the interpretation of John and Mary, while each places its restrictions on they.

3.8.3 Moved Predicates

We have further seen that the each-conflict effect is retained with predicate movement:

292) *John and Mary know that they have each become fond of each other

293) * John and Mary know how fond of each other they have each become

On the account developed earlier, this is because the predicative nature of the WH-phrase forces each other to take they as its semantic antecedent, not John and Mary. Thus both the reciprocal NP and each are related to the same subject, and each-conflict arises.

(293) has this structure on the layered trace view, with indexing assigned:

(294) * [John and Mary]_i know how fond of each other_i they_i
have each become how fond of each other'_i

(Again, the prime notation is used simply for clarity). Within the layered-trace formulation of reciprocal interpretation, we have two possibilities of assigning values to the I-set -- either each other will be interpreted with respect to John and Mary (its local binder), or each other' will be interpreted with respect to they (its local binder). However, we must bar the first possibility, since otherwise there would be no each-conflict in the sentence.

Since we have two anaphoric NPs in (290), each other and each other', it appears that each conflict arises because each other' is forced to count as the designated member of the I-set (each other, each other'). Then, each other' is the member of the I-set which is initially assigned a value, and whose value becomes that of each member of the I-set.

This is different than what happens in (287-288), since the suspension of each-conflict must, on the layered trace view, be due to the possibility of each other (the member of the T-set contained in the moved WH-phrase) being treated as the designated member of its T-set, thus assigned a value in virtue of being locally bound by John and Mary, and so assigning values to the T-set.

Therefore, the predicate/ non-predicate distinction must be acknowledged in the layered trace view, specifically in the definition of 'designated member of a T-set'. Keeping to the restriction that an anaphor's syntactic and semantic antecedent must c-command it, this leads to the following amendment to (289).

(289)

(A) for X an anaphor, and a member of a T-set I, and Y, a local binder for X, assign values to the members of I as follows:

i) if X = reciprocal, for each a such that $v(a, Y)$, the values of the members of I are those b such that $v(b, Y)$ and $b \neq a$.

ii) if X = reflexive, PRO, or trace, for each a such that $v(a, Y)$, the values of the members of I are those b such that $v(b, Y)$ and $b = a$.

(B) Term X the designated member of its T-set.

(C) for every node ^jN dominating X but not dominating Y, if ^jN bears a thematic grid G, then either G is saturated or there is another node ^kN such that:

(i) ^kN dominates both X and Y, and

(ii) the thematic grid of N^k is saturated, and

(iii) the thematic grids of N^k and N^i are identical in all other respects.

(289C) encodes the same restriction as the theta-compatibility requirement (4iv'), although (289) still requires that the anaphor be c-commanded by Y. With the addition of (289C), only each other' can be the designated member of its T-set in (294), leading to each-conflict.

Thus the layered trace approach to anaphor connectivity, which might be viewed as an alternative to the chain-accessibility sequence framework developed in this chapter, must build the basic notion of a theta-compatible path into the semantic component. This will ultimately give rise to the same type of redundancy between the syntactic and semantic constraints on anaphors which we earlier identified for the indexing-theoretic formalization of chain-accessibility (see sections (3.4.3 and 3.4.4)). Similar problems arise in consideration of the suspension of antecedent clash in connectivity examples with moved non-predicates, but I will not consider those in detail here.

The layered trace approach to anaphor binding, formulated in indexing theory, is motivated by the possibility of retaining the strict c-command requirement for anaphor binding, even for cases of connectivity. Although this requirement can be

maintained, it is done at the cost of building into the semantic component the notion of theta-compatible accessibility sequence, and therefore it seems that nothing is to be gained by this approach.

Appendix 1

In this appendix, I give the indexing-theoretic formulation of the licensing condition for anaphors and pronominals, within the framework of chain accessibility.

1) Chain Accessibility Sequence (definition)

S is a well-formed chain accessibility sequence only if:

- i) A is a_1
- ii) a_2 is a projection of the governor of A
- iii) for every pair (a_i, a_{i+1}) , either (1) or (2):
 - 1) a_{i+1} immediately dominates a_i
 - 2) (a_i, a_{i+1}) is a link of a chain
- (iv') for every a_i in S , if a_i bears a thematic grid G , either G is saturated or there is a member of S a_j which bears a saturated grid G' identical to G in all other respects.

8) Chain Accessibility (definition).

B is chain accessible to A through an accessibility sequence $S = (a_1, \dots, a_i, \dots, a_n)$ such that:

B is a sister to some a_j in S .

(3) Binding Theory Compatible Indexing (Definition)

I, a set of assigned indices, is binding-theory compatible wrt. an expression A and an accessibility sequence S for A iff:

- i) for A = an anaphor, A is coindexed under I with an expression B which is chain-accessible to A through S.
- ii) for A = pronominal, A is not coindexed under I with an expression B which is chain-accessible to A through S.

(4) Licensing Condition

for A = anaphor or pronominal, A is licensed only if:

- i) there is a chain accessibility sequence S for A, such that there is assigned a BT-compatible indexing I for (A, S); and
- ii) there is no proper subsequence S_j of S, such that S_j is an accessibility sequence for A, and there is a possible BT-compatible indexing I' for (A, S_j).

Definition (1) is unchanged from that in the linking framework.

66. Equivalently, B is immediately dominated by some a_j in S, and B is not any A_k in S.

Appendix 2: The Problem of Pseudo-Clefts

In the main body of this chapter, the connectivity of specificational pseudo-clefts (SPCs) is briefly considered. I offered an analysis there which asserts that an A' chain associates the focus phrase with the variable inside the subject, thus accounting for the connectivity in terms of chain accessibility. In this appendix, I look at such sentences in more detail, and conclude that (i) my analysis is insufficient; and (ii) two other analyses fail as well to account for all the properties of these constructions. As I will show, drawing heavily upon Chapter 6 of Higgins (1973), SPCs have a variety of properties which distinguish them from all other connectivity constructions. I will not be able to fully account for them, but I hope that the review of these properties will help to delimit the problem for the reader.

This appendix examines the various odd anaphoric properties of such sentences, and suggests three possible ways to view the problem. At this point, none of the analyses is fully satisfactory, but the discussion lays out the nature of the problem in more detail than in the chapter itself.

As Higgins and Akmajian note, PCs (of the approximate form in (1)) have two quite distinct interpretations, the predicational and the specificational. I will concentrate for the most part in this section on PCs in which the focus phrase is an AP,

Following this analysis, the structure of the PC in (2) giving rise to the predicational interpretation will be:

(5) [what [John is]] is
i

[_{AP} t [_i [_{AP} [_{A'} important [to me]]]]]

The free relative subject is raised from inside the small clause; its trace is theta marked by important. Let us assume for concreteness that the free relative is dominated by an NP node (ignoring the obvious problem that there is no N head).

What is the internal structure of the free relative of (5)?
Two possibilities come to mind.

(6) [what [[John [is [

CP j C' IP I' AP

e]]]]

j

(7) [what [[John [is [CP j C' IP k I' AP [t [e]]]]] k AP j

That is, it is possible that the what either binds an EC which constitutes the entire AP, as in (6), or it binds an EC which constitutes a subpart of the AP, as in (7).

(8) is grammatical, and presumably may have the structure (9) (cf discussion in this chapter).

8) How nice do you consider John?

9) [CP AP A' i C' do IP I'
[consider John t]]]]
VP AP i AP i

This means that WH-movement can move a subpart of AP, leaving its subject behind, exactly as in (7). Further, in (7) but not (8), the non-theta-marked phrase John binds a theta-marked trace, fulfilling the theta-criterion. I therefore assume that (7) is at least a possible structure (and a grammatical one).

Thus the full structure of (5) may be taken to be (10).

(10) [what [[John [is [
 CP j C' IP k I' AP
 [t [e]]]]] is [t
 k AP j AP i
 [[important [to me]]]]
 AP A'

Two fundamental questions about specificational PCs must be confronted. First, what permits the specificational interpretation to exist at all? Second, what is the proper account of connectivity in such structures? How is it related to the specificational interpretation?

In the next section, I summarize the syntactic properties of SPCs, none of which are shared with predicational pseudo-clefts (PPCs). I will then turn to possible accounts of SPCs, and the connectivity and other effects which occur therein.

3.9 Properties of SPCs

3.9.1 Anaphora

First, anaphors inside the focus phrase may have as antecedents those NPs which locally c-command either the variable inside the subject clause (12), or NPs which locally c-command intermediate traces associated with the WH phrase.

11) what John saw was a picture of himself

what the men admire are each other's plays

12) what John wants Mary to paint are pictures of himself

what Joyce and Shaw believed I like are each other's plays

The latter cases exemplify the 'multiple binding-domain effet' discussed at length in chapter three. (The grammaticality of such examples was challenged by Szabolcsi (1986), but for the several informants I consulted they are acceptable.)

Second, pronominals shallowly embedded within the focus must be disjoint in reference from NPs locally c-commanding the variable in the subject.

13) what John is is proud of him (disjoint reference forced)

14) what John is is important to him (coreference permitted
only on the predication interpretation)

Third, R-expressions in the focus must be obviative from pronouns (or anything else) c-commanding the variable:

15) a. what he is is proud of John (disjoint reference forced)

b. what he believes is that John is intelligent (disjoint
reference forced)

- 16) what he is is important to John (coreference permitted only on the predicational interpretation)

Let us consider this last set of facts, which seem to indicate that condition C applies to force disjoint reference between John and he even though the pronoun does not c-command John. We adopt the position that Condition C applies at S-structure (and perhaps at LF as well), following Chomsky (1981). This means either that in the structure of (15), Condition C is violated, or else (15) is transformed into an LF structure which violates Condition C.

Suppose we adopt the analysis of SPCs suggested in the thesis, in which the focus is the head of a chain terminating in the variable. We might then suppose that the Condition C effects are due to the fact that since he c-commands the variable, John must be obviative from it. In the terms of chapter 3, there will be a chain accessibility sequence through which he is accessible to John, and Condition C will be put in the following rough way:

- (17) An R-expression is obviative from any chain-accessible NP.

However, examples like (18) suggest that this is incorrect:

- (18) Which pictures that John bought does he like t ?

Here, he is chain-accessible to John, yet coreference is allowed.

As Howard Lasnik observes, coreference is barred in the

following case:

19) how proud of John is he ?

Given the raising analysis of predicative sentences, (19) might have the structure (20) or (21):

20) [[t [how[proud of John]] is
CP AP i AP i
[he [e]]]
i AP

21) [[how[proud of John] is
CP AP i
[he [[t [e]]]]
i AP i AP

In (21), Condition C in its standard formulation is violated, but it is not in (20), which we argued is an available structure. Further, the same lack of coreference is present in (22):

22) a. how proud of John do you consider him?

b. how proud of John does he believe Mary to be ?

there is no possibility that in (22a) a trace is present inside the moved AP. And in (22b) even if there were a trace inside the AP, it would be coindexed with Mary, not John. Therefore, some extension of Condition C is required.

In Chapter 3, I introduced the Theta-Compatibility Requirement (4iv') on accessibility sequences, to account for the fact that an anaphor inside a moved predicate must be bound to the lowest potential antecedent. We can account for (22), and for (19), if we supposed that Condition C is to be formulated in terms of

Since the focus is not a predicate, the accessibility sequence can go directly up through the tree to the matrix node, and the pronoun will not be accessible to the R-expression.

In addition to these facts concerning anaphor, pronoun, and R-expression relations, pseudo-clefts also have several other properties which require attention. First, a pronoun in the focus position can be understood as a variable with respect to a QNP which c-commands the gap in the subject:

25) what every man likes is his car

26) what every man does for a living is earn his pay

27) what some guy did yesterday was eat his cornflakes

Finally, a negative element in the subject can license polarity any in the focus, again if the negative element c-commands the gap in the subject (thanks to Richard Larson for this observation):

28) what John didn't do was sell any of his cars

29) what John didn't buy was any pictures of Fred

Recall the fact that pseudo-clefts whose focus position is an AP headed by an adjective like important is ambiguous between a specificational and a predicational interpretation. The presence of polarity any, or a QNP-dependent pronoun, in the focus, forces the specificational reading:

30) [what this fact isn't] is important to any analysis of metrical stress

31) [what Mary isn't] is annoying to any of her enemies

32) [what every man is] is important to his mother

This kind of connectivity does not, apparently, occur in other constructions in which anaphor connectivity occurs. In (33), the truly bound-variable interpretation of the pronoun is not available (see Chapter 5 for discussion), and in (34) any is not licensed:

(33) which of his poems did every poet read?

(34) * Pictures of anyone, John didn't buy.

3.9.2 Other facts

Higgins 1976 (chapter 6) notes a number of other odd properties of pseudo-clefts. First, he notes that the entire construction is frozen in the sense that none of its constituents can be deleted, moved, or moved out of.

The focus position of the PC is an extraction island:

35) *who is what John is proud of ?

36) who is what John is important to t? (only the predicational reading)

Further, the whole focus phrase itself cannot be moved, overtly or covertly:

37) ??who thinks that what John is is proud of which woman?

38) who thinks that what John is is important to which woman?
(only Pred interpretation)

Finally, Subject-Aux inversion by itself is impossible:

39) *Is what John is proud of Mary ?

40) Is what John is important to Mary? (pred. reading only)

The facts in (35-6) might reduce to this latter fact, for questions, but Topicalization indicates the same islandhood of the focus:

- 41) Of Mary, John is proud
- 42) *Of Mary, what John is is proud
- 43) To Mary, what John is is important

The wh-clause of the PC cannot be further moved, as Higgins notes.

- 44) *It was what John was that was proud of Mary
- 45) It was what John became that was important to Mary (only pred. reading)

This latter fact may be related to the fact that (as Howard Lasnik observes) the what-clause cannot raise:

- 46) *what John is seems proud of Mary
- 47) what John is seems important to Mary (Pred only)

The copula cannot delete:

- 48) what John is is important and what Bill is is interesting
(both readings)
- 49) what John is is important and what Bill is interesting
(only predication)

Finally, there are odd constraints on the distribution and interpretation of modals and adverbs in PCs.

The PC copula cannot be preceded by a modal, and the focus cannot be adverbially modified:

CASs. (19) and (22) can be distinguished from (18) if Condition C is put in the following way:

- (23) For an R-expression X, X must be obviative from any expression Y unless there is an accessibility sequence extending from X to the matrix IP node through which Y is not accessible.

This way of formulating Condition C also blocks coreference in (24):

- (24) *_i [_i how [_i proud [_i of John]]] does he
 IP AP A' PP
think that you are t_i?

By the theta-compatibility requirement, the CAS must extend from the WH-moved AP to the trace, and from that point on up to the matrix IP node, and he is accessible through this CAS. The shorter sequence {John, PP, A', AP, IP} is ill-formed, since it fails to meet (4iv').

Returning now to the pseudo-cleft case in (15) and (16), observe that (on the hypothesis that a chain relates the focus phrase to the variable inside the subject) the only well-formed chain accessibility sequence connecting John to the matrix node also contains the variable, since the focus is a predicate. Through this sequence, the pronoun is accessible, hence the disjoint reference.

Even this account, however, fails to account for the fact that (15b) is ungrammatical.

- (15b) what he _i believes t is that John _i is intelligent.

50) *what John is may be proud

51) *what John is is probably proud

Negation on the copula is possible only with an implicit contrastive reading:

52) what John is isn't proud of Mary... (but of Bob)

The modal should has two interpretations, the deontic (expression obligation or responsibility), as in (53), and what Higgins terms the emotive, as in (54):

53) John should was his car.

54) It's a pity that John should have washed his car.

The emotive should appears in the immediate clause embedded under an emotive or factive predicate, as in (54). When it is more deeply embedded only the deontic reading is possible:

(55) It's a pity that you said that John should have left early

However, as Higgins notes, should occurring inside the subject of a SPC (but not a PPC) can have the emotive interpretation:

56) It's a pity that Mary should be so proud of Bob

57) It's a pity that what Mary should be is so proud of Bob

58) It's a pity that what Mary should be is important to Bob

The emotive reading of should occurs in (58) only on the Specificational interpretation of the PC.

I believe that the same facts extend to negative polarity:

59) John didn't say that anyone bought a car

60) John didn't say that what anyone did was buty a car

61) John didn't say that what anyone is is important
to Mary

The predication reading of (61) seems completely impossible.

This concludes my summary of the facts; I now turn to three possible ways of approaching the problem, acknowledging that none of them sufficiently handles all the facts.

3.10 Three Accounts

The three accounts I wish to explore are the following: literal reconstruction, in which the focus phrase is put into the position of the gap in the subject, with the various anaphora conditions stated over LF representations; chain formation, along the lines of what was suggested in the text of this chapter; and predication, in which the wh-clause is taken as predicated of the focus, with the binding conditions extended to cover predication structures as well as chain structures.

3.10.1 Literal Reconstruction

In Chapter 2, I argue against Reconstruction as a general approach to anaphor connectivity. The arguments given have chiefly to do with questions, demonstrating that the restricted quantification and operator-variable relations of questions are disturbed in an undesirable way if reconstruction occurs. This

chapter took the most liberal view of reconstruction, namely that it is a consequence of the same type of movement, substitution, and adjunction operations which occur elsewhere in the grammar.

Suppose we take the position that in a pseudo-cleft like (64), the structure can remain exactly as it is, or the focus phrase can be moved into the gap. This derives two representations, (65) and (66).

(64) [what[John is]] is [t [
 i AP i AP
 [important [to me]]]]
 A'

(65) [what[John is]] is [t [
 i AP i AP
 [important [to me]]]]
 A'

(66) [what[John is [t [
 AP i AP
 [important [to me]]]]] is]
 A' i

(65) is no problem; we may assume that the theta-role of important is discharged in the normal manner of being assigned to the trace t which the adjective governs. This gives us the _i predicational interpretation.

For (66), we would have to assume that both what and is delete in the SS-LF mapping.

In order to get the basic meaning of the specificational

pseudo-cleft to come out right, the theta-role of important must be assigned to John; assuming the raising analysis of AP structures, then the trace in the AP in (66) must bear the index of John, not that of what John saw as indicated. Assuming the preservation of indexing across levels of representation, then the trace must be coindexed with John at S-Structure as well. In other words, the specificational interpretation of (69) must have the S-Structure (70), which gives rise to the partial LF (71).

(69) what John is is important to me

(70) [what [John is]] is [t [
 j AP j AP
 [important [to me]]]]
 A'

(71) [what [John is [t [
 j AP j AP
 [important [to me]]]]] is
 A'

Here the first problem arises. The chain (John, t) exists to fulfill the theta-criterion; by the projection principle, the chain must be well-formed at SS as well (it is obviously well-formed as LF). This means, apparently, that the trace must be base-generated inside the focus phrase, and furthermore that the theory of chains must be put so as to license the chain (John, t) at SS, even though there is no c-command between the two members.

The advantage of this reconstruction approach to PCs is that the facts noted earlier are simplified; the condition governing

GNP-pronoun relations, for example, would apply to the LF representation of (72), namely to (73):

(72) [what [every man is]] is [t [important [to his mother]]]]
 A' AP j AP

(73) [what[every man is [t [
 $\begin{matrix} j & & AP & j & AP \\ [& important & [to & his & mother]]]]]] is \\ A' \end{matrix}$

This avoids having to say anything special about pseudo-clefts in the statement of that condition. The correlation of the specificational interpretation and the possibility of anaphoric connectivity in pseudo-clefts is explained; binding can only occur when the focus is reconstructed into the gap, and when this occurs the structure cannot receive the predicational interpretation.

However, the problem surfaces anew in having to broaden the theory of chains so as to allow the existence of the chain (every man, t).

In chapter 2, I additionally gave an argument against reconstruction as a general solution to connectivity in questions on the basis of the multiple binding-domain effect seen in cases like (74):

which picture of himself did John think that Bill liked ?

We noted above, on the basis of (12), that the effect occurs in pseudo-clefts as well, calling into question the adequacy of the

reconstruction account of connectivity in PCs. After reconstruction, the structures will be (75):

12)

what John wants Mary to paint are pictures of himself

what Joyce and Shaw believed I like are each other's plays

75) (what) John wants Mary to paint pictures of himself (are)

(what) Joyce and Shaw believed I like each other's plays(are)

In these structures, we expect an SSC effect, prohibiting the higher subject from binding the anaphor.

For these two reasons, it appears that the reconstruction approach is problematic. We now note a third. The subject of a PC can contain a parasitic gap:

76) [what John painted t without liking e] were pictures of himself

The connectivity here indicates the specificational interpretation, which, on hypothesis, is to be treated by lowering the focus phrase to the position of the gap. In this type of case, we have two gaps, hence we'd need to copy the focus into both gaps:

77) John painted pictures of himself without liking pictures of himself]

Here a problem arises. Parasitic gap constructions can (with some decrease in grammaticality) have lexical subjects, as in (78).

78) [what John painted after Mary commissioned] were pictures of himself

This will be transformed into the LF (79):

79) John painted pictures of himself after Mary
commissioned pictures of himself]

But (79) violates the binding theory, since one anaphor is free in its binding domain. These problems noted, I will now address one of the objections posited by Higgins (1973) against a transformational derivation of SPCs.

Higgins notes that while (80) is acceptable, the corresponding declarative (81) is not:

80) what I like about John is his tie

81) *I like John's tie about him

In a theory in which Deep Structure feeds rules of semantic interpretation, this poses a serious problem for the notion that (80) is transformationally derived from (81), for two reasons. The reason is that the pseudo-cleft transformation must be optional, since otherwise all declarative sentences would have to be pseudo-clefted. Higgins observes that if the ungrammaticality of (81) is attributed to a Deep structure constraint, then (80) should be ungrammatical as well, since (on the transformational view) they share a common Deep Structure.

In the current framework, of course, the ill-formedness of (81) must still be explained, and whatever rules out (81) must not also rule out (80).

We observe the following facts: sentences similar to (80) are perfectly acceptable if the post-verbal position is occupied by

a quantifier, a WH-trace, a WH-in-situ phrase, or (I believe) the trace of a heavy-NP-shifted NP or topicalized phrase.

82) I don't like anything about Fred

I like everything/few things/nothing/ something about fred

83) What do you like about Fred ?

84) who likes what about Fred?

85) ?I like about Fred the fact that he's such a pleasant singer

? His hat, I like about John

On the literal reconstruction view of PCs, (80) is transformed at LF into (81). This means that the characterization of the paradigm in (80)-(85) cannot be stated over LF representations.

The S-Structure generalization emerges as (86):

(86) At SS, the post-verbal position of like must be occupied by a non-referential phrase.

If we were to adopt a non-reconstruction analysis of SPCs, in which the LF representation of (81) is isomorphic to its S-Str representation, we can state the descriptive generalization as follows:

(87) At LF, the post-verbal position of like must be occupied by an A' bound empty category.

Given the facts noted above which argue against the literal reconstruction view, we reject it, and proceed to the other two possibilities.

3.10.2 Chain Formation

The position I explored in Barss (1984), following Cinque (1982), and which I incorporated into this chapter, is that there is a chain headed by the focus phrase which terminates in the subject-internal gap. Thus for an SPC like (88), the chain structure will be as in (89):

88) what John saw e was [a picture of himself]

89) ([a picture of himself], what, e)

Under this analysis, the connectivity of pseudo-clefts follows as a result of the existence of this chain, within the framework of chain accessibility developed in Chapter 3.

Uncontroversially, (what, e) forms a chain as well, formed by WH-movement. The controversial aspect of this analysis is the extension of the chain to include the focus phrase. Accepting this approach requires abandoning the view that chains are formed purely through movement, and requires a careful statement of the conditions on chain well-formedness.

Two of the most widely (although not universally) acknowledged well-formedness conditions on chains are the c-command condition and the binding or antecedence condition. Let us take the latter to require that:

- (90) In any chain C, for every link (a_i, a_{i+1}) ,
a_{i+1} must be anaphorically dependent upon a_i.

This requires the linking structure (91), or the indexing structure (92):

91) [what [John was e]] was [proud of himself]

92) [what _i [John saw e _i]] was [a picture of himself _i]

If we suppose that indices or links can be freely assigned, then no problem arises here.

The c-command condition requires that:

(93) In any chain C, for every link (a_i, a_{i+1}),
a_i must c-command a_{i+1}.

Let us now consider the anaphoric properties of SPCs in light of the proposal that there is a well-formed (extended) A' chain associating the focus with the subject-internal WH-phrase and trace. On this view, an SPC is something like a Tough-movement sentence:

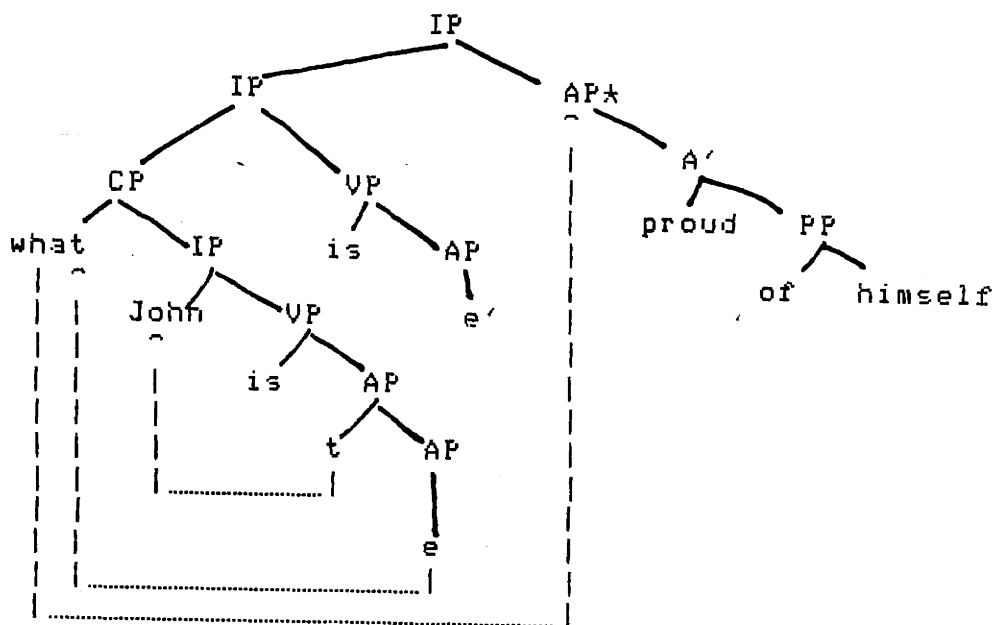
94) [Pictures of himself] are easy [for John] [OP [PRO to like t]]

In Chapter 4, I discuss such structures, and argue that there is a composed chain (pictures of himself, OP, t) formed by adding

the subject to the movement chain (OP, t). John controls PRO, and PRO is chain-accessible to himself through the trace t; thus PRO may be the antecedent of the anaphor. This basic difference between a SPC and (94) is that the operator in (94) is non-overt, and the operator in (91) is overt.

By (93), the focus in (91) must c-command what. This will require that the focus extrapose to IP-adjunction, as in (95), where the AP* has moved from the position marked by e' (this extraposition is obligatory, since e does not itself c-command what).

(95)



Here, AP* c-commands what, and so the chain (AP*, what, e) is well-formed by (93). Himself is chain accessible to t in virtue of this chain (t c-commands e, and himself is contained within AP*).

Now we will consider which of the properties reviewed above are or are not treated correctly in this analysis. First, the anaphor and pronominal connectivity facts are accounted for (11-13), including the multiple binding-domain effects of (12) (recall that the literal reconstruction view cannot capture these latter effects).

The fact that the focus phrase is an island to extraction (35-36) is accounted for by the extraposition analysis (95), since AP* will not be L-marked after movement, and any extraction must then cross one barrier. However, this would class the ungrammatical cases (35) and (36) together with weak subjacency violations, which seems incorrect; these cases are absolutely unacceptable.

The Condition C connectivity effects in SPCs do not hold in other constructions exhibiting anaphor connectivity, like questions and Topicalization. While (96) is ungrammatical if John and he are taken to corefer, (97) and (98) are acceptable.

96) * [what he likes t] are those pictures that John painted yesterday

97) which of the pictures that John painted yesterday does he like t?

98) The pictures that John painted yesterday, he really likes t.

As observed earlier, there is connectivity in SPCs with respect to QNP-pronoun binding and negative polarity licensing (25-32), but this does not occur in other chain structures. (cf 33-34).

The WH-clause of a SPC cannot be further moved, as noted above.

(99) a. [pictures of himself] are [what John bought t]

b. *[what John bought t], pictures of himself are e.

This might be taken as support for the chain analysis, since in (99b) what is not c-commanded by pictures of himself, in violation of (93). However, the AP of a Tough-movement structure can be Topicalized:

(100) a. John is easy [OP [PRO to like t]]

b. [easy [OP [PRO to like t]]], John certainly is .

The contrast between (100b) and (99b) indicates that the treatment of SPCs as involving chain composition (as in Tough-constructions) is incorrect.

As further evidence, consider the fact that there can be no gapping in SPCs (cf. 49). there can be gapping in Tough-constructions:

(101) John is easy to please, and Mary hard to like.

We noted above that a negative polarity item inside the subject of an SPC can be more deeply embedded with respect to its negative licensor than in other constructions, even PPCs (59-61). This fact does not extend to Tough-constructions:

(102) John said that Fred isn't hard to give anything to

(103) *John didn't say that Fred is hard to give anything to

Tough-constructions behave like standard complementation

structures in this regard; the polarity any must be in the immediate clause embedded under negation.

Finally, we note that the presence of parasitic gaps in SPCs is not problematic (as it was for the reconstruction view), since the two gaps will be present at LF, and are licensed by WH-movement internal to the WH-clause.

We therefore conclude that although the chain analysis can account for certain data (the multiple binding domain effect, parasitic gaps) which could not be accounted for under the literal reconstruction view, it nonetheless fails to account for most of the connectivity effects in SPCs, some of which can be accounted for by Reconstruction.

3.10.3 Predication

An alternative analysis for Tough-movement has often been proposed (see Williams (1980), Chomsky (1982)), in which the AP containing the empty operator is predicated of the subject. Although the AP apparently does not assign a theta-role to the subject, it nonetheless contains a variable which is bound by an operator with no lexical content. On the Predication view, this constitutes an open position, qualifying the AP to act as a predicate.


Williams (1983) proposes much the same treatment for SPCs. In this analysis, the WH-clause is taken to be the predicate (open

in virtue of containing a trace bound by a WH-phrase which is not semantically an operator), predicated of the focus phrase.

104) [a picture of himself] was [what John saw t]


 Predication

105) [what John saw t] was [a picture of himself]


 Predication

Let us now see what the Predication view of connectivity would look like, and which phenomena it might account for.

In the framework of chain accessibility sequences developed in this chapter, a CAS for an anaphor can have a link (A, B) if B dominates A, or A and B are members of a well-formed chain. We can extend this, permitting (A,B) to be a link of a CAS if A is the subject of a predicate P, which has open position B.

Thus, on the predication view of SPCs, the chain accessibility sequence for himself in (106) will be {himself, PP, N', NP*, e, VP, IP}. The link (NP*, e) is permitted, with our extension of the definition of CAS sketched above, since NP* is the subject of the predication and e is the open position in the predicate.

106) [a [picture [himself] was
 NP* N' PP
 [what[John [saw e]
 CP IP VP

This approach might extend as well to cases like (107), in which we also get connectivity (see Higgins (1973), Barss (1984)).

107) [John's gift] was to himself

Gift is a nominalization of the verb give, which assigns three thematic roles: Source (John, in (108)), Theme (a book), and Goal (to Mary).

108) John gave a book to Mary

When nominalized, verbs need not assign their thematic roles. We thus assume that in (107), the expression John's book has an undischarged Source thematic role. We can additionally suppose that the NP acts as a predicate, predicated of to himself. the predicate is open with respect to a position in the head noun's theta-grid.

Thus the connectivity of (107) might be brought under the extension of the definition of CASs sketched above (take 1 to be the Source thematic role, and 2 to be the Goal role).

107') [NP<1x, 2, 3> John's [N'<1, 2, 3> gift] was [PP to himself]

The CAS for the anaphor is {himself, PP, 2, N', NP}. The link (PP, 2) is licensed by the fact that PP is the subject of the predication, and 2 is the open position involved in the predication.

This account can take care of the anaphor and pronoun connectivity in SPCs, as well as the existence of parasitic gaps. However, most of the other facts concerning SPCs receive no satisfactory solution in this approach either. For example,

in Tough-constructions --treated the same way as SPCs in this approach - negative polarity any inside the Subject cannot be licensed by a negative element c-commanding the open position:

109) a. John didn't pay attention to anyone's mother

b. * [anyone's mother] is easy [for John [PRO to not pay attention to t]

110) I convinced [noone to tell stories about anyone

* [stories about anyone] are hard [PRO to convince [noone to tell t]

Similar remarks hold for the other connectivity facts in SPCs. The result of the predication analysis is virtually the same as for the chain analysis (with the exception of (107), which receives no treatment in the chain analysis). If we treat SPCs as involving chain composition, we justify this on the basis of other constructions with chain composition -- only to run into the problem that the connectivity of SPCs is, to a large extent, confined to SPCs. If we adopt an alternative account of TM construction and the like as involving Predication, we can treat SPCs in this analysis, only to run into exactly the same problem. To abstract away from the theoretical accounts, specificational pseudo-clefts have properties which appear to be uniquely confined to this one construction.

I simply cannot offer a thorough and satisfying solution to the intriguing set of problems posed by SPCs. Given this, and given that the chain analysis is at least as successful as any other treatment which occurs to me in accounting for anaphor and

pronoun connectivity in SPCs, I keep to the chain analysis suggested in this chapter (and by Cinque (1982), Barss (1984)). Until the properties of SPCs are more fully accounted for, it would be unwise to make modifications to the CAS framework developed here. With the exception of pseudo-clefts and the related type of equative in (107), the CAS framework is capable of handling the complex set of connectivity facts in all other connectivity constructions.

Chapter Four: On the Content of Empty Categories

In the preceding two chapters, we have focused on a particular type of construction, schematically characterized in (1), and attempted to develop a framework in which the grammaticality of the construction, as well as other properties of it, will follow. The presence of the empty category is of considerable importance, because it is only when there is a chain having both the EC and the container of the anaphor as members that the sentence is grammatical; thus the ungrammaticality of the sentence in (3).

1) [...anaphor...] ...NP...EC...]
 |.....|
 c-command

2) which pictures of himself does John like EC?

3)* [each other's parents] promised the men [EC to leave]

In this chapter, we examine more closely the properties of empty categories, specifically their featural content. A range of data will be presented which calls into question the adequacy of the typical GB analysis of ECs, according to which ECs manifest the features +/- anaphor and +/- pronominal, features which determine their referential interpretation and (to some extent) their distribution as well. Much of the data concerns structures such as (1-3); we observe that the ECs in such

structures do not exhibit the behavior which they are predicted to have by the typical GB characterization of ECs and their features.

In light of these facts, I will present two different hypotheses, and suggest reasons to prefer one over the other. I shall draw extensively upon the analyses of ECs (and their anaphoric properties) presented by Brody (1984, 1985), and Chomsky (1981, 1985, 1986). Additionally, I shall rely on the view of 'reconstruction' developed in the preceding chapters. The first hypothesis to be advanced is the Feature Match Hypothesis, which proposes that the features of a trace matches those of the head of its chain. This hypothesis will be motivated and discussed in sections 1 through 7. Certain problems will be observed concerning this hypothesis, which will lead to a somewhat more radical proposal, namely that traces bear no features at all. This will be termed the No Features Hypothesis. The adoption of this proposal bears on the issue of whether indexing or linking is to be preferred as the notation for referential dependence.

In motivating and defending the No Features Hypothesis (NEH), I will develop and extension of the Local Binding Condition on chains discussed in the previous chapter. My extension, which I will term the Chain Obviation Condition, accounts for a variety of facts concerning the behavior of the A-position empty category associated with a left-peripheral phrase, facts which

receive no satisfactory treatment in the standard GB view of empty categories. This condition, among other things, allows the elimination of Condition C as a separate principle of the grammar. Together with the NFH, this obviation condition on chains allows the elimination of the category 'R-expression' from the theory.

In the subsequent sections, a variety of properties of parasitic gaps will be examined, in light of the NFH. Topics to be addressed in this section include: strong and weak crossover effects (and the lack of them) in parasitic gap structures; the lack of anaphor connectivity in parasitic gaps, and the import of this for the analysis of connectivity developed in Chapters 2 and 3, and for the characterization of PGs, Tough movement, and other constructions as involving the movement of a non-overt operator; and the conceptual organization of the binding theory. We will consider the proposals of Lasnik and Higginbotham that obviation is a separate relation from anaphoric dependence, and argue that the parasitic gap data support this view.

It will be seen that the crossover phenomena will motivate the adoption of Higginbotham's Accessibility Condition on operator-dependent pronominals. The final section of the chapter demonstrates that the anti-c-command condition on parasitic gaps derives straightforwardly from the Chain Obviation Condition.

4.1 Review: Empty Category features in GB

A number of the subsystems and principles of government and binding theory have to do with the properties, and distribution, of empty categories. The theory of binding characterizes the possibilities of referential dependence of all governed NPs, including the ECs NP-trace, WH-trace, and the empty pure pronominal pro. The theory of binding also gives the distribution of PRO to a large extent, forcing it to appear in only ungoverned positions. The theory of control, although still a largely unspecified subsystem, seeks to characterize the referential properties of PRO. The Empty Category Principle (ECP) places tight restrictions on the appearance of non-pronominal ECs, and theta-theory also forces a local relation between NP-trace and a local binder. There is no specific 'theory' of empty categories, in the way that there is a theory of Case; there is no autonomous subsystem within the grammar which deals with all and only the properties of ECs. Rather, the term 'theory of ECs' is generally used (as I shall use it) to refer to all the different principles and subsystems of the grammar which are of relevance to the syntactic appearance of ECs.

Within this part of the theory, we can distinguish (at least) four questions. First is the question of existence: are there

empty categories, or is it necessary to formulate a version of GB without them, having the same empirical coverage?

It is important to ascertain that a GB theory with ECs is empirically adequate, internally consistent, and learnable. I assume that the answer to this is yes. I further assume the existence of empty categories (although it is quite possible that one could construct a version of GB lacking them).

I shall assume that the theory of GB, with its commitment to the existence of ECs, is a workable and useful framework, and proceed on to more detailed, theory-internal matters. This brings us to the second question.

This is the question of the characteristic internal properties ECs have. The most obvious of these, of course, is that they lack phonetic content¹. Within the domain of syntactic properties, the most commonly adopted position is that ECs bear some assignment of the features +/- anaphoric, +/- pronominal. I shall refer to these as binding features.

This is not, of course, the only position suggested in the literature (cf. Brody (1985), Bouchard (1984)). It is however, by far the dominant one, and it is this position which I will assume as I develop the argument of this chapter, and it is that

1. But do not lack phonological effects, as the various contraction phenomena of English illustrate. See Chomsky (1981), Jaeggli (1981) for discussion

position with which I will contrast in some detail the position I shall propose. I will call this the standard analysis of ECs.

These two binary features give rise to four distinct internal assignments. Below I give the NP-types associated with each feature assignment in the standard view.

4. 'Standard' GB Theory:

<u>Features</u>	<u>Categories</u>
+A -P	NP-trace; himself, each other (pure anaphors)
-A +P	pro; him, she (pure pronominals)
-A -P	WH-trace; John, the man (R-expressions)
+A +P	PRO: no overt counterpart (due to the Case Filter)

These binding features have a well-defined role: certain well-formedness conditions (the binding conditions, the ECP) apply to expressions bearing certain binding features. Condition A applies to (governed) expressions bearing the feature [+A], and so forth. The assertion that these features are manifested on ECs predicts that they will have the same potential referential dependencies as overt categories with the same feature assignments, since the binding conditions apply to features categories regardless of their phonological content.

It does appear that this approach to ECs goes a long way toward characterizing the structural relations between ECs and thier antecedents, thus providing strong support for the feature theory. The following are typical examples.

5. a. John_i was arrested t_i
 John_i likes himself_i
- b. *John_i seems it appears t_i to have left
 *John_i thinks it appears to himself_i that mary is here
- c. *He_i likes John_i
 *who_i does he_i like t_i ?
- d. John_i likes him_i
- e. John tried_{S'} [PKU to leave]_S

NP-trace and overt anaphors have a similar distribution, both having to be locally c-commanded by a coindexed NP (5a,b). This follows from the classification of both as +A, since on this classification both will be subject to Condition A.

Similarly, in the treatment of Strong Crossover proposed in Chomsky (1981), both examples in (5b) are out for the same reason. Both John and the A-position trace are classed as R-expressions, and by Condition C cannot be A-bound.

Finally, the complementarity in distribution of overt anaphors and pronominals (cf. (5a,d)) is formally captured in the binding theory through the complimentary Conditions A and B. If ECs manifest binding feature assignments, then we ought to expect anaphoric pronominals. Such categories, eg. PKU, will be theoremtically restricted to ungoverned contexts. Thus the

existence of ECs occurring in ungoverned structural contexts, as in (5e), is predicted by the assumption that ECs bear binding features.

ECs have also been demonstrated to bear gender and number features (Davis 1981). I shall assume this to be correct, without discussion.

Certain NPs have been asserted to be arguments, others to be non-arguments. Within EC theory, it is supposed that NP-trace is a non-argument, while PRO, pro, and WH-trace are arguments. It should be mentioned that these four do not exhaust the standard assortment of ECs; prominently among ECs whose featural content is unclear are empty operators, and intermediate trace.

One final question, which emerges from the question of what features ECs have, is how the ECs get those features. Most of the debate on this question focuses on how ECs are assigned binding features. Indeed, the term 'assigned' is also prejudiced, since some theories have it that binding features are intrinsically part of the EC, and checked at some level of representation. Below, I give a short summary of several of the proposals on this topic.

Modes of Feature Assignment

a) Contextual or Functional Definition (Chomsky 1982):

Features are assigned to ECs at S-Structure, by an algorithm:

EC is a variable (-A-P) if it is in an A position and is locally A'-bound.

who does Bob like t ?
 i j i
 -A
 -P

EC in an A position which is not a variable is an anaphor (+A).

John was arrested t
 i i
 +A
 -P

EC in an A position that is not a variable is a pronominal if it is free or bound by an NP with a distinct theta-role.

I tried [EC to leave]
 i i
 +A
 +P

- b) Intrinsic Assignment (Chomsky (1985)): ECs have the features they have due to intrinsic properties; EC left by A-movement is an anaphor, EC left by A-to-A' movement is a variable; PKU is a lexically specified EC.
- c) Random Assignment (Brody(1984)): Randomly assign features; other principles of grammar will filter out improper assignments.

*John saw e

4.2 What are The Binding Features of ECs?

Of the four questions outlined above, I would like to focus on the second, namely what the binding features of ECs are. (From this discussion will emerge a general position on the third question as well.)

As mentioned earlier, the theory of ECs which asserts that they are assigned binding features, and that their distribution

is largely predictable on this basis, makes strong, empirically verifiable predictions. Let us consider NP-trace.

NP-trace is classified as a +anaphor, -pronominal EC. When it appears in governed positions, it is therefore subject to condition A of the binding theory. Since the ECP, on its formulation in 'Barriers', requires all non-pronominal ECs (that is, NP-trace and WH-trace) to be governed, this means that NP-trace will always be subject to Condition A².

Much of the first two chapters of this thesis is concerned with what appear to be exceptions to Principle A, involving overt anaphors. The sentences in (6a) and (6b) are representative examples. It is therefore of great interest to see whether non-overt anaphors display the anaphor connectivity effect. On the basis of the grammaticality of example (6c)³, it may be concluded that they do.

6. Condition A problems:

a. [Pictures of himself]_i , John_k likes e_i a lot._k

b. [Which pictures of himself]_i does John_k like e_i ?_k

2. Or, in the terminology of Chomsky (1985), the licensing condition. NP-trace is predicted to have the same distribution as overt anaphors, as far as the binding theory goes; its distribution is additionally constrained by the ECP.

3. This example was brought to my attention by Samuel Epstein. Its implications for the theory of ECs are to be discussed in forthcoming coauthored work.

c. How likely[t to win]] is[John [e]]?
AP S i S i AP k

In chapter 2, it was argued that the effect in (6a,b) is not due to an expanded definition of c-command or binding (under which John would actually bind himself in these representations, satisfying Chomsky's formulation of Condition A), but rather requires a reformulation of Condition A itself. Thus the grammaticality of (6c), as opposed to the wildly ungrammatical (7), indicates that the classification of the EC in (7c) as an anaphor is correct. Such classification subsumes the treatment of (6c) under the analysis of the grammaticality of (6a,b). If the trace of John in (6c) is an anaphor, then the grammaticality of the sentence follows from the revision of Condition A I proposed in Chapter 3. (See later in this chapter, and chapter 5, for further discussion, and an alternative view).

(7) *the man who John was likely that Mary loves believed
EC to have left yesterday

(cf. :the man who it is likely that Mary loves
believed that John left yesterday)

By the same type of argument, the Strong Crossover effect in (8) is taken to be evidence that WH-variable is an R-expression.

8. Strong Crossover: *who did he see t ?

9. Condition C: *she* *admires* *John*

Treating the EC in (8) as an K-expression, subject to Condition

C, allows the ungrammaticality of (8) to be classified together with that of (9), each a violation of Condition C, which prohibits the A-binding of R-expressions.

However, two types of evidence have been brought forth in the literature which suggest that this is not a correct assimilation. First is the existence of examples of what we might call complex crossover, those examples like (10) first observed by Higginbotham (1980).

10) * [which pictures of [which man] _i] does he like _j t_j ?

These examples ought plausibly to be treated together with (8), yet in (10) there is no condition C violation. This leads Higginbotham (1983) to conclude that complex crossover cases like (10) have nothing to do with the principle that bars (9); rather, they are ungrammatical due to a failure to meet a necessary condition on the binding of pronouns by quantificational antecedents. (See Chapter 2 for review of Higginbotham's argument, and see later in this chapter for a review and discussion of Higginbotham's Accessibility Condition, which characterizes when a pronominal can be understood as a variable bound by an operator.) Since this condition rules out (8) as well, the treatment of WH-trace as an R-expression is partially undermined. This calls into question the assignment of binding features to WH-trace.

The second reason to perhaps doubt the treatment of WH-trace

as an R-expression comes from the need to complicate the statement of Condition C in order to allow cases of empty operator movement (see discussion of (17)) below. Such complication is unnecessary for cases where the R-expression is a name, as in (9).

In the next section, I will introduce data which suggests that the possibility of referentially associating an A' bound EC -- an EC which on typical accounts is considered to be an R-expression -- with another category varies depending on the binding features of the head of the EC's chain. The immediate conclusion is that we cannot simultaneously maintain that such an EC (illustrated below) is always -A-P, and that the binding interactions between EC and NP* below are strictly due to the binding features associated with the EC. I shall then explore possibilities which will capture these facts.

```

11)      [ X ]      ... NP* ... EC
          i
          \...../
          c-command

```

4.3 Suppression of Strong Crossover

Consider again structures of the form of (6a), repeated below, except that the anaphor constitutes the entire dislocated phrase. (12a) is such an example, as is (12b).

- (12) a. It's himself who John likes e
 i i i i
 b. It's himself who he likes e
 i i i i

The examples in (12) are perfectly grammatical. Like the examples of (6), and the numerous examples of connectivity discussed in chapters 2 and 3, these present a problem for Condition A of the binding theory, a problem which was addressed in those chapters, and resolved by re-formulating the licensing condition for anaphors (and pronominals).

However, these sentences also pose a problem for Condition C of the binding theory, which is reproduced in (16-17).

English has three constructions which permit an anaphor to appear by itself in a non-argument position: clefts, Topicalizations, and pseudo-clefts, as in (13-15).

- 13) it was himself who John saw (in the mirror)
 14) Himself, John saw e
 15) what John saw was himself

The reformulation of Condition A proposed in Chapter 3 allows the anaphor to be referentially dependent on John, since John locally c-commands the object trace, and (plausibly) himself A' binds this trace.

However, there is a further problem posed for the binding theory by such examples as (13-15): the empty category is presumably a variable, hence an R-expression, and is locally A-bound by John, in violation of Condition C, in either of the two formulations suggested by Chomsky.

16) Condition C (Chomsky 1981): An R-expression is A-free.

17) Condition C: (Chomsky 1985) An R-expression is A-free (in the domain of the head of its chain).

The additional clause is added to (17) to allow such sentences as (18) to be grammatical. Here, the variable is A-bound by John, in violation of (16). If the additional clause of (17) is adopted, the trace is allowed to be A bound by John since John is outside the c-command domain of the head of the chain (OP,e).

18) John is easy [OP [_i PKU to please t _i]]

(17) additionally predicts the inadmissability of the interpretation for (18) on which PKU is controlled by John (i.e. 'John is easy for John to please John'), since this interpretation would have PKU coindexed with John and with the variable; since PRO is inside the domain of the operator, it illicitly A binds the variable.

(19) * John is easy [OP [_i PRO to please t _i]]

The schema ruled out by (17) is (20):

(20) ...X ..._i ...Y ..._i e]]

where: Y = A position
X = A' position
(X,...,e) is a chain
Y c-commands e

This schema is manifested by example (13-15); John is inside the domain of himself, and the operator in SPEC of CP position,

4.3.1 Binding at Different Levels of Representation: A Non-Solution

It might be supposed that dividing the binding principles between different levels of representation, so that different conditions apply at SS and LF, might allow the retention of the conditions as formulated. In (14), for example, it might be suggested (following Chomsky 1982) that the Topic phrase himself is coindexed with the variable only at LF, through a reindexing algorithm, such as predication⁴. This contradicts the

4. Chomsky (1982; fn. 11) proposes a rule of Predication, which takes the form of a reindexing rule. He assumes that the Bijection Principle, which accounts for weak crossover effects, applies in the post-S-Structure component, ordered prior to Predication. Thus the BP will apply to (and not be violated by) the representation (i), which will be converted, via Predication, into (ii).

(i) the man_i who his_j mother loves t_i

(ii) the man_i who his_i mother loves t_i

Chomsky notes, following an observation of Edwin Williams', that this mechanism seemingly would give rise to a way out of Condition B violations. In (iii), the left-dislocated NP John can corefer (and indeed, must) with one of the pronouns, but not both. This is explained as a Condition B violation if (iv) is the LF structure.

(iii) John_i, he_i likes him_i

(iv) John_i, he_i likes him_i

However, as Williams notes, nothing prevents (v) from being the LF representation, converted to (iv) through predication.

conclusion of Chapter 2 that Condition A must apply at S-Structure, but let us suspend this conclusion for the moment.

The essential task, on this line of reasoning, would be to allow Principle C to apply earlier than this reindexing, thus to an S-Structure representation of (14) in which John and e have different indices. At SS, the indexed representation would be (24):

24) Himself_i OP John_j saw e_i

i and j would be equated through predication at LF. Since the NPs John_i and e_j are contraindexed when Condition C applies, the

R-expression e_j will be A-free, and the structure in (24) will

satisfy condition C⁵ ^{Chomsky (1982)} cites a similar argument, due to Edwin Williams, which shows that the index-rewriting procedure would allow the suspension of Condition B effects in such sentences as (i).

However, consider now what this approach would have to say about (25):

(v) John_i, he_i likes him_j

Chomsky concludes that the Binding Theory must re-apply after the index-rewriting process has applied.

5. Chomsky (1982, fn. 11).

25) * John, he likes e (with John and he coreferent)

The indexing (26) would be permitted at SS, later rewritten to (27) at LF.

26) John he likes e (SS)
 i j i

27) John he likes e (LF)
 i i i

Since in order to rule in (14), we are positing Condition C as an S-Structure condition which crucially does not apply at LF, the indexing in (26) is the only one evaluated by Condition C. Since in this representation no K-expression is A bound in any local domain, Condition C cannot rule it out. Similarly, if we adopt Higginbotham's Condition C*, (18) is permitted, since only after Predication rewriting is the K-expression e determined by the structure to share a value with ⁶he.

6. Predication, on this view, would involve deriving (i) from (ii) by linking OP to the Topic:

i) Himself, [OPC John likes e]]
 | |
 | |
 | |

ii) Himself, [OPC John likes e]]
 ↑ ↑
 | |
 | |

Alternatively, we might suppose that the problem with (25) is that the pronoun he is illicitly locally A' bound; in at least some cases, it appears that this is not permitted (see Epstein 1985, and the references cited there). However, given the grammaticality of (28), it is clear that (i) this proposed prohibition is not in fact descriptively correct, and (ii) it is at least not applicable to (25) and the like.

28) Himself , he likes e
 i i i

It therefore is the case that the explanation of the ungrammaticality of (25) is not to be sought in placing the binding conditions at different levels, but rather in the exact nature of ECs.

It appears (this will be discussed in detail shortly) that the grammaticality of structures of type (13-15), instantiating schema (11), is due to a complex interaction of the categorial type of phrase X, and the c-command and locality relations between Y and the empty category. In the discussion below I will examine several such cases.

I will in particular explore the possibility, following suggestions of Higginbotham 1985, that the binding conditions divide into two distinct sets, a set of dependency relations

C* will apply to the S-structure (i).

(Condition A, the first clause of the Accessibility Condition (AC) of Higginbotham 1983), and a set of obviation Conditions (Condition B, C, and the latter clause of the AC).

The minimal contrast between (14) and (25) indicates the following descriptive statement to be correct:

(29)

In the configuration (11), e is subject to Condition C when X is an R-expression; e is not subject to Condition C when X is an anaphor.

We emphasize that the proper treatment of these facts must distinguish the two structures (at least partially) by the binding type of X. This is so since the sentences are structurally identical, and the local environment of e is the same; in both cases e is A' bound by a phrase in SPEC, e is casemarked and theta-marked, and e is left under A' movement.

We now shall consider the third case, that where X is a pronominal; as we shall see, EC behaves as a pronominal in these cases.

4.3.2 More Oddities

First observe that (30) is ungrammatical, on the interpretation in which him and John are understood to be coreferential.⁷

7. All such examples are slightly marginal for some speakers, due perhaps to a general dislike of topicalizing pronouns. I

30) * It's him who John likes t
 i i i i

This might be taken to indicate that t in this example is behaving like the EC in (25); that is, as an R-expression, which is prohibited by Condition C from being coindexed with John.

Condition C places no restrictions on the depth of embedding of the R-expression with respect to its A binder; the sentences in (32) are just as ungrammatical as those in (31) (coreference assumed for all examples).

(31) a. *It's John who he likes

 b. *He likes John

(32) a. *It's John who he thinks you should vote for

 b. *He thinks you should vote for John

However, crucially, when the EC associated with a topicalized pronoun is more deeply embedded than in (30), it may be grammatically A bound.

(33) It's him who John said Mary loves t with all her heart
 i i i i

This indicates that the EC in (30, 33) is behaving as a pronoun, not as an R-expression. The grammaticality judgements of (30) and (33) mirror those of (34) and (35).

 have somewhat idealized the indicated judgements in these examples; the crucial contrasts, however, are present, and quite apparent.

(34) * John likes him
 i i

(35) John said Mary loves him with all her heart
 i i

This might be taken as evidence that literal reconstruction -- lowering of the clefted phrase into the position of the gap -- occurs in these examples, prior to the application of Conditions A, B, and C. However, this possibility was rejected in chapter 2, for a number of reasons .
8

8. To quickly add two more, consider what LF representation we would arrive at by such lowering in cases like (i).

i) It was him whol John saw eJ (SS)
 IP

ii) it was who [John saw him] (LF)
 IP

In this derived representation, we have three constituents - it, was, and who -- which play no role in anything. It might be supposed that the first two can be deleted at LF, in that they have little semantic content, but what of the WH-phrase in COMP?

Secondly, it has often been observed that cleft sentences are semantically distinct from their declarative counterparts in that they place focus upon one element. Although it is possible to examine the syntactic properties of such sentences in isolation, any full theory of language must give full recognition of the role of focus in language, and a syntactic analysis such as the lowering account just sketched badly fails to do so. To give one unfortunate consequence, a clefted or Topicalized QNP must take scope over the entire CP in which its variable occurs, crucially wide scope over any QNP which occurs in that CP. (iii) has interpretation (iv), not (v):

iii) Someone, everyone likes (due to Gueron(1982))

iv) [Ex:person(x)][every y:person y] [y likes x]

v) [Ex:person(x)][every y:person y] [y likes x]

Given this result, we must acknowledge that whatever principles of anaphora give rise to the facts exhibited in (30, 33) and (31a, 32a) must be defined over representations in which the EC occurs, that is, structures isomorphic in crucial respects to (11).

To add one final example, consider again sentences (36) and (37):

36) It's himself who John likes e
i i i i

37) It's himself [who [John [[said [e' [Mary
 i CP i IP i I' VP CP* i IP*
 [[loves e]]]]]]
 I'* VP* i

The grammaticality of (37) indicates that the EC is not subject to Conditions B or C. The question now is whether the EC is subject to Condition A of the binding theory, as revised in the preceding chapter.

This is actually difficult to ascertain. If the EC is subject to Condition A, there will have to be an accessibility sequence through which an antecedent is minimally accessible to it. For

Yet, (vi) has both interpretations (iv) and (v). A lowering theory would derive (vi) from (iii), prior to interpretation (in the GB model), thus obliterating the distinction.

vi) Everyone likes someone

This type of fact is taken up in detail in chapter 5.

example (37), John would have to be minimally accessible to the EC. John is in fact minimally accessible to the EC, through the sequence $\langle e_i, e'_i, CP^*, VP, I', IP \rangle$. However, we observe that it is impossible to tell whether the EC is really subject to Condition A, or whether it is an artefact of the structure that there will always be an accessibility sequence through which the NP is minimally accessible. We observe that John is already accessible to himself through the same chain accessibility sequence. Let us examine this point.

The question at hand is whether the bottommost EC in a representation like (38) is subject to Condition A.

(38) [anaphor] ...NP ...e' ...e
 i i i i

In order for the structure to be grammatical, there must be an accessibility sequence through which an antecedent for the overt anaphor is minimally accessible. Let us suppose that the chosen antecedent is NP_i. This antecedent will be accessible through

some accessibility sequence S, which will include e'_i. Since e_i and e'_i are members of the same chain, NP_i will always be minimally accessible to e_i as well as to the overt anaphor.

Thus if the structure is grammatical, the overt anaphor will have to have an antecedent in accordance with Condition A; it therefore will be impossible to ascertain whether e_i is itself

subject to condition A or not.

To summarize the descriptive results:

(39)

- a) The A-position EC associated with an A' position R-expression or overt operator is subject to Condition C.
- b) The A-position EC associated with an A' position pronoun is subject to Condition B.
- c) The A-position EC associated with an A' position anaphor is subject to neither Condition B nor Condition C (and possibly not to condition A either).

The standard theory of ECs predicts that the ECs in all such cases will be subject to Condition C.

With the assumption that the theory of ECs must be adjusted to capture these facts, and that the relevant syntactic representations are of the form (11) (not LF structures deformed by lowering so as to eliminate the EC), we now proceed to develop a preliminary account.

4.4 The Feature Matching Hypothesis

Let us suppose that the descriptive generalizations in (39) are fairly close to being formally correct, and that the ECs in such structures as (31a, 32a), (30, 33), and (37) differ in their binding features (that is, differ in the assignment of +/-A, +/-P).

Suppose we were to adopt the following characterization of EC feature values:

(40)

The Feature Match Hypothesis (FMH)

In a chain $C = (NP_1, \dots, NP_i, \dots, NP_n)$, every pair of NPs must match in features $+/-A, +/-P$.

We assume that the binding features of lexical NPs are intrinsically specified; overt anaphors are $+A-P$, overt pronominals are $-A+P$, and overt names and operators are $-A-P$. Thus, derivatively, all ECs in a chain headed by an overt category share the features of the head of the chain. (This raises the problem of the features assigned to PRO, which always heads its chain; we turn to this important matter in (section 4.7)).

We will first consider examples of ECs in A' chains in light of this hypothesis.

Consider the import of the FMH for such sentences as (8, 31a), considered in the previous section. I repeat these examples below, with the features which will be present on the ECs following the FMH indicated:

(41) *who_i did he_i see t_i ?
-A-P -A-P

(42) *It's John_i who_i he_i saw e_i
-A-P -A-P

(43) It's himself who John likes e
 i i i i
 +A-P +A-P

(44) It's him who John likes e
 i i i i
 -A+P -A+P

(45) It's him who John said Mary loves t with all her heart
 i i i i
 -A+P -A+P

(46) It's himself who John thinks Mary likes e
 i i i i
 +A-P +A-P

47) John was arrested t
 i i
 -A -A
 -P -P

The FMH assigns the features -A-P to the EC in a case of overt operator movement, such as (41), thus ruling out the sentence as a condition C violation, exactly as in the standard account. The EC in (42) is assigned the same features; this forces the EC to be subject to condition C, correctly accounting for the fact that it cannot be A-bound (within the domain of who).

For cases like (44, 45) we saw that the EC must be free in its governing category (GC), but may be bound by an NP outside the GC. The FMH hypothesis assigns the EC the features of a pure pronominal, correctly predicting its anaphoric behavior.

Finally, for cases like (43,46), the EC is assigned the features +A-P, making it subject to Condition A. As we saw earlier, whether this turns out to be true is indeterminable, but the characterization makes no false empirical predictions.

4.5 The FMH and A Chains

Let us now consider the FMH and A chains. (47) is a simple passive sentence, in which John and e form an A chain.

47) John_i was arrested t_i

The typical GB account proposes that t_i in (47) is an anaphor, subject to condition A of the binding theory; John is its local antecedent, satisfying the condition.

The FMH hypothesis, on the other hand, assigns to the trace the features of the head of its chain (John), assigning it the features -A-P.

48) John_i was arrested t_i
 -A-P_i -A-P_i

Now arises a problem. This structure violates Condition C of the binding theory, since an R-expression (the trace) is A-bound within the domain of the head of its chain.

In this respect, the sentence is predicted to be

ungrammatical, like the Strong Crossover example (42).

How can this problem be resolved? Note a distinction between the two examples. In (42), the illicit binder of the EC is John, and is exterior to the chain containing the trace (eg. not a member of that chain). On the other hand, in (48), the illicit binder is part of the same chain as the trace.

Let us now adopt an auxiliary hypothesis.

- 49) Auxiliary Hypothesis: The Binding Theory regulates coindexing relations (referential dependencies) between two NPs only when they are members of different chains.

Now, for the passive sentence (48), the R-expression t is bound by John, but since they are members of the same chain this binding is not ruled out by Condition C.

In the next section, we will discuss two sets of cases relevant to the evaluation of the FMH and (49). The first set comprises cases in which the Binding Theory is thought crucially to obtain within chains; the adoption of (49) will require that we find some other means of dealing with these cases. In the second part, we will show that although the FMH is overly restrictive without (49), the auxiliary hypothesis serves to prevent the FMH from going wrong in several cases.

4.6 The Auxiliary Hypothesis

4.6.1 Chain-Internal Binding

Consider the following schematic structure, where X occurs in an A position.

(50)



By the definition of bind assumed in GB theory, X (A-)binds Y in this structure. The import of (49) is that this binding relation is evaluated for well-formedness by Conditions A, B, and C of the binding theory only if X and Y are members of distinct well-formed chains, what I will call chain-external binding. If X and Y are members of the same chain, (what I shall refer to as chain-internal binding), Conditions A, B, and C do not apply to their relation. On the other hand, the Binding Theory of Chomsky (1981, 1985), which has no such condition as (49), applies to this binding relation regardless of the chain relation between them.

There are two configurations in which chain-internal binding is assumed to take place in standard GB theory. The first is the Chomsky (1985) treatment of improper movement; the second is

the obligatorily local relation between NP-trace and its antecedent, given in the typical account by treating NP-trace as an anaphor. These are discussed below.

4.6.2 Improper Movement

It has been observed that raising cannot proceed from the subject position of certain predicates which can take infinitival complements, as in (51).

(51) *John_i is illegal [t_i to leave]

(meaning: 'it is illegal for John to leave')

The adjective illegal can take infinitival complements, as in (52):

(52) It is illegal [EC to leave]

Here, the EC has no antecedent, and has the typical interpretation of arbitrary PRO; it is assumed that this is one instance of PRO_{arb}. On the typical account of ECs, PRO is

excluded from governed positions, hence the subject position of (52) and (51) is ungoverned, permitting PRO to occur there.

Now the question arises as to why the sentence (51) is ungrammatical. Aoun (1982) suggests that 'S' breaks a chain'. meaning that in a well-formed A-chain no link (X,Y) can have an S' node intervening. The complement subject position of (51/52) is ungoverned, indicating that illegal takes an CP complement:

(53) John_i is [illegal [CP [t_i to leave]]]
 AP IP

On Aoun's account, the chain (John, t_i) is ill-formed, due to the intervening CP (S') node. This account is rather stipulatory, however.

Chomsky's 1985 treatment of such ungrammatical constructions rules them out as due to a conspiracy between the ECP and Condition C of the Binding Theory. The trace t_i has to be properly governed; since it occurs in the subject position of an IP, and is not governed by the adjective, the only possible candidate for a proper governor would be a trace in SPEC of CP.⁹ This means we would have an S-Structure like (54):

54) John_i is [illegal [t'_i [CP [t_i to leave]]]
 AP IP

Here, the trace t'_i in SPEC position properly governs the subject trace t_i. It may then delete prior to LF (following the

 9. If the trace were governed by the adjective, as would happen if the complement were IP, not CP, it could be properly governed by the adjective, exactly as Chomsky(1986) suggests for such raising adjectives as likely.

i) John_i is likely [t_i to leave]
 i IP

Agreement between John and likely gives the adjective the status of an antecedent; it governs the trace as well, satisfying the ECP. See Chomsky (1986) for details.

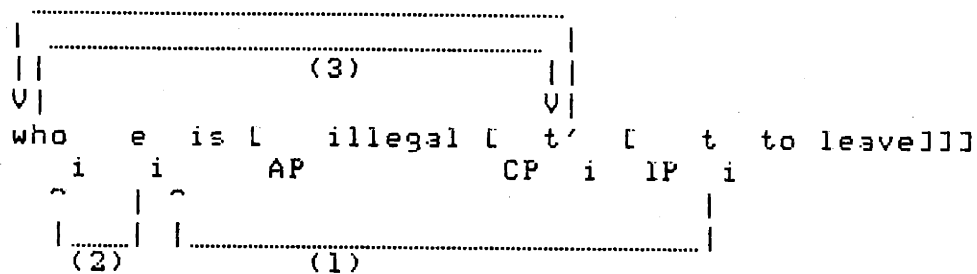
an R-expression. This is part of the intrinsic feature characterization of EC binding features. Thus, direct movement of John from subject position to subject position will leave _i behind an anaphor trace, but there will be no proper governor for the trace. Movement of John through SPEC position will leave behind an R-expression, violating Condition C. This movement moves from an A position to an A' position, then to an A position; this is what has been descriptively known as 'improper movement'. Chomsky's treatment of it reduces the prohibition against such movement to independent principles of the grammar.

However, there is a derivation of such constructions¹¹ which will apparently leave an anaphor in the embedded subject position, and leave a trace in the SPEC position to govern the subject trace. The derivation proceeds as follows:

11. This derivation builds upon a simpler one, which moves John from subject position to subject position, down into the SPEC position, and back into the higher subject position. Such a derivation might be problematic, since movement from an A' position to an A position never grammatically occurs. The derivation in the text, suggested by Samuel Epstein, avoids this problem; all movements are of the normal type, A-to-A, A-to-A', or A'-to-A'.

(56)

(4)



The movement moves from subject position to subject position, leaving an anaphor t_i . The WH phrase then moves to the embedded SPEC position, and then back up to matrix SPEC position, leaving a trace in the lower SPEC position. The intermediate trace then is available to properly govern the subject trace, and the A chain (e_i, t_i) will satisfy Condition C, since t_i is (on this derivation) not an R-expression.

The Strict Cycle Condition is not violated; movements (1) and (2) clearly meet it, and movements (3) and (4) all take place on the highest cycle, therefore there is no violation.

There is, however, an entirely different treatment of such ungrammatical cases of raising as illustrated in (51), provided within the approach to empty categories and chain structure advanced by Michael Brody (1985). This characterization of chain structure will be of use in the second section of this chapter as well.

Brody observes that there is a complementary distribution

between PRO and trace: PRO is ungoverned, and trace is governed. Brody further points out that within GB theory this is only a descriptive generalization, one which is entirely accidental within the formal system of principles constraining the distribution of ECs. The reason for this is that the two are given by autonomous modules. The fact that PRO is ungoverned is a theorem following from the Binding Theory (under the characterization of PRO as an anaphoric pronominal), while the requirement that trace be governed follows from the ECP, not from the Binding Theory. He suggests that this complementary distribution should be formally captured. This treatment (outlined in the next paragraph) has the implication that the distribution of PRO is not to be given by the Binding Theory. If this move is correct, then the status of PRO as an anaphoric pronominal is undermined, since this classification has no other purpose than to ensure the distribution of PRO. In turn, this undermines the classification of ECs in terms of the features +/- A +/- P.

This latter point is quite important. If Brody's proposal is adopted, then it is undesirable to retain the two binary binding features, because we would expect every array of features to be manifested. This means there will be an empty anaphoric pronominal, and it will be restricted to ungoverned positions by both Brody's account and by the Binding Theory, an unwelcome redundancy. This result is extremely interesting, since it implies that ECs do not bear binding features at all;

and this is the ultimate conclusion I shall reach in this chapter, motivated by other considerations.

Brody observes that PRO always heads its A-chain, while NP-trace never heads its A-chain. PRO is always governed, while trace is always governed. He seeks to unify these facts. His account may be briefly summarized as follows.

Both PRO and NP-trace are, in his system, pure anaphors. By the theta criterion, every well-formed chain contains one and only one argument. Assume that ECs are optionally taken to be arguments, and also assume that for independent reasons, the argument of a chain must be in its head position. This permits the descriptive generalization to be put this way: an EC heads its chain iff it is ungoverned. Brody's system is somewhat more complex, but it has this result. In the next section of this chapter I review Brody's proposal somewhat more fully.

Consider now an "improper movement" sentence like (51), assuming no trace in SPEC position.

(51) John is illegal [CP [IP [EC to leave]]]

By the theta criterion, the chain (John, EC) must exist. But this is an ill-formed chain, since EC is in an ungoverned

and its chain-internal antecedent satisfies its binding requirements; this is not possible given the condition (49).

4.6.3 NP-trace as an Anaphor

The second instance of chain-internal binding in the standard view of ECs is the analysis of the distribution of NP-trace, assumed to be an anaphor and so subject to Condition A of the Binding Theory.

4.7 Chomsky (1985) and Condition (49)

In the preceding subsection, we have discussed various structures in which condition (49) causes the BT to not apply. Chomsky (1985) makes a rather similar suggestion, although in consideration of rather different structures.

- (57) A binding relation between an argument and a non-argument is not subject to Binding Theory
(Chomsky 1985; ex. (189))

Chomsky is concerned with the distribution of the pleonastic there in English, as in (58):

- 58) there was a man in the garden
 i i

Chomsky assumes that there is a nonargument, and occurs in a Casemarked position, while a man is an argument, and occurs in a theta-marked position. They constitute an expletive-argument.

13

pair. The expletive is always accompanied by an argument . Chomsky notes, following Burzio (1981), that there is always a very local relation between the members of the expletive argument pair, mirroring the relation between a moved NP and its trace. Given this, as well as the case and theta-role properties of the pair, Chomsky suggests that a pair such as (there, a man) constitutes something like an A-chain, what he terms a CHAIN. The locality relation between the two members is then attributed to conditions on chain well-formedness. The two members are coindexed at S-Structure, for reasons to be reviewed below. Chomsky further notes that if (there, a man) constitutes a chain, and is subject to the requirement that for any chain link (a, b) a must bind b, then this would seem to be a

14

Condition C violation . Chomsky suggests (57) (his (189)) as a way out of this dilemma; although there binds a man in (58), there is a non-argument, and so the binding is not ruled out by Condition C.

Chomsky then observes that "...as it stands, (189) rules out application of binding theory to NP-movement: the trace is a non-argument bound by the head of its chain, typically an

13. In English; other languages (German, Dutch, for example) have structures with isolated pleonastics, as in the German 'es werde getanzt' (lit. 'it was danced'). See Perlmutter (1978), Burzio (1981) for discussion.

14. this essential point is made by Safir (1982), among others; see Safir's work for extensive discussion

argument'. This leads him to suggest a modification of (189).

(59)

(=(190)) A binding relation between an argument and a Case-marked non-argument is not subject to binding theory.

In NP-movement, the NP-trace is not case-marked, hence (59) doesn't apply, allowing the BT to apply to NP-trace.

There is another option to consider within the framework of Chomsky 1985. First, (59) could be made directional, as in (59'):

59') The binding of an argument by a non-argument is not subject to the binding theory.

This now permits the structure (58) to be grammatical, in spite of the fact that there binds the R-expression a man, while permitting NP-trace to be subject to the Binding Theory. This departs from the basic notion represented as (49) in the framework being developed in this chapter.

Second, Chomsky elsewhere in (1985) suggests that the cooccurrence of pleonastics and arguments is no accident. He proposes a general, and powerful, constraint, the Principle of Full Interpretation (FI), which requires that every element present at LF (and at other levels of representation) 'must be licensed... None may be disregarded (p. 98)'. This gives rise to the following treatment of pleonastics:

The requirement that expletive-argument pairs share the properties of chain links, so far stipulated, should follow from binding theory. This result would follow if expletives are not permitted to

appear in LF representations. That is, at LF all CHAINS are chains; only arguments or their traces appear in A positions -- a rather natural requirement given the role of LF as representing the contribution of the language to semantic interpretation. We may assume that the expletive A can be eliminated, in accordance with the condition on recoverability of deletion, only if A is replaced by a coindexed element B, hence by movement of B to the position occupied by A, forming a chain¹⁵
 (B,e)."

This derives the LF (60) from (58):

60) A man_i was e_i in the garden

The resultant chain (a man, e) contains a trace which in

Chomsky's framework is subject to Condition A¹⁶. Chomsky's

15. This quote from Chomsky (1985) actually ends with "...forming a chain (A, e)." I assume this to be a typographical error.

16. Saddy(1985) observes that the element there appears elsewhere in English as a pronominal prepositional phrase:

- i) John went to Boston, but Mary didn't go there.
- ii) John went to Boston, and Mary went to New York.
- iii) *John went to Boston, and Mary went New York.

He then proposes, as an alternative to Chomsky's analysis, that the pleonastic in (iv) is replaced by the PP, not by the NP:

iv) There was [a man] [in the room]

v) LF: [In the room] was [a man] t

Further support for this analysis comes from scope facts. (vi) and (vii) are essentially synonymous, while (viii) differs in interpretation.

(vi) There wasn't a man in the room (vii) In the room wasn't a man (viii) A man wasn't in the room

Basically, (viii) asserts that there was a man such that he was

Note that this might provide a solution to the problem raised by the coindexation of there and the NP [a man] in Chomsky's analysis, since on Saddy's analysis the there will be coindexed at SS with the PP. This plausibly is to be permitted, since the following appear grammatical:

x) John put [a book which belongs [on the table]
there
i

There are two problems with Saday's analysis. First, (xi), (xii) have no PP, thus there is nothing to replace the pleonastic at LF:

Second, cases parallel to (vi) and (vii) cease to be equivalent with respect to scopal interpretations when the PP contains a quantified NP:

(xiii) is true in the situation where every room but one have a man in it, while (xiv) is false. This indicates that (xiii) may have the scope ordering (xv) or (xvi), while (xiv) may have only (xvi).

(xvi) every room x [not [a man y in x]]

- 361 -

analysis then is committed to the view that NP-trace is subject to Condition A of the Binding theory. In a typical case, the antecedent which locally binds the NP-trace anaphor is a member of the trace's chain, as in (61).

61) John_i was arrested t_i

The post-verbal trace is an anaphor, subject to Condition A of the binding theory. Further, the NP which binds it in satisfaction of the binding theory is John, the head of the chain. By (49), this should not count as satisfying condition A.

so that the scope orderings for both ought to be equivalent. In particular, (xv) ought to be impossible as an interpretation for (xiii). The PP moves to subject position, but the quantifier has scope within negation. If negation is not moved, we would have this LF:

(xvii) [[in t_i] [not_k [every room]_i [[a man]_j [t_j t_k]]]]

The variable t_i is not within the scope of its quantifier, hence the structure is illicit. If negation is assigned sentential scope through QR, we have this LF:

(xviii) [not_k [every room]_i [in t_i] [[a man]_j [t_j t_k]]] But this leaves unexplained why both interpretations are not available for both sentences. I therefore conclude that the scope facts must be treated independently of the pleonastic replacement, and assume Chomsky's analysis (under which the NP moves to replace the pleonastic). On the interpretation of there-insertion sentences, see Williams (1983) and Higginbotham (1985). On interesting scope problems with pleonastics, see also Davis (1981) and Aoun (1985).

4.7.1 The Redundancy of ECP and Condition A

One of Chomsky's (1985) motivations for revising of Condition A of the Binding theory (subsumed under the licensing condition) is to eliminate a redundancy between Condition A (as formulated in Chomsky (1981)) and the ECP for such examples as (62).

62) John_i seems [that [t_i left]]

(62) violates the ECP, in that t_i is neither lexically properly governed, nor antecedent governed by John. Additionally, on the Chomsky (1981) formulation of Condition A, the anaphor is not bound in its GC, since the GC is the embedded clause.

On the (1985) formulation of the BT, on the other hand, the GC for the anaphor is the matrix clause. This is so since there is no BT-compatible indexing for the anaphor within any smaller domain than the matrix S. Therefore, this sentence fulfills Condition A, while still violating the ECP. Thus the redundancy is eliminated, for this example. We shall see shortly that, within the 'Barriers' approach to the ECP, the ECP places tight restrictions on the structural relation between an NP-trace and its antecedent, and that Condition A places no further restrictions on the relation. That is, for any pair (X, NP-trace), if X is the antecedent of the trace, and the structural relation between X and the trace satisfies the ECP,

then this relation also satisfies Condition A (but not conversely). Condition A in this sense is always redundant with the ECP; its effects for NP-trace are always masked by the ECP. This suggests that Condition A should never apply to NP-trace. Before proceeding to this discussion, however, a few remarks are in order about example (62).

Elsewhere in (1985), Chomsky proposes a condition on CHAINS (NP-trace chains, and expletive-argument pairs):

(Chomsky's ex.171)

If $C = (a_1, \dots, a_n)$ is a maximal CHAIN, then a_n occupies its unique 0-position and a_1 its unique Case-marked position.

Assuming, following Chomsky, that expletive argument pairs constitute chains at LF of the normal type, we revise Chomsky's (171) to the Chain Position Condition:

Chain Position Condition (CPC)

If $C = (a_1, \dots, a_n)$ is a maximal chain, then a_n occupies its unique 0-position and a_1 its unique Case-marked position.

Observe that in (62), the chain (John, t) must be formed, to satisfy the Theta Criterion. This chain violates the CPC, since the trace occupies a Casemarked position and so a_1 is not the unique Casemarked position of the chain. Thus the sentence is ruled out by the CPC, in addition to the ECP, and the (1981)

binding theory. This means that although the (1985) revision of the BT eliminates one redundancy, it leaves another.

(62) can be modified, however, if we place *t* in a caseless position. There are three types of example to consider: (63), (64), and (65).

(63) *John_i seems_i that it is certain [*t*_i to be nice]

(64) *John_i seems [there to have been arrested *t*_i]

(65) *John seems that [*t*_i to have left]

We can reject (65) as being ungrammatical for another reason, namely that that complementizers never co-occur with infinitival clauses, as (66) illustrates.

(66) * John wants that PRO to be elected

Now consider (64). Condition A is apparently not violated, since pleonastics induce weak SSC violations, as the following shows (see Chomsky (1981)).

(67) The men believe it to be obvious that pictures of each other will go on sale

Chomsky (1986) analyses (64) as an ECP violation. Lasnik and Saito (1984) propose that theta-marking by a lexical head is an instance of proper government; Chomsky uses the ungrammaticality of cases similar to (64) to argue that this proposal is incorrect; he advocates a reduction of the ECP to antecedent-government. Thus the trace in (64) must be

antecedent governed in order to satisfy the ECP.

In Chomsky's system, an NP-trace will be antecedent-governed through a rather extended notion of antecedence which adds agreement to referential dependence. For a simple case like (68), the trace *t* is inside a barrier (the VP) which does not contain the trace's antecedent, John. However, the participle agrees with the subject, and the participle is inside the VP. Taking agreement to be syntactically represented as coindexing, this means that the participle is coindexed with the trace; since the participle also c-commands and governs the trace, the trace is antecedent properly governed by the participle.

(68) [John [was [arrested *t*]]]
 IP i I' i VP i i

For example (63), the indexed representation is (69):

(69) John seems that it is certain [*t* to
 i i j j i
be nice]

Here, the adjective certain (the governor of the trace) is not coindexed with the trace, hence does not properly govern it.

Thus Chomsky argues, the sentence violates the ECP¹⁷, but does not violate other principles (other than Subjacency, violations of which give less severe ungrammaticality than is present in (69)).

17. Lasnik and Saito (1984; fn 56) discuss a similar example (*John seems it is certain to leave), demonstrating that the complete unacceptability of the example indicates that it violates the ECP, not simply Subjacency.

Let us now consider whether this is correct. Chomsky (1985) proposes that the relation between a pleonastic and its associated argument mirrors the relation between a NP-trace and its antecedent because the pleonastic is eliminated at LF; the elimination takes place when the argument moves to the position of the pleonastic. This movement leaves a trace:

70) There was a man in the room

LF: [a man] was t in the room
 i i

Given the proposals of Chomsky (1986), the trace is subject to the ECP, and the satisfaction of the ECP will be effected through the agreement of the subject with the verb.

The pleonastic it is always associated with a post-verbal CP constituent with an overt complementizer (see Chomsky and Lasnik (1977)), as in (71):

- (71) a. It is illegal for Mary to leave
 b. It is likely that Mary will leave
 c. *It is illegal Mary to leave
 d. It was assumed that Bob is tall

Observe now that in (63) there is no post-verbal CP, since the infinitival construction is an IP. Therefore, it is plausible to suggest that this structure is ungrammatical for this reason, giving the result that we cannot determine whether its ill-formedness is due to a lack of proper government of the trace.

This leaves us with (64). (64) also contains a pleonastic,

and considerations similar to those just discussed for (63) raises the possibility that this structure violates other conditions of grammar.

(64) John_i seems there to have been arrested t_i

By Chomsky's assumptions, the only possible proper governor of the trace is arrested, which must therefore be coindexed at S-Structure with the trace to be an antecedent of it. This means that (if the ECP is to be satisfied) the indexation must be as in (72):

72) John_i seems_i there_i to have been arrested t_i i_i

Since the pleonastic will have to be eliminated at LF by substituting t for it, the coindexation between there and t is forced by this consideration as well.

The theta criterion is satisfied only if there is a well-formed chain containing John and the trace. There are two possibilities: (John, t) and (John, there, t). Several authors (Chomsky 1981, Rizzi 1982, Lasnik 1985) have suggested a local

18

binding condition (LBC) on A chains¹⁸. We saw in the previous chapter that Rizzi (1982) explains quite puzzling facts about

18. Chomsky gives the condition as one of four properties of chains; he does not argue for it as a well-formedness condition on chains. Rizzi (1982) and Lasnik (1985) do argue that the LBC is a necessary condition on the well-formedness of A chains.

clitic movement in terms of this condition. The adoption of the local binding condition on chains entails that the chain (John, t) is ill-formed since there, not John, is the local binder of the trace.

So, let us consider (72), taking the only possible chain to be (John, there, t). The coindexation between the pleonastic and t will cause the trace to be properly governed, hence the ECP is satisfied, if this indexing is licit ¹⁹.

Through the elimination of the pleonastic, the S-Structure (72) will be converted to the LF representation (73).

73) John_i seems t'_i to have been arrested t_i i

Both traces are properly governed in this representation. Hence we suppose that the ill-formedness of (72) is due to a property of the S-Structure configuration. Possible candidates include ECP (it may be that the trace t_i is not properly governed), or conditions on pleonastics, or it may be that the chain (John, there, t) is ill-formed. Recall that Chomsky's

19. If the indexing were different, as in (i), then the chain (John, t) would be well-formed, since John would be the local binder of t.

i) John_i seems there_i to have been arrested t_j j i

However, both the ECP and the requirement that pleonastics are eliminated at LF would be violated. See below for discussion.

claim is that nothing is violated in this example except the
20
ECP .

Consider the possible D-structures for (64)/(72). There are
at least two.

74) seems there to have been arrested John

75) seems e to have been arrested there

(74) can be mapped into the S-Structure (64) by movement of
John, from embedded object position to matrix subject position.
Assuming, following Chomsky, that there cannot be accidental
coindexation of John and the pleonastic, this means that the
indexed representation must be:

76) John_i seems there_j to have been arrested t_i

That is, the ban on accidental coindexing will block the indexed
representation in (72) from being derived from the D-Structure
(74). The SS (72) can arise only if the prohibited 'accidental
coindexation' takes place.

In (76), the ECP is violated. However, (76) is, on the
proposals concerning pleonastics presented by Chomsky (1985,
1986), ill-formed for another reason: there is no element to
replace the pleonastic (the trace will not suffice, since it is
not coindexed with there). Thus the SS (76) is ill-formed both

20. Except for Subjacency, which can be ignored, since the
example is far worse than a Subjacency violation.

because of the ECP, and because the pleonastic cannot be eliminated at LF.

Now suppose that (75) is the D-Structure of (64), with John inserted into the subject position by generalized transformation. Again assuming there to be no accidental coindexing, this gives rise to the indexed S-Structure representation (77):

77) John seems there to have been arrested t
i j i

But this representation violates the theta criterion. Assuming the local binding condition on chains, there is no way to associate John with t in a well-formed chain, since John does not bind the trace.

We must now ask whether there are any structures which show that NP-trace cannot be properly governed in virtue of being theta-marked by a head; both of Chomsky's (1986) examples, given here as (63) and (64), plausibly violate independent conditions on the distribution of pleonastics. There is at least one such example, (due to Mark Baker):

78) * John_i seems that it_j was told t_i [that Mary is beautiful]_j

This structure has the appropriate form: the trace is in object position, and is caseless (which must be so, given the Chain Position Condition); by Burzio's generalization, the subject is

theta-less, hence a pleonastic is present; and the pleonastic has an associated argument, crucially distinguishing it from the cases considered earlier. No condition on pleonastics is violated; hence Chomsky is correct in proposing that NP-trace cannot meet the ECP by being theta-marked by a lexical head; it satisfies the ECP only under antecedent government²¹.

Now re-consider the discussion of (74-77). Given the conclusion that NP-trace must be antecedent governed, these representations violate the ECP. We also saw that they violate conditions on pleonastics and the theta-criterion, in that the indexed structure (72) has no well-formed D-Structure; it is not even derivable. The reason for this is the 'ban on accidental indexing', which prevents the pleonastic in the S-Structure corresponding to (74) from picking up the index of the chain (John_i, t_i). This is an odd constraint, since the theory of indexing assumed freely assigned indices. This suggests that this ban ought to be given up. (This means that (72) will be generable). Of course, if given up, it is unclear what rules out (72); this structure meets the ECP, the theta-criterion, and the conditions on pleonastics. This problem is taken up in the next section.

21. Epstein (forthcoming) argues that an EC can satisfy the ECP if it is casemarked by a lexical head, but not (necessarily) antecedent governed. Even if this argument is correct, Chomsky's conclusion about NP-trace is essentially correct: NP-trace can satisfy the ECP only if it is antecedent governed.

4.7.2 The Ban Against Accidental Coindexing

Recall that the ungrammaticality of Chomsky's example (64) is, on his view, to be attributed to the ECP.

64) *John seems there to have been arrested t

The trace would be properly governed, if the pleonastic were coindexed with John and the trace; in such circumstances, the participle would acquire the index of the pleonastic (under agreement), and would then properly antecedent govern the trace. This is the indexed representation which would yield well-formedness:

72) John seems there to have been arrested t
i i i i i

Chomsky proposes a ban against accidental coindexing, which blocks the indexing of (72), forcing (73), which violates the ECP.

73) John seems there to have been arrested t
i j j i

As we saw earlier, (73) is out for independent reasons, namely that there is no way to replace the pleonastic at LF, given the constraint on recoverability of deletion. Thus, we showed, the pleonastic problem masks the effects of the ECP.

Let us consider the formal status of the ban on accidental coindexings. Within Chomsky's theory of pleonastics, (79) is

the D-structure of the indexed S-Structure (80).

79) There was a man in the room

80) There was a man in the room
 i i

This indexing is given by the general mechanism of coindexing, which is, in the GB system, entirely free between A positions. This gives rise, for example, to the coindexing in (81-2).

81) John saw himself
 i i

82) John saw himself
 i j

The theory of binding rules out (82).

It is apparent that this general schema of free coindexing will perfectly well generate the representation (72); John and the trace are coindexed under movement, and the pleonastic, being in an A position, is freely coindexed with these NPs. It therefore would seem that the 'ban' on accidental coindexing must refer specifically to pleonastics. Further, this ban must be put so as to allow the free coindexation in (80). It is difficult to see how such a ban could be formally stated.

I would like to suggest that we need not have any such ban. Rather, the indexed representation (72) is ruled out if we adopt two proposals: Brody's (1985) caselinking theory, which places conditions on the government and Case-marking relations which hold of heads and non-heads of chains; and Rizzi's (1982) Local Binding Condition on Chains (LBC), reviewed in the previous

chapter.

To quickly review, the LBC holds that for any well-formed chain, each link (A,B) must be such that A is the closest binder of B in the representation. In order to satisfy the ECP, (64) must have the indexing (72). The LBC forces the chain to be (John, there, t). As we shall see in a moment, Brody's principle excludes such a chain.

Brody supposes that Case is a property that an NP has (or doesn't have) intrinsically; all overt NPs have Case, all ECs do not have Case. Tensed INFL has Case. Certain verbs have Case as well; these correspond to the verbs which assign Case, in the more standard view of Case (Chomsky (1981), Burzio (1981), and many others). NPs must occur in a structural position compatible with their inherent case properties; approximately, if an NP heads its chain, it must be governed by a governor (INFL or verb) that has Case, if the NP is overt; and if empty, the NP must occur in an ungoverned position. The converse requirement holds of NPs which do not head their chains: if empty, they must be governed, and by governors (adjectives, passive participles) which do not have case. As a consequence of Brody's theory, if a non-head of a chain is overt (and so has Case), then there will be no position it can occur in compatible with its case requirements. This will exclude the chain (John, there, t), since an overt NP (there) occurs in it, in a non-head position. This has the result of ruling out (72), since the LBC

forces this chain to exist.

We now review Brody's exact formulation of these notions. Brody defines two conditions on chains, (84) and (85):

83) Definition of Case-Linking

NP* is Case-linked iff:

- i) NP* is governed by some governor G and has Case.
- or ii) NP* is ungoverned and has no Case.

In case (i), we say that NP* is case-linked to G.

84) The Case-Linking Condition (CLC)

NP* is case-linked iff NP* is the head of its chain.

85) The Case-Matching Condition (CMC)

If NP* is case-linked to G, then G has a case that matches that of NP*.

overt NPs always have Case; ECs do not have Case; most active participles have case; raising verbs (seem, appear), and passive participles do not.

Now consider there in (72).

72) John seems there to have been arrested t
 i i i i i

By the Local Binding Condition, the only well-formed chain which associates John with the theta-position t is (John, there, t). There occurs in non-head position, hence by (84) must not be Case-linked. To be non-case-linked, an overt NP (which therefore has Case) must meet the negation of (83i) or (83ii). That is, either not[NP* is governed by some G and has Case], or

not[NP* is ungoverned and has no Case].

For an NP X, X meets the negation of (83i) or (83ii) in one of six circumstances.

- (86)
- a. X is ungoverned and has Case.
 - b. X is governed and has no Case.
 - c. X is ungoverned and has no Case.
 - d. X is governed and has Case.
 - e. X is governed and has no Case.
 - f. X is ungoverned and has no Case.

We must reject (f), since it one of the definitions of Caselinking (83ii). We must also reject (86a), since it is the other definition of Caselinking. The element under consideration is there, which is overt and thus has Case. Therefore, it cannot under any circumstances meet (86b, c, e, or f). This leaves (86d): in order to appear as the non-head of a chain, a pleonastic must be ungoverned.

Suppose there is governed by some G, as occurs in (72), where there is governed by seem. In this circumstance, there is Case-linked by (83i); this violates the CLC.²² Thus (72) is ruled out by the CLC, without recourse to Chomsky's proposed ban on accidental indexing.

Now consider the other case, where there meets the

22. In essence, this forces there to be the head of its chain. But if this occurs, then the CLC is satisfied, but the CMC is violated: there has Case, but the verb to which it is Case-linked has none.

87) A man was illegal [[there' to be t outside]]
 i CP IP i i

88) A man was illegal [[t to be t' outside]]
i CP IP i i

But now observe that pleonastics are replaced at LF under Chomsky's treatment. A pleonastic which does not head its chain will be replaced by a trace, as in (88), where *t* occurs in the position occupied by there' at S-Structure. Since this element is an EC, it must meet the ECP. Since it is ungoverned, it

23. If we were to suppose that the CLC and the CMC apply at LF, then (88) would be blocked in another way: the trace t' , being ungoverned, is case-linked, hence must head its chain. This means that the only well-formed chains are (a man) and (t' , t), both of which violate the theta-criterion and the Chain Position Condition (64). This indicates a redundancy between the ECP and

Thus we see that adopting Brody's theory (which is independently motivated by the complementary distribution of PRO and NP-trace), taken together with the LBC on chains, and Chomsky's treatment of pleonastics, allows us to dispense with the ban on accidental coindexation. This is just as well, given the difficulty (or impossibility) of formally specifying such a constraint.

4.7.3 Other Cases

Consider again example (78), which we saw violates the ECP, and does not violate any condition on pleonastics.

the CLC. The redundancy is not total: in Baker's example (78), the CLC is satisfied with respect to the chain (John, t), but the ECP is violated. I explore this redundancy further in work in progress.

24. Observe that the element moved into the pleonastic at LF is in a configuration of proper government at S-Structure. If we adopt the Lasnik and Saito (1984) gamma-assignment system of ECP, then we derive (ii) from (i):

- (i) John_i is illegal [there_i to be outside t_i]
+G
- (ii) (i) John_i is illegal [t_i to be outside t'_i]
+G +G

The problem is that the trace t moved into there in the SS-LF mapping bears the feature +G at S-Structure. If it is allowed to carry its feature along with it, then (ii) is derived, and this LF representation meets the ECP. This indicates that either the feature is assigned to positions, not categories, or that the feature assignment framework needs to be reevaluated and altered, or that +G is assigned to NP-trace only at LF.

78) * John_i seems that it_j was told t_i [that Mary is beautiful]_j

By Chomsky's original line of reasoning, there is a redundancy to be eliminated: both the ECP and Condition A of LGB are violated here. This is one of the motivations for Chomsky's (1985) revision of Condition A.

However, now consider a simple passive.

(89) John was arrested t

For such examples, the ECP and Condition A are again redundant. Given that the ECP places tighter restrictions on the relation of a trace to its antecedent than Condition A would²⁵, and that these additional restrictions are empirically motivated, this suggests that Condition A should simply not apply to NP-trace.

In the FMH system, this will generally be true. The only instance where an NP-trace is +A-P is when the head of its chain is a lexical anaphor, as in:

90) [John_i [believes [himself_i]]]
 IP i I' VP IP* i
 +A-P

 [to have been [arrested t]]]]]
 I'* VP* i
 +A-P

25. The ECP, on the 'Barriers' formulation, will require that the trace be bound by the subject of its clause; Condition A requires only that a trace is in the minimal domain in which it could be bound, which permits the anaphor to seek an antecedent in the next clause up, in some circumstances.

Here, the trace is governed and bears the features +A-P, and thus is subject to Condition A. Observe that the auxiliary hypothesis (49) has an interesting result here: the head of the chain, himself, cannot act as the antecedent required by Condition A, since the trace and the anaphor are part of the same chain.

However, the overt anaphor will have to independently satisfy Condition A. The NP John is minimally accessible to the anaphor himself, through the sequence {himself, IP*, VP, I', IP}.

Condition A, as revised in Chapter 3 (ex. 7), requires that the trace also have an antecedent minimally accessible to it; and (49) removes the head of the chain from consideration. Within the framework developed earlier, there is an accessibility sequence through which John is minimally accessible to the trace: {t, himself, IP*, VP, I', IP}. This sequence is well-formed, since t and himself are members of the same chain.

(To the extent that the FMH can be maintained, this is independent evidence for the chain-accessibility framework of Chapter 3.)

We have essentially the same results for topicalized anaphors.

Consider example (91), from Lasnik and Saito (forthcoming).

91. [John [[thinks[[that
 IP I' VP CP C'
 [himself,[Mary [[likes t]]]]]]]]
 IP* IP* i I'* VP* i

Lasnik and Saito suggest that such topicalization is adjunction

26
 to S (to IP). First consider the anaphor in the structure.
 The anaphor is bound by John, the subject of the matrix
 clause. himself is subject to the binding theory, requiring it be
 governed and to have a minimally accessible antecedent 27 28.

John is such an antecedent, accessible through the sequence
 {himself, IP*, C', CP, VP, I', IP}. (Additionally, the lower

 26. The EC is subject to the ECP; in keeping with the Barriers
 formulation, there must be a VP-adjoined trace in such sentences
 as well. I will suppress mention of this additional trace in
 the text.

27. Recall that in the accessibility framework, the binding
 theory (chapter 3, exx. (4-7)) applies to all anaphors, not just
 governed ones. The anaphor will be forced to have a governor,
 by the definition (4) of chain accessibility sequence. Further
 evidence that the IP-adjoined position is governed comes from
 the impossibility of coreference between John and him in (i).

(i) John thinks that him, Mary likes t

On the FMH, the trace will have the features -A+P, and so will
 be subject to Condition B. The trace is A-free within IP,
 satisfying Condition B. Within the Chomsky (1985) theory, the
 pronoun him will have to be free in its GC only if it is
 governed. The GC for an element adjoined to IP is the next
 highest IP, (since the pronoun is not included in the lower IP).
 The obligatory disjoint reference in (i) indicates, within the
 context of Chomsky's theory, that the pronoun is itself
 governed. Within the theory developed in Chapter 3, the pronoun
 must be governed in order that it be licensed. Therefore, on
 either view, the IP-adjoined position is governed.

28. Himself is governed by the INFL of the embedded clause,
 since INFL m-commands the IP-adjoined anaphor.

subject, Mary, is accessible; I will ignore this in the following discussion, as it is irrelevant).

This is consistent with the following data:

92. a. * [John's mother] thinks that [himself , [Mary
 i S *i* S
loves t]]
 i

b. * John_i said that Susan thinks that [himself_S , [Mary_i S_S loves]]
t]]
i

Now consider the trace left by topicalizing himself. By the FMH the trace will be an anaphor. It is in a governed position, hence capable of having a well-formed accessibility sequence. Mary is accessible to it, but is not a sufficient antecedent. Himself is not a possible binder, since it is in an A' position, and it is a member of the trace's chain, and is thus not available, by (49).

John is accessible to the EC in virtue of the chain relation between himself and t. This permits the sequence to be well-formed, and John is minimally accessible through it. For examples (92), John is not minimally accessible to the anaphor.

PRO cannot be topicalized:

93) *John thinks that [PRO [Mary likes t]]

The standard analysis of ECs will treat the trace in (93) as the same as the trace in (91), presumably an R-expression. Note

that this causes problems, reminiscent of (14), in why this R-expression is not in violation of Condition C for an example like (94):

94) John thinks that [herself_i [Mary_i likes t_i]]

Treating the object EC as an anaphor in all cases is insufficient, since (95) is grammatical, where there is no A-position NP accessible to the EC:

95) John thinks that Bob_i [Susie_j likes t_i]]

The FMH gets the right results, treating the EC in (97) as an anaphor, but the EC in (96) as an R-expression, predicting Condition C effects for (96) but not (97):

96)* John thinks that Bob_i [he_i likes t_i]]

97) John thinks that himself_i [he_i likes t_i]]

But now re-consider the sentence in (93). By the FMH, this will be ruled out by the PRO theorem even if the Topicalized PRO

29

is in an ungoverned position. The reason for this is that the trace will bear the binding features of its chain-head, and so will be an anaphoric pronominal as well. (Further, recall from Chapter 3 that there can be no anaphoric pronominals with the definition of chain-accessibility sequence developed there.

29. We noted above, however, that the IP-adjoined position is governed.

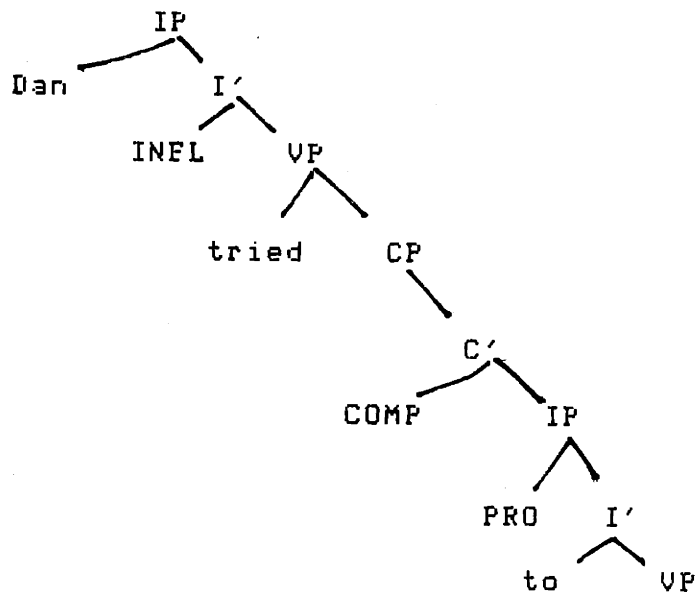
Thus, independent of the features assigned to the trace in (93), it is ungrammatical simply for having PRO. Let us suspend this aspect of the general framework, since it turns on a rather delicate part of a definition, which might be reformulable.).

The standard analysis of ECs, although incapable of making the correct predictions concerning the facts of (96) and (97), can rule out (93) if the Topicalized PRO is in a governed position. To contrast the two analyses, the standard analysis forces Topic position to be governed, in order to rule out the example. The FMH allows Topic to be ungoverned, ruling (93) out as a PRO theorem violation with respect to the object EC.

Within the Barriers definition of government, PRO will indeed be governed in (93).

In a typical infinitival, like (98), PRO is ungoverned. IP includes PRO, and is not an L-marked maximal projection, and hence is a BC for PRO. CP becomes a barrier by inheritance, and so the higher verb cannot govern PRO.

98)



In the topicalization structure (99), the adjoined NP PRO will be governed by the higher verb. Since IP does not include PRO, it is not a potential BC for PRO; CP is thus not a barrier by inheritance. CP is L-marked by the verb, hence does not become

30

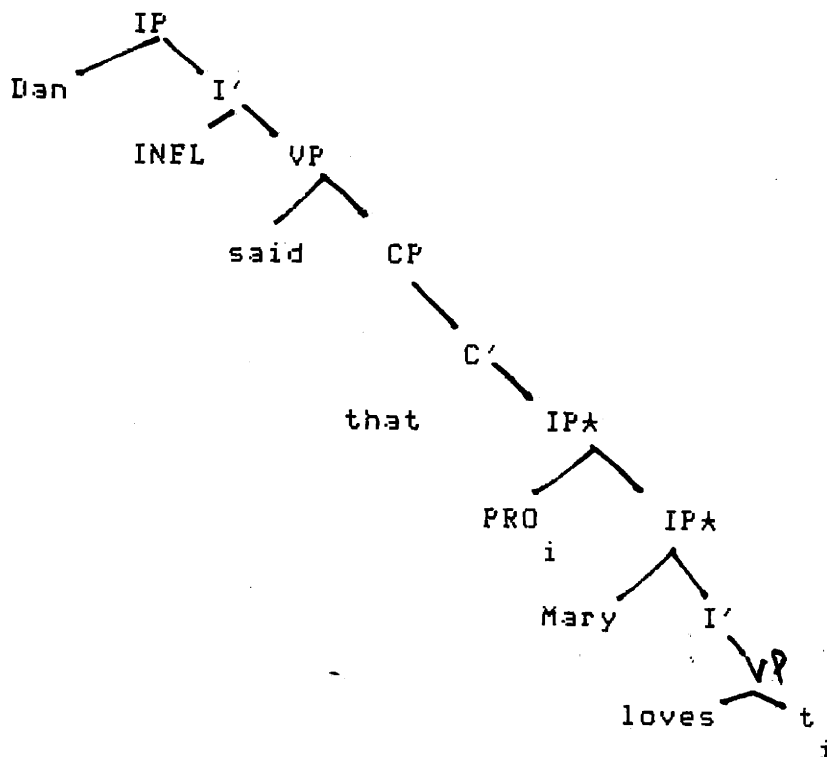
a barrier the other way either .

30. Note that extraposition creates a problem:

1) * John [_{VP} said yesterday] that [PRO [Mary likes t]].

The extraposition removes the PRO from the m-command domain of the verb. Note that the FMH rules this out straightforwardly.

99)



However, PRO presents another problem, to be addressed in the next section, which entails abandoning the FMH and the auxiliary condition (49).

4.7.4 Passivized PRO

PRO cannot be topicalized, as we saw above; such examples as (93) were ruled out by appealing to the FMH, which gives rise to a PRO theorem violation in the trace.

However, PRO can be passivized:

100) John tried [PRO_i to get arrested t_i]

By the FMH, there will be a PRO theorem violation, as the

31

feature assignment will be (34) :

101) John tried [PRO to get arrested t]
 i i
 +A+P +A+P

This then contradicts the FMH as presented.

There are several ways out of this problem. First, we might suppose, as has been suggested by Bouchard and others, that PRO is not excluded from governed positions, but from casemarked positions. However, given the evidence against this position, I will not adopt it.

Alternatively, we might suppose that the FMH holds only of A' chains, and that within an A chain, such as the one comprising (PRO, t) in (101), the EC is an anaphor, as supposed in the standard analysis. The reader will note that all the evidence motivating the FMH was of the form ((12-15);(25); (30),(33)), in which the EC is the tail of an A' bar position. The auxiliary condition (49) was added to keep the FMH from making false predictions with A chains.

This is a conceptually undesirable move, for two reasons. First, it makes a fundamental separation of A chains and A' chains, and leaves it unexplained. Why should movement of a copnstituent to an A' position leave behind a trace which is a

31. I am grateful to Juan Uriagereka for pointing out this problem.

copy of its features, while A movement always leaves an anaphor?

Secondly, it would retain the essential redundancy between Condition A and the ECP in predicting the distribution of NP-trace. The motivation for considering NP-trace as an anaphor comes strictly from its distributional similarity with overt anaphors. The theory of binding, capturing this fact, was developed prior to the development of the ECP. Factually, overt anaphors appear in positions in which NP-trace cannot, undermining the idea that the distribution of NP-trace is explained by appealing to Condition A.

Thus, if we were to restrict the FMH to A chains, then we would have to either retain the redundancy in A chains, or suppose that the EC of an A chain bears no binding features. This again is an odd result; why should an A' bound trace have the features of the head of its chain, while an A trace has no features whatsoever?

Therefore, it would appear, in light of the grammaticality of (100), that we should seek an alternative to the FMH. This alternative will have to account for the facts motivating the FMH in the first place, while avoiding the problems and redundancies of the FMH. The next section presents such an alternative, which holds first that ECs (including PRO) do not bear any binding features, and that linking is to be preferred as the notation for anaphoric dependence.

4.7.5 A Redundancy between (49) and Linking

Consider now the ungrammatical sentence below:

102) * [Himself [was [arrested t]]]
 IP i I' VP i

Condition A requires that, for every anaphor A, there be an NP B such that B is minimally accessible to A through a well-formed accessibility sequence, and that A be linked to B (or coindexed with B). If we were to keep to the standard analysis of ECs, which considers t in (67) to be an anaphor, himself would be the antecedent, accessible through the sequence {t, VP, I', IP}.

Since the NP himself is a sister to one of these nodes, it is accessible to the anaphoric trace.

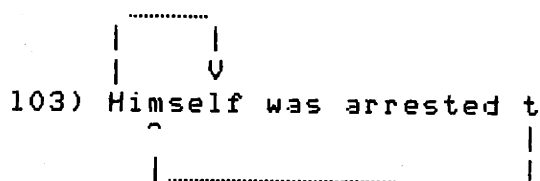
However, it would seem that himself is also predicted to be an acceptable antecedent for itself. Given that $(\text{himself}_i, t_i)$ is a well-formed chain, the accessibility sequence $\{t_i, \text{VP}, I', \text{IP}\}_i$ is well-formed, and himself is accessible through it to the governed anaphor himself. If we formulate the theory of binding in terms of indexing, himself can act as its own antecedent.

It would be easy to augment Condition A in (ex.(7), ch.3) with the additional requirement that A be distinct from B. However, this is a stipulation added to one condition of the binding theory, having little empirical motivation except to block such

examples as (102).

Now consider the import of (49) for this example. (49) states that no chain-internal binding is visible to the binding theory. For an anaphor, this means that the anaphor must have a chain-external antecedent, minimally accessible to it. (49) prevents himself from acting as its own antecedent, as obviously the anaphor is in the same chain as itself. Therefore, no stipulative addition to Condition A need be made; instead, the sentence is ungrammatical due to the interaction of the binding theory with a generally motivated constraint on the application of all binding principles.

Note that there is another way to block (49), within linking theory. Suppose we take the Linking formulation of Condition A, as developed in chapter 3. In (102), himself is minimally accessible to himself. This means that the following linking satisfies Condition A:



But this linking violates an independent principle of linking theory: no element may be anaphorically dependent upon itself (Higginbotham (1983)). (See chapter 1 for review).

Thus it would seem that (49) is redundant with linking theory, as far as such examples as (102) go. (49) is additionally

needed, as a supplement for the FMH, for problems posed by such examples as (47) for the FMH, where no principle of linking theory is of assistance.

Therefore, this might be taken as an argument against linking theory. Alternatively, it might be taken as evidence for linking theory, motivating the abandonment of the FMH and (49) in favor of some other analysis which yields the correct results for examples like ((12-15);(25); (30),(33)).

We have now seen two reasons to seek an alternative to the FMH and (49), one empirical, the other theory internal. In the next section we develop such an alternative.

4.8 Do ECs Have Features?

In this section, I will introduce the No Features Hypothesis as an alternative to the Featuring Matching Hypothesis (FMH) proposed earlier. As we saw, the FMH is necessarily supplemented by the Condition (49), which draws a distinction between chain internal and chain external binding. We argued that the FMH is empirically superior over the 'standard' view of ECs, in that it allowed an explanation of data not explained under the standard view. We did see, however, that certain problems arose, most seriously in the treatment of PRO. The NFH will be able to capture the facts motivating the FMH, without the problems posed by it.

To quickly review, the FMH was motivated by a descriptive generalization: a case-marked EC in an argument position, A' bound by a left-peripheral noun phrase NP_X, acts like an R-expression when NP_X is an R-expression or overt operator; and like a pronominal when NP_X is a pronominal. When NP_X is an anaphor, it is either the case that EC is subject to Condition A of the Binding theory (as revised in Chapter 3), or is not subject to any of the three binding conditions. (We saw that in this latter case it is impossible to tell.) The standard view of ECs predicts that the EC in all cases acts like an R-expression. The descriptive generalization is supported by

the following data:

106) NP* is an R-expression

- a. *John_i , he_i admires t_i
- b. *John_i , he_i thinks that Mary admires t_i

107) NP* is a pronominal

- a. *Him_i , John_i likes t_i
- *It's him_i who John_i likes t_i
- b. Him_i , John_i thinks Mary likes t_i
- It's him_i who John_i thinks Mary likes t_i

108) NP* is an Anaphor

- a. *Himself_i , John_i likes t_i
- *It's himself_i who John_i likes t_i
- b. Himself_i , John_i thinks Mary likes t_i
- It's himself_i who John_i thinks Mary likes t_i

Two sets of problems then arose, both having to do with A chains. In the first, the EMH by itself predicts that NP-movement (raising or passive) of an R-expression or a pronominal should violate conditions C or B, respectively. This is so because the EC will be bound both within the domain of its chain, and within its governing category, since the ECP forces

the chain-internal antecedent of the EC to be structurally very close to the EC. The relevant cases are in (109)-(111).

109) John was arrested t
 i i
 -A-P -A-P

110) He was arrested t
 i i
 -A+P -A+P

111) John believes himself to have been arrested t
 i i i
 +A-P +A-P

(109) is predicted by the FMH to violate Condition C, and (110) should be out by Condition B. The FMH does permit the example (111) to be grammatical, since in this case (and in only this case) the EC is an anaphor.

In order to resolve this problem, the auxiliary condition (49) was introduced.

(49) Chain-internal binding is irrelevant to the Binding Theory

(49) has the additional result of predicting the ungrammaticality of (112).

112) *Himself was arrested t
 i i

Within the framework of the revised licensing condition for anaphors developed in the preceding chapter, this sentence is predicted to be grammatical, independent of which view of ECs we adopt: the standard view, or the FMH. However, by adopting the auxiliary condition (49), himself cannot function as its own antecedent, since it is a member of its own chain.

It was observed earlier that (49) is redundant with a core principle of linking theory, the one which prevents any element from being anaphorically dependent upon itself. When the NFH is developed, this redundancy is eliminated, since (49) will be abandoned.

The second set of problematic cases have to do with the distribution of PRO. The FMH is motivated by the facts of (106-108), where an EC can act as a nonanaphoric non-pronominal (R-expression), as a nonanaphoric pronominal (subject to Condition B), or an anaphoric non-pronominal. The FMH follows the standard view of ECs in presuming that such behavior is to be characterized by the assignment of features +/-A +/-P to ECs (what I have referred to as binding features). Given this, the existence of an anaphoric pronominal (PRO) is predicted. Both the standard view and the FMH view assigns PRO these features:

113) John tried [PRO to sneeze]
+A+P

As observed earlier, the exact reformulation of the licensing condition for anaphors and pronominals I gave in Chapter 3 (as (4-7)) actually rules out the existence of PRO, since all anaphors and all pronominals must be governed in order to be licensed. Let us suppose for the moment that (3-7) could be reformulated to eliminate this problem, and pursue the status of PRO within such an augmented theory. Thus, let us suppose that an anaphoric pronominal is allowed to exist in the FMH system.

The problem arises when PRO is passivized or raised. The standard view characterizes the trace as an anaphor, assigning it the features +A-P, as in (114). The FMH, on the other hand, assigns the trace the same features as the head of its chain, as in (115).

114) John tried [PRO_i to be arrested t_i]
 +A+P +A-P

115) John tried [PRO_i to be arrested t_i]
 +A+P +A+P

The sentence (114) is predicted grammatical within the standard view, as the trace will be locally bound (by PRO), in accordance with Condition A of the Binding theory. This is, however, redundant with the ECP, as we saw earlier, raising the question as to whether NP-trace ought to have any binding features at all.

The array of features present on the trace in (115) raises a very serious problem for the FMH. The EC in object position -- the trace of the moved PRO -- has the same binding features as PRO, and is governed. It is therefore subject to both Condition A and Condition B, which cannot both be satisfied; hence the sentence violates the PRO theorem.

Recall that the Topicalization of PRO, which is impossible, is ruled out within the FMH as a violation of the PRO theorem. In sentence (116), the EC bound by PRO is assigned the features

+A+P, and thus violates the PRO theorem.

116) *John_i thinks that [PRO_i [Mary loves t_i]]
(eg. 'John_i thinks that Mary loves him_i')
i i i

Once again, it appears that A' bound ECs have binding features, but A-bound ECs do not.

Observe that the redundancy between Condition A and the ECP in the distribution of trace, a redundancy present in the standard view of ECs, occurs only with respect to A chains, that is, NP-trace. In (117), John serves both as the antecedent which satisfies the trace's requirements as an anaphor, satisfying Condition A, and as the antecedent which permits the trace to be properly antecedent governed. Thus the redundancy.

117) John_i was seen t_i
i i

In the case of A' bound trace, however, the redundancy does not exist. Consider (118).

118) who_i do you [t'_i [like t_i]]
i i i

Here, the argument trace t_i is properly governed by its chain internal antecedent t'_i. However, on the standard view of ECs,

as well as on the FMH view, the trace is an R-expression, and is not subject at all to Condition A. If the A' bound trace is associated with an anaphor, as in (119), it is subject to Condition A on the FMH, but the antecedent which satisfies it as

far as Condition A goes is not the same antecedent as the one responsible for it being properly governed.

119) Himself_i, John [t'_i [likes t_i]]
 i VP i VP i

Therefore, we might adopt the following position:

(120) Only ECs occurring as non-heads of A' chains (eg. WH-traces) bear binding features.

This means that NP-trace has no binding features, which is an acceptable result, since its distribution will be given entirely by the ECP. If we adopt this position, the auxiliary Condition (49) can be eliminated, since it was introduced only to prevent the FMH from ruling out cases of A-movement as in (109-111). (49) can actually be derived from (120), since the binding theory only applies to elements which bear features.

Thus we are still able to predict the range of facts in (106-108); we eliminate the problems raised by (109-111); we correctly rule out Topicalized PRO (see (116)), while avoiding the most serious problem for the FMH, passivized PRO (see (115)).

We must still confront two empirical questions. The ungrammaticality of (112) was subsumed under (49) in the FMH, preventing an anaphor from satisfying its own binding requirements. (120) allows us to eliminate (49), but then the problem of how to rule out (112) arises. In the discussion of (49), it was noted that the application of (49) to (112) is

Finally, consider this example:

123) He_i thinks that John_i was arrested t_i

By the formulation of Condition C given by Chomsky (1985), this ought to be grammatical, since the R-expression John is free in the domain of the head of its chain, the chain being (John,

³
t). This suggests that Chomsky's Condition C must be modified

⁴
as follows :

124) An R-expression must be A-free (in the domain of the head of its A' chain)

This now correctly rules out (123). Crucially, it fails to rule out cases of improper movement like (121), since the R-expression in subject position is in an A-chain with John, not

apply to governed anaphors and pronominals, and to all R-expressions.

3. I am grateful to Elaine McNulty for pointing out this argument to me. Howard Lasnik observes that the same problem is raised for an even simpler case, (i):

(i) * He_i saw John_i

John is free in the domain of the head of its chain.

4. That is, an R-expression must be free in the domain of the head of its A' chain, if it has one; otherwise it must be entirely A-free. The longest expansion of the condition, following standard formalism in syntax and phonology, applies first; if it cannot, the shorter expansion applies.

5
an A' chain .

Thus we suppose that Brody is correct, and that what is amiss in cases of improper movement is movement from an ungoverned position. By Brody's condition, the trace will have to head its chain, giving rise (in (121)) to the chains (John) and (EC). This violates the theta criterion, as John has no θ role.

Within Brody's theory, it is desirable that PRO not be considered an anaphoric pronominal. If this were the case the restriction of PRO to ungoverned positions would be given twice: By his condition, and by the Binding Theory. Indeed, one of the chief motivations of his condition is to eliminate the PRO theorem, by allowing the feature +/- Pronominal to be eliminated for ECs. Brody considers both NP-trace and PRO to be pure anaphors. He notes (fn.23), however, that it is possible to consider ECs to be featureless, with the locality between trace and its antecedent given by some other condition than the binding theory (he suggests the Bounding Theory). We have seen this possibility arise earlier, in noting that the ECP restricts

5. Further, it allows such ungrammatical cases as (i):

(i) He thinks that John , I like t
i i i

John is free in the domain of the head of its A' chain, namely in the domain of John. This suggests that if Condition C is to be maintained as a binding condition, it must be even further restricted. In the text, I develop an account of 'Condition C' effects which eliminates the condition as a separate well-formedness condition.

the appearance of NP-trace to a subclass of the positions in which Condition A lets it appear. So, let us suppose that neither NP-trace nor PRO have binding features, and that ECs are optionally taken to be arguments. By the CLC, an ungoverned EC must head its chain, a governed EC must not head its chain.

We now have arrived at the following position: all and only A' dependent ECs have binding features, and these features match those of the head of the A' chain. We have made crucial use of Brody's theory, yet departed from his major proposals in supposing that NP-trace and PRO do not have any features, and A' dependent ECs have some assignment of both anaphoric and pronominal features. In Brody's system, there is no feature +/- pronominal for ECs.

Given the Feature Matching Hypothesis, this means that there can only be three types of feature assignments to ECs: +A-P, -A-P, and -A+P. The 'gap' in this paradigm is the lack of an anaphoric pronominal, that is, a counterpart to PRO. There can be no such empty category in this extension of the FMH system, since the only ECs which have features are those which are non-heads of A' chains (by (120)). If there were a +A+P empty category, it would have to be a non-head, since it bears features; but by Brody's condition, an EC is a non-head of its chain iff it is governed. But, if the +A+P empty category were governed, it would be subject to both Conditions A and B of the BT, and would be ruled out. Thus the lack of an anaphoric

pronominal appears to follow from a conspiracy of the feature match hypothesis, Brody's condition, and the Binding Theory.

Consider this result, however. In the Chomsky (1981, 1985) formulations of the Binding Theory, Conditions A and B apply only to governed elements; this is why, in the standard theory

of ECs, PRO can exist, but only in ungoverned positions .

6. In examining the role of government in the BT, we need to examine a crucial case, where a pronoun exceptionally casemarked by a verb must be disjoint in reference from that verb's subject, while an anaphor in that position can be bound by the verb's subject:

i) John believes himself to be intelligent
 i i

ii) *John believes him to be intelligent
 i i

Thus it appears that the matrix clause is the GC for both elements.

Within the (1981) theory, this result was achieved by incorporating accessibility of subject into the definition of governing category. In the infinitival clause [X to be intelligent], there is no accessible subject, and so the GC is the higher clause.

In the (1985) theory of binding, which eliminates the notion of 'accessibility of subject', the result is achieved differently. For the anaphor, there is no BT-compatible indexing in the infinitival IP which could satisfy the Binding Theory, and so the matrix IP becomes the GC for the anaphor. There is a BT compatible indexing within this lower IP when the subject is a pronoun, since in any CFC there is a BT-compatible indexing for a pronoun. What forces the matrix clause to be the GC for the pronoun -- and hence be the domain of disjoint reference -- is that there is no governor for the pronoun in this lower clause. The GC for any element must include its governor, and the governor of the pronoun is the verb believe. Once we are in a domain containing the verb, we are in a new CFC, which must then comprise the matrix clause. Therefore, Condition B applies to the pronoun in the matrix IP, forcing it to be disjoint in

In the framework developed in Chapter 3, the binding theory applies to all anaphors and pronominals, and requires them to be governed. Thus the accessibility-sequence formulation of the BT blocks the appearance of an anaphoric pronominal in any context. Therefore, there is a redundancy between this consequence of the Chapter 3 binding theory, and the conspiracy noted in the preceding paragraph.

How might this redundancy be eliminated? By either set of principles, there cannot be a licensed anaphoric pronominal. The determination of the binding domain for overt pronominals requires that government be specified in the BT; therefore, let us keep the first result, namely that PRO cannot be licensed because it will violate the Binding Theory, and seek to eliminate the second. The second way of ruling out the appearance of an anaphoric pronominal is to adopt Brody's Case-Linking Condition, the Binding Theory, and the assumption (120), which holds that ECs occurring as non-heads of A' chains bear binding features. Both Brody's condition and the Binding theory are independently motivated. Therefore, I suggest that what is to be abandoned is (120). We have earlier argued that ECs which are not non-heads of A' chains do not bear binding features. Thus, abandoning (120) means that no empty categories

reference from the NP John. Thus, in contrast to the 1981 theory, the 1985 theory must make reference to government, independently of the PRO theorem. As noted in Chapter 3, this result carries over to the revised BT I develop there.

bear binding features.

4.9 The No Features Hypothesis

We can achieve all the right empirical results (cf. 109-111, 114-115)) if we assume (120), which holds that there is a fundamental difference between chain-internally A' bound EC on the one hand, and NP-trace on the other. The former bear binding features, while the latter do not. We follow Brody in assuming no featural distinction between PRO and NP-trace; their complementary distribution (governed or ungoverned) follows from Brody's condition, and the relation between NP-trace and its chain-internal antecedent follows from the ECP.

- (120) Only ECs occurring as non-heads of A' chains
bear binding features.

Although (120) appears to be empirically adequate, in covering all the data not treated under the standard view of ECs, it seems conceptually odd. Why should A' bound ECs bear binding features, while other ECs do not? Further, we have just seen that the appearance of a redundancy motivates abandoning (120).

This leads us to explore an alternative hypothesis, namely that no ECs bear features. The basic phenomena to be accounted for, still, are (106-108), and further data of this type involving parasitic gap constructions, to be addressed in the

next section.

125) The No Features Hypothesis (NFH)

Only overt elements bear binding features

Most of this discussion has been put in terms of indexing theory. I will now adopt linking theory, as it provides a quite simple means of defending (125).

Let us consider the paradigm in (106-108), repeated here:

106) NP* is an R-expression

- a. *John_i , he_i admires t_i
- b. *John_i , he_i thinks that Mary admires t_i

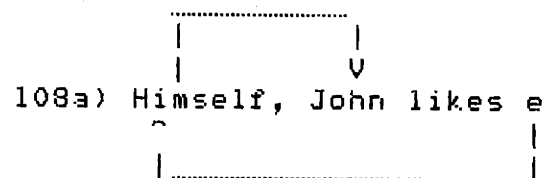
107) NP* is a pronominal

- a. *Him_i , John_i likes t_i
- *It's him_i who John_i likes t_i
- b. Him_i , John_i thinks Mary likes t_i
- It's him_i who John_i thinks Mary likes t_i

108) NP* is an Anaphor

- a. *Himself_i , John_i likes t_i
- *It's himself_i wha John_i likes t_i
- b. Himself_i , John_i thinks Mary likes t_i
- It's himself_i who John_i thinks Mary likes t_i

Indexing theory provides (on the intended interpretations) exactly one indexing for each example. Linking theory provides more than one. Consider (108a) for example. (I assume that e must link to himself, under movement; see Higginbotham (1983)).



This is the linking allowed under the formulation of Condition A within linking theory advanced in Chapter 3. Condition A forces an anaphor to link to a minimally accessible antecedent; John is minimally accessible, through the sequence {himself, e, VP, I', IP} and so the link shown satisfies Condition A.

No further linkings are permitted. Linking the EC to John is not permitted, by Higginbotham's condition R, repeated below along with other principles of the linking theory of binding.⁷ Since the EC has its interpretation fixed by the anaphor to which it is linked, it cannot further link to any other position.

Principles of Linking (Higginbotham 1983)

- 126) (R) The interpretation of an expression is given in one and only one way.

7. (129ii) is abandoned by Higginbotham in later work; it plays no role in my discussion below, and I include it basically for completeness.

127) Anti-C-Command Condition: No element may c-command its antecedent.

128) *D(X,X) [no element may be dependent upon itself].

definition: X is dependent upon Y if: (i) Y is contained in an antecedent of X, or (ii) for some Z, X is dependent on Z, and Z is dependent on Y.

129) Definition of Antecedent: Y is an antecedent of X iff (i) or (ii):

i) X is linked to Y, or for some Z, X is linked to Z and Y is an antecedent of Z.

ii) X c-commands Y, and X and Y share an antecedent.

129') Constraint on linking: linking is free between A positions, and automatic under movement.

Similarly, the linking of John to either e or to himself is barred by (126), since John already has its entire interpretation lexically fixed, thus preventing it from acquiring any other interpretation through anaphoric dependence. I shall understand the major consequence of (R) to be that traces cannot link to more than one expression, and that names cannot link at all.

Note that the linking just suggested for (108a) implies that the trace itself has no A-binding requirements. Earlier, we saw

8
that it is impossible to tell, in indexing theory, whether a trace A' bound by an anaphor is subject to Condition A, or is

8. See discussion of (108) at the beginning of this section.

not subject to any of the conditions of Binding. Within linking theory, we arrive at the latter possibility.

Now let us consider the ungrammatical structures of (107), which, we argued, are to be subsumed as Condition B effects under the licensing condition for anaphors and pronominals. We may state Condition B, following the discussion of Chapter 3, as in (130). ((130) follows from the full formalization given in Chapter 3, but due to the complexity of that condition I give (130) as a summary of the consequence of the full binding theory. I shall shortly revise (130) as a condition on obviation, following (Higginbotham (1985) and Lasnik (1976), but the present formalization is sufficient for the immediate discussion.):

130) Condition B (Linking Version)

A pronominal P may not link to X, if X is minimally chain-accessible to P.

X is minimally chain accessible to Y iff:

- (i) X is chain accessible to Y through some sequence S_1 , and
- (ii) there is no proper subsequence S_2 of S_1 , such that S_2 is a chain-accessibility sequence for Y, and some Z is chain-accessible to Y through S_2 .

(107a) has the partial linking in (131), effected through movement:

*Him, John likes t

(132)

.....
| |
| U
*Him, John likes t
|
|

When the EC is more deeply embedded, it is possible to have coreference between John and the pronoun, as we saw in (107b). I defer full discussion of this example until the next section.

but let us briefly consider it.

(107b) Him, John thinks Mary likes t

We observed that the contrast in grammaticality between (107a) and (107b) might be viewed as being due to t acting as a pronominal, motivating the assignment of the features -A+P to the EC, forcing it to be disjoint in reference from any c-commanding NP in its governing category. We have now abandoned the assignment of features to ECs, so this contrast must be accounted for differently.

Within Linking theory, we can achieve the right result without assigning features to the EC. Again, through movement, we have⁹ the partial linking (133).

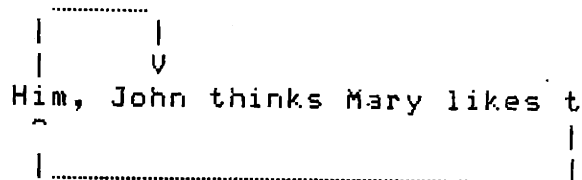
(133)

Him, John thinks Mary likes t
|
| |

Here, John is chain accessible to the pronominal, but not minimally accessible; thus the linking of him to John, as in (134), is grammatical.

9. The watchful reader will notice that I have suppressed mention of any intermediate traces in these examples. Strictly speaking, there will be three intermediate traces, and so there will be no direct link from t to the topicalized pronoun. I address this immediately in the next section.

(134)



This allows coreference.

Let us finally turn to the case of (106), in which the left-peripheral NP is an R-expression.

106) NP* is an R-expression

- a. *John, he admires t
- b. *John, he thinks that Mary admires t
(out with coreference between John and he)

For these, we have the preliminary linkings in (135)¹⁰:

(135)

- a. *John, he admires t
 |
 |
 |
- b. *John, he thinks that Mary admires t
 |
 |
 |

Linking of John or the EC to anything, including the pronoun, is barred, for the reasons shown in the other examples. The only way in which coreference can be established between he and John

10. These are not entirely accurate; the EC will be linked to John through a series of links, given that there are several intermediate traces in the sentences. I return to this shortly. The simplified structures presented here serve to illustrate the basic line of argument.

is to link as in (136)¹¹ :

(136)

a. *John, he admires t

b. *John, he thinks that Mary admires t

What precludes this linking? It might be suggested that pronouns can never be linked to non-argument positions, but this would seem to be falsified by the following:

137) John, I think he's a nice guy

Clearly, here, the pronoun is dependent upon John, and this

11. Linking the pronoun to the trace violates the anti-c-command condition (127). At the end of this section, I will offer a reformulation of the Local Binding Condition on Chains, which accounts for the examples for which it was motivated (the incompatibility of reflexive clitics and raising constructions), as well as ruling out both the linking in (136) and the linking of the pronoun to the trace in (136'):

(136')

a. *John, he admires t

This will allow us to eliminate the anti-c-command condition from the theory. developing this reformulation of the LBC is the main concern of the remainder of this section.

should be represented by the linking in (138):

(138) John, I think he's a nice guy
 |
 |.....|

Further, it has been observed (Chomsky (1982, fn. 11)) that there is a suppression of weak crossover in restrictive relatives where the head is definite:

139) the man [who [his mother said I should meet t]

 'the man x such that x's mother said I should meet x'

However, there is a retention of Strong crossover in restrictive relatives:

140) *the man who he likes t

 not: 'the man x such that x likes x'

I follow Higginbotham (lectures) and Sells (1982) in presuming that the essential fact about such relative clauses as (139) is that the head must be independently referential, i.e. non-quantificational. I return to the linking relations in such structures shortly. When the head is quantificational, as in (141), there is a weak crossover effect; (141) cannot mean 'every man x such that x's mother loves x'.

141) every man who his mother said I should meet t

Observe as well that such left-dislocation examples as (137) are

12

not possible with a quantificational NP in left position :

142) * (as for) every man, I think he's a nice guy

Thus we need to capture the following array of facts concerning A' dependent pronouns: a pronoun can be anaphorically dependent upon a referential NP in an A' position (137-8, 139'), unless it c-commands a trace also linked to the referential NP (140), and our original case, (136)); and a pronoun cannot be directly

13

linked under any circumstances at LF to a quantificational (non-referential) NP (141, 142). The latter case, but not the former, is explained by a condition proposed by Higginbotham. In the next section, we review and discuss this condition, which accounts for some of the facts noted above. We will see, however, that it fails to rule out (136). We will then return to the core case (136); we will develop a view under which what is wrong with that sentence is that the chain relating the topic John to its trace is ill-formed, due to a violation of a

12. This fact is noted and discussed in Epstein (1983)

13. For an example like (i), the relevant representation is the LF (ii):

(i) [Every man thinks he is nice]
S

(ii) [[every man] [t thinks [he is nice]]]
S S

? | ^
| |

In (ii), the pronoun is not linked to the quantificational NP [every man], but to its trace.

generalized Locality Condition on Chains. This reformulation of the LBC will also allow us to eliminate (127) as a separate principle, and will allow us to derive the anti-c-command condition on parasitic gaps.

4.9.1 Review of The Accessibility Condition

I adopt Higginbotham's (1983) Accessibility Condition, which constrains relations between pronouns and quantifiers; it specifies the restricted syntactic circumstances under which a pronoun can be understood as a variable dependent upon a quantifier. In (143), for example, he or him can be understood as a variable with respect to every man, but not in (144).

143) Every man thinks that he is handsome

Some picture of every man annoys him

144) some of his friends annoy every man

which pictures of every man does he especially like ?

The definitions necessary for the condition are summarized as (145).

(145) A formal variable is an empty category occurring in an A-position, and linked to a non-argument (phrase in a A' position); the operator to which the formal variable is linked is its binder. A sequence (v_1, \dots, v_n) of formal variables is a V-chain if each variable v_i is contained in the binder of v_{i+1} .

The structure (147) (Higginbotham's (59)) is the LF representation of the S-Structure (146) (Higginbotham's (58)),

(I) P is accessible to y if v c-commands P; and
 $\quad \quad \quad n$

(II) P is not accessible to y if P c-commands v .
 $\quad \quad \quad n$

Consider the contrast between (150a), (151a), and (152a). On the interpretation in which the pronoun is a variable dependent on every man, (150) is fine; (151) has the weak ungrammaticality of a Weak Crossover violation (cf. '?who does his mother love?'), and (152) has the strong unacceptability of a Strong Crossover violation (cf. '*who does he like?'). The (b) examples are the LF representations.

(150) a. Some picture of every man annoys him

b. [every man][Some picture of t'] [t annoys him]

^		^	
.....		

(151) a. His mother likes some picture of every man

b.[every man][Some picture of t'] [his mother likes t]

^		^	
.....		

(152) a. He likes some picture of every man

b.[every man][Some picture of t'] [he likes t]

^		^	
.....		

On the intended interpretation, the pronoun is dependent upon the formal variable t. The pronoun P therefore must be accessible to this trace. The maximal V-chain for this EC is (t, t'). In (150), the condition for accessibility is met; t c-commands P, and P does not c-command t'. In (151), since t

does not c-command P, P is not accessible to the formal variable t, and so the reading where P is interpreted as a variable is blocked. Similarly, both parts of the accessibility condition are violated in (152); t' fails to c-command P, and P does c-command t. (152) is more strongly ungrammatical than (151); this is presumably due to the fact that both clauses (I) and (II) of the AC are violated in (152), while only (I) is violated in (151).

Now let us return to our cases, (135-142).

4.9.2 A' Dependent Pronouns

Higginbotham's Accessibility condition straightforwardly rules out the example (141). The head of the relative clause is quantificational, i.e. an operator, and hence the pronoun must be accessible to e in order for the bound-variable reading to be licit. Since e does not c-command the pronoun, the pronoun is not accessible; hence (141) is a case of weak crossover, like (151).

For the example (142), there is no formal variable associated with the quantificational NP (recall that formal variables must be empty categories). There are two possibilities for ruling this type of structure out. First, it could be supposed that since there is no formal variable associated with the quantifier, the accessibility condition is violated: there is no y for the pronominal to be dependent upon. Secondly, it could

be supposed that this structure violates the ban on vacuous quantification, which requires that every quantifier bind a syntactic variable. This position could only be maintained if pronouns are excluded from the class of items which count as syntactic variables, which seems doubtful, given the partial acceptability of resumptive pronoun structures in English:

153) ??Which man do you believe the claim that Mary wonders whether he's a competent mayor?

WH phrases may marginally bind pronouns when the pronouns are inside islands.

(142) does not improve when the pronoun is embedded within an island:

154) * Every man, I wonder whether he's a good son

Apparently either the accessibility condition or the resumptive strategy distinguishes between WH phrases and quantificational NPs.

Either way, the unacceptability of (142) can be partially accounted for. What is crucial for us is that the simple coreference of (138) is distinguished from the operator-pronoun relation of (142).

The suppression of weak crossover in relative clauses with definite (and thus non-quantificational) heads, as in (139), also falls out from Higginbotham's system. The weak crossover effect is due to a violation of clause (1) of the accessibility

condition. Since the head of the relative clause in (139) is not quantificational, the accessibility condition is simply not involved.

Considering (139), two linkings might be possible: link the pronominal directly to the head, or to the EC.

155) the man [who [his mother said I should meet t]
[]

156) the man [who [his mother said I should meet t]

This leaves two cases to be distinguished:

(157) NP...pronoun...
 (grammatical, corresponding to (138))

(158) $\begin{array}{c} \text{NP} \dots \text{pronoun} \dots \text{t} \\ \text{NP} \dots \text{pronoun} \dots \text{t} \end{array}$ (ungrammatical, corresponding to (136), (106))

(Where in both structures NP is definite.)

Configuration (158) is not subsumable under the Accessibility Condition, since NP here is not quantificational, and since the pronoun is not dependent upon t anyway. Configuration (158) is instantiated in the case (136), to which we return.

Consider for a moment the typical strong crossover example (159), which cannot mean 'which man x is such that x likes x ?':

159) which man does he like t?

Higginbotham (1983) considers this example, noting that it has two linking structures.

160) [[which man] does[he like t]]?

161) [[which man] does[he like t]]?

The linking in (161) violates the Accessibility Condition. By the definition of dependence (128), he is dependent upon the formal variable t, and thus must be accessible to it, which it

14
is not .

The linking in (160) does not violate the AC, since he is not dependent upon the variable, by definition (128); the AC is not

14. Higginbotham further notes that (161) also violates the anti-c-command condition (127). More complex examples can be constructed, in which he is linked to a variable which it does not c-command; these examples are no better than (161), indicating that the AC is separately involved.

(i) *which picture of which man does he like? (with the bound variable reading of the pronoun).

(ii)LF: [which man] [which picture of t] [he likes t']

invoked in this case.

Recall that Higginbotham proposes a constraint on the assignment of linkings: linking is free between A positions, and obligatory under movement. Thus an expression X can link to Y only if they both occupy A positions, or if there is a movement relation between them. In (161), which man is in an A' position, so the linking between it and the pronoun cannot have been assigned except under movement. But since there is no movement relation between the two phrases (the pronoun is not the trace of the WH-phrase), the linking is ungrammatical.

However, two reasons to reject this way of ruling out (160) emerge from the kinds of structures we are examining in this thesis. First, in a left-dislocation sentence like (162), John is apparently in an A' position, and there certainly is no movement relation between John and he, and so the linking in (162) is indistinguishable from that between which man and he in (160).

(162) John, I think he's a nice guy

┌──────────┐
└──────────┘

By Higginbotham's line of argument, this linking should be ungrammatical (or unassignable). It might be possible to suggest that there is a hidden preposition in (160), so that the structure is syntactically 'as for John,...', but this somewhat questionable rebuttal will not extend to our second case.

Observe that in cases of topicalization, the topic phrase is in an A' position. When this phrase is an anaphor, it will link to an argument:

(163) I think that himself, John likes

|
|.....|

Here, this is a linking between an A' position and an A position; it too should be unassignable, since linking is free, in Higginbotham's framework, only between A positions.

Therefore, we must find some way of excluding (160) other than appealing to this constraint on linking; apparently, linking must be entirely free, between positions of any type.

(160)'s linking structure is just like that of (158), except that the left-peripheral phrase in (158) is not quantificational. As we've seen, the AC does not rule out (160), and therefore it is likely that we can rule the two out together.

What I would like to suggest is that the linking represented in the schema (158) (exemplified by (136) and similar structures) can be ruled out by a generalization of the linking formulation of the Local Binding Condition on chains (Rizzi (1982), Lasnik (1985)), to cover A' chains as well as A chains. I explore and develop this account in the next section. A consequence of this proposal will be that we can explain the ungrammaticality of the examples in (106) while maintaining the

No Features Hypothesis.

4.9.3 Generalizing the Local Binding Condition

We have discussed structures of the form (157) and (158), which differ in grammaticality. The purpose of this discussion is to advance a general principle on chains which will rule out (158).

(157) NP...pronoun...
 |
 |.....| (grammatical, corresponding to (138))

(158) NP...pronoun ... t
 |
 |.....| (ungrammatical, corresponding to (136) (106))

The obvious difference between (157) and (158) is the presence of a trace in the latter case. This trace forms a chain with NP in (158), exemplified in (136a).

(136)
 |
 |
 |
 a. *John, he admires t
 |
 |.....|

Suppose we take the fundamental distinction between the two linking structures to be that the presence of the pronoun in (158) interferes with the well-formedness of the chain (NP, t).

In chapter 3, we reviewed Rizzi's (1982) argument that a Local

Binding Condition constrains the well-formedness of A chains. The constraint is repeated as (164); one of Rizzi's crucial examples is (165).

(164) $C = (a_1, \dots, a_n)$ is a chain iff, for
 $1 \leq i \leq n$, a_i is the local binder of a_{i+1} .

(i.e. (X, Y) is a link of a well-formed chain only if X is the local binder of Y).

(165) * Gianni _i si _i e stato affidato t' _i t _i

In order to satisfy the theta criterion, (Gianni, t) must be a chain; but it is ill-formed, since Gianni is not the local binder of its trace t. Rather, both si and its trace t' are closer binders of Gianni's trace.

Within linking theory, Rizzi's example has this representation at S-Structure:

(166) * Gianni si e stato affidato t' t

The LBC can be recast in terms of linking:

(167) (X, Y) is a link of an A chain only if:

- (i) Y is linked to X, and
- (ii) if Z is linked to X, then Z does not c-command Y.

(167ii) captures the sense of 'local binder'. In (166), (Gianni, t) is not a well-formed A chain, since although t is linked to Gianni, si is also linked to Gianni, and si c-commands t.

Abstractly, the illicit linking in (166) is very similar to the illicit linking in (158). They differ in three respects. In (166), the intervening element is a reflexive clitic, while in (158) it is a pronominal. We may assume that this is irrelevant. In (166), but not (158), there is another trace linked to the intervening binder; we may assume this to be irrelevant as well. The final, significant, difference is that in (166), the relevant chain (Gianni, t) is an A chain, while in (158) the chain (John, t) is an A' chain. Let us then suppose that we can generalize (167) to cover chains of either type:

(169) $C = (a_1, \dots, a_n)$ is a well-formed (A or A') chain only if:

(i) for every a_i in C, a_i is linked to a_{i+1} ; and

(ii) if Z is linked to a_i , then Z does not c-command a_{i+1} .

This generalization of the local antecedence condition on chains suffices to rule out the general schema (158); in (136a) the chain (John, t) is ill-formed, since he is linked to John, and he c-commands t.

(158) was the last case motivating the Feature Match Hypothesis. (169) rules out (136) without making any reference to binding features on the empty category of (136) (α in the $i+1$ condition (169)). We therefore have succeeded in maintaining the No Features Hypothesis, and we have accounted for all the data which motivated a reconsideration of the standard view of empty categories.

The remainder of this section is devoted to certain modifications of (169).

4.9.4 On Deriving the Anti-C-Command Condition on Linking

In this section, I would like to suggest that we can reformulate (169) to rule out coreference in examples like (170).

(170) He likes John

In indexing theory, the coreferential interpretation of (170) has just one syntactic representation, (171). This is ruled out by Condition C of the binding theory, since the R-expression John is A-bound.

(171) $\begin{matrix} *He & \text{likes} & John \\ i & & i \end{matrix}$

In linking theory, as Higginbotham (1983) notes, this

interpretation of (170) has two linking structures .

(172) He likes John

~
| |

(173) He likes John

~
| |

(172) is ruled out by Condition R (see (126)); names have their interpretation lexically determined, hence they cannot be referentially dependent. Higginbotham (1983) rules out the linking in (173) by positing the anti-c-command condition I repeated as (127) Higginbotham (1985) departs from this condition, instead ruling out (173) as a violation of his Condition C*.

(174) Condition C*

R-expressions are obviative from every c-commanding expression in an A-position.

(175) Definition: X and Y are obviative iff the structure in which they occur does not determine them to share

16

a value .

15. The string (170) of course has another representation, where neither NP is linked to the other; this fails to determine coreference. Other possible linkings (link John to John, link each NP to the other at the same time, etc.) are ruled out by the other principles of linking.

16. Higginbotham follows Lasnik (1976), (1981) in supposing that obviation (or disjointness of reference, in Lasnik's terms) and anaphoric dependence (coreference) are fundamentally distinct relations of grammar. I essentially agree with this viewpoint. It should be noted that the approach to such cases as (173) advanced here is not inconsistent with this view, as will be

In (173), the structure determines he and John to share a value (John), since they are linked. Hence (173) violates (174), since he c-commands John. (Formulating the binding theory in terms of obviation has certain advantages, chiefly in consideration of Condition B effects; I shall review Higginbotham's supportive arguments later). The restriction in (174) that R-expressions need be obviative only from A-positions, not A'-positions, is important, since otherwise the linking in (176) is illicit:

(176) $\begin{array}{c} \text{---} \\ | \\ \downarrow \\ \text{himself John likes t} \\ | \\ \text{---} \end{array}$

discussed to some extent later in the text. It should be noted, however, that there is one class of examples which receives a natural treatment in Lasnik and Higginbotham's theories, but about which I have little to say. This is the paradigm in (i) - (iii) (see Lasnik (1976) for much discussion)).

- (i) After John walked in, the bastard sat down.
- (ii) John, I really like the bastard.
- (iii) *John thinks that the bastard is really nice (on the coreferent reading of John and the bastard)

Such epithets may be anaphorically dependent upon either a non-c-commanding expression in an A position (i), or upon a c-commanding A' position (ii), but not upon a c-commanding A position. Within a theory that recognizes obviation and dependence, the facts are easily accounted for: epithets are similar to names in that they must be obviate from every c-commanding A position, but differ from names in that they may be anaphorically dependent. Higginbotham's (1985) treatment of standard Condition C examples like (173) (his Condition C*) extends to cases like (iii), but mine does not. I hope to explore this problem for my treatment in future work.

Now, observe that the condition ruling out (173) crucially makes reference to c-commanding elements; a representation like (173) is only barred because he c-commands John. So, (177) is grammatical:

177) [His mother] loves John
 | |
 |.....|

The extended local binding condition on chains (169) also singles out c-command as a crucial factor; a chain is only disrupted if Z c-commands Y. (178) is grammatical; each other and t are both linked to they, but since each other doesn't c-command t the chain (they, t) is well-formed¹⁷.

178) They seem to [each other] [t to be nice]
 | | |
 |.....| |
 |.....|

It is not implausible that two separate principles, say (169) and (174), both separately make use of c-command, one of the core relations of grammar. However, it should be noted that we can collapse (169) and (174) together.

 17. This is supported by the fact that (i) is better than (ii):
 (i) ? It seems to him that John is nice (ii) * He thinks
 i i i
 that John is nice
 i

This fact is noted and discussed in Lasnik (1985). The slight unacceptability of coreference in cases like (i) is discussed in Barss, Johnson and Saito (in progress).

There is a certain conceptual similarity between (169) and (174). (174) requires that R-expressions be obviative, while (169) requires that a chain internal expression be obviative within the domain of its chain-internal binder, obviative from everything but this binder.

Suppose we collapse the two as follows:

(179) The Chain Obviation Condition

For a chain $C = (a_1, \dots, a_n)$, and for an expression Z where Z is not a member of C and Z is in an A position, if Z is dependent upon a_i , then Z does not c-command any member a_j of C .

That is, if some expression is dependent upon part of a chain, it cannot c-command any part of the chain. In essence, I am claiming that chains are obviative from every c-commanding A-position expression (but crucially not every c-commanding A' expression, as the grammaticality of (23) shows). For case (166), take si to be Z , Gianni to be a_i , and the trace t to be a_j . The relevant chain is (Gianni, t) . si is dependent upon Gianni, and c-commands t , in violation of (179). Similarly, the trace t' of si is dependent upon Gianni, and t' c-commands t , in violation of (179).

For (173), take he to be Z , and take John to be both a_i and

a . This also violates (179).

j

In both cases, an A-position element c-commands part of a chain upon which it is anaphorically dependent. If this line of argument is correct, then we can eliminate reference to R-expressions entirely from the binding theory.

It may be desirable for another reason to collapse (169) and (174) together. Consider (180).

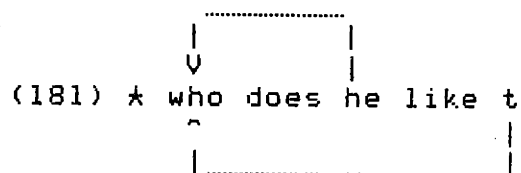
180) * He thinks Dan was arrested t
 | |
 | | | |

In our formulation of (179), the hierarchical relation between X and Z is not specified. This means that either Z is within the c-command domain of X, as in Rizzi's example (166), or X is within the c-command domain of Z, as in (180) (take he to be Z, Dan to be X, t to be Y). By our formulation, even in (180), the chain (Dan, t) is ill-formed, since he is linked to Dan, and c-commands Dan and t. (180) is redundantly out by Condition C*, since this binding condition requires that Dan be obviate from every c-commander. Collapsing Condition C* and (169) together into (179) eliminates the redundancy.

We could of course keep Condition C* and (169) separate, and specify that (169) only holds when Z is within the c-command domain of X, but this seems undesirable, given the possibility of collapsing (169) with (174).

4.9.5 Strong Crossover

Another structure which manifests the configuration blocked by (179) (and (169)) is that in (181), a simple case of Strong Crossover.



Higginbotham (1983) notes that this linking structure does not violate the accessibility condition, since the pronoun is not dependent upon the formal variable (cf. the discussion above of (160) and (161)). He proposes that the linking of the pronoun to the operator is prohibited, since who is in an A' position, and linking is free only between A positions. We have shown evidence that this restriction on linking ought to be abandoned.

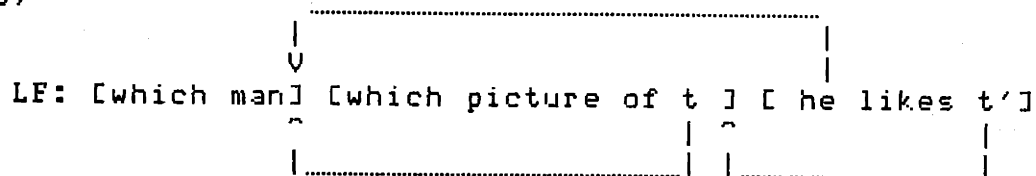
Our generalized locality condition on chains (179) blocks this linking, since the chain (who, t) will be ill-formed.

However, there are more complex cases of strong crossover which cannot be subsumed under (169). In Higginbotham's example (182), he cannot be understood as a variable dependent on which man.

(182) which pictures of which man does he like t?

Suppose at LF the pronoun links directly to which man.

(183)

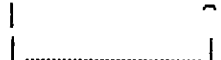


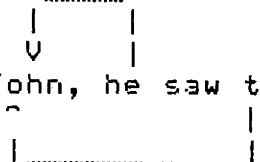
Here, (169) is not violated. The chain headed by the antecedent of he is (which man, t), and the pronoun does not c-command either one of these elements. The accessibility condition does not block this either, since he is not dependent upon t, and therefore the fact that he c-commands t' is irrelevant. two options are possible. We might modify Higginbotham's original proposal that linking is free only between A positions, and say that X may freely link to Y only if Y is in an A position. This allows free linking in (108), although it leaves the linking in (162) as a counterexample. On this view, the linking of he to which man in (181) and (183) is illicit. This will then be redundant, for case (181), with (179) (or (169)).

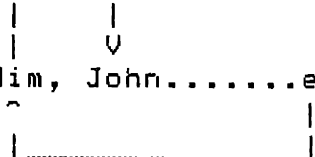
Alternatively, we might strengthen the accessibility condition (AC) by requiring that if a pronoun is to be understood as a variable bound to an operator Q, then it must link to the formal variable bound by Q. Thus in (183) linking he to the WH-phrase directly is insufficient to determine the bound variable interpretation of he. This allows linking to be free between all positions, allowing (108) and (162). Again, there will be a redundancy with respect to (181): both the AC and (179) rule out the linking.

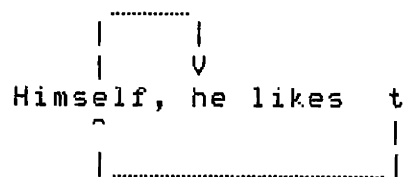
4.9.6 On Indexing

We have now arrived at an adequate account within linking theory of the contrasts in these three representations:

184) * He saw John


185) *John, he saw t


186) Him, John.....e (grammatical)


Himself, he likes t


Linking theory crucially allows us to distinguish (185) from (186), in a way that indexing theory -- which does not encode antecedence directly -- does not.

Indexing theory assigns to them identical indexing structures:
 (187)

NP ... NP ...t
 i i i
 \ / \ /
 c-command c-command

Since the richer notation of linking theory permits a distinction to be made between (185) and (186) in the configuration of its dependencies, we take this as another reason to prefer linking notation over indexing.

4.10 Condition B

We now return to the correct formulation of Condition B in linking theory, in particular focusing on examples like (107). Condition B is that principle of anaphora which rules out the coreferential reading between John and him in (188).

188) John likes him

Theories of binding differ as to what they assert is the phenomenon occurring in such examples. The strongest position is that advocated by Lasnik (1976), (1981), who argues convincingly that pronouns are required to be disjoint in reference from c-commanding NPs in their governing categories. As Lasnik (1981) emphasizes, 'disjoint in reference' is not to be equated with 'non-coreferent'. The indexing theory of binding proposed in Chomsky (1981) is most easily interpreted as a somewhat weaker position, holding that syntactic structures (the indexed phrase marker) are ill-formed if they determine a pronoun to be coreferent with a locally c-commanding NP. It is apparent that the GB indexing system is too weak to handle

certain cases of prohibited pronominal coreference involving plurals, as Lasnik (1981) demonstrates. The theory developed in Higginbotham (1985), which I assume here, is intermediate between the two, although more similar to Lasnik's than to Chomsky's in recognizing two quite distinct relations, dependence (i.e. coreference) and obviation (the equivalent of Lasnik's disjointness of reference).

Higginbotham (1983, 1985) has pressed the point that what is to be ruled out is allowing John to serve as the syntactic antecedent of the pronoun in this example; in fact, the two NPs may corefer, but only if the reference of the pronoun is determined in a way other than by linking up the pronoun to

18
John . Thus it is this structure which is to be ruled out:

18. In Higginbotham's (1985) treatment, a pronoun P must be obviative with respect to every NP X in the governing category of P which c-commands P. X and Y are obviative if the structure in which they occur does not determine them to share a value (i.e. does not determine P to overlap in reference with X). This leaves the possibility open that X and P do share a value, if that is determined by information other than what is present in the syntactic representation. Lasnik's (1976) analysis of examples like (188) imposes a stronger restriction: a pronoun P must be disjoint in reference from X, if X c-commands P in P's governing category. This disjointness of reference must hold regardless of the way in which the two are determined to share a value.

To give one example (suggested by Howard Lasnik) in which the two views make different empirical predictions, consider the following discourse:

(i) Does anyone love Mary?

(ii) Well, Mary loves her.

(189)

John likes him

 |
 |
 |

Higginbotham posits this case:

(191) John thinks he likes him

Here either pronoun, but not both, may have John as an antecedent.

The linking in (192) is ruled out straightforwardly by Condition B:

The example (ii) seems to tolerate an interpretation in which her refers to Mary. What we might say is that her picks up its reference from Mary in the first sentence (i), thus being obviative from Mary in the second example. (her is obviative, by Higginbotham's definition, from Mary in (ii), since no part of the structure (ii) determines them to share a value.) Lasnik's analysis of pronominal coreference, on the other hand, forces Mary and her to be disjoint in reference, thus blocking the interpretation for (ii) which seems at least marginally available. This would seem to be evidence in favor of Higginbotham's view. However, this interpretation, for me, is only really available when Mary is read with prominent stress. I have no explanation for this, but because of this fact, and the marginality of this reading, I hesitate to base too much upon it. I adopt Higginbotham's position in the text, but recognize that Lasnik's may actually be correct, if (ii) can be explained.

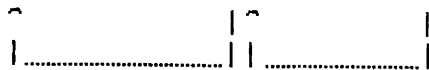
Lasnik's position would seem to be supported by the impossibility of coreference in (iv).

(iii) There's John.

(iv) John likes him

It seems then that there are discourse factors which allow or disallow such coreference. Spelling out a formal theory of such factors is crucial to a full theory of structurally-influenced reference.

192) John thinks he likes him



Here, the pronoun him has a c-commanding antecedent in its governing category (the embedded clause).

In the terms of the chain-accessibility framework developed in chapter 3, he is minimally chain-accessible to him, and so him cannot be linked to he.

Now consider (193).

193) John thinks he likes him



Here, the governing category for the pronoun him is the embedded clause; since him has no antecedent within this clause, there is no violation of condition B.

Thus Condition B must be cast so as to prohibit a pronoun from being c-commanded by an NP which the linking structure determines to share a value with the pronoun. Higginbotham

19

proposes that this is to be put in terms of obviation :

19. See Lasnik (1976); the Disjoint Reference Condition posed there is strongly similar to (194) (although see previous footnote for a possible difference). Both proposals treat cases which receive no complete treatment in Chomsky (1981), like the impossibility of understanding the referent of him to be included in the (plural group) denoted by they in (i).

(i) They like him

X is obviate from Y iff the structure in which they occur does not determine them to share a value.

(194) Condition B*: A pronoun P is obviate from a c-commander Z, where Z is within the governing category for P.

In addition to (193), there is another crucial case which shows that Higginbotham's formulation, in terms of obviation, is to be preferred over one put simply in terms of linking, like (7) of Chapter 3 (see also the final section of this chapter). There is another way in which the structure in which a pronoun occurs can determine it to share a value with another: first and second person pronouns are inherently referentially restricted, and their very appearance in a sentence determines their value (see also Lasnik (1976), (1981)). Higginbotham observes that the contrast between (194) and (195) can be assimilated to that

20

between (196) and (197) in his condition B* .

See also Lasnik (1981) for illuminating discussion of this issue, and the treatments of it available in the 'On Binding' and GB frameworks.

20. The essential point here is that the two pronouns are required to be obviative, although presumably no linking is involved. Thus it would be a mistake to equate 'obviate from' with 'not linked to', since linking X to Y is only one of the ways that X and Y can be determined by the syntactic structure to share a value (i.e. to overlap in reference). Cases like (194)-(196) receive a unified treatment in indexing theory in Lasnik (1981). Lasnik's point is that these cases cannot be handled in an indexing theory which assigns only one index to an expression, as the GB system does; rather, he argues, these examples (and others) are evidence that the richer indexing system of 'On Binding' is well-motivated.

194) *We like me

195) We think that Mary likes me

196) *John likes him

|
|.....|

197) John thinks Mary likes him

|
|.....|

Let us then accept Higginbotham's proposal that the licensing condition for pronouns is to be put in terms of obviation (with obviation understod as in (175)), and adapt it to the chain-accessibility framework of chapter 3. Consider the contrast between (198) and (199) (cf. 107).

198) Him, John thinks Mary loves e

199) Him, Mary thinks John admires e

Coreference between him and John is acceptable only in (198).

We observed that this looks descriptively like the pronoun is lowered into the position of the gap prior to the application of Condition B. We have argued that this is a descriptive generalization, since this type of lowering is not possible in the general case. Further, we have argued that the empty category itself bears no binding features, so this contrast is not explained by supposing that Condition B applies to the EC.

The basic idea I suggest (as an extension of Condition B*) that a pronominal must be obviate from any NP which is minimally chain-accessible to it. More formally, we could reformulate (6)

and (7) of chapter 3 to ensure that a structure is BT-compatible with a pronoun only if the pronoun is obviate with respect to every chain-accessible NP in the structure. A pronominal P will be licensed only if it is obviate from every minimally chain-accessible NP.

(6') Binding Theory Compatible Linking (Definition)

L, a set of assigned links, is binding-theory compatible wrt. an expression A and an accessibility sequence S for A iff:

- i) for A = an anaphor, A is linked under L to a antecedent B which is chain-accessible to A through S.
- ii) for A = pronominal, A is obviative under L from every expression B which is chain-accessible to A through S

(7') Licensing Condition

for A an anaphor, A is licensed only if:

- i) there is a chain accessibility sequence S for A, such that there is assigned a BT-compatible linking L for (A, S) ; and
- ii) there is no proper subsequence S_j of S, such that S_j is a chain-accessibility sequence for A, and there is a possible BT-compatible linking L' for (A, S_j) .

for P a pronoun, P is licensed only if:

- i) for every chain accessibility sequence S for P, (such that there is no proper subsequence S_j of S, such that S_j is a chain-accessibility sequence for P, and there is a possible BT-compatible linking L' for (P, S_j)), there is assigned a BT-compatible linking L for (P, S) .

Basically, the licensing conditions for anaphors works exactly as in the Chapter 3 formulation: given an anaphor, there must be one (out of potentially many) chain-accessibility sequence such that the anaphor is linked to an NP which is accessible to the anaphor through the sequence. (The relevant sequences have to be minimal, in the sense that they cannot have subsequences which have the necessary BT-compatible properties.)

For pronouns, there is a stronger condition: there can be no minimal chain accessibility sequence S for the pronoun such that the pronoun is determined by the syntactic representation to share a value with an NP accessible to the pronoun through S.

It might be objected that the licensing conditions as I have formulated them fail to have the complete symmetry of the licensing condition Chomsky (1985) proposes (a symmetry which carries over to my licensing condition (7) in Chapter 3). Whereas Chomsky's is a single licensing condition for anaphors and pronominals, I have given two separate ones, which are similar but not identical. This would appear to be motivated by the facts. The binding theory is satisfied, for an anaphor, if the anaphor is linked to any NP accessible to it through any minimal chain-accessibility sequence. On the other hand (as I shall show in a moment) the binding theory is satisfied for a pronoun only if the pronoun is obviate from every NP accessible

21

to it through any of its local chain accessibility sequences .

Now we will examine the operation of the licensing condition for pronominals, as reformulated. In (196), John is minimally accessible to the pronoun through the sequence {him, e, VP, I', IP}, hence the linking in the example is ill-formed.

Cases like (198-9) present a very interesting problem. The theory of successive movement holds that the fronted pronoun has been moved through a series of positions; if each position is marked by a trace, then the surface structure representation of

22

(198) is something like (198') (indices used for convenience only):

```

(198') Him [      John [      [      t'''
           i IP*      i I'* VP*      i
           [ thinks [      t'' [      [      Mary
           VP*      CP      i C' IP
           [      [      t' [      loves t ]
           I' VP      i VP      i
  
```

The relevant chain is {him, t''', t'', t', t}. There are a number of chain accessibility sequences for the pronoun, including S1 = {him, t, VP, I', IP} and S2 = {him, t''', VP*,

21. The last section of this chapter demonstrates this point even more forcefully. I introduce evidence that there are certain accessibility sequences -- constructed through parasitic gaps -- through which pronouns must be obviate, although they cannot serve to locate an antecedent for an anaphor.

22. This exact structure is perhaps incorrect. The basic point to be made here is that the existence of intermediate traces -- whatever their exact position -- poses a problem for the formulation of condition B.

I'*, IP*}. John is accessible to the pronoun through only one of these, namely S2. We might either conclude that Condition B is satisfied if there exists one accessibility sequence through which the pronoun is obviate from its antecedent (S1 in this case), or we might suppose that only accessibility sequences involving the theta-marked trace need be taken into account. On this latter view, since John is not minimally accessible to the pronoun through S1 (the sequence constructed through the theta-marked trace) Principle B is satisfied.

We can reject the first possibility immediately, when we consider the contrast between (198) and (199).

(199') * Him [i IP* Mary [[t''' [i VP* thinks
 i IP* I'* VP* i VP*
 [t'' [[John [
 CP i C' IP i I'
 [t' [respects t]
 VP i VP i

In (199), John may not be coreferent with the pronoun, while in (198) this coreference is allowed. There is a chain accessibility sequence {him, t''', VP*, I'*, IP*}, and John is not accessible to the pronoun through it; yet coreference is blocked, showing that Condition B is violated.

This leaves us with the second option: for some reason, it is the chain accessibility sequence which includes the theta-marked trace t which matters. For the pronoun to be conjoint in reference with some NP, that NP must not be minimally accessible to the pronoun through a chain accessibility sequence including

the theta-marked trace.

There are two ways in which this fact can be analyzed. The first is to accord specialized status to the variable trace; this treatment is suggested in Barss (1984), which considers similar contrasts. We might revise the licensing condition for pronouns accordingly, along the lines of the following:

(200)

A pronoun must be obviate from any NP which is minimally accessible to it, through a sequence including the A-position EC bound by the pronoun.

(200) requires that the pronoun be obviate from John in (199'), since here John is minimally accessible through the sequence {him, t, VP, I', IP}, and t is the A position trace. By contrast, John is not minimally accessible to the pronoun through such a sequence in (198'). It is accessible through a sequence containing the A position trace, the sequence S3 = {him, t, VP, I', IP, C', CP, VP*, I'* , IP*}, but not minimally accessible, since Mary is accessible through a proper subsequence of S3. John is minimally accessible through the sequence {him, t'', CP, VP*, I'* , IP*}, and through {him, t''', VP*, I'* , IP*}, but neither of these contains the A position trace, and so by (200) him need not be obviate from John.

(200) is, of course, rather suspicious, since it leaves unexplained why the lowest trace is the only one relevant for this condition (why not the highest one?); and it is further odd that Condition A should allow an antecedent to freely seek an

antecedent which is local to any of the traces in a chain, while condition B is restricted. There is additionally another problem. In the representation (199'), because there are several intermediate traces, there are several minimal chain accessibility sequences. In particular, there are the two $S_1 = \{ \text{him}, t, \text{VP}, I', \text{IP} \}$ and $S_2 = \{ \text{him}, t'', \text{VP}, I', \text{IP} \}$. John is accessible through S_1 , and Mary is accessible through S_2 , while John is not. What we must ask is the following: suppose him is linked to John. Since there is a chain accessibility sequence through which John is not accessible, why isn't this linking allowed? why must pronouns be disjoint in reference from an NP if the NP is accessible through any sequence, while anaphors can be linked to an NP which is accessible through one, but not necessarily every, sequence? This is precisely where the complementarity of the distribution of anaphors and pronouns breaks down, and which led to the separation of the licensing conditions as in (7'') (cf chapter 3, ex. (7)).

This crucially distinguishes pronouns from anaphors, since in the case of anaphors if the anaphor is linked to an NP accessible to it through one minimal CAS, it need not be linked to an NP accessible through any other CAS. Thus there is a fundamental asymmetry between the multiple binding domain effect for anaphors, and for pronouns. We shall now show that (200) derives from the Projection Principle, taken together with the

licensing condition for pronominals given above (7').

In addition to its specialized status in the calculation of obviation for pronouns, the A position trace is distinguished from all the other traces in the chain in another way: it is the only one required by the Projection Principle. Our tentative formulation of (200) was based on the assumption that (198') is the relevant representation for (198), and (199') for (199); from that representation emerges the idea that the lowest trace has specialized status with respect to obviation. What we will now show is that there is no reason to adopt this as a primitive fact. Rather, the specialized status of the object trace in (198)-(199) can be derived.

Lasnik and Saito (1984) (following suggestions by David Pesetsky) propose that movement need leave a trace only when the presence of that trace is forced by independent conditions of grammar, chiefly the Projection Principle. Following this line of argument, we observe that only t in (198') and (199') is necessarily present for the purposes of fulfilling the Projection Principle; all the other traces are optional, as far as this principle goes. (If Chomsky's (1986) formulation of the ECP is correct, then t' is additionally obligatory at S-Structure, since it serves to gamma-mark t in accordance with the ECP.)

Therefore, the sentence (198) may actually have this S-Structure configuration:

(201) Him [John [[thinks [[]
 IP I' VP CP C'
 [Mary [[t' [loves t]
 IP I' VP VP

In this structure, no trace is present outside the c-command domain of Mary. Therefore, for this structure, the only minimal chain accessibility sequences for the topicalized pronoun will be {him, t, VP, I', IP} and {him, t', VP, I', IP}. Crucially, John is not chain-accessible to the topic pronoun through either sequence. (Mary is minimally accessible through both.) Therefore, the licensing condition for pronouns (7') will not be violated, hence coreference is permitted.

Similarly, (199) must have an S-Structure representation in which t is present, by the Projection Principle. There will be a chain accessibility sequence {him, t, VP, I', IP}, and John is accessible to the pronoun through this sequence. Since (7') requires pronouns to be obviative from any NP minimally accessible to the pronoun through any chain accessibility sequence, him cannot link to John in this example. Thus the difference in coreference possibilities in (198) and (199) is accounted for.

The stipulative formulation (200) can now be abandoned, in favor of (202), an informal summary of the licensing condition for pronouns given in (7'):

(202) A pronominal is obviate from any NP minimally chain accessible to it.

Now, no specialized status is accorded the A position trace. It has a specialized role in determining pronominal coreference, but this role is derived from the Projection Principle: it is the only trace necessarily present at S-Structure in (198), hence only accessibility sequences containing it are relevant to Condition B.

4.10.1 At What Level Does Condition B Apply?

In Chapter 2, we suggested that Condition A applies at S-Structure. Although the formulation of Condition B is similar to the formulation of Condition A, this is not evidence that Condition B applies at S-Structure; it is perfectly possible that it applies only at LF, at both levels of representation ²³.

It is clear what type of evidence could show whether Condition B applies solely at SS, or solely at LF.

Above, we showed that the A position trace in (198-9) has specialized status in determining coreference possibilities between the topic pronoun and NPs c-commanding the pronoun's trace. This fact was derived from the fact that the trace is necessarily present to fulfill the Projection Principle, which the other traces are optional. This is true at S-Structure, and

23. In the next chapter, I also reconsider the strength of the evidence adduced in favor of the position that Condition A holds at S-Structure.

at LF.

There is, however, another type of trace which is necessarily present at SS, but which may delete prior to LF: the trace which properly governs a subject trace. Consider (203):

203) who do you think [t' [t left]]?
 CP IP

The EC t' must be present at SS, since it is the only possible proper governor for t. Within the gamma-feature framework of ECP, t' can delete in the LF component, giving rise to the LF representation (204)(see Lasnik and Saito (1984)) for details).

204) who do you think [[t left]]?
 CP IP
 +G

Now consider a variant of (203-4), in which the A' phrase is a pronoun, not a WH-phrase.

205) It's him who John predicted t will win
 It's him who John thinks t might have passed the test
 It's him who Mary told John [t might win the race]
 It's her who Mary thinks t almost got arrested

On my judgement, there is no possibility of coreference between John and him in (205), in contrast to (206), in which the pronoun is associated with an object position trace, and (207), in which it is associated with a more deeply embedded subject position. In both cases, coreference between him and John is permitted.

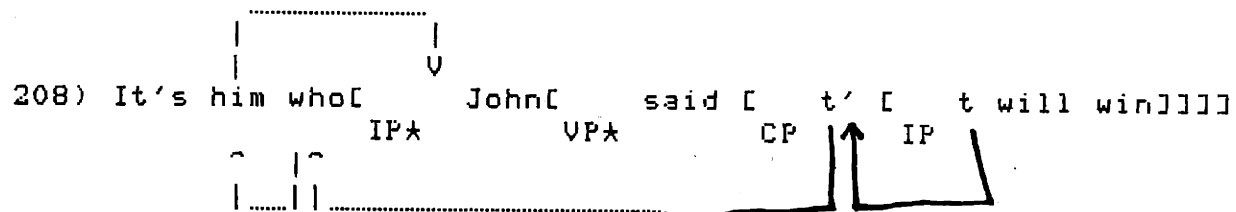
206) It's him who John thinks Mary likes t

207) It's him who John said Mary thinks [t might have passed the

test/ will win

It's him who John knows Mary said [t might have passed the test

Let us consider the structure of a representative example.



As noted above, t' is necessarily present at S-Structure, to gamma mark the trace t. t' may then delete in the LF component.

The judgement on this example indicates that the link between John and the pronoun is illicit, in violation of (202).

Consider now the two accessibility sequences which make John accessible to him: S1 = {him, t, IP, CP, VP*, I'*, IP*}, and S2 = {him, t', CP, VP*, I'*, IP*}. John is not minimally accessible through S1, since there is a proper subsequence of it meeting the definition of accessibility sequence (S3 = {him, t, IP}). Therefore, the accessibility sequence which participates in blocking coreference between John and him is S2, which is defined through the intermediate trace t'.

Thus, Condition B must apply at a level of representation at which t' in (208) is present; by the preceding argument, it is present only at S-Structure, showing that Condition B holds (at least) at this level of representation.

However, another class of examples (suggested by Jim Higginbotham) casts some doubt on this result. Higginbotham observes that when the topic pronoun is in nominative case, then coreference is marginally permitted in sentences structurally identical to (208):

209) It's _i he who John _i thinks t will win

If condition B applied at S-Structure, it should preclude coreference in this example, since it is identical to (208) in structure; only the case of the pronoun differs.

Kayne (1982), and Kiss (1985), propose that certain verbs may assign Case to a trace in the Specifier (or COMP) position of their complement. Developing this suggestion, we might suppose that in (209), Case is assigned to the chain (he, who, t', t) at position t, giving the chain nominative Case (which overtly shows up on the pronoun); while accusative Case is assigned to the chain by the verb think, at position (t'), in example (208). The overt case marked on the topic pronoun will then match the case assigned to the chain. Let us suppose that if a chain is well-formed, the Case of the head of the chain must match the Case assigned to some member of the chain.

Descriptively, we can distinguish the difference in coreference possibilities in (208) from that in (209) in terms of Case: the pronoun has to be obviate from an NP (John) which is minimally accessible to it, through a sequence which includes

the case-marked trace. How may we incorporate this formally?

Suppose, contrary to the proposal above, that Condition B applies only at LF. Let us further suppose that at LF, a chain must contain its case-matched position. Thus, in (209), the trace t' can delete in the LF component, since its presence is not forced by the casematching condition; the case relevant to casematching is t. In (208), t' is necessarily present at LF, since this is the position at which the chain receives Case. While the judgements are very delicate, I believe that they support this conclusion.

To perhaps contradict this line of argument, however, consider this example (judgements reflect the interpretation where him and John are coreferent):

210) It's him who John thinks Mary said left

211) * It's him who Mary thinks John said t won the race

These ought to be equivalent in grammaticality, if the hypothesis just advanced that condition B holds only at LF is maintained. The reason is that the chain has accusative Case, so that, by hypothesis, there must be an accusative-marked trace in the representation at LF. Since the movement in both cases proceeds through the SPEC position of the complement of think, a trace in this position this could be the one which satisfies the case-matching requirement. Thus, (211) could have this LF structure:

211) * It's him who Mary thinks[It' John said[[t won the race]]]

Crucially, it would appear that the SPEC position of the complement of said can be empty at LF (It must have a trace present at S-Structure, however, for the ECP). Therefore, no coreference between him and John ought to be possible, since there is no CAS through which John is minimally accessible to the pronoun.

Because of the delicacy of the judgements, I leave this issue as unresolved, although the approximate structure of the problem is somewhat defined.

4.10.2 A note On Clitics

It is apparently a fact that in the core case of cliticization, the clitic is attached to the verb which assigns its trace a thematic role. This fact is partially derived in the standard theory of ECs, by treating clitic trace as an anaphor; cliticization is thus assimilated to NP-movement.

Within the Barriers theory of ECP, the distribution of NP-trace is given by the ECP, restricting the distribution of an NP-trace and its (chain-internal) antecedent to a subset of the pairs of positions in which an overt anaphor and its antecedent may grammatically occur. In this sense, the ECP 'masks' the effect of Condition A as applied to NP-trace. This undermines the justification for treating NP-trace as an anaphor, in that

doing so does no work in the theory, and it further raises doubt as to whether this treatment is correct, in that Condition A is purely redundant with the ECP. We have further developed a new account of EC features, the No Features Hypothesis, which has as a desirable consequence the elimination of this redundancy.

We have further followed the Barriers view of ECP, in that proper government is restricted to antecedent government. Hence, a clitic trace must be properly antecedent governed.

Therefore, we can (and perhaps must) derive the landing site of a clitic from the ECP. Recall the problem for NP-trace posed in Barriers: since the NP lands outside the VP, and the VP is a barrier, there is a failure of government between the NP and its trace. This problem is circumvented by supposing that agreement between the subject and INFL ultimately places the subject's index on the verb, internal to the VP; hence the trace is properly governed, in this extended way.

Clitics do not agree with their verbs; hence they themselves must land in a position from which they can govern their trace. This means that clitics cannot move entirely outside the VP, and must adjoin to the verb, or to the VP.

A particularly nice consequence of this treatment, which abandons the idea that clitic traces are anaphors, is that it avoids a problem inherent in the previous story: that anaphors must be A bound, thus the antecedent (the clitic) must be in an

A position. A positions are defined as those positions to which a theta-role can be assigned; it seems then rather implausible to suppose that the landing site for clitics is an A position. The ECP account of the distribution of clitic trace allows it to be the case that clitics are not in A positions.

4.11 Reconstruction and parasitic gaps

In this section, I will examine two properties of parasitic gaps, in light of the discussion and proposals of this chapter and the previous ones. The first issue is the apparent opacity of parasitic gap constructions with respect to anaphor connectivity, and the fact that the type of obviation or disjoint reference effect seen in (106) and (107) persists in parasitic constructions. The implication of these and related facts for the conceptual organization of the binding theory will be discussed.

The second issue is the fact that parasitic gaps are apparently subject to an anti-c-command condition, in that the A-position WH-trace which they are associated with cannot c-command them. I will derive this fact from the Chain Obviation Condition on chains proposed above.

4.12 Opacity in Parasitic Constructions

We have seen in earlier examples that there is a multiple binding domain effect with successive cyclic movement of a WH-phrase containing a shallowly embedded anaphor. The anaphor may be grammatically referentially dependent upon any NP which c-commands one of the traces of the chain within the GC for that trace.

212) which pictures of himself did John say t that Bob feels^t that Dan liked t?

In our terms, there are (at least) three chain accessibility sequences for the anaphor, and the three antecedents are accessible to the anaphor through them.

Oddly, the same does not hold true with parasitic gaps. Chomsky (1986) proposes that parasitic gaps are related to the head of the licensing gap's chain through another type of chain, what he terms a 'composed chain'. The intent of this section is to examine the ways in which such composed chains participate in connectivity phenomena.

Consider example (213). This type of contrast was first noted, I believe, by Kearney (1982); see also Barss (1985), Haik (1985), Chomsky (1986) for discussion.

(213)

- a). which pictures of himself did John [t' [paint t] without Mary [e' [liking e]]]?
b). *which pictures of himself did Mary [t' [paint t] without[OPC John [e' [liking e]]]]?

There is a chain (which pictures of himself, t' , t) formed through movement of the overt WH phrase. Following Chomsky (1986), there is a second chain formed in the DS-SS mapping by movement of the parasitic operator internal to the adjunct phrase. This chain is (OP, e' , e). As we saw earlier, any of the three positions internal to the first chain may be a site from which an antecedent for the anaphor is calculated.

However, the sharp contrast between (213a) and (b) indicates that neither trace of the parasitic chain may function in this way; it is as if the parasitic gap is invisible with respect to anaphor connectivity² There are a variety of ways in which this odd asymmetry might be dealt with. If we adopt the two-operator

1. That is, either of the traces in the first chain -- the one formed through movement of which pictures of himself -- is capable of extending the chain accessibility sequence for the anaphor.

2. The parasitic gap can of course itself bind an anaphor, as in (i):

- (i) which women did you talk to without introducing t to each other's husbands?}

The point made in the text is simply that the chain accessibility sequences for the anaphor in cases like (213b) cannot be extended through the parasitic gap.

analysis proposed by Chomsky, we might suppose that there is a restriction in the reformulated Condition A, such that only what we might call the anaphor's core chain - formed through movement of the phrase containing the anaphor -- may be accessed by the algorithm which defines chain accessibility sequences. In Chomsky's (1986) treatment of parasitic gaps, there are two core chains formed in the derivation of an S-Structure like (213b): (which pictures of himself, t', t) and (OP, e', e). These two chains are combined together, in a composed chain, at S-Structure, with the constraint that OP must be subjacent to the licensing gap ³ t .

Chomsky (1986) defines composed chains as follows ⁴ (in a later section, I will modify this definition):

(his (130))

if $C = (a_1, \dots, a_n)$ is the chain of

3. I follow conventional terminology in referring to the WH-trace whose presence licenses the presence of the parasitic gap as the real or licensing gap, and in some examples I follow the practice of noting the licensing gap as t and the parasitic one as e.

4. The composed chain is subject to a well-formedness condition: the real gap cannot c-command the parasitic gap. Chomsky suggests that this might derive from Condition C. In the system of binding developed in this thesis, there is no Condition C; facts analysed as Condition C effects in Chomsky's theory (as well as others which receive no treatment in Chomsky's theory) are treated as arising from the Chain Obviation Condition here. In a later section, I will demonstrate that the anti-c-command requirement on parasitic gaps derives as well from the COC.

the real gap, and $C = (b_1, \dots, b_m)$ is the chain of the parasitic gap, then the 'composed chain' $(C, C') = (a_1, \dots, a_n, (b_1, \dots, b_m))$ is the chain associated with the parasitic gap construction and yields its interpretation.

The parasitic chain, on this view, will be associated with the anaphor's core chain only through chain composition, and so is (by my hypothesis) unavailable for the purposes of extending CASS for the anaphor contained inside the head of the first core chain. Thus, in the definition of chain accessibility sequence given as (4) in Chapter 3, only core chains are relevant.

If we were instead to adopt the proposals of Cinque (1984,85), who claims essentially that a parasitic gap is actually a resumptive pro (non-overt pronominal), we might suppose that the anaphor will be blocked from reconstructing into the parasitic gap, because there is no well-formed chain associating the two. This will thus be similar to the impossibility of reconstructing into a PRO, as in (214):

214) *which of each other's parents promised the men [PRO to leave]?

The contrast in (213a,b) is consistent with any approach to parasitic gaps which asserts a fundamental asymmetry between the two gaps at the level of representation at which Condition A applies.

These facts are not consistent with any approach to parasitic

gaps which asserts that the 'licensing gap' and the parasitic gap are identical at the level of representation at which Condition A applies. Such analyses include Chomsky's (1982) theory, in which the parasitic gap and the licensing gap are distinct at DS but identically variables at SS, each locally A' bound by the WH phrase; Haik's (1985) analysis, in which the two are formed through an (extended) process of across-the-board extraction; and Kayne's (1982) treatment, in which both gaps are viewed as variables, locally bound to the same overt operator, as indicated below:

(214)

local A' binding

[which picture of himself] did John see t without Mary liking e

local A' binding

local A' binding

*[which picture of himself] did Mary see t without John liking e

local A' binding

Therefore, any theory of anaphor connectivity apparently will have to involve an analysis of parasitic gaps which does impart this type of fundamental asymmetry to the two gaps at the level of representation at which Condition A applies. For this discussion I shall assume without modification Chomsky's (1986)

2-operator account, and suppose that the licensing chain and the parasitic chain join together into an extended or composed chain.

We note that the asymmetry evinced in (213) will now provide a test for the 'parasitic' nature of any given gap. It is not always the case that a parasitic gap occurs in a position from which overt movement is prohibited, as (Engdahl (1983)) observes. In the example below (215), both gaps appear in positions from which extraction to the matrix SPEC of CP position can occur, as (216) and (217) indicate.

215) which man did you warn e that the police were about to arrest e'?

216) which man did you warn e that the police were about to arrest Mary?

217) which man did you warn Mary that the police were about to arrest e?

Thus it is of some interest to ascertain which gap in (215) is the parasitic gap. By the 'Barriers' theory, the parasitic gap must be the one embedded inside the CP complement. This is so since if the matrix direct object e were the parasitic gap, then (i) there would be no position for the non-overt operator to
5
reside, and (ii) the structure would violate Condition C of the

5. Chomsky's analysis requires that for all types of parasitic gap constructions, there is an S-Structure landing site for the non-overt operator. Such examples as (i) are quite problematic on this view:

binding theory, since the trace e' would be A-bound (by e) within the c-command domain of its operator (which man).

This receives strong supportive evidence from facts concerning the behavior of e and e' with respect to anaphor connectivity. We expect that the real gap, but not the parasitic one, will provide a site for reconstruction. This is exactly the case, as we may observe in the following examples.

- (218) a. which of each other 's friends did John warn Mary that
the men were about to kill t ?
b.* which of each other 's friends did John warn t that
the men were about to kill e ?

In (218a), the trace in lowest object position (inside the CP complement) is a real gap, that is, one associated through movement with the WH phrase containing the anaphor. As we expect, this exhibits typical connectivity.

The situation is different in (218b). A trace occurring in the same position as the one in (218a) cannot provide a reconstruction site. This is apparently because, following the "Barriers"-theory designation of this gap in (218b) as a

- (i) [who did [[friends of e] [talk about t]]]
 CP IP NP I' VP

The gap inside the subject is the parasitic gap. The problem is that, on standard assumptions about the structure of NPs, it is not clear that there is a position in which UP could reside.

parasitic gap, this EC is only associated with the overt WH phrase containing the anaphor through 'chain composition', not through movement. Anaphor connectivity actually provides a relatively theory-neutral test for the parasitic status of an EC.

We still must construct an explanation for why this occurs, that is, why a trace bound at S-Structure by a WH-phrase containing an anaphor can only serve as a site for connectivity if it is related to the WH phrase through movement, i.e. in a core chain. (220) is at least descriptively correct .

6. Adriana Belletti has observed that extraction of an NP from an adjunct clause is more acceptable than extraction of a PP.

- (i) ? which man did you go to New York without speaking to ?
- (ii) *To which man did you go to New York without speaking?

Browning (in progress) observes that (iii) is as completely unacceptable as (iv).

- (iii) * which pictures of herself did John go to New York before Mary saw e?
- (iv) * which pictures of herself did John paint t after Mary commissioned e?

Browning assumes that in (iii) which pictures of herself has been directly extracted from the position marked by e, and argues that the ungrammaticality of this example shows that the ungrammaticality of (iv) has nothing to do with the presence of the parasitic operator. If this line of argument is correct, we will have to abandon the approach reflected in (220).

However, there is another accounting of (iii) available, which allows us to maintain (220) and collapse the treatment of (iii) and (iv) together. Luigi Rizzi (p.c.) has suggested that the contrast in (i) and (ii) is related to the fact that parasitic gaps are always NPs, never PPs. He suggests that there is no direct extraction from within the adjunct in (i); rather, which

obligatory, then the licensing gap t_j in (223) must, in the derivation of LF, be replaced by $[e_{NP\ i} se[mother]_j]$, as in (222). If the parasitic gap remains unaffected, then the LF representation (224) is derived¹⁰.

223) [who $se[mother]_j$ did you meet t_k [without
 $PRO_{i,k}$ liking e_j]]

224) who [did you meet [$e_{NP\ i,k}$ $se[mother]_j$]
 [without [$PRO_{i,k}$ liking t_j]]

This representation is ungrammatical, since t_j is A' free.

Since t_i is a variable, the structure violates the prohibition against free variables (see Chapter 5 of this thesis for discussion). Apparently, 'double lowering' is involved: both the real gap and the parasitic gap must be eliminated through lowering.

However, we observe that the type of strong crossover effect

 lowering, if at all.

10. Again, we observe that the derived representation (224) would appear to violate the ECP, in that e_i is separated from its closest (and only) antecedent who_i by two barriers.

which occurs in (221) also occurs internal to the parasitic gap construction.

225) * [who_i se[mother]]_j did you meet t_k [before
[he_i described e_j]]

If the parasitic gap remains unaffected by the lowering process hypothetically involved in ruling (222) out as a Condition C violation, then (225) should be grammatical, since the LF representation so derived will be (226):

226) who_i did you meet [e_k se[mother]]_i [before
[he_i described e_j]]

Thus if Lowering is responsible for the ungrammaticality of (222), (and so by extension, it ought to be responsible for the ungrammaticality of (225) as well) it must lower the remnant of [who_i se[mother]]_j left after raising of who_i, that is [e_i se[mother]]_j, into the parasitic gap as well. This means, in essence, that the lowering is not simple substitution-- it is impossible to substitute one phrase for two traces-- but must involve morphological copying in the derivation of LF. The correct representation of (225) at LF will be:

227) who_i did you meet [e_k se[mother]]_i [before
[he_i described [e_i se[mother]]_j]]

This representation now violates Condition C, as desired.

Now consider the implication of this conclusion for parasitic gap sentences with anaphor connectivity.

If LF Lowering obligatorily copies the remnant of the WH phrase into the parasitic gap, then the derived LF representation for (228) will be something like (229)¹¹.

228) [which picture of [himself]] did John see
 NP i j
 t without Mary liking e
 i k i

229) [which x] did John see [x picture of [himself]]
 NP j i
 without Mary liking [x picture of [himself]]
 k i

This then creates two copies of the anaphor himself at LF. Recall that this lowering treatment is part of an analysis which also includes the proposal that Condition A holds of LF, and not of S-Structure. Thus, the binding condition evaluates the LF (229) for well-formedness. The first himself is bound in its governing category, by John.

However, the second anaphor is not bound in its GC by an

 11. We remind the reader that the conclusion of chapter 2 was that no such extraction of which is indeed possible, for independent reasons, and thus that no such lowering is feasible. Here I ignore this problem, showing that even were this type of lowering to be possible, undesirable results arise. It is not at all clear that such an LF representation as (229) is well-formed, in that its constituent structure does not match up with the S-Structure constituency. cf. Chapter 2 for discussion.

acceptable antecedent, since the only c-commanding NP within the GC for the anaphor (taking [Mary liking pictures of himself] to be the GC) is Mary. Thus the lowering treatment fails to account for the grammaticality of (228).

Suppose Condition A were modified, so that two anaphors which are copies of a single anaphor at S-Structure -- as the two anaphors in (16) are -- need not both be bound locally; rather, Condition A would be loosened so that even if only one of the two is locally bound the representation is ruled grammatical. This is the approximate form of the 'layered trace' treatment. We have earlier, in chapter 3, presented justification for rejecting the layered trace view.

Further, this type of modification will fail to distinguish (228) from the ungrammatical (213), represented below.

213b.* which picture of himself did Mary see t without John
liking e

This sentence will have, on the lowering view, the LF representation (230).

230) [which x] did Mary see [x picture of [himself]]
NP j i
without John liking [x picture of [himself]]
k i

Here, exactly as in (229), there are two overt anaphors, each a copy of the same S-Structure anaphor, where one is locally bound in accordance with Principle A, and the other is not. The

proposed modification permits (228-9) to be grammatical, yet improperly permits ((213b), (230)) to be grammatical as well.

We therefore conclude that the asymmetry between 'real' gaps and parasitic gaps with respect to anaphor connectivity provides another argument against treating connectivity with an analysis which involves LF lowering of the anaphor to the position of the variable. Apparently, something like (220) is to be recognized as a formal constraint on the operation of the binding theory.

4.12.2 The lack of Sloppy Identity

We now note a final opacity-type effect seen in connectivity sentences with parasitic gaps. (See Haik 1985 for similar remarks). As is well known, some types of anaphora, in particular VP-anaphora, permit a sloppy identity reading with anaphors, as in the example below.

(231) John saw himself before Bill did

meaning: (John saw John before Bill saw Bill)

Suppose this is due to the replacement of [did] by the preceding VP (as suggested by Williams 197), Haik (1985), among others), so that the LF representation of (231) will be (232).

232) John saw himself before Bill saw himself

The sloppy reading is expected now, since the two anaphors will have to each be locally bound, satisfied only when John and Bill each bind an anaphor. It seems a reasonable hypothesis, then,

that if any operation in the S-Structure - LF mapping actually copies an anaphor, then the copy must be independently bound.

We note that there is no such sloppy reading available with parasitic gaps, as below.

233) which pictures of himself did Bill look at t while
Sam paid for e?

This does not have the sloppy interpretation which might be answered by something like (234).

234) His wedding pictures. (meaning that Bill looked at Bill's wedding pictures while Sam paid for Sam's wedding pictures.)

Rather, (233) is simply a question about pictures of Bill. Therefore, given the hypothesis just sketched about copied anaphors, it would seem that no such copying is (or can be) involved in the derivation of the LF representation of (233). But this is precisely what is asserted to occur in the reconstruction treatment of anaphor binding, as extended to constructions like (228)-(229). If this line of argument can be maintained, it demonstrates that not only is such lowering as occurs in the derivation of the LF (229) from the S-Structure

(228) not obligatory, it is not even possible¹².

12. Of course, it is possible that some other treatment of sloppy identity in VP-anaphora constructions can be formulated which would be restricted in its application to VP-anaphora. If this were the case, then my argument concerning the implications of the lack of sloppy identity in parasitic gap constructions would not go through.

4.12.3 Obviation and Parasitic Gaps

We shall see below that although (220) is descriptively correct for anaphor binding, it is not the case for the obviation conditions governing possibilities of coreference involving pronominals and R-expressions (e.g. the licensing condition for pronouns (7'), and the Chain Obviation Condition (179), which I formulated to rule out a variety of ill-formed structures, including Condition C effects involving names). The rest of this section is devoted to an examination of this fact.

Consider (235).

(235) it was him that Mary recognized t because he had described e
so well

The fact about this example is that the pronoun he in the adjunct must be disjoint in reference from the clefted pronoun him. Since the parasitic gap is coreferential with ¹³ the clefted pronoun, the parasitic gap is disjoint in reference from the subject pronoun he.

This contrasts with the interpretation of the next example, in which the empty category may be coreferential with the pronominal (although it need not be).

13. More precisely, the parasitic gap has its value determined by the clefted pronoun.

(236a)

it was himself that John nominated before he voted for e
PRO voting for e

It appears, then, that although the parasitic gap may not be the reconstruction site for anaphor binding, it nonetheless participates in the suppression of condition C effects in such examples. To clarify, we have seen that there is a lack of condition C effects in examples like (236b), which was accounted for in an earlier section. Basically, the binding type of the left-peripheral phrase is what determines the presence or absence of necessary disjointness of referential value between he and t in such cases; when a name is substituted for the anaphor, as in (236c), the disjointness is forced.

(236) b. Himself, he likes t

c. John, he likes t (out on the coreferential reading)

Similarly, it is the fact that the ultimate antecedent of the parasitic gap in (236b) is an anaphor which allows the parasitic gap to be coreferential with the pronominal subject of the adjunct clause. Therefore, although the parasitic gap cannot participate in the calculation of an antecedent for the anaphor, nonetheless the relation between the parasitic gap and the anaphor is what permits coreference between the parasitic gap and the c-commanding pronominal.

Basically, the head of the licensing chain in (235) and (236a) is a pronoun or an anaphor; this fact determined whether the parasitic gap may corefer with the subject of the adjunct. Thus

it appears that the parasitic gap is subject to the obviation condition governing pronouns if and only if the head of its licensing chain is a pronominal.

Further evidence for this is seen in the next example:

(237)

- a. Him, John claimed Mary liked even though he criticized e
- b. it was him that John claimed Mary liked even though he knew she hated
- c. it was him that John claimed Mary liked even after PRO hearing everyone laugh at e

In the first case, but not the second or the third, the parasitic gap must be disjoint in reference from the pronominal

14

subject of the adjunct clause .

14. In an example like (i), the adjunct may be interpreted as modifying the matrix clause, or the embedded clause.

- (i) John said that Fred had left [before PRO eating breakfast]

We presume that the adjunct is structurally attached either as a constituent of the matrix clause or of the embedded clause. In the previous case, PRO will be controlled by John, in the latter case by Fred (in each case, by the minimally c-commanding NP).

We note that the point of attachment of an adjunct clause can be disambiguated when the structure is approximately of the form in (237). Consider the following examples.

- ii) It was him that John said Mary should vote for t
[after PRO nominating e]
(PRO cannot be John)
- iii) It was him that John said Mary liked without [PRO saying that she loved e] (ambiguous attachment)
- iv) It's him who John said Mary had insulted [while PRO describing e]
(with coreference between him and John, the PRO must be

This indicates, then, that the extended or composed chain formed through parasitic gap licensing is relevant to chain accessibility for pronouns; a pronoun must be obviate from every NP which is minimally chain accessible to the pronoun, even through a chain accessibility sequence defined through the parasitic gap. This is then a fundamental asymmetry between anaphor binding and pronominal obviation. In light of the revised licensing conditions for pronouns given as (7') above, an accessibility sequence for a pronoun can have a link (X, Y) where X and Y are two members of a composed chain, even though

 controlled by Mary.)

v) It's him who John said Mary had insulted t [while PRO describing Jimmy] (PRO can be controlled by either John or Mary).

In all cases take him to be coreferential with John. In all cases, the parasitic gap has the left-peripheral pronoun as its antecedent, and so the pronoun must be obviative with respect to any NP minimally chain accessible to it through a chain accessibility sequence constructed through the parasitic gap. Therefore, in (ii), John cannot control PRO, since PRO is minimally chain accessible to him. When the parasitic gap is more deeply embedded, as in (iii), PRO is no longer minimally accessible to the pronoun, hence the sentence is ambiguous; PRO can be controlled by John or Mary. (iv) like (ii), has only the interpretation where Mary controls PRO. When the parasitic gap is eliminated in (v), the sentence once again becomes ambiguous.

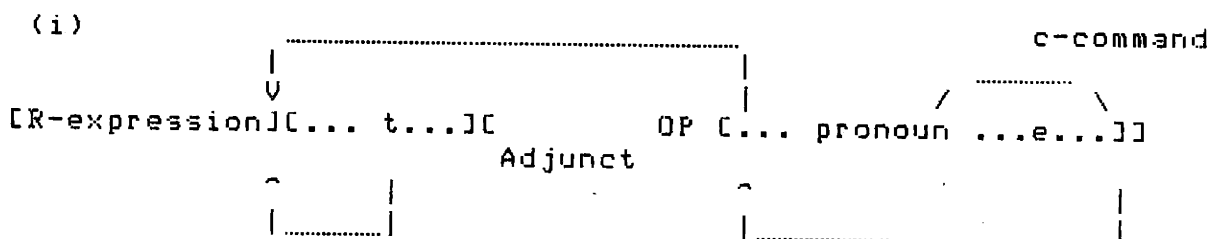
Such facts raise interesting questions for the theory of natural language parsing. Consider the contrast between (iv) and (v). From the point of view of parsing, the adjunct can be parsed as a constituent of the main clause only if it does not contain a parasitic gap, or if the topicalized pronoun and John are interpreted as disjoint in reference. It would appear then that the parser must hold both possible attachments (matrix or embedded clause, with matrix subject or embedded subject control of PRO) in memory, rejecting one when the parasitic gap is reached.

- 239) a. It's John who Mary voted for t after he nominated Sue
 b. It's John who Mary voted for t after he asked someone to
 nominate Sue

We explained the lack of such coreference in cases like (240) by postulating the Chain Obviation Condition.

(179) The Chain Obviation Condition (COC)

For a chain $C = (a_1, \dots, a_n)$, and for an expression Z where Z is not a member of C and Z is in an A position, if Z is dependent upon a_i , then Z does not c-command any member a_j of C.



- (ii) It's John who Mary kissed t after he said that Julie thinks that Bob said that she should marry e
 (iii) It's John who Mary kissed t after he said that Julie thinks that Bob said that she should marry Peter

Coreference is possible in (iii). Again this is quite interesting from the point of view of the theory of parsing. In contrast to the cases with left-peripheral pronouns, sentences of the form (i) will be ill-formed with coreference no matter how deeply embedded the parasitic gap is. Coreference between John and he is permitted in (iii) because the adjunct contains no parasitic gap. Thus coreference can be assigned to the two NPs only after the parser has determined whether the overall structure has a PG or not, or holds both (indefinitely long) parses in memory until the end of the sentence is reached.

(240) * John, he likes t

(241) a. *Gianni si e stato affidato t t' (Rizzi (1982))

b. * He likes John

c. * He thinks John, Mary likes t

d. John promised himself [PRO to get a sandwich]

In (240), there is an A' chain (John, t). Since, on the coreferential reading, he is linked to John, he cannot c-command any member of the chain; but since in (240) he does c-command the trace, the CUC is violated.

Apparently, this also applies to composed chains, as we see in considering the impossibility of coreference between John and he in (238). In this structure, he is dependent on John, and thus on the extended chain containing John and the parasitic gap.

Since in (238) he c-commands the parasitic gap, the chain obviation condition is violated.

The COC is a quite general principle, also ruling out the illicit Romance structures in which a clitic is bound by a derived subject (241a), and cases like (241b) which were analysed by Chomsky (1981, 1985) as Condition C violations, and by Higginbotham (1985) as a violation of Condition C*. As we argued earlier, the COC is an appealing constraint, since it allows us to account for the grammaticality of (236b) and the like (improperly ruled out by Condition C), the ungrammaticality of (241b), and the ungrammaticality of (241c), which as argued above causes problems for the formulation of Condition C.

The COC is quite deliberately formulated in terms of chains. In (241d), since no chain associates John and PRO, the fact that himself is linked to John and (presumably) c-commands PRO will not violate the COC.

Thus it would seem that the obviation facts of (238) and (237a) indicates that the left-peripheral phrase and the parasitic gap are related through a chain. This gives support to Chomsky's proposals that a parasitic gap's core chain and the licensing chain are united in a composed chain.

The picture that emerges is that composed chains are not relevant to the calculation of antecedence possibilities for anaphors, but they are relevant to the determination of

However, we must question the truth of the assertion expressed in (220). Anaphor connectivity occurs in other constructions in English involving empty operators or composed chains, specifically Tough-constructions (242) and too-adjectival constructions (243). The following are grammatical.

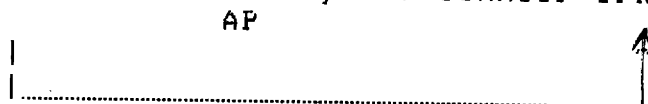
- Presumably, since the NPs John and John and Mary are contained inside the AP in these examples, they do not c-command the anaphors inside the subjects. Thus, we assume that the anaphor does not link directly to the prepositional object.

.....

- 485 -

composed chains do participate in the construction of chain accessibility sequences. If this is the case, then PRO would be accessible, in (242), to the anaphor through the CAS {himself, PP, N', NP, t, VP, I', IP}. PRO is accessible to the anaphor through this sequence, and thus the linking in (245) satisfies the binding theory.

245) Pictures of himself are [easy for John[OP LPRO to like t]]



Crucially, the link (NP, t) of the CAS is permitted because the two expressions are co-members of the composed chain.

The grammaticality of these examples indicates that composed chains of the form found in Tough-constructions are relevant to anaphor connectivity. The asymmetry between (213a) and (213b) indicates that composed chains of the type found in parasitic constructions do not participate in the formation of CASs for anaphors. This is a most striking fact, one which I have no explanation for. It is hoped that this peculiar fact can be derived from other properties of empty operator constructions, but at the present time I do not have an adequate account.

Nonetheless, the facts in (213b) and ((235), (237a)) lead to a very interesting conclusion: one type of composed chain fails to participate in the calculation of CASs for anaphors, yet does participate in the calculation of CASs for pronouns. Further, parasitic gaps are quite active in the operation of the COC, as

we saw in (238).

We speculate that this difference gives additional evidence that the two relations of anaphoric dependence and obviation must be formally recognized by the theory of binding. It is the nature of dependence that an element X (like an anaphor) which is required by its intrinsic nature to have an antecedent fulfills the binding theory if it acquires one antecedent. We may conclude on the basis of the evidence presented in this chapter, as well as the discussions of (Lasnik (1976), (1981), and Higginbotham (1985)), that intrinsically obviative elements, like pronouns or (as I have argued) chains, are forced to be obviate from any expression accessible to it. It is therefore not surprising that there are domains in which obviativity is necessary but dependence is blocked.

We will now discuss further supportive evidence for the notion that we should understand the facts presented above as indicating a real distinction between obviation and dependence. The relevant structures are parasitic gap sentences with operator-dependent pronouns.

We observed above that in examples like (225) (repeated here) the Strong Crossover effect occurs; in this example, the pronoun he cannot be understood as a semantic variable bound to the operator who.

225) whose mother did you meet t before he described e ?

* [who selmother]] did you meet t [before
 i j k j
 [he described e]]
 i j

In the indexed structure given, there is no Condition C violation, thus this is a type of SCO effect which could only reduce to Condition C as formulated in Chomsky ((1981), (1982), (1985)) if the operator who were moved out of the constituent [who selmother]] , and then [e selmother]] were lowered to the position of t in the derivation of LF.

We presented arguments above that such a derivation is impossible. Further, we note that the same SCO effect is seen in examples like (246a), in which such reconstructive lowering is clearly impossible, since the entire overtly moved phrase is an operator which must remain in scope position (see Higginbotham (1980) for details).

(246)

- a. which pictures of which man did you buy after he painted ?
- b. for which x, x a man, for which y, y pictures of x, did you buy y after x painted y?

(246a) cannot have interpretation informally given as (246b). Now consider the relevant portion of Higginbotham's Accessibility Condition (see (148)).

- (247) A pronoun P is not accessible to the formal variable v bound by an operator O if P c-commands v of the v-chain headed by v.

(248)

- (145) A formal variable is an empty category occurring in an A-position, and linked to a non-argument (phrase in a A' position); the operator to which the formal variable is linked is its binder. A sequence (v_1, \dots, v_n) of formal variables is a V-chain if each variable v_i is contained in the binder of v_{i+1} .

(246) contrasts in grammaticality with the minimally different (249) (the two are identical except for the object NP in the adjunct clause). (249) is accepted by most speakers as perfect (although some speakers find it to have the status of a weak crossover violation).

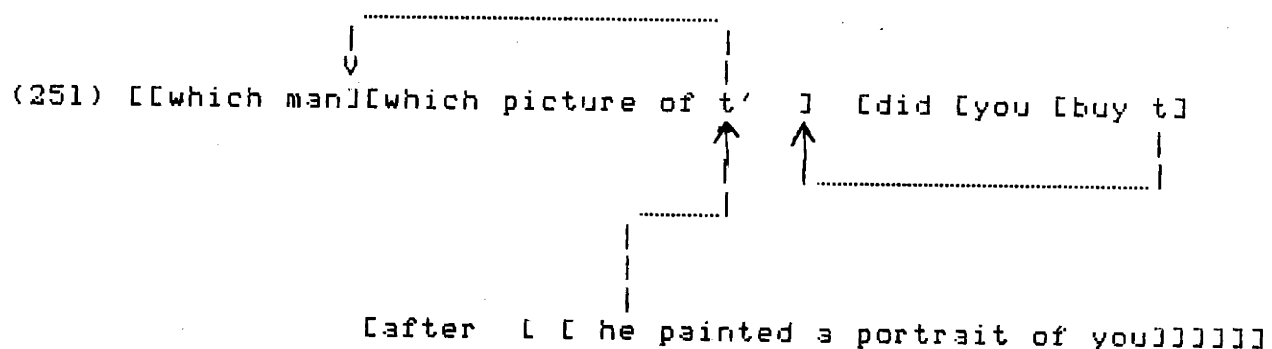
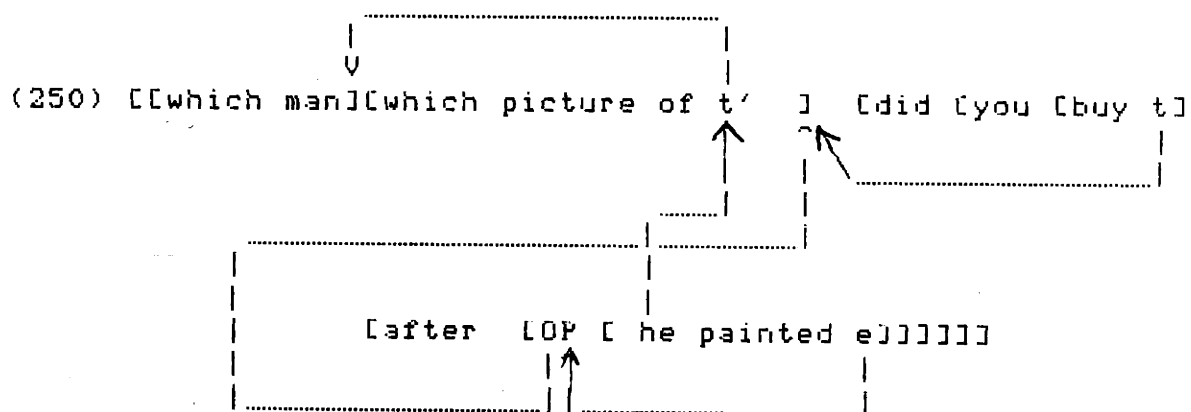
(249) a. which pictures of which man did you buy
after he painted a portrait of you?

b. for which x, x a man, for which pictures y of x, did
you buy y after x painted a portrait of you?

Therefore, it would appear that the strong ungrammaticality of (246) (on the reading where the pronoun is understood as a bound variable) is due to the fact that the pronoun c-commands the parasitic gap. If the SCO effect is to be accounted for generally by the Accessibility Condition, then the parasitic gap must count as v_n of the variable chain associated with the formal variable v bound by which man.

Let us now consider how this works. In (250) is the LF representation of (246) which would give rise to the illicit interpretation. I assume that parasitic operators are present

at LF (but see Browning (in progress) for an interesting alternative), and I further assume that the parasitic operator is not linked to the licensing gap, but rather to one of the A' positions in the licensing chain. I defend this position in a later section of this chapter. (For reasons of simplicity, I give (250) without any intermediate traces, so OP is directly linked to which picture of which man). (251) is the LF representation for (249a) giving the bound variable interpretation (249b).



In (251), he is dependent upon the formal variable t', which is bound by the operator which man. The maximal v-chain headed by t' is (t', t). Since he does not c-command t, (247) is not

violated, and the pronoun is at least potentially accessible to t' . What we see in the parasitic gap construction (250) is that it is not sufficient that the pronoun does not c-command the trace t ; in order to be accessible to the formal variable t' , the pronoun must not c-command the parasitic gap e (which covaries with t). Thus it would appear that the parasitic gap is also the terminal element (v_n) of a v-chain headed by t .

This is reminiscent of the obviation effects discussed earlier. Just as a topicalized pronoun must be obviate from every NP accessible to it through any accessibility sequence, so we must take every v-chain in a structure into account when checking to see whether a pronoun can be bound to an operator as a variable. The second part of the Accessibility Condition (148) takes the form of an obviation condition, blocking certain interpretations under which two expressions share values when they are in a particular syntactic configuration. This is consistent with the observation made above: parasitic gaps are active in invoking obviation conditions, even though they are not active in the determination of antecedents for anaphors contained in the head of the licensing chain.

Higginbotham (1983) observes (drawing upon examples from Jacobsen (1979)) that an effect similar to what I observe above occurs with restrictive relatives. There is a crucial contrast between (252a) and (253a).

dependent upon the assignment of values to the NP [his wife], which is in turn dependent upon assignment of values to his (since here his is to be understood as a variable bound to the WH-phrase operator). Thus her is dependent (in the formal sense of (256), from Higginbotham (1983)) upon the WH-phrase's trace t', and so the Accessibility Condition is invoked. The exact same holds for the pronoun she in the second example.

(256) Definition of Dependence

X is dependent upon Y if:

- (i) Y is contained¹⁶ in an antecedent of X, or
- (ii) for some Z, X is dependent on Z, and Z is dependent on Y.

Y is an antecedent of X if X is linked to Y, or if X is linked to some W, and Y is an antecedent of W (see (129i) above). Thus by (256), an expression is dependent upon its antecedents, and upon an constituents of an antecedent.

By the definition of dependence, the pronouns her and she are dependent upon the formal variable t' in the two examples, and so for the bound variable interpretation to be available, the pronouns must be accessible to this formal variable. We now repeat the Accessibility Condition:

16. Not necessarily properly contained; thus if Y is an antecedent of X, then X is dependent upon Y itself, in addition to any constituents of Y.

(148) The Accessibility Condition

If a pronoun P is dependent on a variable v , then P must be accessible to v . Let C be the longest V -chain (v_1, \dots, v_n) such that v_1 [the head of the V -chain] is v and the binder of v_n does not contain P .

(I) P is accessible to v if v c -commands P ; and

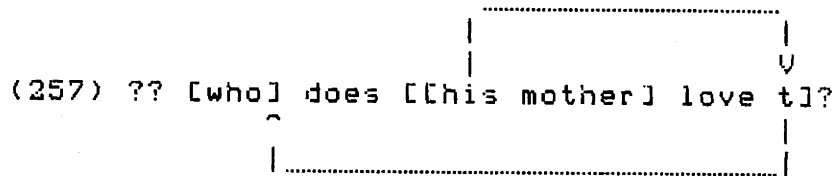
(II) P is not accessible to v if P c -commands v .

(145) A formal variable is an empty category occurring in an A -position, and linked to a non-argument (phrase in a A' position); the operator to which the formal variable is linked is its binder. A sequence (v_1, \dots, v_n) of formal variables is a V -chain if each variable v_i is contained in the binder of v_{i+1} .

As Higginbotham observes, the semantics of restrictive relatives allows us to consider each of the traces in (254) and (255) to be occurrences of the same semantic variable. For the cases in (254) and (255), then, we may take each trace to be v_n of a variable chain.

In (254), the AC is met. The pronoun her is accessible to the formal variable upon which it depends, since it is c -commanded by one of the traces (t) and c -commands neither. In (255), however, both clauses of the AC are violated. First, neither trace c -commands the pronoun, and second, the pronoun c -commands the trace t. Recall that in Higginbotham's framework a weak

crossover violation arises when clause I is violated, but clause II is satisfied, as in (257).



A Strong Crossover violation arises only when clause II is violated. Thus, Higginbotham concludes, the complete unavailability of the bound variable interpretation in (255) is evidence that clause II of the AC is violated.

The facts about Higginbotham's examples fit with our characterization of the behavior of parasitic gap constructions with respect to dependence and obviation conditions, and our suggestion that the anaphora conditions formally partition into these two quite different types of condition. The first clause (clause I) of the AC is a dependence condition, specifying under what circumstances an element can have its value fixed by another. Clause II is an obviation condition, specifying the circumstances under which one element (a pronoun) cannot have its value fixed by another (the formal variable). For multiple occurrences of the formal variable, it is sufficient that only one occurrence of the variable c-command the pronoun in order that clause I be satisfied. However, for clause II to be satisfied, it is not sufficient that the pronoun fails to c-command one of the formal variables; it must fail to c-command all of them.

Similarly, in our ungrammatical example (225), it is not sufficient that the pronoun is accessible to t in virtue of not c-commanding it. The pronoun c-commands the parasitic gap e, and this violates the AC.

We may replicate cases similar to (254) with parasitic gaps. The fact about (254) is that the pronoun can be understood as a variable, dependent upon the formal variable bound by the WH-phrase¹⁷. This variable has two occurrences, t and t'. The pronoun must be c-commanded by one of these, and must not c-command any of them.

(258) is a type of parasitic gap structure discussed in Barss (1984b) and Longobardi (1985), in which the parasitic gap is inside an adjunct occurring between the subject and VP of the clause.

(258) which paper did John even before receiving e
decide to review t?

If a pronoun is contained inside such an adjunct and there is no parasitic gap in the adjunct, we have a weak crossover

effect¹⁸, as in (259).

17. That is, the value of the pronoun, on each variable assignment to the sentence, is that of the NP his wife; this NP contains a pronoun interpreted as a variable bound by the WH-phrase. Thus the value of her changes with each variable assignment, and so functions as a variable.

18. That is, when we try to interpret the pronoun as a variable bound by the WH-phrase. (259) is of course fine when the

- (259) ?? which paper did John even before receiving a letter
from its author decide to review t ?

The addition of a parasitic gap in the adjunct improves the sentence, if the parasitic gap c-commands the pronoun:

- (260) a. which paper did John even before receiving e
from its author decide to review t ?
b. ?? which pictures of which man did Mary even
before telling his wife about say t were ugly?

This is expected, given the characterization of pronoun accessibility above. On the bound variable reading, the pronoun in (260a) is dependent upon the formal variable bound by which paper; following Higginbotham, we take both the real gap t and the parasitic gap e to be occurrences of the same formal variable. As we have seen, the Accessibility Condition is satisfied when only one of the occurrences c-commands the pronoun, as the parasitic gap does in (260a). When the pronoun is not c-commanded by the parasitic gap, as in (260b), the weak crossover effect remains. Similar contrasts may be found in the (a) and (b) examples in (261) and (262).

- (261) a. ??which pictures of which man did [Mary say[t were ugly]
[even before PRO showing a Rembrandt to him]]?
b. which pictures of which man did [Mary say[t were ugly]
[even before PRO showing e to him]]?

pronoun refers to something else. In all the examples that follow, I give the judgement for the example on the bound variable interpretation of the pronoun.

(262) a. ??which pictures of which man did Mary [even before
showing a Rembrandt to him] say t were ugly?

b. which pictures of which man did Mary [even before
showing PG to him] say t were ugly?

Further, given our characterization of Clause II as an
obviation condition, and our observation that parasitic gaps
participate in the imposition of obviation, we should expect a
Strong Crossover violation if the pronoun c-commands either the
parasitic gap or the real gap, but not necessarily both. This
is in fact what occurs.

(263) * which pictures of which man did Mary [even before
he saw e] say [t were ugly]

(263) is completely ungrammatical on the interpretation where he

19

is a variable bound by which man .

19. However, weak crossover is suppressed in parasitic
constructions of the form (2a) (I believe this fact was first
noted by Howard Lasnik).

(1) a. who did his father talk to t?

b. for [which x: person(x)] did [x's father] talk to x?

(2) a. who did you meet t after his father talked to e?

b. for [which x: person (x)] did you meet x after
[x's father] talked to x?

(2a) can have the interpretation (2b), but (1a) cannot
(grammatically) receive the interpretation (1b).

As H. Lasnik notes (1985 lecture, MIT), this is apparently a
counter-example to Koopman and Sportiche's (1982) proposal that

4.12.4 The Relation Between the AC and the COC

We have proposed a generalized obviation condition on chains, and argued that it is a general principle which rules out a variety of ungrammatical constructions, including those which are ruled out by Chomsky's Condition C and Higginbotham's Condition C*. We have thus argued for the elimination of

20

Condition C (or C*) from the grammar . In this subsection, we

the Weak Crossover effect is to be viewed as a violation of their Bijection Principle, which holds that an operator must locally bind one and only one variable. In (1), the Bijection Principle rules out (1b) as an interpretation, since who binds two variables. However, the BP also incorrectly rules out (2) on interpretation (2b), since here the operator locally binds three variables (the two gaps and the pronoun).

Since it is apparently a fact that parasitic gaps cannot be c-commanded by the licensing gap, the grammaticality of (2a) on the relevant reading raises a problem for the Accessibility Condition. The pronoun in (2a) is clearly not accessible to the PG, hence it must be accessible to the real gap. But, due to the independent anti-c-command requirement, t does not c-command e. Therefore, t does not c-command his, and so the pronoun is not accessible to it. We would then expect a weak crossover violation. I have no explanation for this fact. The same occurs in (3), where his can be bound as a variable to which man.

- (3) which pictures of which man did Mary destroy t even
before telling his wife about PG?

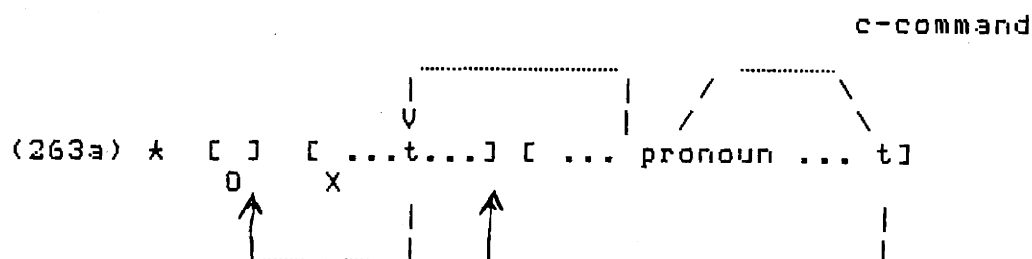
Barss, Johnson and Saito (in progress) examine this problem.

20. A potential problem for this move arises when we consider anaphoric epithets. As Lasnik (1976) shows, epithets like the bastard, the little idiot and the like share with names the fact that they cannot corefer with a c-commanding pronoun:

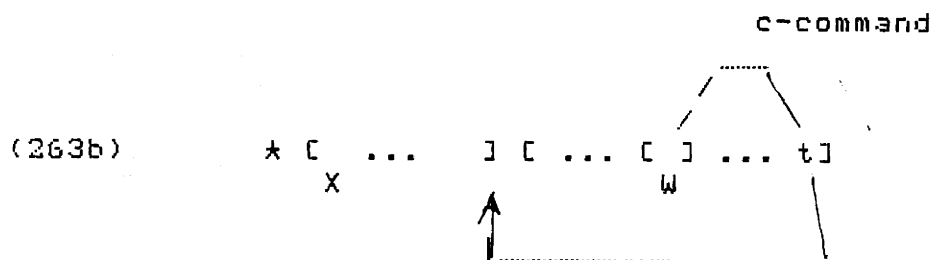
- (i) *He thinks the bastard is smart
 i i

be quite similar configurations. Clause II of the AC rules out

21
the schema (263a) .



Where 0 is an operator, t its formal variable, X a container of t, and t' the formal variable of X. The Chain Obviation Condition rules out structures like (263b).



The COC rules out a representation in which there is a chain with members Y and Z (not necessarily distinct; cf. discussion

21. We earlier argued for a strengthening of Higginbotham's condition, to require that if a pronoun is to be understood as a variable bound by an operator, it must be linked to the trace (formal variable) of the operator. This prohibits the pronoun from directly linking to the operator at LF. Higginbotham proposed that such direct linking was to be barred for another reason (and so would not be under the purview of the AC), namely that free linking (linking not produced under movement) was only between A-positions. We have suggested that this constraint should be given up, allowing free linking between all positions. Therefore, in the system as modified here, the schema given in (263a) is the only one which could give rise to the interpretation where the pronoun functions as a variable bound (in the semantic sense) by 0.

of (173)), and there is some expression R which is dependent upon Y and c-commands Z. Clause II of the AC rules out a representation in which there is a chain with link (X, Y), and there is a pronoun P which is dependent on X or a subconstituent of X and c-commands Y. Clearly, these are closely enough related that some attempt should be made to subsume Clause II of the AC under the more general well-formedness condition, repeated here.

(179) The Chain Obviation Condition (COC)

For a chain $C = (a_1, \dots, a_n)$, and for an expression Z where Z is not a member of C and Z is in an A position, if Z is dependent upon a_i , then Z does not c-command any member a_j of C.

Suppose we were to revise (179) as follows (added part underlined).

The Chain Obviation Condition (Revised)

For a chain $C = (a_1, \dots, a_n)$, and for an expression Z where Z is not a member of C and Z is in an A position, if Z is dependent upon a_i or upon a subconstituent of a_i , then Z does not c-command any member a_j of C.

If this reformulation were correct, then Clause II of the Accessibility Condition could be eliminated. Unfortunately, the revision cannot be maintained, as it would rule out perfectly

grammatical cases like (264a,b).

(264) a. [Pictures of John]₁, he likes t₁

b. [Pictures of John]₁ [[annoy t₁] him]₂

Consider (264a). The revised COC would rule it out, since there is a chain (pictures of John, t), he is dependent on John (a subconstituent of one member of the chain), and he c-commands t. A similar problem arises for (264b), assuming the Belletti and Rizzi analysis of psych-movement. Thus, it seems that we must keep the COC as formulated in (179), and retain the Accessibility Condition in full. However, we note that the two conditions are similar enough that it is desirable to seek a unification of them.

4.12.5 Crossover in too-Adjective Constructions

Chomsky (1985) discusses such examples as (265) and (266).

(265). John is [too stubborn [PRO to talk to Bill]]

(266). John is [too stubborn [PRO to talk to e]]

He notes that (265) has two quite distinct interpretations, paraphrased as (267).

- (267) a. John is too stubborn for John to talk to Bill
 b. John is too stubborn for anyone (to be able) to talk to Bill

These interpretations arise from the possibility of PRO being controlled by John (yielding (267a)) or being assigned an arbitrary interpretation (267b). Only the arbitrary interpretation is available for (266); this sentence cannot mean that John is too stubborn to talk to himself.

Assuming the presence of an empty operator in such constructions, Chomsky analyses the impossibility of controlled PRO in (266) as arising from Condition C. The indexed representation giving the unavailable interpretation is (268).

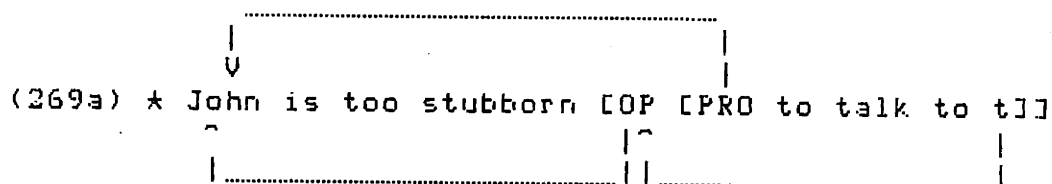
(268) John is [too stubborn] OP
 i AP CP i
 [PRO to talk to t]]
 IP i i

Here t is the R-expression bound by the operator OP; this trace is A-bound within the domain of the head of its chain, violating Condition C.

In this chapter, I have argued that we may eliminate Condition C, in favor of the Chain Obviation Condition. I have further argued, in this section, that the COC holds for composed chains (such as those which unite a parasitic chain and its licensing chain), not only core chains.

Let us suppose that chain composition is not restricted to parasitic gap constructions, but occurs as well in the type of

empty operator construction illustrated in (265-8). This seems a plausible assumption. We assume the following linking structure for (266), on the (unavailable) controlled PRO interpretation.



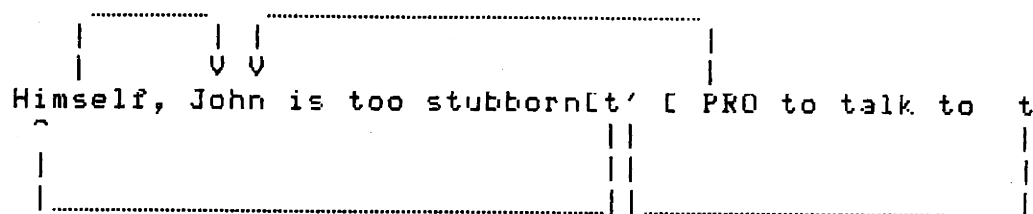
The link between OP and t results from movement of the operator; the link between OP and John indicates the fact that John is the antecedent of the operator. Finally, the link between PRO and John indicates the (illicit) control.

There is a core chain (OP, t) which exists as a result of the movement of OP internal to the CP. By hypothesis, there is a composed chain (John, OP, t), of which the CUC holds.

Thus in this structure PRO is dependent upon an expression (John), and PRO c-commands a member (t) of the composed chain of which John is a member. Thus, the Chain Obviation Condition is violated, and the ungrammaticality of (266) (on the controlled PRO reading) is accounted for by the principle which subsumes Condition C in our system.

We note that (269b) allows controlled PRO, as indicated by the linking assigned.

(269b)



This reading is expected in our system, although it the type of example incorrectly ruled out as a Condition C violation in Chomsky's theory, which treats t as an R-expression, and would be A-bound within the domain of the head of its chain. There is a core chain (himself, t' t) resulting from the movement of himself to left-peripheral position²². No A-position element in the structure is dependent upon any member of the chain (except for t, of course). Therefore, the COC is not violated. In

22. I follow Lasnik and Saito (in progress) in supposing that in cases of embedded Topicalization (i) the Topic phrase is adjoined to S (IP, in the Barriers theory):

(i) I think that Bob, Mary loves t

I further assume that such IP-adjunction is at least an option for matrix Topicalization, as in (269b), where himself will be adjoined to IP. The string might also be assigned a different structure, with himself in a base-generated Topic position, outside the domain of the matrix SPEC, which is filled by a non-overt operator.

(ii) [Himself[OP [John is [too stubborn
CP IP
[t'[PRO to talk to t]]]]]
CP IP

The sentence is, on this analysis, derived through movement of OP to SPEC position. See Chomsky (1977, 1986) for discussion. The point relevant to the text discussion of (269b) is that a direct movement analysis is available for (269b), creating the core chain (himself, t', t).

crucial contrast to the ungrammatical (269a), PRO is not dependent upon any member of the chain.

A similar treatment is available for the example in (269c) (noted by Jim Higginbotham):

(269c)

```

Fred has himself [OP [PRO to blame t]]
^  ^           |  ^   |  ^   |
|  |           |  |   |  |   |
|  |           |  |   |  |   |
|  |           |  |   |  |   |

```

The extended chain here is (himself, OP, t). PRO is coreferent with the trace, and c-commands it within the domain of the head of its chain, and so violates Chomsky's Condition

23

C .

4.13 Deriving The Anti-C-Command Condition on Parasitic Gaps

It is well known that parasitic gaps cannot be licensed by c-commanding subjects.

23. Indexing theory would assign to it the representation (i):

```

(i) Fred has himself [OP [PRO to blame t ]
      i             i   i   i           i

```

The trace is left by movement from an A-position to an A' position, and thus by the theory of EC feature determination assumed in Chomsky (1985) would be an R-expression. The EC is A-bound (by PRO) within the domain of the head of its chain, and so Condition C is violated.

(270) which papers did you file t [before [OP [you had read e]]]?

(271) * which papers [t fell off the table [before [OP [you had read e]]]]?

(272) which papers did John [before reading e] decide to tell his secretary [t were unavailable]?

The grammaticality of (272) (from Barss (1984b)) indicates that subject gaps can license parasitic gaps, as long as they do not c-command them (see Chomsky (1986) for discussion).

In the functional determination theory of empty categories developed in Chomsky (1982), this fact has an elegant explanation. In this approach, an EC is a variable iff its local binder is in an A' position; an EC locally A-bound by an NP with an independent theta-role is functionally determined to

be PRO. In (270) and (272)²⁴, the parasitic gap will be a variable, and the structure grammatical. In (271), the closest binder of the parasitic EC is t, which is in an A position and has a theta-role distinct from that of the EC inside the adjunct clause. Thus in (271) e will be determined to be PRO, and, since it is governed, will violate the PRO theorem. Hence the ungrammaticality of this type of example follows from the theory of ECs.

The theory of functional determination is abandoned in Chomsky (1986) (see also Brody (1984), Epstein (1984)), and so this

24. In the functional determination theory, there will be no non-overt operator in these examples.

explanation of the contrast in (270-272) becomes unavailable. The parasitic gap in all cases will be an R-expression (variable), whose closest binder is the empty operator. Chomsky (1986) suggests several possible ways to account for this fact. He posits (273) as a condition on chain composition:

(273) In a composed chain (C, C'), the head of C' must be 0-subjacent to the final member of C.

(X is n-subjacent to Y iff there are n or fewer barriers for X which do not contain Y).

In the theory of barriers, VP is a barrier, hence in (271) there is a barrier (VP) separating the real gap from the parasitic operator, while in (270) the licensing gap is inside this barrier. In (272) as well, the licensing gap is inside the matrix VP, hence 0-subjacency may hold (see Chomsky (1986) and Barss (1984b) for details).

Chomsky is forced to abandon this way of deriving the anti-c-command condition. The Barriers theory proposes in English verbs raise to INFL (as an instance of head-to-head movement). This has an important effect on the barrierhood status of VP, namely voiding it. Such movement endows INFL with

lexical content²⁵; since INFL also (within the Barriers theory) theta-marks the VP, after verb raising VP is L-marked by INFL,

25. Recall that XP is not a barrier if it is L-marked. It is L-marked if it is theta-marked by a lexical head, which is defined as a head with the features +/-N, +/-V. INFL theta-marks VP, and once the verb has moved to INFL, INFL has lexical features.

and so is no longer a barrier. Thus the verb in its moved position can properly govern its trace, and the VP is no longer

26 a barrier. Thus in (271), after V-raising, the subject trace is the same number of barriers away from the parasitic operator as the object trace in (270) is, and so the 0-subjacency requirement (273) is insufficient to explain the contrast.

27 Chomsky then considers (274) as another means of deriving the anti-c-command condition.

(274) In any chain $C = (a_1, \dots, a_n)$ [including core chains, composed chains, and subsequences of composed chains], such that for every link (a_i, a_{i+1}) a_i c-commands a_{i+1} , C has exactly one Casemarked position and exactly one thetmarked

26. See Chomsky (1986) for details. Consider a raising sentence. In this theory, the raised verb will be coindexed with the derived subject through agreement; the verb will be coindexed with its trace, and so by transitivity of indexing the trace of the verb will be coindexed with the NP-trace. Thus both the verb's trace and the derived subject will be capable of properly governing the NP-trace.

(i) [John [[seems [t [t to be nice]]].
i I' INFL i VP i IP i

Sam Epstein observes that within this theory Subjacency must be taken as a condition on derived representations, rather than on movement. The reason for this is that it is only after verb raising that the VP ceases to be a barrier. If subjacency is a condition on movement, then the verb raising itself will violate Subjacency.

27. A modification of his (147); modifications reflect Chomsky's text comments.

position.

Consider (271). Here, every link of the composed chain (which papers, t, OP, e) is such that the first member c-commands the second, hence the composed chain is subject to the condition (274). The chain has two case-marked and two theta-marked positions (the two traces), hence (274) is violated.

In (270), the composed chain (which papers, t, OP e) has a link (t, OP) such that the first member does not c-command the second. Thus the composed chain is not subject to (274). Rather, only the two core chains are, and they each separately conform to the requirements; each has only one case-marked or theta-marked position.

I shall now show that Chomsky's proposal does not clearly extend to all cases. (274) is intended as an extension of his condition (147) on A chains, quoted as (275) (see also Chomsky (1985) for discussion).

(275) A maximal A chain (a_1, \dots, a_n) has exactly one Case-marked position (namely a_1) and exactly one theta-marked position (namely a_n).
(Chomsky (1986; ex. (147), p. 63)

First, consider the extension of the case requirement to A' chains, which Chomsky suggests are subject to (274). Kayne ((1981), (1983)) argues that verbs may assign Case into the specifier ((in his terms, COMP) position of their complements; Kiss (1984), adopting Kayne's proposal, suggests that A' chains

may receive Case in more than one position. While Chomsky's (147) is well motivated as a condition on A chains, it appears to not be well motivated for A'chains, and by extension, for composed chains.

Now consider the proposed extension of the theta-uniqueness requirement to composed chains. While it is true that the composed chain for (271) contains two theta-marked positions, this is not so for the composed chain which would arise in the ungrammatical (276).

(276) * which man [t was [arrested t'] [before [OP [you
voted for e]]]]?

If the NP trace is not included in the composed chain, then the composed chain will be (which man, t, OP, e). There is only one theta-marked position in this chain, namely e, therefore (274) is not violated in this respect. Suppose the NP-trace is included; then the composed chain will be (which man, t, t', OP, e). This chain has a link (t', OP) which is such that the first member does not c-command the second. Therefore (like the composed chain for the grammatical sentence (270)) this chain will not be subject to (274). Thus I conclude that (274) is not capable of deriving the anti-c-command condition on parasitic gaps.

This leaves us with Chomsky's final proposed way of deriving the anti-c-command condition. Recall that his formulation of condition C of the Binding Theory prohibits an R-expression from

being bound in the domain of the head of its chain. Chomsky suggests that this might be strengthened to hold of composed chains as well. That is, for a composed chain (C, C') , the variable of C' will have to be free in the domain of the head of C . In (271), where the composed chain is (which papers, t , OP , e), the parasitic gap is A -bound by t in the domain of the head of the composed chain, namely which papers.

We have argued in this chapter for the elimination of Condition C, on empirical grounds, and on the basis that it can be replaced by a well-formedness condition (the Chain Obviation Condition) which accounts for a variety of other facts in a unified way. If this line of argument is correct, we cannot adopt Chomsky's account, since our framework does not include the binding principle he appeals to. We can, however, approach the problem in essentially the same way as Chomsky: just as he assimilates the ungrammaticality of (271) to Condition C violations, so will we, accounting for (271) as a violation of the Chain Obviation Condition, which in our system subsumes Condition C.

Parasitic gaps are only licensed by overt A' movement, that is, movement to an A' position (WH-movement, Topicalization, Heavy NP-shift) which occurs in the D-Structure-S-Structure mapping, movement which is thus represented at S-Structure. I assume, on the basis of this fact, that parasitic operators link

This means that we will have the following linking representations for (270) and (271) (intermediate traces omitted for clarity).

(277) [which papers] did you [file t] [before [OP [you had read e]]]?
[]

(278)

* [which papers] [t fell off the table [before [OP [you had read e]]]]?

I depart from Chomsky in the structure assumed for composed chains. Let us suppose that the parasitic chain composes not with the entire licensing chain, but with the antecedent of the non-overt operator and its antecedents. For cases like (277) and (278), then, the composed chain will be (which papers, OP,

28. This means, perhaps unfortunately, that parasitic operators are distinct from the non-overt operators in Tough-constructions and too-adjective constructions, since their antecedents are in A-positions.

e).

This allows us the following definitions of A' chain and composed chain:

(279) $C = (a_1, \dots, a_n)$ is an A' chain only if:

(i) for every $i < n$, a_i is in an A' position, and

(ii) for every link (a_i, a_{i+1}) ,

(a) a_i c-commands a_{i+1}

(b) a_{i+1} is linked to a_i

(280) $C = (a_1, \dots, a_n, b_1, \dots, b_m)$ is a

well-formed composed chain only if:

(i) C is an A' chain (by (279))

(ii) b_1 is a non-overt operator.

Now consider the ill-formedness of the sentence (271), which by hypothesis has the linking (278). This linking violates the Chain Obviation Condition. The composed chain is (which papers, OP, e). As we have seen, composed chains are subject to the COC. The 'real' gap, t, is dependent upon a member of the composed chain, namely which papers, and c-commands both OP and e, which are members of the composed chain. This is what the COC prohibits.

In the grammatical (270), which has the linking (277), there is the same composed chain, and the real gap t is again dependent upon part of the composed chain. However, t c-commands no part of the composed chain, and so the COC is not violated. The same holds for all cases where the real gap does not c-command the parasitic operator, including examples like (272) where the licensing gap is in subject position.

Thus the anti-c-command condition on parasitic gaps is derived from the Chain Obviation Condition.

Chapter Five: NP-Trace, Quantifier Scope, and a Constraint on Movement

In previous chapters, we have examined cases of anaphor binding in which the anaphor occurs at S-Structure (and, I have argued, LF) outside the c-command domain of its antecedent. In this chapter, we will examine structures in which an NP-trace occurs outside the c-command domain of its antecedent. In particular we will examine constraints on interpretation forced by this syntactic configuration when the antecedent is quantificational. The examination of this phenomenon will reveal constraints on the derivation of LF-structures. We shall further argue that the two operations of WH-movement and QR (quantifier scope assignment) are strictly ordered in the LF component. We conclude with remarks on extraposition, first demonstrating (in part following Taraldsen (1981)) that extraposition is a syntactic operation. We suggest that Taraldsen's demonstration that extraposition must precede all instances of WH-movement derives from the ECP, and demonstrate that extraposition cannot occur in the LF component.

We argue that (1) must be considered to be a principle of grammar constraining movement.

- (1) All instances of movement associate two positions X and Y, one of which c-commands

1
the other .

(1) prohibits movement from one position to another, where the extraction site neither c-commands nor is c-commanded by the landing site. Most instances of licit movement are 'upward', i.e. to a c-commanding position. As demonstrated in May (1977), (1985), QR can also move a quantificational NP downward, to a position c-commanded by the extraction site. (1) limits movement to these two cases.

The constraint in (1) will be shown to constrain movement in the S-Structure to LF component, on the basis of the interpretation of sentences like (27) below. In the second section, we discuss extraposition, arguing that it must be considered to be a syntactic movement rule. This will be substantiated by examining facts observed by Taraldsen (1981), having to do with interactions between extraposition and pronominal coreference.

2
We begin by focusing on sentences like that in (2) . In the

1. Equivalently, within the Path theory advanced by Kayne (1983), there must be an unambiguous path between the two positions. I am indebted to Viviane Deprez for bringing this to my attention.

2. I am grateful to Sam Epstein for bringing the grammaticality of the sentence in (2) to my attention, as well as the problems it raises for Condition A of the Binding theory, the theta-criterion, the principles governing the assignment of features to empty categories, and well-formedness conditions on chains. He and I discuss this type of example in forthcoming work. To the best of my knowledge, the grammaticality of (2) is

second section, I shall discuss in particular sentences of the form of (2), but involving quantificational noun phrases, and argue that the interpretations of such sentences gives insight into conditions on syntactic movement.

5.1 Non-C-Commanded NP-Trace

First, some preliminaries. The discussion in this section is partially adapted from Barss and Epstein (in progress).

2) [[How likely[t to win]]
CP AP IP i

is [John [e]?
IP AP i

Here, John is the derived subject of the main clause, having been raised out of the subject position of the clausal complement of the adjective. The D-structure of the example is as in (3):

3) [[] [e [How likely[John to win]]]]
CP IP AP IP

NP-movement converts the sentence to (4), and WH-movement

first noted by Baltin (1982); in Baltin's discussion the important property of the example is that the complement of the adjective is moved into COMP along with the how likely sequence. Baltin's discussion is based within a pre-GB framework, and thus does not note the problems the example raises in the various GB modules. Baltin actually gives the example as (i), without the trace:

(i) how likely to win is John? (his 63)

derives the S-Structure (2).

4)

[[[[John [how likely[t to win]]]]?
CP IP i AP IP i

Consider the simpler case (5), in which the AP is not questioned.

5) [John is [likely[t to win]]]
IP i AP IP i

In the preceding chapter, I developed a theory of empty categories which asserts that NP-trace is not an anaphor, and indeed bears no binding features. Thus the primary principle constraining the trace itself is the ECP. Additionally, the theta-criterion requires that there be a well-formed A chain associating John and the trace. This is so since the trace is a non-casemarked non-argument, and John is a casemarked, non-theta-marked argument. Neither the singleton chain (t) nor the chain (John) satisfies the theta-criterion, but the complex chain (John, t) does.

As for the ECP, the trace in (4) is not theta-marked by a lexical head, nor is it casemarked by a lexical head. It therefore must be antecedent governed. I here follow Chomsky (1986) and assume that the trace is antecedent governed by likely, in virtue of the special relation between likely, John in its derived position, and AGR. This has the desirable consequence that in (2) the trace's proper governor is carried along with it, and so the trace can be properly governed at

S-Structure³

Structures like (2) call into question the Local Binding Condition on chains formulated by Rizzi (1982). Since the LBC requires that in a chain C, for each link (X, Y), X be the local binder for Y, the constraint is violated in (2), for the chain (John, t). This is so since John does not bind the trace at all. I leave it to the reader to verify that the Chain Obviation Condition advanced in Chapter 4 (which subsumes the LBC) is not violated.

We note that the sentence in (2) is rather similar in form to (6), which illustrates the anaphor connectivity effect.

(6) [which pictures of himself] [did John see t]

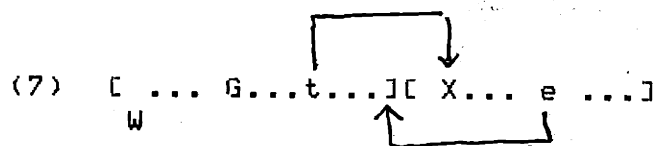
In both cases, we have a dependent element (the trace, or the

3. There is an alternative possibility, which is to assign to (5) and (2) the raising structure assumed for other adjectival constructions in Chapter 3, following Stowell (1981) and Rizzi (1982). This gives us (i) for (5).

i) [John is [t' [likely] t. to win]]
 IP i AP i_{AP} IP i

Here, John will have moved twice. AP is a barrier for t, so John cannot antecedent govern it directly. However, t' is not excluded by the AP, therefore antecedent governs it. Similarly, since AP does not include t', it is antecedent governed by John. For reasons of clarity, however, I keep to the simpler analysis of the phrase structure of raising constructions in the text; all the remarks made there would carry over to structures like (i).

We suggest that this overlap between Condition A and the ECP explains the descriptive parallel between (2) and (6). (2) instantiates the schema (7), where W has moved from the position marked by e, and G is a governor of t, and X is the antecedent of t in a chain (...X, t,...).



- 522 -

in (2). Within this formulation of the ECP, it follows that X must be the subject of minimal clause containing the maximal projection of G. Therefore, the ECP forces NP-movement to be to a locally c-commanding position.

We conclude that the fact that in cases like (5), the fact that the antecedent of the trace c-commands it follows from the ECP, and that the fact that the NP-trace in (2) is not c-commanded by its antecedent needs no special treatment. The analysis of ECs in Chapter 4, with its assertion that NP-trace is not subject to the binding conditions, can be maintained. The apparent parallel between (2) and (6), which might be interpreted as indicating that NP-trace is an anaphor, can be explained as being due to a certain similarity between the ECP and the licensing condition for anaphors.

I now conclude this section. I have attempted to show that the grammaticality of the sentence in (2) requires no particular modifications of the principles of grammar, and that the sentence acquires no exceptional status in virtue of having the structure it has. In the next section, I shall demonstrate that the WH movement of the AP containing the trace seems to restrict the sentence's interpretation in a very striking way: if the EC is the trace of a raised quantificational NP, then the sentence cannot have the lowered quantifier interpretation (in which the QNP has narrow scope with respect to likely). This fact will be attributed to the constraint (1). Before advancing to this, we

will review May's (1977), (1985) analysis of scope assignment in raising constructions.

5.2 Quantifier Scope and Lowering

May (1977) makes the important observation that a sentence like (9) is ambiguous, and (10) is not.

9) A hippogryph is likely [EC to be apprehended]

10) A hippogryph tried [EC to be apprehended]

Both have an interpretation in which the indefinite NP a hippogryph has wide scope, as paraphrased in (11) and (12).

11) there is a hippogryph x such that it is likely that x will be apprehended

12) there is a hippogryph x such that x tried [x to be apprehended]

(9) has an additional interpretation, one which unlike (11) and (12) does not involve commitment to the existence of hippogryphs, in which likely has scope over the QNP. This is paraphrased in (13).

13) It is likely that there will be a hippogryph apprehended

(10) has no such interpretation.

May's treatment of quantifier scope postulates the rule QR, which adjoins a QNP to S (and to VP, in the framework of May (1985)). The rule applies in the S -Structure-LF mapping,

deriving LF representations in which every QNP has been assigned scope, through this adjunction. The scope domain of the QNP is defined as its c-command domain.

The typical case of QR is 'upward', moving the QNP from its S-Structure position X to an adjunction position Y that c-commands X. May observes that this ought not to be stated in the rule itself, since the effect can be derived from other principles.

Consider (14):

14) [_S [every man] thinks [_S Mary is tall]]

There are two S nodes in this example, and thus two possible scope adjunction sites.

Suppose QR adjoins [every man] to the lower S, giving (15):

15) [_S _i e thinks [_S every man [_i Mary is tall]]

This LF representation violates two principles. First, the quantifier binds no variable, violating the restriction against vacuous quantification (c.f. Chomsky (1982)). Second, the EC in matrix subject position -- a theta-marked position -- is free, violating the constraint against free variables⁴. Since these conditions are independently needed for instances of -----

4. Assuming that a theta-marked trace bound by an operator is its variable.

WH-movement, the upward nature of QK need not (and should not) be stated as a part of the rule itself.

Returning to the examples (9) and (10), May's treatment of the additional interpretation of (9) makes crucial use of the difference between the subject positions in the two examples.

9) A hippogryph is likely [EC to be apprehended]

10) A hippogryph tried [EC to be apprehended]

(9) is a typical Raising construction.

The matrix subject position of (9) is a non-thematic position, as is attested by the well-formedness of (16), in which the subject position is occupied by a pleonastic. The ill-formedness of (17) lend support to the analysis of (10) as a sentence in which the matrix subject position is theta-marked.

16) it is likely that a hyppogryph will be apprehended

17) * It is tried that a hippogryph will be apprehended

This difference is further attested by one's intuitions about the meaning of the sentences; in (10), the hippogryph is both trying to do something, and (potentially) being apprehended. In (9), a hippogryph plays no role in the likelihood; what is likely is the truth of the proposition [a hippogryph (will be) apprehended].

The GK analysis of the sentences gives a different characterization to (9) and (10). The EC of (9) is an NP-trace, which marks the D-structure position of a hippogryph. The

matrix position is non-theta marked. The EC of (10) is PRO, which, although referentially dependent upon a hippogryph (like the trace in (9)), bears a distinct theta-role from it.

QR can adjoin a hippogryph to the matrix S in both examples, giving (18) and (19); ((19) indicates the wide scope interpretation of a hippogryph over likely).

18) [a hippogryph [e is likely _i _S _i _S _i _i _i to
be apprehended]]

19) [a hippogryph [e tried [PRO to
_S _i _S _i _i _i _i to
be apprehended]]

Suppose QR applies to adjoin a hippogryph to the embedded S nodes. This derives (20) and (21):

20) [e is likely [[a hippogryph [t to
_S _i _S _i _S _i _i _i to
be apprehended]]

21) [e tried [a hippogryph [PRO to
_S _i _S _i _S _i _i _i to
be apprehended]]

(21) is ill-formed since the EC in matrix subject position does not meet the well-formedness conditions on any EC type. If NP-trace, the the EC is required to be in a chain with an A-position argument; the only candidate is the PRO of the lower clause. The chain (EC, PRO) has two thematic roles, in violation of the theta-criterion. Finally, if the EC is a syntactic variable, it will be required to be A' bound, which it

is not.

Furthermore, since the EC is understood as a variable, it has to be within the scope of its quantifier, but it is not. That is, all semantic variables are subject to the Scope Condition :⁵

(21') At LF, every variable *v* is within the scope of its quantifier.

In (20), as May (1985) argues, the non-thematic status of the EC in matrix subject position allows it to have the same status as the overt pleonastic in (16); he assumes it is an occurrence of pleonastic *pro*, and I will follow his analysis. It, crucially, will not be subject to the Scope Condition. The trace in embedded subject position is theta-marked, and thus functions as the variable.

A desirable consequence of this particular treatment of the scope interaction of *likely* and the raised QNP is that if the complement has another QNP as its object, as in (22), this second QNP may have scope over the raised QNP just in case the raised QNP has scope inside that of *likely*.

22) Someone is likely [t to address every rally]

allowed scope order: likely V E

not allowed: V E likely

This result could not be achieved, as May emphasizes, if we

5. See also Longobardi (1985), Szabolcsi (1986).

were to treat the scope interaction between likely and the raised QNP as a result of raising the predicate likely instead of lowering the QNP. (Further, as he observes, this would require that we allow raising predicates, and only raising predicates, to be scope assigned, leaving it a mystery as to why try cannot be scope-assigned in (10)).

What we will be concerned with in the discussion below is the is the ill-formedness of quantificational of the approximate form of (18), in which:

1. The QNP has been raised out of the subject position of the complement of likely in the D-Structure -- S-Structure component.
2. The AP is WH-moved, or Topicalized.
3. The QNP is understood as having narrow scope with respect to likely.

section(QNP lowering and AP-Raising)

Let us examine each step of the derivation of (25), which is the LF representation related to the interpretation of the string (23) in which someone has narrow scope with respect to likely.

23)[Someone is likely [t to address the
 S i S i
 rally]]

24)[pro is likely [someone to address the
 S i S i

rally]]

25) [pro is likely [someone[e to address the
S i S S i
rally]]

We might consider the derivation to proceed in two steps. First, the QNP moves back to its D-Structure position, deriving (24). Then, it will move from this position and adjoin to the embedded S node, deriving (25). Alternatively, we could think of this as proceeding in one step. (25) might be derived from (23) directly, by moving the QNP from the matrix subject position to a position adjoined to the embedded S; that is, omitting (24) from the derivational sequence. This latter view is the one May adopts.

The step relating (23) to (25), or (24) to (25), is simply QR, moving a QNP from an A-position to an A' position, specifically S-adjunction. As we saw earlier, the typical upward movement of the rule is not part of the rule itself, therefore nothing inherent in the rule blocks 'downward' movement on the second derivation.

If we take the first view, that is, that (25) is derived from (24), before QR applies there is movement from an A position to a c-commanded A-position, deriving (24) from (23). This is the converse of S-Structure NP-movement. No principles of grammar block this, either.

In the derived representation, the (free) empty category in subject position is pleonastic, and by May's assumptions the

fact that it c-commands the variable and the QNP is of no consequence. Crucially, the *pro* need not be c-commanded by the moved QNP.

5.3 Restricted Cases

Now let us consider a sentence like (27), which is just like (2) except that the QNP someone is the derived subject, instead of the non-quantificational NP John.

25) Someone is likely to address every rally

26) How likely is someone to address every rally?

27) How likely [t to address every rally] is someone?

We observe that the ambiguity of (25) is present in (26), but not in (27). Specifically, while (25) has an interpretation in which someone is interpreted within the scope of likely, there is no such interpretation of (27). (This is further supported by the fact that in (25) every rally may have scope over someone, while in (26) it may not). (26), I believe, has the full range of interpretations that (25) has, including those in which someone has narrow scope with respect to likely. I return to cases like (26), which involve extraposition, in a later section in more detail. I focus here on the restricted interpretation of (27).

In (29)-(31) are listed the informal logical representations

for the available and unavailable interpretations.

29. Interpretations of (25).

- a. [some x:person(x)][likely [[every y:rally(y)]
[x to address y]]]
- b. [likely [[some x:person(x)][every y:rally(y)]
[x to address y]]]
- c. [likely [[every y:rally(y)][some x:person(x)]
[x to address y]]]

30. Interpretations of (26)

- a. [some x:person(x)][how likely][every y: rally (y)]
[x to address y]
- b. [how likely][some x:person(x)][every y: rally (y)]
[x to address y]
- c. [how likely][every y: rally (y)][some x:person(x)]
[x to address y]

31. Interpretations of (27)

- a. [some x:person(x)][how likely][every y: rally (y)]
[x to address y]
- b. NOT:
[how likely][[some x:person(x)][every y: rally (y)]
[x to address y]]
- c. NOT:
[how likely][[every y: rally (y)][some x:person(x)]
[x to address y]]

Let us suppose, following May and others, that when likely is understood as having scope over the QNP someone the QNP is adjoined to the embedded S. This means that while the well-formed LF representation (32) may be derived from (25), the LF representation in (33) is not apparently available for (27).

32) e [likely [someone [...]]
 AP S S

33) * [how likely [someone [e ...]]
 S i S i k

 [pro [e]
 i AP k

Given that every rally may have wide scope over someone only when someone is adjoined to the complement S, the lack of both interpretations (31b) and (31c) follows from this prohibition.

Why is (33) ill-formed? There are two possible avenues to explore. (33) might violate one of the conditions that hold of LF representations: the ECP, the Scope Condition, etc. Alternatively, (33) might be well-formed, but the derivation associating (27) with (33) might be illicit. In other words, the unavailability of certain interpretations of (27) might in principle be due either to a violation of a condition on LF representations, or a condition on movement.

In the next sections, I shall show that none of the relevant LF well-formedness conditions are violated in (33), thus supporting the view that the derivation which derives (33) from (27) is not proper. I suggest that the constraint on movement (1) is what rules this out, and show this to have other consequences as well.

Note that I am assuming here that the entire WH-phrase [how likely t to address every rally] must be present in the matrix

SPEC of CP position at LF, i.e. that there is no option of extraposing the S complement to derive an LF like (26). This will be examined in a later section.

5.3.1 Representational Constraints and (33)

Compare (33) and the well-formed (25').

33) * $\left[\begin{array}{c} \text{how likely} \\ S \end{array} \left[\begin{array}{c} \text{someone} \\ i \end{array} \left[\begin{array}{c} e \\ S \end{array} \right] \right] \right]$

$\left[\begin{array}{c} \text{pro} \\ i \end{array} \left[\begin{array}{c} e \\ AP \end{array} \right] \right]$

25') $\left[\begin{array}{c} \text{pro} \\ S \end{array} \left[\begin{array}{c} \text{is likely} \\ i \end{array} \left[\begin{array}{c} \text{someone} \\ S \end{array} \left[\begin{array}{c} e \\ S \end{array} \right] \right] \right] \right]$

address the rally]]

We assume, following May, that none of the LF well-formedness constraints are violated in (25'). The Scope Condition is met, as e_i is within the c-command domain of someone_i . The prohibition against vacuous quantification and free variables is not violated, since someone_i binds a variable (e_i), and the only variable, e_i , is quantifier bound.

We assume that the trace e_i is properly governed as S-Structure, satisfying the ECP along the lines suggested earlier. More neutrally, we suppose that this EC fully satisfies the ECP, whatever its correct formulation.

The pleonastic empty category left in matrix subject position in (33) presumably meets whatever necessary requirements hold of it. The structure cannot be ungrammatical due to the fact that the subject EC is free, since it is free in (25') as well, the (allowed) lowered reading of (25).

It is readily apparent that these conditions are equivalently satisfied in (33). (33) additionally contains the trace of the moved AP, e . This trace is presumably properly governed by the _k WH phrase in SPEC of CP position. It is implausible that the movement of someone from matrix subject position to an adjoined position within the WH phrase interferes with the process of properly governing the AP trace.

I therefore conclude that the LF representation (33) is well-formed, in the sense that it meets those conditions which hold of LF representations. This means that (33) is not an LF which can be associated through a well-formed derivation with (27).

5.3.2 The Interpretations of (26)

Now consider (26). I suppose that the DS-SS derivation of the S-Structure (26) involves extraposing the sentential complement of likely, which originates inside the AP which likely heads. Two such derivations are possible: extraposing the S prior to WH-movement, or after. For reasons to be made clear in a

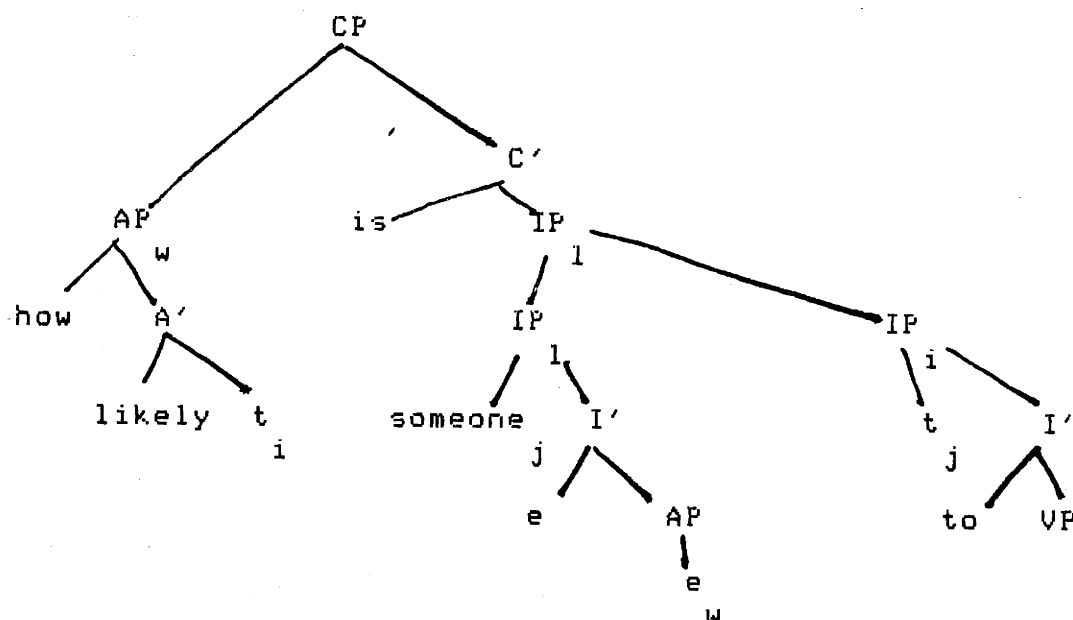
moment, let us assume the first derivation. The position of attachment of [t to address every rally] is unclear. Let us
 S

suppose that the extraposed S is adjoined to the matrix S node, which is the first maximal projection dominating the AP out of which the extraposed constituent moves. Thus, after Extraposition but prior to WH movement the structure will be:

34) [[[[] [Someone is
 CP C' C S S j
 [how [likely t]] [t to VP]]]]
 AP A' i S j i

WH-movement then derives the S-Structure (35):

35) [[how [likely t]] [[]
 CP AP A' i k C' C S
 [Someone is [e] [t to VP]]]
 S j AP k S j i



The multiple scope interpretations work as for (25'), by adjoining someone to IP₁, as in (36), or adjoining it to IP_i, as

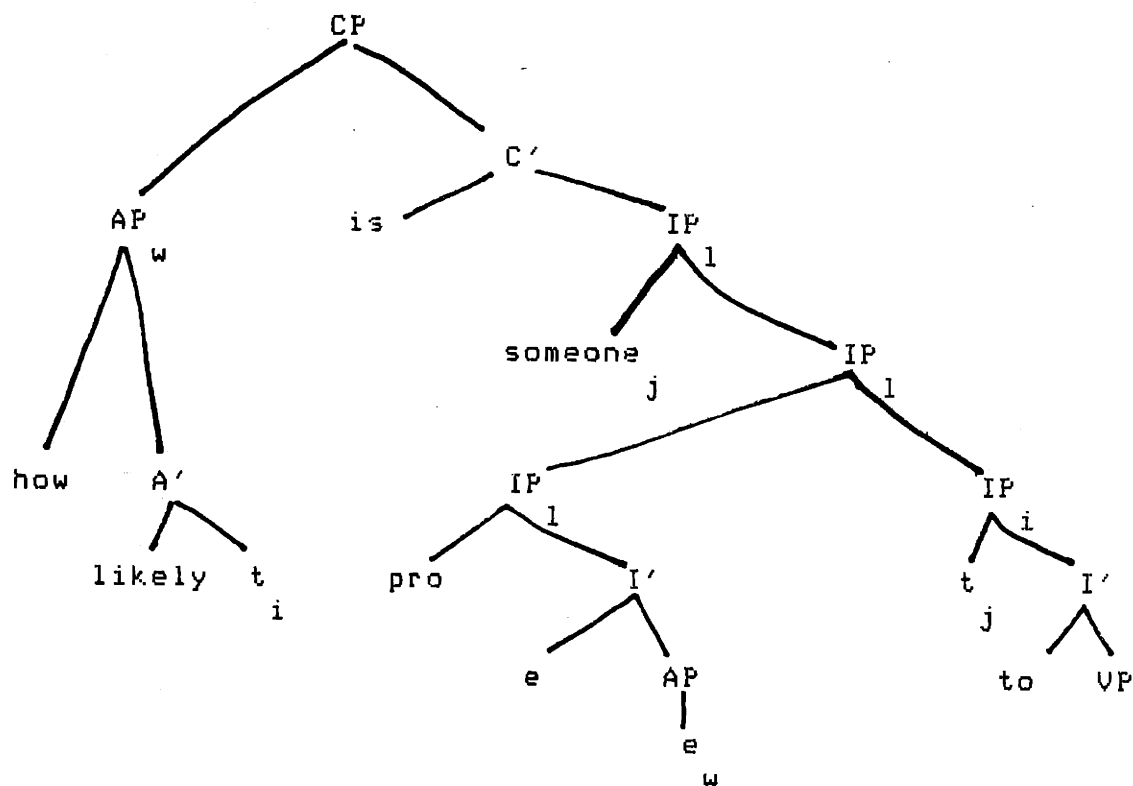
in (37). (Every rally will adjoin to IP_i, and so for the latter

case the QNPs someone and every rally may commute in scope.)

36)

(a) [how likely][[some x:person(x)][every y: rally (y)]
[x to address y]

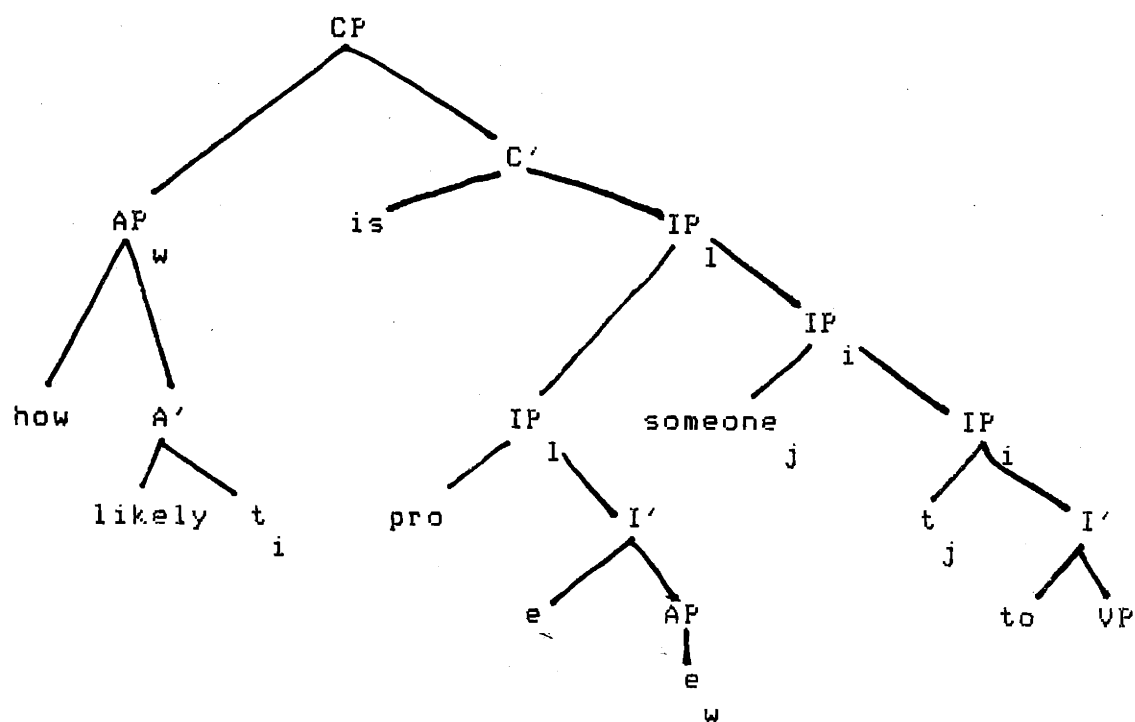
(b) [some x:person(x)][how likely][every y: rally (y)]
[x to address y]



37)

(b) [how likely][[some x:person(x)][every y: rally (y)]
[x to address y]

(c) [how likely][[every y: rally (y)][some x:person(x)]
[x to address y]



In (36), the two quantifiers someone and how likely may commute in scope. In (37), the two quantifiers someone and every rally may commute in scope. Apparently, the presence in (26) of WH movement does not block any of the three scope orderings available for (25).

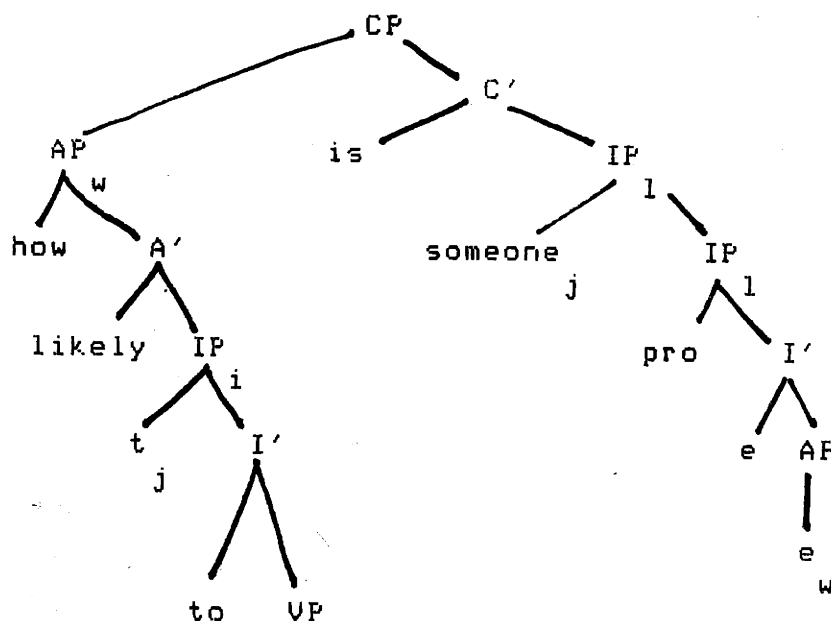
5.3.3 The Interpretations of (27)

As we saw above, only one of the three scope ordering available for (25) (and (26)) is available for (27). The only possible scope ordering is that in which someone has scope over likely, (and over every rally).

There are, in (25), two possible scope adjunction sites--the matrix S node and the embedded S node (the complement clause).

We have seen that, in May's theory, the relative scope ordering of likely and someone is determined by which S node is chosen. When the QNP is adjoined to the matrix S node in an example like (25'), it has scope over the predicate likely, whereas when it is adjoined to the complement S it has scope within that of likely (cf. 18). Apparently, the latter scope adjunction is illicit when the AP has been WH-moved, as in (27). Let us first consider the available ordering, which we suppose is given by this LF:

38)



Recall that it is the theta-marked anaphor t inside the WH-phrase which acts as the variable associated with the QNP someone. By the Scope Condition, this variable will have to be understood within the scope of someone. The interpretation

indeed available for this sentence supports this Condition:
someone must be interpreted with widest scope.

May (1985) demonstrates that a WH phrase in COMP (in his terms) and a subject quantifier can be understood in either scope order. To illustrate, the sentence in (39) has both interpretations in (40).

39) Who_i does everyone love t_i

40. a. [which x: person (x)][every y:person (y)] [y loves x]

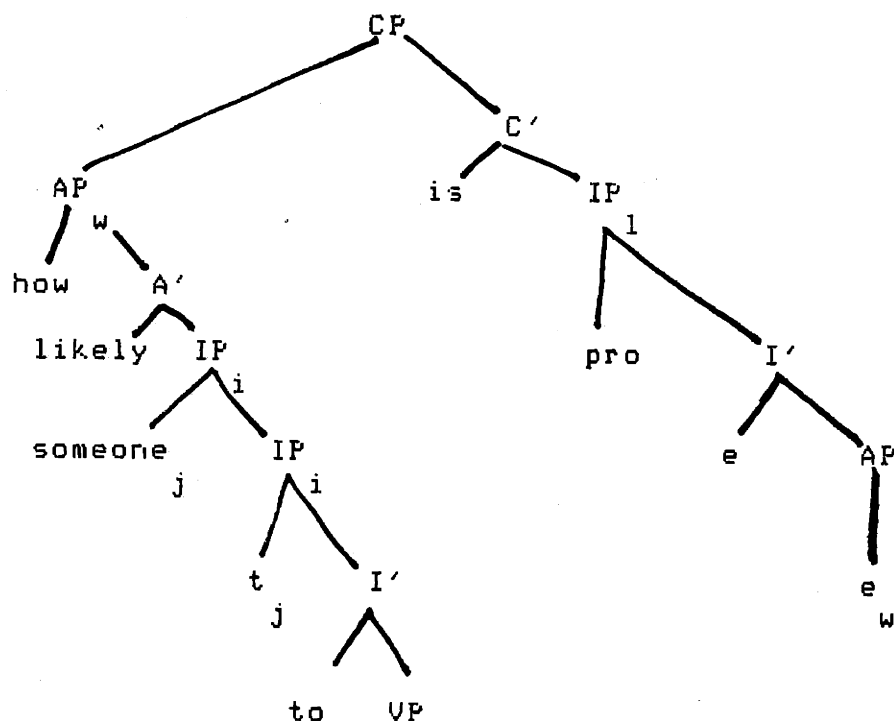
b. [every y:person (y)] [which x: person (x)][y loves x]

(40a) admits a single answer, like 'John', while (40b) admits a distributive answer.

So, we see that in the LF representation (38), the QNP someone and the WH-phrase how likely t to win are in a configuration which permits them to commute in scope. However, since the WH-phrase contains a variable associated with someone, the QNP must be interpreted with wide scope over the entire WH-phrase.

Now, there is another way to get the QNP someone to have scope over its variable t_j, namely to adjoin it to the S node which is the complement to likely. This will require the LF representation in (41).

41)



Taking VP to be 'address every rally', with someone adjoined to the complement S, we will expect that every rally will be able to have wide or narrow scope over someone, and that both QNPs will have scope within that of likely.

These are exactly the two interpretations blocked for (27). So, it appears that (41) is an ill-formed LF representation for (27). (41) is actually a representation of the general form (33). As we observed earlier, apparently none of the LF well-formedness constraints is violated in an LF structure like (33), which leads us to the conclusion that it is the derivation of the structure (41) (from (27)), not the structure itself, which is at fault.

Observe that the derivation of the LF (41) from the S-Structure (27) involves movement which associates the matrix subject position with the S node contained in the WH phrase, moving the QNP which occupies the matrix subject position at S-Str into a position of S-adjunction at LF. This derivation may involve one, or a series of steps; but clearly at least one step of this derivation will associate two positions which are such that neither c-commands the other. By contrast, all the movements deriving (36) and (37) associate positions in which one c-commands the other.

This suggests the constraint (1), repeated here.

- (1) All instances of movement associate two positions X and Y, one of which c-commands the other.

5.3.4 On Answers and Variables

As reviewed above, May has argued that in the configuration (42) the WH-phrase and the IP-adjoined QNP may 'commute in scope', that is, be understood in either scope order.⁶

(42) [WH [QNP [...]]]
 CP IP IP

Below, I shall briefly outline a theory of what constitutes a proper answer to a question, adapted from Higginbotham and May

6. May makes different assumptions concerning phrase structure, and which phrases are maximal, than Chomsky (1986). I will keep in the text to the 'Barriers' framework.

(1982) and Higginbotham (MIT lectures). The reader is referred to Higginbotham and May's work for details.

Suppose for convenience that the domain of discourse is restricted to two individuals, John and Mary. In (44) are answers which are relevant to, or appropriately related to, the question in (43). (45), on the other hand, is not really an answer to the question.

(43) a. Who [left]?

b. for which people x , [x left].

(44) a. John did.

b. Noone did.

c. I don't know if John left, but I know Mary did.

(45) Fish have fins.

In the rough logical form (43b), the variable x ranges over the two individuals John and Mary. There are four possible scenarios:

(46) a. John left, and Mary did not.

b. Mary left, and John did not.

c. Mary and John both left.

d. Neither John nor Mary left.

An answer to the question is a sentence which is inconsistent with at least one of these possible scenarios. (44a) is inconsistent with (46b) and (d). (44b) is inconsistent with (46a, b, and c). (44c) is inconsistent with (44a) and (44d). Thus an answer to a question narrows down the number of possible scenarios which could be true. (45) is thus not an answer,

since it fails to eliminate any of the possible scenarios ⁷.

Now consider the configuration (42), instantiated by (47b), the LF representation of (47a).

47a) what did noone buy t?

47b) [what [noone [t buy
CP i IP j IP j
t]]]?
i

(47a) has only one type of answer, the singular answer, as in 'A pair of green shoes'. This reflects the interpretation of (47b) with noone having narrow scope with respect to what, i.e. 'what is the thing x such that no y bought x?'. There is no possible distributed reading for the question in (47a).

This follows from the theory of answers. Suppose that noone were interpreted with wide scope. The question would then have the approximate meaning 'for no x, tell me what x bought'. But this is a question which by its very form has no particular answer.

Let me explicate this in the terms given above. In the approximate logical representation (48), noone has wide scope over the interrogative.

7. Of course, under the right circumstances (45) could be an answer to (43); if we knew beforehand that John would leave if fish have fins, for example.

48) [no x: x a person] [which y: y a thing] [x bought y]

An (appropriate) answer to a question is a statement which is inconsistent with at least one of the set of scenarios associated with the open sentence in the question. There is simply no scenario possibly associated with (48); the quantifier noone, having wide scope over the WH-operator, restricts the domain of discourse to a set having no members. Therefore, no answer could be inconsistent with a subset of the possible scenarios, since there aren't any to begin with. Thus, the fact that the quantifier noone cannot be understood with wide scope over an interrogative is derived from the theory of answers.

May argues that (49) is ambiguous, with the two quantifiers understandable in either scope order.

(49) everyone loves someone

- a. for every person x, there is some person y such that x loves y.
- b. for some person y, every person x is such that x loves y.

When one of the quantifiers contains a variable bound by the other, the pronoun-binding quantifier must have wide scope:

(50) everyone loves someone who knows him

- a. for every person x, there is a person y who knows x such that x loves y.
- b. NOT: for some person y who knows x, every person x is such that x loves y.

This follows from the Scope Condition, since the pronoun counts as a variable.

Now consider (51), on the reading where the value of he is determined by the quantifier everyone.

51) which picture of him does every man like ?

a. His wedding pictures.

b. Not: John likes his wedding pictures,
Dan his graduation pictures, Bob the
ones of him in France, ...

The question is whether every man must have wide scope over which pictures of him. The answer is apparently no, for two reasons. First, the question forbids a distributed answer, permitting only the relational answer in (51b), where the pictures which each man likes are related to him in the same way as all the other men's pictures are related to them.

Secondly, (52) has a perfectly acceptable relational reading.

52) which pictures of him does no man like ?

a. His high school graduation picture.

If the relational answer were contingent upon noone having wide scope, (52) would be as incoherent as (48).

We assume (see Engdahl (1981) for a different view) that relational answers are acceptable only when Higginbotham's Accessibility Condition (cf. chapter 4) is met in the full answer. That is, (51a) is an abbreviated form of (53), and in (53) the pronoun will be accessible to the QNPs formal variable.

(53) Every man likes his wedding pictures.

No relational answer is available in (54).

(54) Which of his pictures proved that every man is tall?

a. NOT: His wedding picture.

This will be due to the failure of accessibility in (55).

(55) His wedding pictures prove that every man is tall.

Returning now to (51), the distributed reading is impossible when him is taken as a variable bound by every man, although the distributed reading is available if him is independently referential.

This follows from the Accessibility Condition. In order to be understood as a true variable, the pronoun will have to be accessible to the quantifier's formal variable at LF, which it is not:

56) [which pictures of him] [every man] [t likes t']

Here, t' is the formal variable; it does not c-command the pronoun, hence the failure of Accessibility.

Let us return now to cases like (27).

We observe that (57) has no distributed reading, and no relational reading either. The question is unacceptable on any interpretation.

(57) [how likely [t to win]] is noone?

a. NOT: as likely as his competitors

The absence of the relational reading is a puzzle, since relational readings are available for sentences like (52). The two questions are very similar, differing essentially in that the QNP-dependent element in (52) is a pronoun, while in (57) it is a trace. Let us then suppose that relational answers are an entirely distinct phenomenon from true Quantifier-variable relations, and that when NP-trace has a QNP as its antecedent it functions purely as a variable. It is thus subject to the Scope Condition, and so (57) must be interpreted with noone having scope over t.

Two possible interpretations, (58b) and (59b), meet this requirement. They are derived from the LF representations (58a) and (59a).

(58) a. [[how likely[t to win][is [noone
 CP AP IP IP
 [pro [e]]]]]
 IP AP

b. [no person x][for which degree y] [y likely [x to win]]

(59) a. [[how likely[noone [t to win][is
 CP AP IP IP
 [pro [e]]]]]
 IP AP

b. [for which degree y] [y likely[no person x] [x to win]]

In parallel with (48), (58b) is ill-formed by the theory of answers; it has no real answer, hence is incoherent. The LF giving rise to this interpretation is well-formed, however.

(59b), on the other hand, is a perfectly sensible question,

asking the hearer to come of with some degree of likelihood *y* (very, not very, as likely as the Queen Mother) such that it is *y* likely that *noone* wins. However, as we have argued, the derivation of the LF structure (59a) giving rise to this interpretation is illicit, violating the constraint on movement (1).

Returning to (57), we observe that the structure is perfectly acceptable if it is embedded.

(60) I know [[how likely [*t* to win] [*noone* is]]

By the Scope Condition, *noone* will have to be interpreted as having scope over *t*. By our condition on movement, this will force *noone* to adjoin to its dominating IP node (61), not internal to the moved AP (62).

(61) I know [[how likely [*t* to win] [*noone* [*pro* is]]]

(62) I know [[how likely [*noone* [*t* to win]] [*pro* is]]

(61) gives rise to the only interpretation possible for (61), namely 'I know, for no *x*, how likely *x* is to win', essentially an assertion of ignorance on the part of the speaker. (62) would give rise to the interpretation 'I know the degree *x*, such that it is *x* likely that *noone* wins'. This absolutely is not a licit interpretation of the sentence. Although a coherent meaning, it is produced by an LF which, we propose, has no licit derivation.

We further observe that the following, if a bit awkward, is

acceptable.

(63) [[Likely[t to win]], [noone is]].
IP AP IP* IP

(63) again has only the meaning paraphrased in (64), not the meaning (65). (65) would be determined only by an LF (65b) in which noone is adjoined by QR to the IP* inside the Topicalized AP, a movement blocked by (1).

(64) no person x is such that it is likely that x wins.

(65) a. it is likely that noone wins.

b. [[likely [noone [t to
IP AP IP* IP*
win] [pro is]].
IP

(64) is determined by the (licitely derived) LF (64b).

(64b) [[likely [t to win]] [noone
IP AP IP* IP
[pro is]]]
IP

In further support of this proposal, we observe that a quantifier inside the moved AP cannot have wide scope over the QNP raised in the DS-SS mapping out of the AP. For example, consider (65).

66) I know how likely to win every race someone is

This has only the interpretation paraphrasable as (67), not (68). (68) could only be determined by an LF in which someone has lowered inside the AP.

67) I know, for some person x, how likely it is that x won every race.

68) I know how likely it is, for each race y, that some person x won y.

5.3.5 Crossed Binding

Further support for our proposal comes from the prohibition of crossed binding in sentences like (69).

69) how likely to have passed some test assigned to him was every student who took it?

To review, crossed binding is the phenomenon illustrated in (70), where we may take each of the pronouns to be bound as a variable by the other quantifier:

70) [Every student who took it] passed [some test he dreaded]

(Such sentences are discussed in Jacobsen (1977), Higginbotham and May (1982), Haik (1984), May (1985), and the sources cited there.)

This type of reading is available only when the two quantificational NPs have the same scope domains; it is blocked in (71).

(71) Every student who took it said that I had seen some test he dreaded

While grammatical, (71) lacks the crossed reading (specifically, it cannot be bound as a variable by some test he dreaded). This is a consequence of the fact that the embedded quantifier has

the complement clause as its scope domain, while the other has the matrix clause as its domain (after QR).

Higginbotham and May (1982) offer a rule of Absorption, which has the effect of combining two unary quantifiers into a single binary quantifier, if they have the same scope domain. The details of their analysis need not concern us here. The important fact is this:

- (72) For two QNPs Q1 and Q2, where each contains a pronoun, each pronoun may be dependant upon the QNP not containing it iff Q1 and Q2 have the same scope domain at LF.

It has been observed that the crossed binding effect is retained under raising:

- (73) Every student who took it passed some test assigned to him

- (74) [Every student who took it] is likely [t to have
IP*
passed [some test assigned to him]]

Following May (1985), we may suppose that the possibility of crossed binding in (74) is due to the possibility of lowering every student who took it to the lower IP, deriving the LF (75), which meets (72).

- (75) pro is likely[[Every student who took it]
IP*

- [[some test assigned to him] [t to have passed t']
IP* IP*

Now, consider (65), which lacks the crossed binding interpretation of the pronouns.

69) [[how likely] t to have passed [some test
 CP AP IP*
 assigned to him]]] was [every student who took it]]?
 IP

By (72), the only LF representation which could give rise to the crossed binding reading is (76):

76) [[how likely [every student who took
 CP AP IP*
 it] [some test assigned to him]
 IP*
 [t to have passed]]]] was [pro]]?
 IP* IP

Some test assigned to him is assigned scope by adjoining it to IP*. Crucially, the crossed binding interpretation is permitted only if every student who took it is also adjoined to IP*. But, as we have seen, such movement is impossible.

5.3.6 Idioms and Pleonastics

The phenomenon discussed up to this point have had to do with quantifier scope, but other facts also fall under the movement constraint (1).

Both pleonastics and subparts of idioms can raise:

77) There is likely [t to be a man outside]

78) Advantage is likely [t to be taken of John]

Tabs are likely [t to be kept on Mary]

However, if the AP is WH-moved or Topicalized, the sentences become strongly ungrammatical.

- 79) * How likely to be a man outside is there?
- 80) *[[Likely to be a man outside], there was.
- 81) * How likely to be taken of John was advantage?
- *How likely to be kept on mary were tabs?
- 82) *Likely to be taken of John, advantage was
- * Likely to kept on Mary, tabs were.

Earlier in this thesis, we adopted Chomsky's (1985) proposal that pleonastics must literally be replaced at LF by a coindexed argument. Thus (77) has the LF (77').

77') A man is likely [t to be t' outside]

Now consider (79). In the LF component, a man will have to move from inside the phrase in SPEC of CP to matrix subject position: another instance of movement blocked by (1). This movement is of an entirely different nature than QR, demonstrating the generality of (1). Similarly, (1) prohibits the necessary movement for the case of topicalized AP in (80).

Now we turn to the final case, the ungrammatical cases involving idioms in (81) and (82). It is reasonable to suppose that idioms must be reassembled at LF, since the idiom as a whole constitutes a semantic item, not being compositional in meaning. Thus we suppose that in (78), LF movement will

necessarily undo the overt NP-movement, deriving (78')⁸.

78') pro is likely [advantage to be taken of John]

pro are likely [tabs to be kept on Mary]

Now consider the unacceptable strings (81) and (82). In each case, reassembling the idiom will require moving the Raised chunk to a position which neither c-commands nor is c-commanded by its S-structure position. Thus (1) rules out these cases as well.

5.3.7 No LF Extrapolation

The constraint in (1) prohibits deriving (41) from (27). This, we argued, restricts the interpretations of the sentence in (27).

By contrast, the sentence (26), in which extraposition has occurred, seems to have the interpretations available which (27) lacks. I assume that (26) has the structure (83), where [It to address every rally] is adjoined to the matrix clause IP₁ by extraposition.

8. Or more fully, (78''):

78'') pro is likely [pro to be taken advantage of John]

pro are likely [pro to be kept tabs on Mary]

Some speakers do not find (78) to be fully grammatical; all find (81) and (82) to be completely impossible.

83) [[how [likely t]] [CP AP A' i k C'] is] [[Someone [e]] C IP1 IP1 j AP k] t to address every rally]]] IP j i

As another example, consider the contrast between (84) and (85).

(84) Someone_i is likely [t_i to break into every one of these cars]

85) [How likely [t_i to break into every one of these cars]] is [someone_i e]?

(84) quite plausibly be uttered to express the opinion that it is quite likely that each of the cars in question will get broken into. This reflects the scope ordering of elements such that likely has wide scope over both of the QNPs, and every one of these cars has scope over someone. This interpretation is provided by an LF in which every one of these cars is lowered and adjoined to the complement clause, and someone is assigned wide scope over it. Both instances of QR satisfy (1), since one (scope assignment of someone) moves an element to a c-commanding position, the other (scope assignment of every one of these cars) moves to a c-commanded position.

(85) cannot be a question about the general likelihood of it being the case that every car will get broken into. Rather, it is a question about a particular someone, questioning the degree to which it is likely that that one individual will break into

every car. This restriction on interpretation is expected, given (1), since it requires the kind of illicit 'sideways' movement prohibited by this constraint.

Now consider (86), the extraposed counterpart of (85).

(86) how likely is someone to break into every
one of these cars?

[how likely t] [is [[someone
i k IP* IP* j
e] [t to break into every one of these cars]]]
k IP j

(86) seems to have an interpretation upon which every one of these cars has wide scope over someone, thus expressing the question 'how likely is it that for every one of these cars there is someone who will break into it?'. This indicates that there is an LF for (86) in which the two QNPs can commute in scope, and the derivation of this LF meets (1). Adjunction of someone to the extraposed clause itself (as in (87)) is blocked by (1), since the landing site does not c-command (nor is c-commanded by) the extraction site.

(87)
[[how likely t] [is [[pro
i k IP* IP*
e] [[someone] [t to break into
k IP j IP j
every one of these cars]]]]]

However, (1) permits someone to be adjoined to the matrix IP* node, as in (88); in its landing site, it c-commands the

extraction site (marked by pro).

(88)

[[how likely t] [is [[someone] [
 i k IP* j IP*
 [pro e] [t to break into
 IP* k IP j
 [every one of these cars]]]]]]

In this position, someone has scope over its variable t, as
 j
 required by the Scope Condition. Then, every one of these cars
 may be moved by QR to adjoin to the extraposed IP, giving the
 full LF (89).

(89)

[[how likely t] [is [[someone] [[pro e]
 i k IP* j IP* IP* k
 [[every one of these cars] [t to break into t]]]]]
 IP IP j m

Now, the two QNPs are very local to one another, Neither is
 included within either of the IPs, and both are included in CP;
 we presume that this configuration permits the QNPs to be
 interpreted in either scope order (see May (1985)), crucially
 permitting every one of these cars to have scope over someone.
 I sketch the configuration more simply in (90).

(90)

[someone [[...] [every one of
 IP* IP* IP* IP
 these cars [...]]]
 IP

Now, we observe that (89) is the LF representation licitely

derived from the sentence (86), which is identical to (85) except that extraposition has applied to derive (86). The question is the following: what prohibits extraposition from occurring in the SS-LF mapping, applying to (85) to derive (86) as an intermediate representation? If this occurred, then (85) could have the interpretation it does not have, where every one of those cars has wide scope over someone. Indeed, if such extraposition could occur, then (85) would be synonymous with (86), which it is not.

Assume that extraposition is IP-adjunction. Then, to extrapose the IP-complement of likely in (85) to derive (86) would violate the constraint on movement I propose s (1). The IP moves from a position inside the WH-phrase, and adjoins to matrix IP*; neither position c-commands the other.

In the next section, I will examine a more complicated case which leads us to conclude that extraposition, even if it meets

9

(1), cannot occur in the LF component.

Before proceeding to that discussion, however, observe that the

9. Further evidence for this conclusion can be adduced from the contrast in 9i) andf (ii):

- (i) how likely to have passed some test assigned to him was every student who took it?
- (ii) how likely was every student who finished it to have passed some test assigned to him?

Crossed binding is permitted only in (ii).

argument just presented to explain the contrast between (85) and (86) is a powerful argument that extraposition must be a syntactic operation, and cannot be restricted to the 'stylistic' component. The reason for this is that if the extraposition in (86) is not the result of a syntactic operation, then its effects would not be registered at LF: both (85) and (86) would have identical inputs to LF, of the form (85), and neither ought to have the interpretation discussed for (86).

Chomsky (1986) presents examples due to Ross (1983), and claims that they demonstrate that extraposition is not syntactic. I will briefly show that these cases receive a simple treatment in the theory of binding developed in this dissertation.

Chomsky considers the following case:

(91) They_i desired that C_i C_j pictures t_j be painted
[of each other]_{i j}

He argues that if this structure enters the LF component (where, he assumes, the Binding Theory applies), a Condition A violation will arise. The subject pictures t_i will induce an SSC effect, forcing the anaphor to be bound in the lower clause. Since in fact the anaphor's binding domain extends to the matrix clause, the extraposition is not represented at LF.

Observe that Chomsky somewhat overstates the argument, since

it could be assumed that Extraposition is both a syntactic and a stylistic rule; (91) could be seen as arising from a stylistic instance of extraposition, not represented at LF, while my example (86) could arise from actual syntactic movement of the extraposed IP.

However, (91) is easily accounted for within the chain accessibility framework developed here, even if (91) is itself evaluated by the binding theory. There is clearly a chain (PP_j, t_j) . There will then be a CAS for the anaphor {himself, t_j , NP^* , IP^* , VP , IP } (ignoring a few steps). The lower subject is not accessible to the anaphor through this sequence, since it is not a sister to any node in it. Thus the matrix subject the men is minimally accessible to the anaphor, and may licitely serve as its antecedent.

In the next section, we again consider extraposition in the LF component.

5.4 WH-In-Situ and Ordering of Rules at LF

Consider the following (due to Howard Lasnik).

(92) Who thinks that everyone is [how likely [t to win]]?

(93) Who thinks that noone is [how likely [t to win]]?

These seem to have the same status (ignoring the echo

interpretation of the AP) as the examples considered earlier, repeated as (94) and (95).

(94) [how likely [t to win]] is everyone?

(95) [How likely [t to win]]

In (92), as in (94), everyone cannot have narrow scope with respect to likely. (93) is as plain unacceptable as (94).

We attributed the facts concerning (94) and (95) to the constraint (1). It appears, then, that (92) and (93) are likewise restricted.

Suppose that the entire WH-in-situ phrase [how likely [t to win]] is first moved to the matrix SPEC position, then QR applies to assign scope to the quantifiers everyone and noone. In this case, (1) will block adjunction of the QNP inside the WH-phrase.

However, suppose that the quantifier is first assigned scope, then the AP is moved. This derivation is illustrated in (96)-(97), where (96) is derived from (93) by QR.

(96) [CP Who [t thinks that[IP pro is
[how likely[IP* noone [t to win]]]]]]?

(97) [CP [how likely[IP* noone [t to win]]]
Who [t thinks that[IP pro is [e]]]]?

Neither of these movements violates (1); the instance of QR is

that found in quantifier lowering, and the WH-movement is upward. (97) will give rise to an interpretation for (93) that the sentence does not have.

What this suggests is that the two operations of QR (quantifier scope-assignment) and move-WH (interrogative scope assignment) are formally different operations, and are strictly ordered in the LF component. Specifically, we claim that (98) is true:

- (98) All instances of WH-movement precede all instances of QR in the LF component.

This will now bar the derivation (96)-(97). With this derivation excluded, the restricted interpretations of (92) and (93) can be assimilated to those of (94) and (95), all resulting from the restrictions on scope assignment imposed by (1).

There is one more derivation of (92)-(93) to consider. Suppose, prior to WH-movement of the AP, the complement of the adjective is extraposed. This will mean that the lower clause of (93) will have a derivation in the LF component paralleling the S-Structure derivation of (86).

The resulting structure will be (99).

- (99) [CP Who [t thinks that [IP IP noone is [how
likely t]] [t to win]]]?
k IP* i k

(100) [CP Who [t thinks that [IP IP pro is [how

```
win]  ] ] ]?
      k
```

(101) [which person x] [y a degree] [y likely [noone z [z to
win]]]

in order to block this structure (100), we must assume one of two things: either there simply is no extraposition in the LF component, (which will of course rule out this derivation). or else extraposition must follow, never precede, instances of LF H-movement. In the next section, I will demonstrate that V-Structure extraposition-plausibly precedes overt WH-movement, and thus I tentatively conclude that (99)-(100) is illicit because there simply is no extraposition in the LF component.

5.4.1 Review of Taraldsen (1981)

Taraldsen (1981) examines the interaction of relative clause extraposition with WH movement, arguing that extraposition can have an effect on the possibility of coreference between a pronoun and a name, and that these facts lead to the conclusion that (in English) extraposition either precedes all instances of WH-movement, or takes place in a different SS representation.

10

Taraldsen argues for an organization of the grammar .

Following Baltin (1979), Taraldsen adopts the position that extraposition is a syntactic adjunction process, which adjoins a relative clause (RC) to S if the NP out of which extraposition has taken place is the subject; and which adjoins the RC to VP, but not to S, when the RC is extraposed from the object. The following paradigm is given to support this conclusion:

103) a. Although [nobody t] would ride with Fred [who
 NP i
 knew just him] , people would [e] who knew
 i VP
 his brother.

 b. * John call[people t] up [who he has never
 NP i
 met before] , and Bill does [e] who he has
 i VP
 never met before

10. Taraldsen's study is much concerned with the markedness of certain rule interactions, and actually suggests this as the unmarked case. He further suggests that Norwegian exhibits a marked option, which has a somewhat different organization.

c. * John said that he would call [^{NP} people t] ⁱ up
 who are from Boston, and [call people up] [he
^{VP} will [e]] ^S who are from Boston.
^{VP}

In (103a), a VP has been deleted, leaving the RC (which is extraposed from inside the subject NP). This indicates that the RC is attached to some position external to VP, plausibly S. In (103b-c), on the other hand, the RC which has been extraposed from inside the object NP cannot be left behind by VP-deletion or VP preposing. This indicates that this is the only possible landing site for the extraposed CP. This interaction of extraposition with syntactic phenomena indicates that it is to be taken as a syntactic, not a stylistic, rule

Taraldsen then considers the interaction of WH-movement and extraposition. He presents two conflicting sets of evidence. The first comes from the apparent lack of WH-island effects with extraposition, from which he concludes that extraposition occurs at a level of representation at which WH-movement has not occurred.

Secondly, Taraldsen observes that extraposition can affect possibilities for pronominal binding, indicating it to be a syntactic process, in that the output of extraposition must be expressed at the level(s) of representation at which the binding theory, at least Condition C, applies.

The crucial contrast is (104) and (105):

(his (17)) (104) how many people who John knows
well does he like ?
i

(his (18)) (105) how many people does he know
who John likes?
i

Coreference is permitted between John and he only in (104). This indicates that the relative clause in (105) is within the c-command domain of he. Given the assumption that extraposition from an object NP adjoins the moved RC to VP, we obtain the right result.

Taraldsen additionally observes the same contrast in cases of long-distance WH-movement, as in (106) and (107).

(106) (his (39)) How many people do you think he
likes [who John knows really well]?

(107) (his (40)) How many people do you think like
him [who John knows really well]?

Coreference between John and him is possible in (107), but strongly prohibited in (106). This indicates that the relative clause in (106) is within the c-command domain of the pronoun him. Apparently, the landing site of RC extraposition is restricted to the landing site available if WH-movement did not take place, namely the VP immediately dominating how many people who John likes at D-Structure.

Taraldsen observes that we would expect coreference to be possible if the RC could extrapose from the intermediate COMP and adjoin to the matrix VP*. This would leave John outside the

c-command domain of he, as in (108).

(108) [how many people t] [do [you
CP j k IP*
[[think [[he [likes t]]]
VP* VP* CP IP VP k
[who John knows really well]]]]
j

Taraldsen proposes that (108) is blocked by the ordering of extraposition and WH-movement; extraposition cannot apply to a structure affected by WH-movement, hence (108) cannot be derived.

Because of this ordering, the only possible structure for (107) is (109), in which John remains within the c-command domain of he.

(109) [[how many people t] [do [you
CP j k IP*
[think [[he
VP* CP IP
[[likes t] [who John knows really well]]]]]]
VP VP k j

What I would like to suggest is that this necessary ordering of extraposition and WH-movement in the DS-SS mapping can be derived from the ECP and from (1).

Extraposition leaves an empty category, the trace t_i above. Since this EC is neither casemarked nor theta-marked by a lexical head, it must be antecedent governed. Such structures as (108) and (109) appear to violate the ECP, since the trace of

the extraposed relative clause is not c-commanded by the RC, hence cannot be properly governed by it in this representation. In order to allow (109), I will adopt suggestions of Noam Chomsky (p.c.) concerning the ECP. These are introduced in the next section; afterwards, I return to (108) and (109).

5.4.2 Proper Government at Intermediate Levels

Epstein (forthcoming) notes a significant problem for the ECP as formulated in Chomsky (1986). Consider the following S-Structure (110), and its two possible LFs (111) and (112):

110) [CP [who]_i [t_i bought what]_j]?

111) [CP [[what]_j [who]_i] [t_i bought t_j]]?

112) [CP [[what]_j [who]_i] [t_i [t_j ' [bought t_j]]]]?

(111) and (112) differ in the presence or absence of a VP-adjoined trace.

In the Lasnik and Saito (1984) formulation of the ECP, t_i will be antecedent governed by who_i, at S-Structure. The object trace t_j created through LF-movement of what_j will be properly governed by the verb, since the verb both theta-marks and case-marks it; both are instances of proper government. Thus the LF representation (111) will satisfy the ECP.

Recall from the preceding chapter that Chomsky advocates the

elimination of theta- and case-marking by a head as instances of proper government. It follows then that if the object trace t_j is to be properly governed, it must be antecedent governed, e.g. c-commanded and governed by an antecedent.

Epstein's observation is this: in both of the LF representations, there is an ECP violation. First consider (111). Here, what neither governs (since VP is a barrier) nor c-commands the trace t_j .

In (112), the trace t'_j properly governs t_j , so the ECP is satisfied for the object trace. However, what fails to c-command the VP-adjoined trace, hence t'_j is not properly governed, so it violates the ECP. Epstein's solution to this problem is to propose that the complete restriction of proper government to antecedent government is incorrect, and that Case-marking by a lexical head also is an instance of proper government. Thus in (111) the ECP is satisfied. (Chomsky's (1986) proposed restriction is based essentially on the proper government of NP-trace; he essentially argues that theta-government by a head is not proper government. Thus, Epstein's and Chomsky's proposals are basically compatible). See Epstein (forthcoming) for details and many more cases.

N. Chomsky (p.c.) suggests an alternative to Epstein's

proposal. He suggests that (110) can satisfy the ECP if we allow gamma-marking to occur at intermediate representations, in the course of the SS-LF mapping. Thus, on this proposal, (110) is mapped into (113), and the VP-adjoined WH-phrase assigns +gamma to t_j as in (114). The WH-phrase then moves to SPEC, leaving no trace, deriving the grammatical LF (115).

113) [CP [who] [t_j [[what] [bought t_j]]]?

114) [CP [who] [t_j [[what] [bought t_j]]]?
+G

115) [CP [[what] [who]] [t_j bought t_j]]?
+G

Let us now return to Taraldsen's cases of extraposition. We observed above that, within the Lasnik and Saito theory of ECP, the trace t_j in (109) will violate the ECP, since it is not

c-commanded or governed by the extraposed clause.

(109) [CP [how many people t_j] [do [IP* you

VP* think [CP IP he

VP VP likes t_j] [who John knows really well]]]]]]
+G

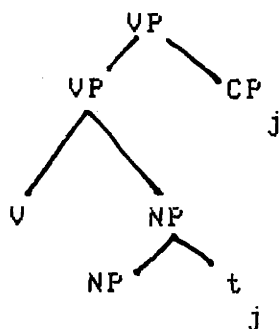
Let us extend Chomsky's proposal concerning (11), and suppose that the trace of a relative clause may also be assigned the feature +G at an intermediate level in the derivation. Thus, there must be one step in the derivation at which the trace is

present, such that it is both c-commanded and governed by the RC. (116) is the first post-D-Structure representation for (109).

(116) [CP [do [IP* you [VP* think [CP IP
 [VP VP likes [how many people t]] [who John
 knows really well]]]]]]
 j k
 j

The relevant part of the structure is as in (117).

(117)



In this structure, NP is L-marked by V, and hence is not a barrier. Since the relative clause (CP) is adjoined to the VP, it is not excluded by it, and VP does not constitute a barrier separating t and CP¹¹. Hence, CP governs its trace, and binds it as well. Therefore, on our assumption, +G can be assigned to

11. We note that this will block extraposition from moving the CP any further, to a position outside VP, since if CP is outside VP this VP will be a barrier, blocking government, hence giving rise to an ECP violation.

the trace by its antecedent CP at this intermediate
 representation. Further movement (as in (109)) may carry the EC
 outside the c-command domain of the CP, with no ECP violation.

Now consider (108), which Taraldsen shows, based on coreference
 facts, is not a possible structure for the sentence (106).

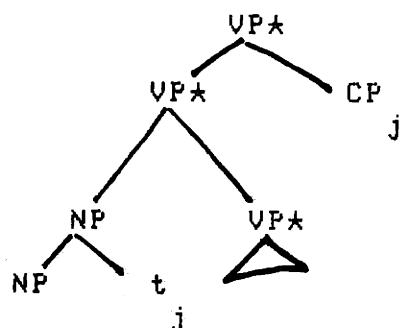
(108) [[how many people t] [do [you
 CP j k IP*

 [[think [[he [likes t]]]
 VP* VP* CP IP VP k

 [who John knows really well]]]]
 CP j

In terms of the theory of successive movement advanced in
 Chomsky (1986), this structure arises through extraposition of
 CP out of the WH-phrase when the WH-phrase is adjoined to the
 matrix VP, VP*. Consider the relevant part of the structure,
 after extraposition:

(118)



In order to assign +G to the trace, CP must c-command and
 govern it. Government is blocked in this case: NP is not

L-marked (in crucial contrast to (117)), hence constitutes a barrier. Therefore, in the structure (108), the trace t_j is not properly governed, hence the sentence violates the ECP.

This derives Taraldsen's proposal that all instances of extraposition must precede all instances of WH-movement. Extraposition creates a trace, and by my extension of Chomsky's (p.c.) proposal, the ECP is satisfied for the trace iff there is one step in the derivation where the trace is antecedent governed. Since extraposition moves the CP out of the NP dominating the trace, the NP will be a barrier for government of t by the CP unless the NP is L-marked; and it is only L-marked in its D-Structure position, before any WH-movement.

To conclude this section, we have defended the position that extraposition is a syntactic process. The grammar does not specify any ordering between the operations of extraposition and WH-movement, but independent considerations, specifically the theory of proper government, force extraposition to precede WH-movement.

5.5 Conclusion

This chapter has examined quantificational structures involving NP-traces which are not c-commanded at S-Structure by their QNP antecedents. We have proposed and defended a

constraint on movement, which not only explains the restrictions on interpretation of such quantificational structures, but also accounts for several other phenomena. We have argued that WH-movement crucially is ordered prior to QR in the LF component, and argued that extraposition is properly viewed as a syntactic adjunction operation, operating in the IS-SS mapping, but not in the SS-LF mapping.

REFERENCES

- Abney, Steven(1986) "Functional Elements and Licensing," presented at GLOW conference, Barcelona.
- (1985) "Functor Theory and Licensing: Toward an Elimination of the Base Component", ms. MIT, Cambridge, MA.
- Akmajian, A. (1971) Dissertation, MIT.
- Anderson, Mona(1984) "Prenominal Genitive NPs," The Linguistic Review 3:1.
- Aoun, Youssef and Robin Clark (1985) "On Non-Overt Operators", ms. USC.
- Aoun, Youssef and D. Lightfoot(1984) "Government and Contraction," Linguistic Inquiry 15:3.
- Baltin, M. (1981) "A Landing Site Theory of Movement Rules", LI.
- Barss, A. (1984) "Chain Binding", ms. MIT.
- Barss, A. (1984b) "Parasitic Gaps and the ECP: The role of VP-SPEC", ms. MIT.
- and Howard Lasnik (1986) "On Double Object Constructions", Linguistic Inquiry.
- Belletti, A. and L. Rizzi (1985) "Psych-verbs and Theta-Theory," talk presented at MIT-Lexicon Seminar.
- Bouchard, D. (1984) On The Content of Empty Categories, Foris.
- Bresnan, Joan(1972) Theory of Complementation in English Syntax, Doctoral Dissertation, MIT, Cambridge, MA.
- (1978) "On the Form and Functioning of Transformations", Linguistic Inquiry.
- Brody, M. (1985) "On The Complimentary Distribution of Empty Categories," Linguistic Inquiry.
- Brody, M. (1984) "On Contextual Definitions and the Role of Chains," Linguistic Inquiry 15.3, 355-380.

- Browning, M. (in progress) "Non-Overt Operators and their Antecedents", MIT.
- Burzio, L. (1981) Intransitive Verbs and Italian Auxiliaries, PhD Dissertation, MIT.
- Chomsky, Noam (1965) Aspects of the Theory of Syntax, MIT Press.
- Chomsky, Noam (1976) Essays on Form and Function
- Chomsky, N. (1977) "On WH-Movement".
- Chomsky, Noam (1980) "On Binding", Linguistic Inquiry.
- Chomsky, Noam (1981) Lectures on Government and Binding, Foris publications, Dordrecht.
- Chomsky, Noam (1982) Some Concepts and Consequences of the Theory of Government and Binding MIT press, Cambridge, MA.
- Chomsky, Noam (1985) Knowledge of Language: Its Nature, Origin and Use, Praeger, New York.
- Chomsky, Noam (1986) Barriers MIT press, Cambridge, MA.
- (1970) "Remarks on Nominalization", in Readings in English Transformational Grammar, Jacobs and Rosenbaum, eds., Ginn and Co., Waltham, MA.
- (1955) The Logical Structure of Linguistic Theory, published (1979) by University of Chicago Press.
- Chomsky, Noam and Howard Lasnik (1977) "Filters and Control", Linguistic Inquiry.
- Cinque, G. (1982)
- Cinque, G. (1984) "A'-Bound pro vs. variable".
- Chung, Sandra (1983) "The ECP and Government in Chamorro", Natural Language and Linguistic Theory, 1.2.
- Couquaux, D. (1981) "French Predication and Linguistic Theory", in Levels of Syntactic representation (May, Koster ed.), Foris.
- Cushing, S. (1982) Quantifier Meanings, North-Holland.
- Davidson, D. (1966) "The Logical Form of Action Sentences," in Davidson (1980) Essays on Actions and Events. Oxford: Clarendon Press.

- Deprez, V. (1985) "On Non-Inflected Perception Verb Complements" ms. MIT.
- (1986) "Movement Constraints and Extraposition Constructions in French", ms. MIT.
- Emonds, Joseph (1976) A Transformational Approach to English Syntax, Academic Press, NY.
- Emonds, Joseph (1985) A Unified Theory of Syntactic Categories, Foris, Dordrecht.
- Epstein, S. (in progress) Dissertation, University of Connecticut.
- Epstein (forthcoming) "On Lexical Proper Government", ms. UConn and MIT.
- (1986) "The Local Binding Condition and LF Chains", Linguistic Inquiry.
- (1984) "Quantifier-pro and the LF Representation of PRO_{arb}", LI.
- Fiengo, R. (1977) "On Trace Theory", Linguistic Inquiry.
- (1980) Surface Structure. Harvard University Press.
- and H. Lasnik (1973) "The Logical Structure of Reciprocal Sentences in English", Foundations of Language.
- Fodor, Jerry (1983) The Modularity of Mind, MIT Press, Cambridge, MA.
- Fukui, N. (forthcoming) MIT dissertation.
- George, Leland (1980) Analogical Generalizations in Natural Language Syntax Doctoral Dissertation, MIT, Cambridge, MA.
- Gillon, B. (1983) Dissertation, MIT.
- Goodall, Grant (1984) Parallel Structures in Syntax, Doctoral Dissertation, UCSD.
- Gueron, J. and R. May (1984) "Extraposition and Logical Form," Linguistic Inquiry 15.1, 1-32.
- Haik, Isabelle (1984) "Indirect Binding", Linguistic Inquiry.
- Haik, Isabelle (1985) The Syntax of Operators, Doctoral Dissertation, MIT.

Hale, Kenneth (1985) "On Nonconfigurational Structures", ms., MIT.

—— (1983) "Warlpiri and the Grammar of Non-Configurational Languages", Natural Language and Linguistic Theory Vol 1.1.

Harris, Zelig (1951), Methods in Structural Linguistics, University of Chicago Press.

Higginbotham, James (1986) "Elucidations of Meaning" ms. MIT.

—— (1985) "On Semantics", Linguistic Inquiry 16:4.

—— (1983) "LF, Binding, and Nminals" Linguistic Inquiry

—— (1981) "Reciprocal Interpretation", Journal of Linguistic Research

—— (1980) "Pronouns and Bound Variables", Linguistic Inquiry

—— and R. May (1982) "Questions, Quantifiers, and Crossing", The Linguistic Review.

Higgins, R. (1973) The Pseudo-Cleft Construction in English, dissertation, MIT (revised 1976).

Hornstein, N. (1984) Logic as Grammar, MIT Press.

Horvath, Julia (1986) Focus in the Theory of Grammar and the Syntax of Hungarian, Foris, Dordrecht.

—— (1981) Aspects of Hungarian Syntax and the Theory of Grammar, Doctoral Dissertation, UCLA.

Huang, C.-T. J. (1984) "On the Distribution and Reference of Empty Pronouns", L.I. 15.4.

—— (1982) Logical Relations in Chinese and the Theory of Grammar, unpublished MIT PhD dissertation.

Jackendoff, Ray (1977) X' Syntax: A Study of Phrase Structure, MIT Press, Cambridge, MA.

Jackendoff, Ray (1972) X' Semantics in Generative Grammar, MIT Press, Cambridge, MA.

Jacobsen, P. (1979) The Syntax of Crossing Coreference Sentences.

- and J. Neubauer (1976) "The Intervention Constraint",
Linguistic Inquiry.
- Jaeggli, Osvaldo (1980) "Remarks on to-contraction," Linguistic Inquiry 11:1.
- Jenkins, I. (1976) Existential Constructions in English, MIT Dissertation.
- Johnson, Kyle (1985) A Case for Movement, Doctoral Dissertation, MIT, Cambridge, MA.
- Kayne, Richard (1983) "Connectedness", in Kayne (1984).
- Kayne, Richard (1984) Connectedness and Binary Branching, Foris Publications, Dordrecht.
- E. Kiss, K. (1985) ms., MIT and University of Budapest.
- Koopman, Hilda and Dominique Sportiche (1982) "Variables and the Bijection Principle," The Linguistic Review 2.2, 139-160.
- Kuno, Susumu (1973) The Structure of the Japanese Language, MIT Press, Cambridge, MA.
- Langendoen, T, and E. Battistella (1981) "The Interpretation of Predicate Reflexive and Reciprocal Expressions in English", in the proceedings of NELS 12.
- Larson, Richard (1985) "Bare NP Adverbs," Linguistic Inquiry 16:4.
- Lasnik, Howard (1985) "A Note on Illicit NP Movement", LI.
- Lasnik, Howard (1976) "Remarks on Coreference", Linguistic Analysis
- (1981) "Two Recent Treatments of Disjoint Reference", Journal of Linguistic Research
- Lasnik, Howard and Joseph Kupin, (1977) "A Restrictive Theory of Transformational Grammar", Theoretical Linguistics 43.
- Levin, Beth (1983) On the Nature of Ergativity, Doctoral Dissertation, MIT, Cambridge, MA.
- Longobardi, G. (1985) "Connectedness, Scope, and C-Command", Linguistic Inquiry.
- (forthcoming) Movement, Scope, and Island Constraints.

- Manzini, M. R. (1983) "Control and Control Theory". Linguistic Inquiry.
- May, Robert (1985) Logical Form, MIT Press.
- May, Robert (1977) The Grammar of Quantification, unpublished doctoral dissertation, MIT.
- Mohanan, K.P. (1983) "Lexical and Configurational Structures" The Linguistic Review, 3.2 1983-84.
- Perlmutter, D. (1978) "Impersonal Passive and the Unaccusative Hypothesis," in Proceedings of Fourth Annual Meeting of the Berkely Linguistics Society, University of Berkely.
- Pesetsky, David (1982) Paths and Categories, Doctoral Dissertation, MIT, Cambridge, MA.
- Picallo, C. (1984) Opaque Domains, unpublished CUNY PhD dissertation.
- Pollock, J. (1985) "I1", ms. University of Paris.
- Postal, P. (1971) Crossover Phenomena.
- Riemsdijk, H. van and E. Williams (1986) Introduction to the Theory of Syntax. MIT Press.
- Reinhart, Tanya (1976) The Syntactic Domain of Anaphora, Doctoral Dissertation, MIT, Cambridge, MA.
- Rizzi, L. (1982a) Issues in Italian Syntax, Foris Publications.
- Rizzi, L. (1982b) "On Chain Formation," unpublished manuscript, Universita Della Calabria..
- Roeper, Thomas (1983) "Implicit Thematic roles in the Lexicon and Syntax" ms., University of Massachusetts, Amherst, MA.
- Ross, John (1967) Constraints on Variables in Syntax, Doctoral Dissertation, MIT, Cambridge, MA.
- Rothstein, Susan (1983) The Syntactic Forms of Predication, Doctoral Dissertation, MIT, Cambridge, MA.
- Saddy, D. (1985) "Some Properties of English Pleonastics", ms. MIT.
- Saito, Mamoru (1985) Some Asymmetries in Japanese and their Theoretical Implications, Doctoral Dissertation, MIT, Cambridge, MA.

- Sells, Peter(1984) Syntax and Semantics of Resumptive Pronouns,
Doctoral Dissertation, U.Mass, Amherst, MA.
- Speas, M. (forthcoming) Categories and Projections in Syntax,
MIT dissertation.
- Sportiche, D. (1981) Structural Invariance and Symmetry in
Syntax, dissertation, MIT.
- Stowell, Tim(1981) Origins of Phrase Structure, Doctoral
Dissertation, MIT, Cambridge, MA.
- (1982) "Subjects Across Categories", The Linguistic Review
2.
- Szabolcsi, Anna(1986) "Operator Binding and Reconstruction",
MIT Working Papers in Linguistics 8
- Taraldsen, T. (1981) "Extraposition and a Class of 'Marked'
Constructions".
- Travis, Lisa(1984) Parameters and Effects of Word Order
Variation, Doctoral Dissertation, MIT, Cambridge, MA.
- Vergnad, J-R. (1982) dissertation, Paris.
- Weisler, S. (1982) "Generalized Binding" UMOPL 8.
- Williams, E. (1977) "Discourse and Logical Form", Linguistic
Inquiry.
- Williams, E. (1980) "Predication", Linguistic Inquiry
- (1983) "Syntactic and Semantic Categories", Journal of
Linguistics and Philosophy.
- (1985) "A Reassignment of the Functions of LF", Linguistic
Inquiry.
- and H. van Riemsdijk (1981) "NP-Structure", The Linguistic
Review
- Zubizarreta, M.-L. (1982) On the Relationship of the Lexicon to
Syntax, unpublished MIT PhD dissertation.